

SOME ASPECTS OF QUANTUM THEORY

John Scales Avery

August 22, 2020

INTRODUCTION¹

I hope that this book will be of interest to students and researchers in mathematics, physics and theoretical chemistry. The first few chapters can be read with ease by anyone with a knowledge of calculus and differential equations. However, some later chapters, and most of the appendices, are more demanding..

Chapter 10 deals with resonance energy transfer and especially with the relativistic treatment of this phenomenon. My fascination with this topic dates back to my Ph.D. thesis work in the early 1960's at Imperial College, which was then a part of the University of London. I had previously been working at the laboratory of Prof. Albert Szent-Györgyi and the Marine Biological Laboratory at Woods Hole, Massachusetts. The problem on which we had been working was a quantum mechanical treatment of the primary process in photosynthesis, where a photon is absorbed, and its energy stabilized. Resonance energy transfer plays a large role in this process². When I started my Ph.D. work in London, I decided to see whether relativistic corrections made a difference³.

My calculations⁴ showed that while the usual non-relativistic treatment leads to transition probabilities that fall off as $1/R^6$, the calculated relativistic transition probabilities had a long-range component that fell off as $1/R^2$. Thus, if we imagine a very large sphere around an excited atom or molecule, the probability that the excitation energy will be transferred to one or another of the acceptors is independent of the size of the sphere! Is this a process that competes with spontaneous photon emission? Or is it an alternative way of treating the joint process of emission and absorption?⁵

Today, sixty years later, I continue to be fascinated by this question. In Chapter 10 experiments are proposed which could demonstrate that resonance energy transfer over macroscopic distance is possible. I am grateful to my son, Associate Professor James Emil Avery of the Niels Bohr Institute,

¹This book makes some use of my previously published book chapters.

²J.S. Avery, Z. Bay and A. Szent-Györgyi, *On the Energy Transfer in Biological Systems*, Proc. Nat. Acad. Sci. (US), **47**, 1742-1744, (1961)

³J.S. Avery, *Resonance energy transfer and related phenomena*, Ph.D. thesis, Imperial College of Science and Technology, University of London, (1965)

⁴J.S. Avery, *Resonance energy transfer and spontaneous photon emission*, Proc. Phys. Soc. (London) **88**, Part 1, (1966)

⁵J.S. Avery, *Use of the S-Matrix in the Relativistic Treatment of Resonance Energy Transfer*. Int. J. Quantum Chemistry, **XXV**, 79-96, (1984)

University of Copenhagen, for his help and advice. He deserves to be listed as co-author of this book. However, I don't want him to be blamed for the book's shortcomings, for example, in case the discussion section of Chapter 10 is seen to be too speculative.

Besides the usual topics, the book also focuses on some aspects of quantum theory that have been of special interest to myself and to my son, James. Among these special areas of interest is the use of 4-dimensional hyperspherical harmonics in reciprocal-space quantum chemistry. We share this interest with Professor Vincenzo Aquilanti and his group at the University of Perugia in Italy⁶. Both James and I have made numerous research visits to Perugia, where we have enjoyed both the wonderful hospitality and great mathematical knowledge of Prof. Aquilanti and his co-workers. I should mention that James has a number of important papers in which he uses hyperspherical harmonics to calculate 3-center and 4-center interelectron repulsion integrals for exponential-type basis sets (ETO's).⁷ James and I are also co-authors of several books on hyperspherical harmonics.^{8 9 10}

My interest in many-dimensional spaces brought me into contact with Professor Dudley R. Herschbach of Harvard University. I have been privileged to visit his brilliant research group many times, and to work closely with Prof. Herschbach and his colleagues for many years.^{11 12}

⁶Aquilanti, V. and Avery, J., *Sturmian expansions for quantum mechanical many-body problems and hyperspherical harmonics*, Adv. Quant. Chem., **39** 72-101, (2001)

⁷*4-center STO interelectron repulsion integrals with Coulomb Sturmians* Avery, James Emil & Avery, J. S., (2018), In : *Advances in Quantum Chemistry*. 76, p. 133-146

⁸*Generalized Sturmians and Atomic Spectra* , by J.E. Avery and J.S. Avery, World Scientific Publishing (2006)

⁹*Symmetry-Adapted Basis Sets* , by J.S. Avery, S. Rettrup and J.E. Avery, World Scientific Publishing Co, (2012)

¹⁰*Hyperspherical Harmonics and their Physical Applications*, by J.E. Avery and J.S. Avery, World Scientific Publishing Co. (2017)

¹¹Avery, J. and Herschbach, D. R., *Hyperspherical Sturmian basis functions*, Int. J. Quantum Chem., **41** 673, (1992)

¹²J. Avery, *D-Dimensional Hydrogenlike Orbitals*, in *Dimensional Scaling in Chemical Physics* , D.R. Herschbach, J. Avery, and O. Goscinski editors, Kluwer Academic Publishers, Dordrecht, Netherlands, (1992),pages 139-164



Figure 1: **Professor Dudley R. Herschbach** accepting the American Institute of Chemistry's gold medal in 2011. He shared the 1986 Nobel Prize in Chemistry for his pioneering contributions to our understanding of the mechanisms of chemical reactions.

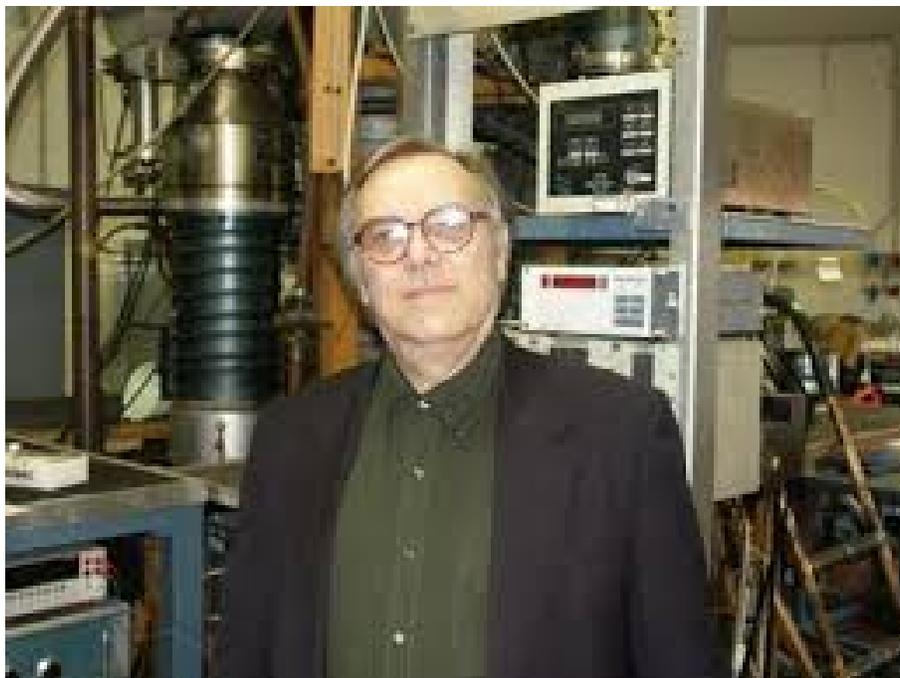


Figure 2: **Professor Vincenzo Aquilanti** (born in 1939). After working at Harvard with Dudley Herschbach, he returned to Italy, where he became the head of the chemistry department at the University of Perugia. He and his group have done pioneering theoretical and experimental work on the mechanism of chemical reactions, using molecular beam techniques. Professor Aquilanti and his group have also developed the use of 4-dimensional hyperspherical harmonics in momentum-space quantum theory, an interest which they share with my son James and myself.



Figure 3: Associate Professor James Emil Avery of the Niels Bohr Institute, University of Copenhagen. He is the author of a number of important papers that uses hyperspherical harmonics to calculate difficult 3-center and 4-center interelectron repulsion integrals for exponential-type orbitals, and is also the co-author of several books on hyperspherical harmonics and generalized Sturmians.

Contents

1	ERNEST RUTHERFORD	11
1.1	Rutherford's model of the atom	11
1.2	The Geiger-Marsden scattering experiment	13
1.3	Rutherford's model of the atom	16
1.4	Informality, enthusiasm and speed	16
1.5	Artificial transmutations of elements	20
2	NIELS BOHR	23
2.1	Christian Bohr's household	23
2.2	Planck, Einstein and Bohr	26
2.3	Atomic numbers	31
2.4	Bohr's Institute of Theoretical Physics	32
2.5	Bohr anticipates the nuclear arms race	38
3	SCHRÖDINGER'S WAVE EQUATION	43
3.1	A wave equation for matter	43
3.2	Felix Bloch's story about Schrödinger	46
3.3	Dirac's relativistic wave equation	46
3.4	Some equations	53
4	HARMONIC POLYNOMIALS AND SPHERICAL HARMONICS	55
4.1	Spherical polar coordinates	55
4.2	The Laplacian operator in spherical coordinates	56
4.3	Homogeneous and harmonic polynomials	57
4.4	Harmonic polynomials and spherical harmonics	57
4.5	An angular integration theorem	61
5	THE SCHRÖDINGER EQUATION FOR HYDROGEN	63
5.1	Separation of the equation	63
5.2	Solutions to the radial equation	64
5.3	Fock's momentum-space treatment of hydrogen	68
5.4	The Pauli exclusion principle and the periodic table	72
5.5	Valence bond theory	80

5.6	Molecular orbital theory	82
5.7	The Hartree-Fock-Roothaan equations	95
5.8	Koopmans' theorem	97
5.9	Electron creation and annihilation operators	99
5.10	Quantum chemistry and the development of computers	100
6	PERIODIC SYSTEMS	119
6.1	The discovery of X-rays	119
6.2	Bragg father and son	121
6.3	J.D. Bernal and Dorothy Crowfoot Hodgkin	125
6.4	The structure of DNA: Molecular biology	128
6.5	Direct and reciprocal lattice vectors	133
6.6	A Hückel calculation for a graphite plane	135
6.7	3-dimensional crystal lattices	136
6.8	Quantum treatment of electrons in crystals	137
6.9	The nearly-free electron approximation	137
6.10	Molecular crystals	140
6.11	Periodic boundary conditions	141
6.12	Homogeneous boundary conditions	142
6.13	Taylor series expansion of the inter-monomer interaction	144
6.14	X-ray diffraction experiments	145
7	HARMONIC OSCILLATORS	147
7.1	Normal modes	147
7.2	Molecular vibrations and rotations	150
7.3	Commutation relations	151
7.4	Phonon creation and annihilation operators	152
7.5	Collections of harmonic oscillators	154
8	THE DIRAC EQUATION	155
8.1	Lorentz invariance and 4-vectors	155
8.2	The Dirac equation for an electron in an external electromagnetic potential	157
8.3	Time-independent problems	158
8.4	The Dirac equation for an electron in the field of a nucleus	159
9	INTERACTION BETWEEN MATTER AND RADIATION	163
9.1	Lagrangian densities for fields	163
9.2	Electromagnetic potentials	165
9.3	Separation of the longitudinal and transverse potentials	170
9.4	Linear polarization	171
9.5	Spontaneous photon emission	172
9.6	Photon absorption	173
9.7	Problems with field theories	175

10	RESONANCE ENERGY TRANSFER	179
10.1	Introduction	179
10.2	Review of the Perrin-Förster theory	180
10.3	A relativistic interaction	181
10.4	The Green's function of the Helmholtz equation	182
10.5	Matrix elements	183
10.6	Transition probability at macroscopic separations	184
10.7	Comparison with spontaneous photon emission	185
10.8	The Perrin-Förster region	186
10.9	A proposed experiment	187
10.10	Discussion: Direct interparticle interaction	189
10.11	Feynman's first seminar at Princeton	190
10.12	The absorber theory of radiation	192
10.13	Karl Schwarzschild's contributions to physics	192
A	THE HISTORY OF COMPUTERS	199
A.1	Pascal and Leibniz	199
A.2	Jacquard and Babbage	201
A.3	Harvard's sequence-controlled calculator	206
A.4	The first electronic computers	207
A.5	Biosemitotics	211
A.6	The Traitorous Eight	214
A.7	Integrated circuits	217
A.8	Moore's law	218
A.9	Self-reinforcing information accumulation	221
A.10	Automation	222
A.11	Neural networks	226
B	GROUP THEORY	239
B.1	Finite Groups	239
B.2	Representations of geometrical symmetry groups	240
B.3	Similarity transformations	241
B.4	Characters and reducibility	242
B.5	The great orthogonality theorem	244
B.6	Classes	248
B.7	Projection operators	249
B.8	The regular representation	251
B.9	Classification of basis functions	252
C	NON-EUCLIDIAN GEOMETRY	257
C.1	Albert Einstein and non-Euclidian geometry	257
C.2	Family background	257
C.3	Special relativity theory	259

C.4	General relativity	260
C.5	Metric tensors	262
C.6	The Laplace-Beltrami operator	266
C.7	Geodesics	271
D	Sturmian basis sets	273
D.1	One-electron Coulomb Sturmians	273
D.2	Löwdin-orthogonalized Coulomb Sturmians	276
D.3	The Fock projection	277
D.4	Generalized Sturmians and many-particle problems	278
D.5	Use of generalized Sturmian basis sets to solve the many-particle Schrödinger equation	279
D.6	Momentum-space orthonormality relations for Sturmian basis sets	280
D.7	Sturmian expansions of d -dimensional plane waves	282
D.8	An alternative expansion of a d -dimensional plane wave	283
E	GENERALIZED STURMIANS APPLIED TO ATOMS	285
E.1	Goscinskian configurations	285
E.2	Relativistic corrections	291
E.3	The large- Z approximation: Restriction of the basis set to an \mathcal{R} -block	293
E.4	Electronic potential at the nucleus in the large- Z approximation	294
E.5	Core ionization energies	297
E.6	Advantages and disadvantages of Goscinskian configurations	299
E.7	\mathcal{R} -blocks, invariant subsets and invariant blocks	300
E.8	Invariant subsets based on subshells; Classification according to M_L and M_s	306
E.9	An atom surrounded by point charges	311
F	MOLECULAR ORBITALS BASED ON STURMIANS	319
F.1	The one-electron secular equation	319
F.2	Shibuya-Wulfman integrals and Sturmian overlap integrals evaluated in terms of hyperpherical harmonics	326
F.3	Molecular calculations using the isoenergetic configurations	332
F.4	Building $T_{\nu'\nu}^{(N)}$ and $\mathfrak{S}_{\nu'\nu}^{(N)}$ from 1-electron components	334
F.5	Interelectron repulsion integrals for molecular Sturmians from hyperspherical harmonics	335
F.6	Many-center integrals treated by Gaussian expansions (Appendix E)	339
F.7	A pilot calculation	343
F.8	Automatic generation of symmetry-adapted basis functions	346

Chapter 1

ERNEST RUTHERFORD

1.1 Rutherford's model of the atom

In 1895, the year during which Roentgen made his revolutionary discovery of X-rays, a young New Zealander named Ernest Rutherford was digging potatoes on his father's farm, when news reached him that he had won a scholarship for advanced study in England. Throwing down his spade, Rutherford said, "That's the last potato I'll dig!" He postponed his marriage plans and sailed for England, where he enrolled as a research student at Cambridge University. He began work at the Cavendish Laboratory, under the leadership of J.J. Thomson, the discoverer of the electron.

In New Zealand, Rutherford had done pioneering work on the detection of radio waves, and he probably would have continued this work at Cambridge, if it had not been for the excitement caused by the discoveries of Roentgen and Becquerel. Remembering this period of his life, Rutherford wrote:

"Few of you can realize the enormous sensation caused by the discovery of X-rays by Roentgen in 1895. It interested not only the scientific man, but also the man in the street, who was excited by the idea of seeing his own insides and his bones. Every laboratory in the world took out its old Crookes' tubes to produce X-rays, and the Cavendish was no exception."

J.J. Thomson, who was interested in studying ions (charged atoms or molecules) in gases, soon found that gaseous ions could be produced very conveniently by means of X-rays. Rutherford abandoned his research on radio waves, and joined Thomson in this work.

"When I entered the Cavendish Laboratory", Rutherford remembered later, "I began to work on the ionization of gases by means of X-rays. After reading the paper of Becquerel, I was curious to know whether the ions produced by the radiation from uranium were of the same nature as those produced by X-rays; and in particular, I was interested because Becquerel thought that his radiation was somehow intermediate between light and X-rays."

"I therefore proceeded to make a systematic examination of the radiation, and I found

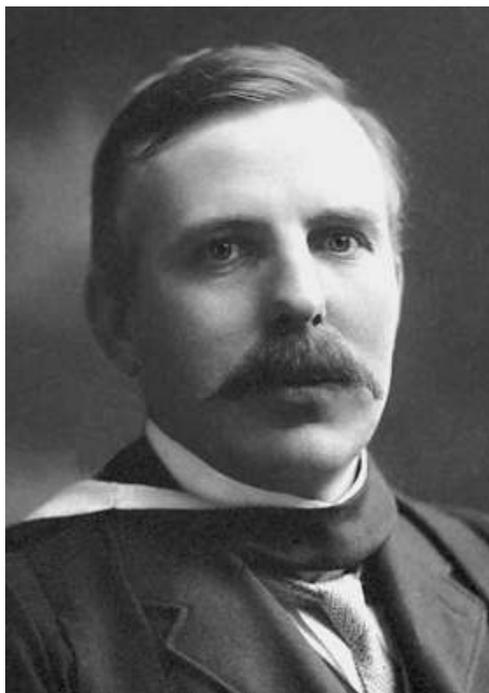


Figure 1.1: **Rutherford receiving the 1908 Nobel Prize in Chemistry.**

that it was of two types - one which produced intense ionization, and which was absorbed by a few centimeters of air, and the other, which produced less intense ionization, but was more penetrating. I called these alpha rays and beta rays respectively; and when, in 1898, Villard discovered a still more penetrating type of radiation, he called it gamma-radiation.”

Rutherford later showed that the alpha-rays were actually ionized helium atoms thrown out at enormous velocities by the decaying uranium, and that beta-rays were high-speed electrons. The gamma-rays turned out to be electromagnetic waves, just like light waves, but of extremely short wavelength.

Rutherford returned briefly to New Zealand to marry his sweetheart, Mary Newton; and then he went to Canada, where he had been offered a post as Professor of Physics at McGill University. In Canada, with the collaboration of the chemist, Frederick Soddy (1877-1956), Rutherford continued his experiments on radioactivity, and worked out a revolutionary theory of transmutation of the elements through radioactive decay.

During the middle ages, alchemists had tried to change lead and mercury into gold. Later, chemists had convinced themselves that it was impossible to change one element into another. Rutherford and Soddy now claimed that radioactive decay involves a whole series of transmutations, in which one element changes into another!

Returning to England as head of the physics department at Manchester University, Rutherford continued to experiment with alpha-particles. He was especially interested in the way they were deflected by thin metal foils. Rutherford and his assistant, Hans Geiger (1886-1945), found that most of the alpha-particles passed through a metal foil with only

a very slight deflection, of the order of one degree.

1.2 The Geiger-Marsden scattering experiment

In 1911, a young research student named Ernest Marsden joined the group, and Rutherford had to find a project for him. What happened next, in Rutherford's own words, was as follows:

“One day, Geiger came to me and said, ‘Don't you think that young Marsden, whom I'm training in radioactive methods, ought to begin a small research?’ Now I had thought that too, so I said, ‘Why not let him see if any alpha-particles can be scattered through a large angle?’ I may tell you in confidence that I did not believe that they would be, since we knew that the alpha-particle was a very fast, massive particle, with a great deal of energy; and you could show that if the scattering was due to the accumulated effect of a number of small scatterings, the chance of an alpha-particle's being scattered backward was very small.”

“Then I remember two or three days later, Geiger coming to me in great excitement and saying, ‘We have been able to get some of the alpha-particles coming backwards’. It was quite the most incredible event that has ever happened to me in my life. It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you.”

“On consideration, I realized that this scattering backwards must be the result of a single collision, and when I made calculations, I found that it was impossible to get anything of that order of magnitude unless you took a system in which the greater part of the mass of the atom was concentrated in a minute nucleus.”

“It was then that I had the idea of an atom with a minute massive center carrying a charge. I worked out mathematically what laws the scattering should obey, and found that the number of particles scattered through a given angle should be proportional to the thickness of the scattering foil, the square of the nuclear charge, and inversely proportional to the fourth power of the velocity. These deductions were later verified by Geiger and Marsden in a series of beautiful experiments.”

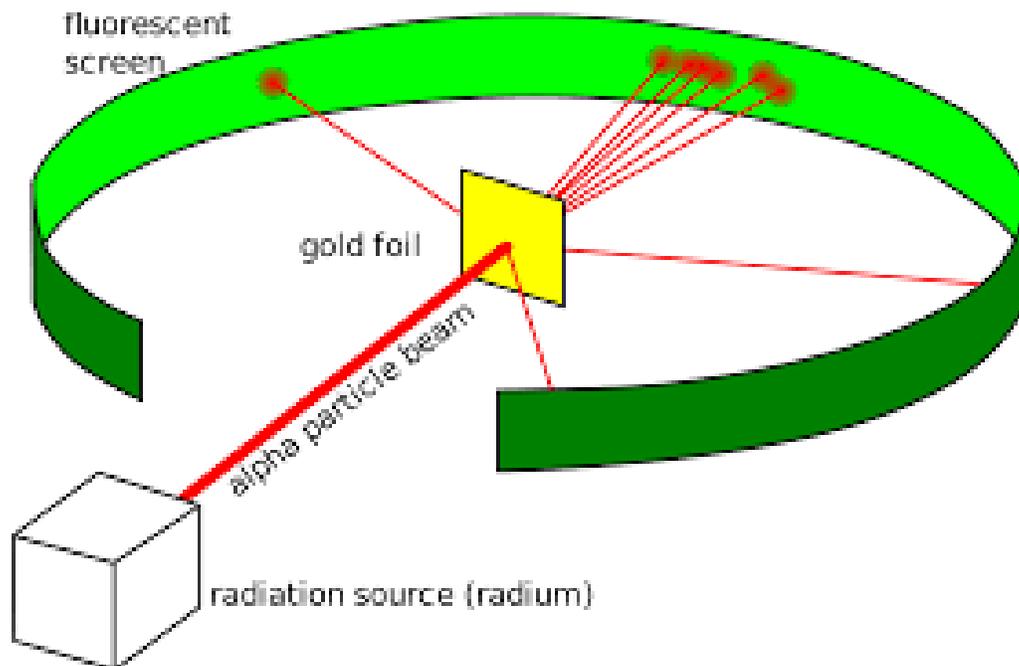


Figure 1.2: The Geiger-Marsden scattering experiment. To Rutherford's great surprise, the experiment showed that some of the alpha particles were scattered backwards. After treating the problem mathematically, Rutherford concluded that most of the mass of an atom must be concentrated in a very small, positively-charged nucleus, around which the much lighter electrons circulate in orbits.

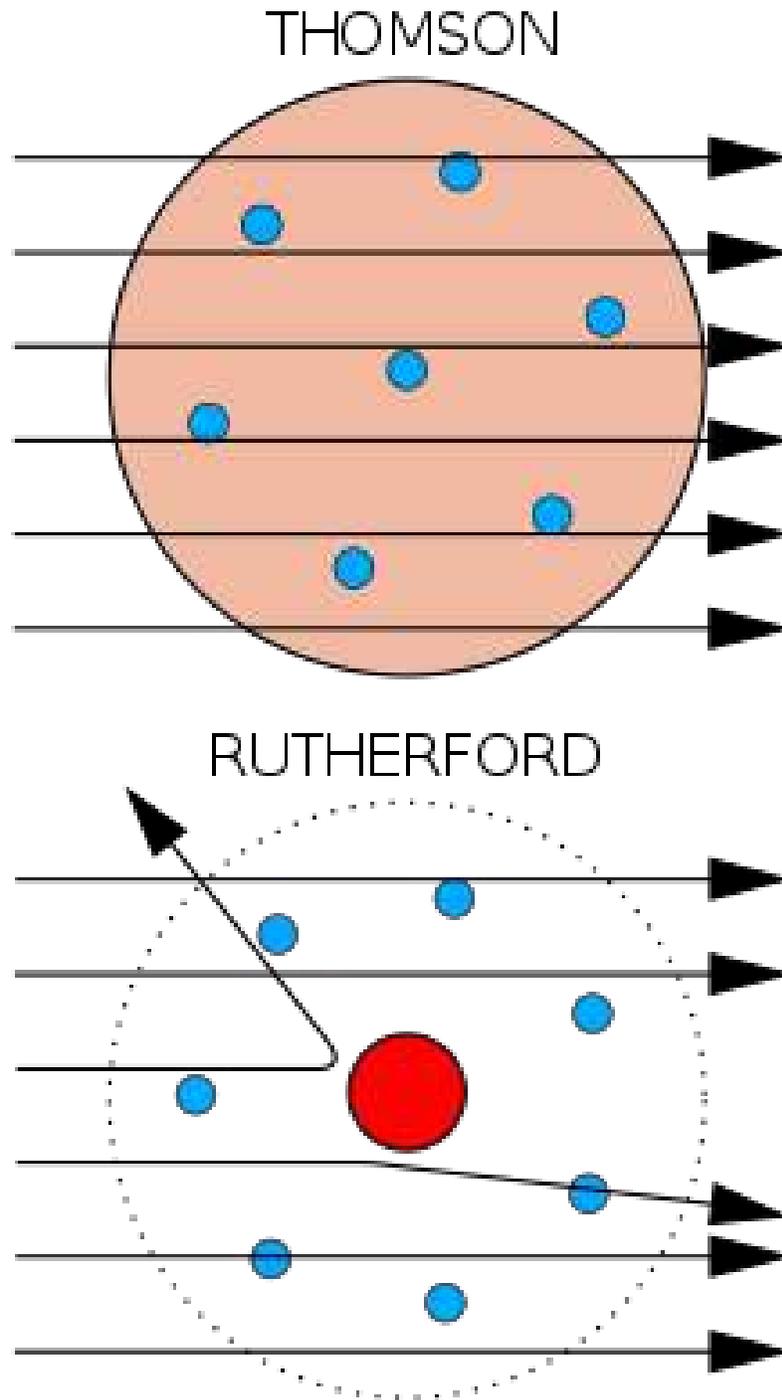


Figure 1.3: In Thomson's model of the atom, the electrons were embedded, like raisins in a pudding, in a diffuse background of positive charge. The Geiger-Marsden experiment forced Rutherford to propose a new model to account for the observed back-scattering of alpha particles.

1.3 Rutherford's model of the atom

According to the model proposed by Rutherford in 1911, every atom has an extremely tiny nucleus, which contains almost all of the mass of the atom. Around this tiny but massive nucleus, Rutherford visualized light, negatively-charged electrons circulating in orbits, like planets moving around the sun. Rutherford calculated that the diameter of the whole atom had to be several thousand times as large as the diameter of the nucleus.

1.4 Informality, enthusiasm and speed

Rutherford's model of the atom explained beautifully the scattering experiments of Geiger and Marsden, but at the same time it presented a serious difficulty: According to Maxwell's equations, the electrons circulating in their orbits around the nucleus ought to produce electromagnetic waves. It could easily be calculated that the electrons in Rutherford's atom ought to lose all their energy of motion to this radiation, and spiral in towards the nucleus. Thus, according to classical physics, Rutherford's atom could not be stable. It had to collapse.

Rutherford's laboratory was like no other in the world, except J.J. Thomson's. In fact, Rutherford had learned much about how to run a laboratory from his old teacher, Thomson. Rutherford continued Thomson's tradition of democratic informality and cheerfulness. Like Thomson, he had a gift for infecting his students with his own powerful scientific curiosity, and his enthusiastic enjoyment of research.

Thomson had also initiated a tradition for speed and ingenuity in the improvisation of experimental apparatus - the so-called "sealing-wax and string" tradition - and Rutherford continued it. Niels Bohr, after working with Rutherford, was later to continue the tradition of informality and enthusiasm at the Institute for Theoretical Physics which Bohr founded in Copenhagen in 1920.

Most scientific laboratories of the time offered a great contrast to the informality, enthusiasm, teamwork and speed of the Thomson-Rutherford-Bohr tradition. E.E. da C. Andrade, who first worked in Lenard's laboratory at Heidelberg, and later with Rutherford at Manchester, has given the following description of the contrast between the two groups:

"At the Heidelberg colloquium, Lenard took the chair, very much like a master with his class. He had the habit, if any aspect of his work was being treated by the speaker, of interrupting with, 'And who did that first?' The speaker would reply with a slight bow, 'Herr Geheimrat, you did that first', to which Lenard answered, 'Yes, I did that first'."

"At the Manchester colloquium, which met on Friday afternoons, Rutherford was, as in all his relations with the research workers, the boisterous, enthusiastic, inspiring friend, undoubtedly the leader but in close community with the led, stimulating rather than commanding, 'gingering up', to use a favorite expression of his, his team."

Although Rutherford occasionally swore at his "lads", his affection for them was very real. He had no son of his own, and he became a sort of father to the brilliant young men in

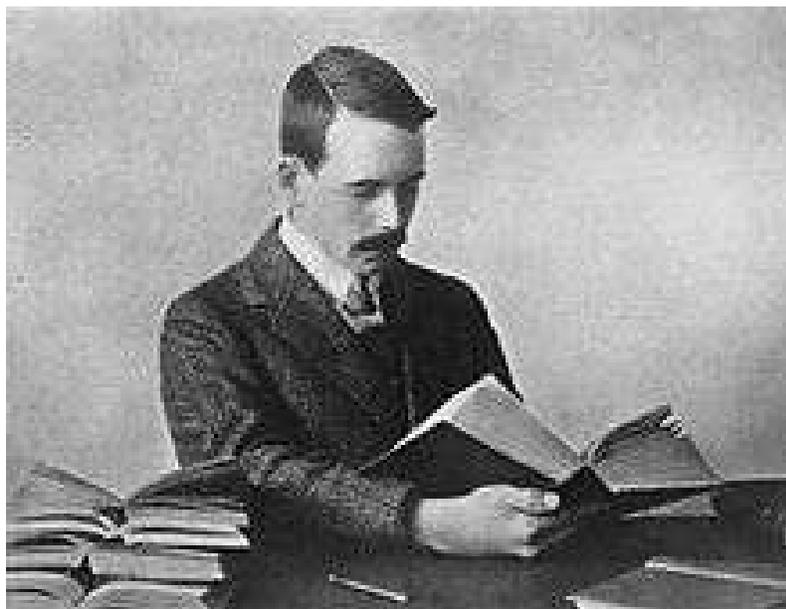


Figure 1.4: **Henry Moseley (1887-1915).**

his laboratory. Their nickname for him was “Papa”. Such was the laboratory which Harry Moseley joined in 1910. At almost the same time, Moseley’s childhood friend, Charles Darwin (the grandson of the “right” Charles Darwin), also joined Rutherford’s team.

After working on a variety of problems in radioactivity which were given to him by Rutherford, Moseley asked whether he and Charles Darwin might be allowed to study the spectra of X-rays. At first, Rutherford said no, since no one at Manchester had any experience with X-rays; “and besides”, Rutherford added with a certain amount of bias, “all science is either radioactivity or else stamp-collecting”.

However, after looking more carefully at what was being discovered about X-rays, Rutherford gave his consent. In 1912, a revolutionary discovery had been made by the Munich physicist, Max von Laue (1879-1960): It had long been known that because of its wavelike nature, white light can be broken up into the colors of the spectrum by means of a “diffraction grating” - a series of parallel lines engraved very closely together on a glass plate.

For each wavelength of light, there are certain angles at which the new wavelets produced by the lines of the diffraction grating reinforce each other instead of cancelling. The angles of reinforcement are different for each wavelength, and thus the different colors are separated by the grating.

Max von Laue’s great idea was to do the same thing with X-rays, using a crystal as a diffraction grating. The regular lines of atoms in the crystal, von Laue reasoned, would act be fine enough to fit the tiny wavelength of the X-rays, believed to be less than one ten-millionth of a centimeter.

Von Laue’s experiment, performed in 1912, had succeeded beautifully, and his new



Figure 1.5: Sir Charles Galton Darwin (1887-1962), grandson of the “right” Charles Darwin.

technique had been taken up in England by a father and son team, William Henry Bragg (1862-1942) and William Lawrence Bragg (1890-1971). The Braggs had used X-ray diffraction not only to study the spectra of X-rays, but also to study the structure of crystals. Their techniques were later to become one of the most valuable research tools available for studying molecular structure.

Having finally obtained Rutherford's permission, Moseley and Darwin threw themselves into this exciting field of study. Remembering his work with Harry Moseley, Charles Darwin later wrote:

“Working with Moseley was one of the most strenuous exercises I have ever undertaken. He was, without exception, the hardest worker I have ever known... There were two rules for his work: First, when you started to set up the apparatus for an experiment, you must not stop until it was set up. Second, when the apparatus was set up, you must not stop work until the experiment was done. Obeying these rules implied a most irregular life, sometimes with all-night sessions; and indeed, one of Moseley's expertises was the knowledge of where in Manchester one could get a meal at three in the morning.”

After about a year, Charles Darwin left the experiments to work on the theoretical aspects of X-ray diffraction. (He was later knighted for his distinguished contributions to theoretical physics.) Moseley continued the experiments alone, systematically studying the X-ray spectra of all the elements in the periodic system.

Niels Bohr had shown that the binding energies of the allowed orbits in a hydrogen atom are equal to Rydberg's constant, R (named after the distinguished Swedish spectroscopist, Johannes Robert Rydberg), divided by the square of an integral “quantum number”, n . He had also shown that for heavier elements, the constant, R , is equal to the square of the nuclear charge, Z , multiplied by a factor which is the same for all elements. The constant, R , could be observed in Moseley's studies of X-ray spectra: Since X-rays are produced when electrons are knocked out of inner orbits and outer electrons fall in to replace them, Moseley could use the Planck-Einstein relationship between frequency and energy to find the energy difference between the orbits, and Bohr's theory to relate this to R .

Moseley found complete agreement with Bohr's theory. He also found that the nuclear charge, Z , increased regularly in integral steps as he went along the rows of the periodic table: Hydrogen had $Z=1$, helium $Z=2$, lithium $Z=3$, and so on up to uranium with $Z=92$. The 92 electrons of a uranium atom made it electrically neutral, exactly balancing the charge of the nucleus. The number of electrons of an element, and hence its chemical properties, Moseley found, were determined uniquely by its nuclear charge, which Moseley called the “atomic number”.

Moseley's studies of the nuclear charges of the elements revealed that a few elements were missing. In 1922, Niels Bohr received the Nobel Prize for his quantum theory of the atom; and he was able to announce at the presentation ceremony that one of Moseley's missing elements had been found at his institute. Moseley, however, was dead. He was one of the ten million young men whose lives were needlessly thrown away in Europe's most tragic blunder - the First World War.

1.5 Artificial transmutations of elements

During the First World War, Rutherford's young men had joined the army, and he had been forced to spend most of his own time working on submarine detection. In spite of this, he had found some spare time for his scientific passion - bombarding matter with alpha particles. Helped by his laboratory steward, Kay, Rutherford had studied the effects produced when alpha particles from a radium source struck various elements. In a letter to Niels Bohr, dated December 9, 1917, Rutherford wrote:

“I have got, I think, results that will ultimately have great importance. I wish that you were here to talk matters over with me. I am detecting and counting the lighter atoms set in motion by alpha particles, and the results, I think, throw a good deal of light on the character and distribution of forces near the nucleus... I am trying to break up the atom by this method. In one case, the results look promising, but a great deal of work will be required to make sure. Kay helps me, and is now an expert counter. Best wishes for a happy Christmas.”

In July, 1919, Bohr was at last able to visit Manchester, and he heard the news directly from his old teacher: Rutherford had indeed produced artificial nuclear transmutations! In one of his experiments, an alpha-particle (i.e. a helium nucleus with nuclear charge 2) was absorbed by a nitrogen nucleus. Later, the compound nucleus threw out a proton with charge 1; and thus the bombarded nucleus gained one unit of charge. It moved up one place in the periodic table and became an isotope of oxygen.

Bohr later wrote: “I learned in detail about his great new discovery of controlled, or so-called artificial, nuclear transmutations, by which he gave birth to what he liked to call ‘modern Alchemy’, and which in the course of time, was to give rise to such tremendous consequences as regards man’s mastery of the forces of nature.”

Other scientists rushed to repeat and extend Rutherford’s experiments. Particle accelerators were built by E.O. Lawrence (1901-1958) in California, by J.H. van de Graff (1901-1967) at the Massachusetts Institute of Technology and by John Cockcroft (1897-1967), working with Rutherford at the Cavendish Laboratory. These accelerators could hurl protons at energies of a million electron-volts. Thus, protons became another type of projectile which could be used to produce nuclear transmutations.

Suggestions for further reading

1. McKown, Robin. *Giant of the Atom, Ernest Rutherford*. Julian Messner Inc, New York, (1962).
2. McLintock, A.H. *Rutherford, Sir Ernest (Baron Rutherford of Nelson, O.M., F.R.S.)*. An Encyclopaedia of New Zealand (1966 ed.). Te Ara - The Encyclopedia of New Zealand, (18 September 2007).
3. Heilbron, J. L. *Ernest Rutherford and the Explosion of Atoms*. Oxford: Oxford University Press, (2003).
4. Dacey, James. *What was Rutherford’s greatest discovery?*. Physics World, (1 September 2011).

5. Aaserud, Finn Kokowski, M. (ed.). *Niels Bohr's Mission for an 'Open World' (PDF)*. Proceedings of the 2nd ICESHS. Cracow. pp. 706-709, (2006).
6. Aaserud, Finn; Heilbron, J. L. *Love, Literature and the Quantum Atom: Niels Bohr's 1913 Trilogy Revisited*. Oxford: Oxford University Press, (2013).
7. Bohr, Niels. *Nobel Prize Lecture: The Structure of the Atom (excerpts)*. In French, A. P.; Kennedy, P. J. (eds.). *Niels Bohr: A Centenary Volume*. Cambridge, Massachusetts: Harvard University Press. pp. 91-97, (1985) [1922].
8. Cockroft, John D. *Niels Henrik David Bohr. 1885-1962*. Biographical Memoirs of Fellows of the Royal Society. 9: 36-53, (1 November 1963).
9. Favrholt, David. *Niels Bohr's Philosophical Background*. Copenhagen: Munksgaard, (1992).
10. Faye, January. *Niels Bohr: His Heritage and Legacy*. Dordrecht: Kluwer Academic Publishers, (1991).
11. Heilbron, John L. *Bohr's First Theories of the Atom*. In French, A. P.; Kennedy, P. J. (eds.). *Niels Bohr: A Centenary Volume*. Cambridge, Massachusetts: Harvard University Press. pp. 33-49, (1985).
12. Heisenberg, Elisabeth. *Inner Exile: Recollections of a Life With Werner Heisenberg*. Boston: Birkhäuser, (1984).
13. Hilberg, Raul. *The Destruction of the European Jews. 2*. New Haven, Connecticut: Yale University Press, (1961).
14. Jammer, Max. *The Conceptual Development of Quantum Mechanics*. Los Angeles: Tomash Publishers, (1989).
15. Jones, R.V. *Most Secret War*. London: Hamilton, (1978).
16. Kieler, JÅrgen. *Resistance Fighter: A Personal History of the Danish Resistance*. Translated from the Danish by Eric Dickens. Jerusalem: Gefen Publishing House, (2007).
17. Kragh, Helge. *The Theory of the Periodic System*. In French, A. P.; Kennedy, P. J. (eds.). *Niels Bohr: A Centenary Volume*. Cambridge, Massachusetts: Harvard University Press. pp. 50-67, (1985).
18. Kragh, Helge. *Niels Bohr and the quantum atom: the Bohr model of atomic structure, 1913-1925*. Oxford: Oxford University Press, (2012).
19. Medawar, Jean; Pyke, David. *Hitler's Gift: The True Story of the Scientists Expelled by the Nazi Regime*. New York: Arcade Publishing, (2001).
20. Pais, Abraham. *Niels Bohr's Times, In Physics, Philosophy and Polity*. Oxford: Clarendon Press, (1991).
21. Powers, Thomas. *Heisenberg's War: The Secret History of the German Bomb*. New York: Knopf, (1993).
22. Rhodes, Richard. *The Making of the Atomic Bomb*. New York: Simon and Schuster, (1986).
23. Richardson, W. Mark; Wildman, Wesley J., eds. *Religion and Science: History, Method, Dialogue*. London, New York: Routledge, (1996).
24. Rife, Patricia. *Lise Meitner and the Dawn of the Nuclear Age*. Boston: Birkhäuser, (1999).

25. Rozental, Stefan. *Niels Bohr: His Life and Work as Seen by his Friends and Colleagues*. Amsterdam: North-Holland, (1967).
26. Stadtler, Bea; Morrison, David Beal; Martin, David Stone. *The Holocaust: A History of Courage and Resistance*. West Orange, New Jersey: Behrman House, (1995).

Chapter 2

NIELS BOHR

2.1 Christian Bohr's household

Christian Bohr (1855-1911) was appointed professor of physiology at the University of Copenhagen in 1886. In this position, he made a number of important discoveries connected with respiration in mammals, including what is now known as the “Bohr effect”, i.e. the tendency of high concentrations of CO_2 and of H^+ ions to increase the efficiency of hemoglobin in releasing oxygen. Christian Bohr was also the teacher of August Krogh, who later won a Nobel Prize in Medicine and Physiology.

Christian Bohr's wife, Ellen Adler Bohr, belonged to a wealthy Jewish banking family, and Niels Bohr was born in the impressive multi-story Adler mansion that still stands today near one of Copenhagen's canals opposite the Danish Parliament. During the time that Niels and Harold Bohr were growing up, this house was the meeting place for many of Copenhagen's leading intellectuals, and the boys were allowed to attend meetings where scientific and philosophical questions were debated. This upbringing contributed to the fact that both Niels and Harold later became famous in their respective fields, physics and mathematics.

The Bohr family has produced outstanding scientists for four generations. Besides Christian, Niels and Harold Bohr, there is also Niels' son Aage, who shared a Nobel Prize in Physics for his work on the excited states of nuclei. Aage's sons, Vilhelm and Thomas, are also outstanding scientists.

Having been brought up in a highly intellectual household, Niels Bohr's scientific abilities developed early. In 1905, when Niels was 20, a gold medal competition was announced by the Royal Danish Society of Sciences and Letters. The challenge was to investigate a method for determining the surface tension of liquids. The method had been proposed earlier by Lord Raleigh, and it involved measuring the frequency of oscillations on the surface of a water jet. After working in his father's laboratory, making his own glassware to produce elliptical water jets, and presenting his results together with a mathematical analysis, Niels Bohr won the gold medal.



Figure 2.1: **Christian Bohr (1855-1911), the father of Niels and Harold Bohr. He was Professor of Physiology at the University of Copenhagen.**



Figure 2.2: Niels Bohr (1885-1952) as a young man.

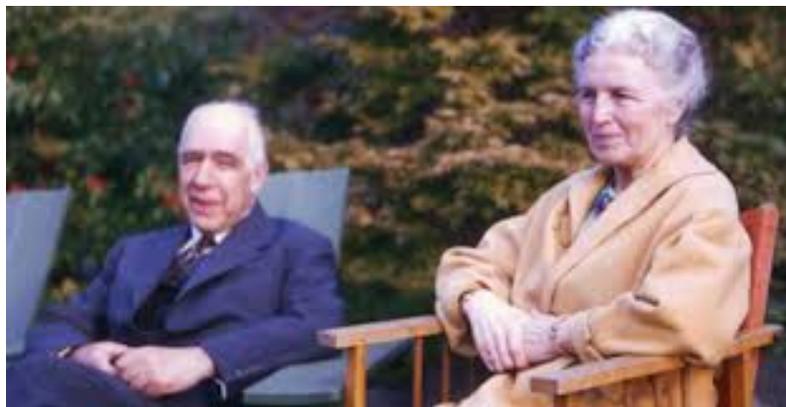


Figure 2.3: Niels Bohr and his wife, Margrethe.

2.2 Planck, Einstein and Bohr

According to the model proposed by Rutherford in 1911, every atom has an extremely tiny nucleus, which contains almost all of the mass of the atom. Around this tiny but massive nucleus, Rutherford visualized light, negatively-charged electrons circulating in orbits, like planets moving around the sun. Rutherford calculated that the diameter of the whole atom had to be several thousand times as large as the diameter of the nucleus.

Rutherford's model of the atom explained beautifully the scattering experiments of Geiger and Marsden, but at the same time it presented a serious difficulty: According to Maxwell's equations, the electrons circulating in their orbits around the nucleus ought to produce electromagnetic waves. It could easily be calculated that the electrons in Rutherford's atom ought to lose all their energy of motion to this radiation, and spiral in towards the nucleus. Thus, according to classical physics, Rutherford's atom could not be stable. It had to collapse.

Niels Bohr became aware of this paradox when he worked at Rutherford's Manchester laboratory during the years 1911-1913. Bohr was not at all surprised by the failure of classical concepts when applied to Rutherford's nuclear atom. Since he had been educated in Denmark, he was more familiar with the work of German physicists than were his English colleagues at Manchester. In particular, Bohr had studied the work of Max Planck (1858-1947) and Albert Einstein (1879-1955).

Just before the turn of the century, the German physicist, Max Planck, had been studying theoretically the electromagnetic radiation coming from a small hole in an oven. The hole radiated as though it were an ideally black body. This "black body radiation" was very puzzling to the physicists of the time, since classical physics failed to explain the frequency distribution of the radiation and its dependence on the temperature of the oven.

In 1901, Max Planck had discovered a formula which fitted beautifully with the experimental measurements of the frequency distribution of black body radiation; but in order to derive his formula, he had been forced to make a radical assumption which broke away completely from the concepts of classical physics.

Planck had been forced to assume that light (or, more generally, electromagnetic radiation of any kind) can only be emitted or absorbed in amounts of energy which Planck called "quanta". The amount of energy in each of these "quanta" was equal to the frequency of the light multiplied by a constant, h , which came to be known as "Planck's constant".

This was indeed a strange assumption! It seemed to have been pulled out of thin air; and it had no relation whatever to anything that had been discovered previously in physics. The only possible justification for Planck's quantum hypothesis was the brilliant success of his formula in explaining the puzzling frequency distribution of the black body radiation. Planck himself was greatly worried by his own radical break with classical concepts, and he spent many years trying unsuccessfully to relate his quantum hypothesis to classical physics.

In 1905, Albert Einstein published a paper in the *Annalen der Physik* in which he applied Planck's quantum hypothesis to the photoelectric effect. (At that time, Einstein was 25 years old, completely unknown, and working as a clerk at the Swiss Patent Office.)



Figure 2.4: Niels Bohr and Albert Einstein in a photo by Paul Ehrenfest. Public domain, Wikimedia Commons

The photoelectric effect was another puzzling phenomenon which could not in any way be explained by classical physics. The German physicist Lenard had discovered in 1903 that light with a frequency above a certain threshold could knock electrons out of the surface of a metal; but below the threshold frequency, nothing at all happened, no matter how long the light was allowed to shine.

Using Planck's quantum hypothesis, Einstein offered the following explanation for the photoelectric effect: A certain minimum energy was needed to overcome the attractive forces which bound the electron to the metal surface. This energy was equal to the threshold frequency multiplied by Planck's constant. Light with a frequency equal to or higher than the threshold frequency could tear an electron out of the metal; but the quantum of energy supplied by light of a lower frequency was insufficient to overcome the attractive forces.

Einstein later used Planck's quantum formula to explain the low-temperature behavior of the specific heats of crystals, another puzzling phenomenon which defied explanation by classical physics. These contributions by Einstein were important, since without this supporting evidence it could be maintained that Planck's quantum hypothesis was an *ad hoc* assumption, introduced for the sole purpose of explaining black body radiation.

As a student, Niels Bohr had been profoundly impressed by the radical ideas of Planck and Einstein. In 1912, as he worked with Rutherford at Manchester, Bohr became convinced that the problem of saving Rutherford's atom from collapse could only be solved by means of Planck's quantum hypothesis.

Returning to Copenhagen, Bohr continued to struggle with the problem. In 1913, he found the solution: The electrons orbiting around the nucleus of an atom had "angular momentum". Assuming circular orbits, the angular momentum was given by the product of

the mass and velocity of the electron, multiplied by the radius of the orbit. Bohr introduced a quantum hypothesis similar to that of Planck: He assumed that the angular momentum of an electron $m_e v r$ in an allowed orbit, (multiplied by 2π), had to be equal to an integral multiple of Planck's constant h . The lowest value of the integer, $n=1$, corresponded to the lowest allowed orbit. Thus, in Bohr's model, the collapse of Rutherford's atom was avoided.

$$\begin{aligned} m_e v r &= n \hbar & n &= 1, 2, 3, 4, \dots \\ \hbar &\equiv \frac{h}{2\pi} \end{aligned} \quad (2.1)$$

Here m_e is the mass of the electron, v is its velocity and r is the radius of its classical orbit around the nucleus, so that $m_e v r$ is the classical expression for the angular momentum of an electron in a circular orbit. In Bohr's quantization of angular momentum, h represents Planck's constant, which had been introduced by Max Planck to explain the frequency distribution of black body radiation.

Bohr calculated that the binding energies of the various allowed electron orbits in a hydrogen atom should be a constant divided by the square of the integer n ; and he calculated the value of the constant to be 13.5 electron-Volts. This value fit exactly the observed ionization energy of hydrogen. After talking with the Danish spectroscopist, H.M. Hansen, Bohr realized with joy that by combining his formula for the allowed orbital energies with the Planck-Einstein formula relating energy to frequency, he could explain the mysterious line spectrum of hydrogen.

$$\begin{aligned} E &= -\frac{1}{2} m_e v^2 = -\frac{1}{2} \frac{Z e^2}{r} \\ r_n &= \frac{n^2 \hbar^2}{Z e^2 m_e} \\ E_n &= -\frac{Z^2 e^2 m_e}{2 \hbar^2 n^2} \\ \Delta E &= \frac{Z^2 e^2 m_e}{2 \hbar^2} \left(\frac{1}{n_1} - \frac{1}{n_2} \right) \end{aligned} \quad (2.2)$$

When Niels Bohr published all this in 1913, his paper produced agonized cries of "foul!" from the older generation of physicists. When Lord Rayleigh's son asked him if he had seen Bohr's paper, Rayleigh replied: "Yes, I have looked at it; but I saw that it was of no use to me. I do not say that discoveries may not be made in that sort of way. I think very likely they may be. But it does not suit me." However, as more and more atomic spectra and properties were explained by extensions of Niels Bohr's theories, it became clear that Planck, Einstein and Bohr had uncovered a whole new stratum of phenomena, previously unsuspected, but of deep and fundamental importance.

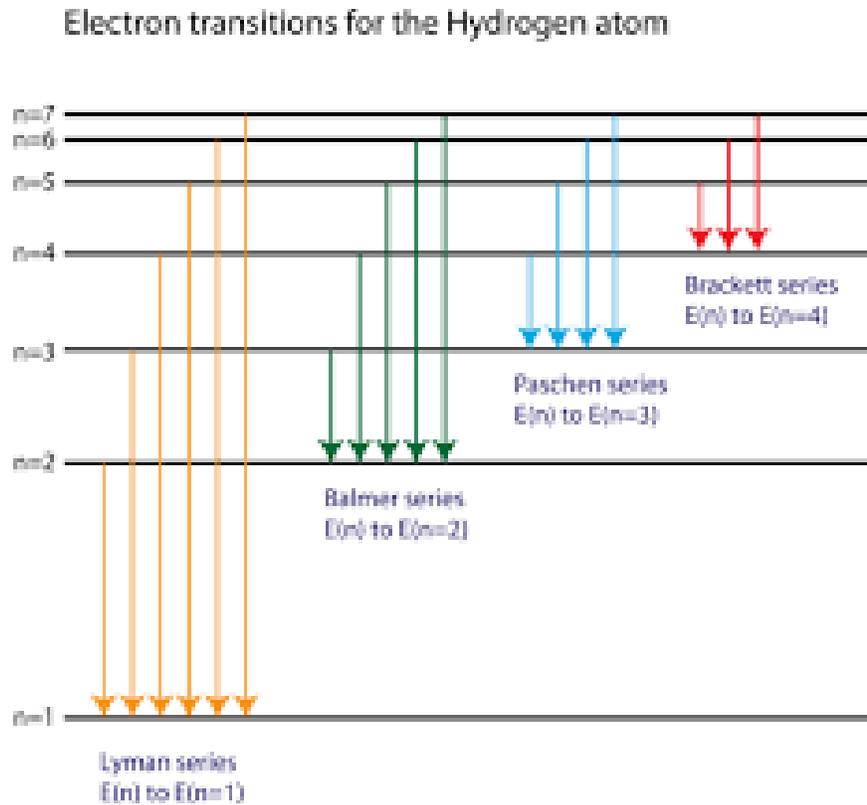


Figure 2.5: Line spectra of hydrogen. The figure show the emission spectrum energies associated with transitions between allowed energies. These transition energies fitted perfectly with Niels Bohr's calculations. As more and more atomic spectra and properties were explained by extensions of Niels Bohr's theories, it became clear that Planck, Einstein and Bohr had uncovered a whole new stratum of phenomena, previously unsuspected, but of deep and fundamental importance.



Figure 2.6: Another photo of Bohr and Einstein by Ehrenfest.

2.3 Atomic numbers

Bohr's atomic theory soon received strong support from the experiments of one of the brightest of Rutherford's bright young men - Henry Moseley (1887-1915). Moseley came from a distinguished scientific family. Not only his father, but also both his grandfathers, had been elected to the Royal Society. After studying at Oxford, where his father had once been a professor, Moseley found it difficult to decide where to do his postgraduate work. Two laboratories attracted him: the great J.J. Thomson's Cavendish Laboratory at Cambridge, and Rutherford's laboratory at Manchester. Finally, he decided on Manchester, because of the revolutionary discoveries of Rutherford, who two years earlier had won the 1908 Nobel Prize for Chemistry.

Rutherford's laboratory was like no other in the world, except J.J. Thomson's. In fact, Rutherford had learned much about how to run a laboratory from his old teacher, Thomson. Rutherford continued Thomson's tradition of democratic informality and cheerfulness. Like Thomson, he had a gift for infecting his students with his own powerful scientific curiosity, and his enthusiastic enjoyment of research.

Thomson had also initiated a tradition for speed and ingenuity in the improvisation of experimental apparatus - the so-called "sealing-wax and string" tradition - and Rutherford continued it. Niels Bohr, after working with Rutherford, was later to continue the tradition of informality and enthusiasm at the Institute for Theoretical Physics which Bohr founded in Copenhagen in 1920.

Niels Bohr had shown that the binding energies of the allowed orbits in a hydrogen atom are equal to Rydberg's constant, R (named after the distinguished Swedish spectroscopist, Johannes Robert Rydberg), divided by the square of an integral "quantum number", n . He had also shown that for heavier elements, the constant, R , is equal to the square of the nuclear charge, Z , multiplied by a factor which is the same for all elements. The constant, R , could be observed in Moseley's studies of X-ray spectra: Since X-rays are produced when electrons are knocked out of inner orbits and outer electrons fall in to replace them, Moseley could use the Planck-Einstein relationship between frequency and energy to find the energy difference between the orbits, and Bohr's theory to relate this to R .

Moseley found complete agreement with Bohr's theory. He also found that the nuclear charge, Z , increased regularly in integral steps as he went along the rows of the periodic table: Hydrogen had $Z=1$, helium $Z=2$, lithium $Z=3$, and so on up to uranium with $Z=92$. The 92 electrons of a uranium atom made it electrically neutral, exactly balancing the charge of the nucleus. The number of electrons of an element, and hence its chemical properties, Moseley found, were determined uniquely by its nuclear charge, which Moseley called the "atomic number".

Moseley's studies of the nuclear charges of the elements revealed that a few elements were missing. In 1922, Niels Bohr received the Nobel Prize for his quantum theory of the atom; and he was able to announce at the presentation ceremony that one of Moseley's missing elements had been found at his institute. Moseley, however, was dead. He was one of the ten million young men whose lives were needlessly thrown away in Europe's most tragic blunder - the First World War.



Figure 2.7: Niels Bohr with his sons at their summer house in Tisvilde.

2.4 Bohr's Institute of Theoretical Physics

In 1916, Niels Bohr was appointed professor of theoretical physics at the University of Copenhagen, a post that had been created especially for him. The following year, in 1917, he started to raise money for the construction of a new institute in which his new department could be housed. The project received large contributions from the Danish government and the Carlsberg Foundation, and from wealthy Danish businessmen. Bohr himself designed the building, which opened in 1920.

During the period when Hitler's Nazi party was coming to power in Germany, Bohr was able to offer a refuge at his Institute of Theoretical Physics to many important physicists who could no longer remain in Germany. Those to whom Bohr gave refuge included Guido Beck, Felix Bloch, James Franck, George de Hevesy, Otto Frisch, Hilde Levi, Lise Meitner, George Placzek, Eugene Rabinowitch, Stefan Rozental, Erich Ernst Schneider, Edward Teller, Arthur von Hippel and Victor Weisskopf. Because of this, because of Bohr's dynamic and inspiring presence, and because he was able to continue the tradition of informality, enthusiasm and speed which characterized J.J. Thomson's Cavendish and Rutherford's Manchester laboratories, Bohr's institute became the world's most important center for theoretical physics, especially during the 1930's.

Bohr was tirelessly energetic. He liked to discuss his ideas in dialogue with one of the bright young men at his institute, putting forward an idea, and expecting a counter-argument to be thrown back. It was like a game of ping-pong. In this way, a new idea could be tested by exploring all of its consequences.

When a new scientist arrived at his institute, Bohr liked to invite the newcomer to accompany him on a two-day walking tour to his summer house in Tisvilde, about 50

kilometers north of Copenhagen. In his autobiographical book “Physics and Beyond”, Werner Heisenberg describes such a two-man tour together with Bohr. This custom allowed Bohr to get to know both the personality and the potential scientific contributions of the new arrival. It also allowed Bohr to get some exercise and to keep himself in good physical condition.

The Nazi occupation of Denmark

On 9 April, 1940, Nazi Germany invaded and quickly occupied Denmark. The Germans explained that their purpose was “to protect Denmark from a British invasion”. During the first three years of occupation the Germans allowed the Danish government, police force and army to exist. However, in 1943, after extensive sabotage actions by the Danish resistance movement, the German policy changed and became much harsher.

Shortly after this sudden change, the Danes became aware that their Jewish population was in danger of being arrested and sent to concentration camps. Luckily it was possible for Danish citizens to organize a secret rescue operation, in which almost all members of Denmark’s Jewish community escaped to Sweden in small boats. Among them were Niels Bohr and his son Aage.

Niels and Aage Bohr fly to England

After some time in Sweden, where he helped to organize aid for Jewish refugees from Denmark, Niels Bohr and his son Aage flew to England in a small aircraft. It flew at a high altitude in order to avoid observation. Niels Bohr’s oxygen mask did not fit properly because of his unusually large head, and he became unconscious. Luckily this was noticed before anything very serious happened.



Figure 2.8: The Institute of Theoretical Physics, established by Niels Bohr at the University of Copenhagen. Today it is known as the Niels Bohr Institute



Figure 2.9: Another view of the Niels Bohr Institute.



Figure 2.10: Aage Bohr (1922-2008), one of Niels and Margrethe Bohr's sons. Together with Ben Mottelson, he was awarded the 1975 Nobel Prize in Physics for developing a successful theory of the excited states of nuclei.



Figure 2.11: Ben Roy Mottelson (born in 1926), who shared the 1975 Nobel Prize in Physics with Aage Bohr. Although now very old, he still comes in to work at the Niels Bohr Institute.



Figure 2.12: George de Hevesy (1885-1966), co-discoverer of the element Haffnium, and pioneer of the use of radioactive tracer elements in biochemistry. He received the Nobel Prize in Chemistry in 1943 for work which he performed at the Niels Bohr Institute. The name “Haffnium” is derived from the Latin name for Copenhagen.

2.5 Bohr anticipates the nuclear arms race

After escaping from Denmark to Sweden in a fishing boat in 1943, Niels Bohr and his son Aage flew to England, and then to Los Alamos in the United States, where work on a nuclear bomb was in progress. In 1943, a special intelligence unit called “Aslos” had been set up to determine how far German work on a nuclear bomb had progressed. Advanced units, entering mainland Europe after D-Day, interviewed captured German scientists and found that the German program had never come near to producing a nuclear bomb.

The news that the Germans would not produce atomic bombs was classified as a secret. Nevertheless, it passed through the grapevine to the scientists working on the atomic bomb project in America; and it reversed their attitude to the project. Until then, they had been worried that Hitler would be the first to produce nuclear weapons. In 1944, they began to worry instead about what the American government might do if it came to possess such weapons.

At Los Alamos, Niels Bohr became the center of discussion and worry about the ethics of continued work on the bomb project. He was then 59 years old; and he was universally respected both for his pioneering work in atomic physics, and for his outstandingly good character.

Bohr was extremely worried because he foresaw a postwar nuclear arms race unless international control of atomic energy could be established. Consequently, as a spokesman for the younger atomic scientists, he approached both Roosevelt and Churchill to urge them to consider means by which international control might be established.

Roosevelt, too, was worried about the prospect of a postwar nuclear armaments race; and he was very sympathetic towards Bohr’s proposals for international control. He suggested that Bohr travel to England and contact Churchill, to obtain his point of view.

Churchill was desperately busy, and basically unsympathetic towards Bohr’s proposals; but on May 16, 1944, he agreed to a half-hour interview with the scientist. The meeting was a complete failure. Churchill and his scientific advisor, Lord Cherwell, spent most of the time talking with each other, so that Bohr had almost no time to present his ideas.

Although he could be very persuasive in long conversations, Bohr was unable to present his thoughts briefly. He wrote and spoke in a discursive style, similar to that of Henry James. Each of his long, convoluted sentences was heavily weighted with qualifications and dependent clauses. At one point in the conversation, Churchill turned to Lord Cherwell and asked: “What’s he talking about, physics or politics?”

Bohr’s low, almost whispering, way of speaking irritated Churchill. Furthermore, the two men were completely opposed in their views: Bohr was urging openness in approaching the Russians, with a view to establishing international control of nuclear weapons. Churchill, a defender of the old imperial order, was concerned mainly with maintaining British and American military supremacy.

After the interview, Churchill became worried that Bohr would give away “atomic secrets” to the Russians; and he even suggested that Bohr be arrested. However, Lord Cherwell explained to the Prime Minister that the possibility of making atomic bombs, as well as the basic means of doing so, had been common knowledge in the international

scientific community ever since 1939.

After his disastrous interview with Churchill, Niels Bohr carefully prepared a memorandum to be presented to President Roosevelt. Realizing how much depended on its success or failure, Bohr wrote and rewrote the memorandum, sweating in the heat of Washington's summer weather. Aage Bohr, who acted as his father's secretary, typed the memorandum over and over, following his father's many changes of mind.

Finally, in July, 1944, Bohr's memorandum was presented to Roosevelt. It contains the following passages:

"...Quite apart from the question of how soon the weapon will be ready for use, and what role it will play in the present war, this situation raises a number of problems which call for urgent attention. Unless, indeed, some agreement about the control of the new and active materials can be obtained in due time, any temporary advantage, however great, may be outweighed by a perpetual menace to human society."

"Ever since the possibilities of releasing atomic energy on a vast scale came into sight, much thought has naturally been given to the question of control; but the further the exploration of the scientific problems is proceeding, the clearer it becomes that no kind of customary measures will suffice for this purpose, and that the terrifying prospect of a future competition between nations about a weapon of such formidable character can only be avoided by a universal agreement in true confidence..."

Roosevelt was sympathetic with the ideas expressed in this memorandum. In an interview with Bohr, he expressed his broad agreement with the idea of international control of atomic energy. Unfortunately, the President had only a few months left to live.

Roosevelt's successor, Harry Truman, had not known about the existence of nuclear weapons before taking office, and he was cautiously feeling his way. Meanwhile, General Leslie Groves, the military commander of the Los Alamos project, was very anxious to get credit for ending World War II, rather than being blamed for wasting billions of dollars of the taxpayers' money. It was easy for Groves to convince Truman to give the order to drop bombs on Hiroshima and Nagasaki. Thus Bohr's efforts to prevent this tragedy failed, and the postwar nuclear arms race which he anticipated still casts a dark shadow over the future of human civilization and the biosphere.

Suggestions for further reading

1. McKown, Robin. *Giant of the Atom, Ernest Rutherford*. Julian Messner Inc, New York, (1962).
2. McLintock, A.H. *Rutherford, Sir Ernest (Baron Rutherford of Nelson, O.M., F.R.S.)*. An Encyclopaedia of New Zealand (1966 ed.). Te Ara - The Encyclopedia of New Zealand, (18 September 2007).
3. Heilbron, J. L. *Ernest Rutherford and the Explosion of Atoms*. Oxford: Oxford University Press, (2003).
4. Dacey, James. *What was Rutherford's greatest discovery?*. Physics World, (1 September 2011).

5. Aaserud, Finn Kokowski, M. (ed.). *Niels Bohr's Mission for an 'Open World' (PDF)*. Proceedings of the 2nd ICESHS. Cracow. pp. 706-709, (2006).
6. Aaserud, Finn; Heilbron, J. L. *Love, Literature and the Quantum Atom: Niels Bohr's 1913 Trilogy Revisited*. Oxford: Oxford University Press, (2013).
7. Bohr, Niels. *Nobel Prize Lecture: The Structure of the Atom (excerpts)*. In French, A. P.; Kennedy, P. J. (eds.). *Niels Bohr: A Centenary Volume*. Cambridge, Massachusetts: Harvard University Press. pp. 91-97, (1985) [1922].
8. Cockroft, John D. *Niels Henrik David Bohr. 1885-1962*. Biographical Memoirs of Fellows of the Royal Society. 9: 36-53, (1 November 1963).
9. Favrholt, David. *Niels Bohr's Philosophical Background*. Copenhagen: Munksgaard, (1992).
10. Faye, January. *Niels Bohr: His Heritage and Legacy*. Dordrecht: Kluwer Academic Publishers, (1991).
11. Heilbron, John L. *Bohr's First Theories of the Atom*. In French, A. P.; Kennedy, P. J. (eds.). *Niels Bohr: A Centenary Volume*. Cambridge, Massachusetts: Harvard University Press. pp. 33-49, (1985).
12. Heisenberg, Elisabeth. *Inner Exile: Recollections of a Life With Werner Heisenberg*. Boston: Birkhäuser, (1984).
13. Hilberg, Raul. *The Destruction of the European Jews. 2*. New Haven, Connecticut: Yale University Press, (1961).
14. Jammer, Max. *The Conceptual Development of Quantum Mechanics*. Los Angeles: Tomash Publishers, (1989).
15. Jones, R.V. *Most Secret War*. London: Hamilton, (1978).
16. Kieler, JÅrgen. *Resistance Fighter: A Personal History of the Danish Resistance*. Translated from the Danish by Eric Dickens. Jerusalem: Gefen Publishing House, (2007).
17. Kragh, Helge. *The Theory of the Periodic System*. In French, A. P.; Kennedy, P. J. (eds.). *Niels Bohr: A Centenary Volume*. Cambridge, Massachusetts: Harvard University Press. pp. 50-67, (1985).
18. Kragh, Helge. *Niels Bohr and the quantum atom: the Bohr model of atomic structure, 1913-1925*. Oxford: Oxford University Press, (2012).
19. Medawar, Jean; Pyke, David. *Hitler's Gift: The True Story of the Scientists Expelled by the Nazi Regime*. New York: Arcade Publishing, (2001).
20. Pais, Abraham. *Niels Bohr's Times, In Physics, Philosophy and Polity*. Oxford: Clarendon Press, (1991).
21. Powers, Thomas. *Heisenberg's War: The Secret History of the German Bomb*. New York: Knopf, (1993).
22. Rhodes, Richard. *The Making of the Atomic Bomb*. New York: Simon and Schuster, (1986).
23. Richardson, W. Mark; Wildman, Wesley J., eds. *Religion and Science: History, Method, Dialogue*. London, New York: Routledge, (1996).
24. Rife, Patricia. *Lise Meitner and the Dawn of the Nuclear Age*. Boston: Birkhäuser, (1999).

25. Rozental, Stefan. *Niels Bohr: His Life and Work as Seen by his Friends and Colleagues*. Amsterdam: North-Holland, (1967).
26. Stadtler, Bea; Morrison, David Beal; Martin, David Stone. *The Holocaust: A History of Courage and Resistance*. West Orange, New Jersey: Behrman House, (1995).

Chapter 3

SCHRÖDINGER'S WAVE EQUATION

3.1 A wave equation for matter

In 1926, the difficulties surrounding the “old quantum theory” of Max Planck, Albert Einstein and Niels Bohr were suddenly solved, and its true meaning was understood. Two years earlier, a French aristocrat, Louis de Broglie, writing his doctoral dissertation at the Sorbonne in Paris, had proposed that very small particles, such as electrons, might exhibit wave-like properties. The ground state and higher excited states of the electron in Bohr’s model of the hydrogen atom would then be closely analogous to the fundamental tone and higher overtones of a violin string.

Almost the only person to take de Broglie’s proposal seriously was Albert Einstein, who mentioned it in one of his papers. Because of Einstein’s interest, de Broglie’s matter-waves came to the attention of other physicists. The Austrian theoretician, Erwin Schrödinger, working at Zürich, searched for the underlying wave equation which de Broglie’s matter-waves obeyed.

Schrödinger’s gifts as a mathematician were so great that it did not take him long to solve the problem. The Schrödinger wave equation for matter is now considered to be more basic than Newton’s equations of motion. The wave-like properties of matter are not apparent to us in our daily lives because the wave-lengths are extremely small in comparison with the sizes of objects which we can perceive. However, for very small and light particles, such as electrons moving in their orbits around the nucleus of an atom, the wave-like behavior becomes important.

Schrödinger was able to show that Niels Bohr’s atomic theory, including Bohr’s seemingly arbitrary quantization of angular momentum, can be derived by solving the wave equation for the electrons moving in the attractive field of the nucleus. The allowed orbits of Bohr’s theory correspond in Schrödinger’s theory to harmonics, similar to the fundamental harmonic and higher overtones of an organ pipe or a violin string. (If Pythagoras had been living in 1926, he would have rejoiced to see the deepest mysteries of matter

explained in terms of harmonics!)

Bohr himself believed that a complete atomic theory ought to be able to explain the chemical properties of the elements in Mendeléeu's periodic system. Bohr's 1913 theory failed to pass this test, but the new de Broglie-Schrödinger theory succeeded! Through the work of Pauli, Heitler, London, Slater, Pauling, Hund, Mulliken, Hückel and others, who applied Schrödinger's wave equation to the solution of chemical problems, it became apparent that the wave equation could indeed (in principle) explain all the chemical properties of matter.

Strangely, the problem of developing the fundamental quantum theory of matter was solved not once, but three times in 1926! At the University of Göttingen in Germany, Max Born (1882-1970) and his brilliant young students Werner Heisenberg and Pascal Jordan solved the problem in a completely different way, using matrix methods. At the same time, a theory similar to the "matrix mechanics" of Heisenberg, Born and Jordan was developed independently at Cambridge University by a 24 year old mathematical genius named Paul Adrian Maurice Dirac. At first, the Heisenberg-Born-Jordan-Dirac quantum theory seemed to be completely different from the Schrödinger theory; but soon the Göttingen mathematician David Hilbert (1862-1943) was able to show that the theories were really identical, although very differently expressed.



Figure 3.1: Bust of Erwin Schrödinger in the courtyard arcade of the main building, University of Vienna.

3.2 Felix Bloch's story about Schrödinger

There is an interesting story about Erwin Schrödinger's derivation of his famous wave equation. According to the solid state physicist Felix Bloch, Peter Debye was chairing a symposium in Zürich, Switzerland, at which de Broglie's waves were being discussed. At one point during the symposium, Debye said: "Well, if there are waves associated with every particle, there must be a wave equation." Then, turning to Schrödinger, he said: "You, Erwin. You're not doing anything important at the moment. Why don't you find the wave equation obeyed by de Broglie's waves?"

During the following weekend, the whole group started off for a skiing trip. "Come with us, Erwin!", they said, but Schrödinger replied: "No, forgive me, I think I will stay here and work." By the end of the weekend he had derived his famous non-relativistic wave equation. He had first tried a relativistic equation (now known as the Klein-Gordon equation), but had rejected it because he believed that the equation had to be first-order in time.

Later, Felix Bloch asked Peter Debye, "Aren't you sorry that you didn't derive the wave equation yourself, instead of giving the job to Schrödinger?" Debye replied wistfully, "At least I was right about the need for a wave equation, wasn't I?"

3.3 Dirac's relativistic wave equation

In 1928, P.A.M. Dirac derived a relativistic wave equation that was first-order in time. To do this, he made use of a set of four anticommuting matrices. Solutions to the Dirac equation in the absence of external fields also obey the Klein-Gordon equation, which is second-order in time, the equation that Schrödinger first tried and then abandoned. Dirac's relativistic equation explained for the first time many details of the spectrum of hydrogen, but critics complained that it predicted the existence of negative energy states, and they asked, "Why don't the positive energy electrons fall down into these states?" Dirac replied "Because the negative energy states are all occupied." 'But then', the critics said, "an extremely energetic photon could create an electron-hole pair!" "Keep looking", Dirac answered, "and you will find that it sometimes happens." Thus, an astonishing consequence of Dirac's relativistic wave equation was the prediction of the existence of antimatter!

Years passed. Then, in 1932, the physicist Carl David Anderson observed in a cosmic ray photographic plate an event that confirmed Dirac's prediction of the existence of antimatter. A highly-energetic photon was annihilated, and converted into an electron-antielectron pair. The antielectron was given the name "positron". Since that time, the antiparticles of other particles have been discovered, created in high-energy events where a photon is annihilated and a particle-antiparticle pair created.



Figure 3.2: Carl David Anderson in 1936.



Figure 3.3: Louis Victor Pierre Raymond, duc de Broglie, (1892-1987).



Figure 3.4: Heisenberg in 1933

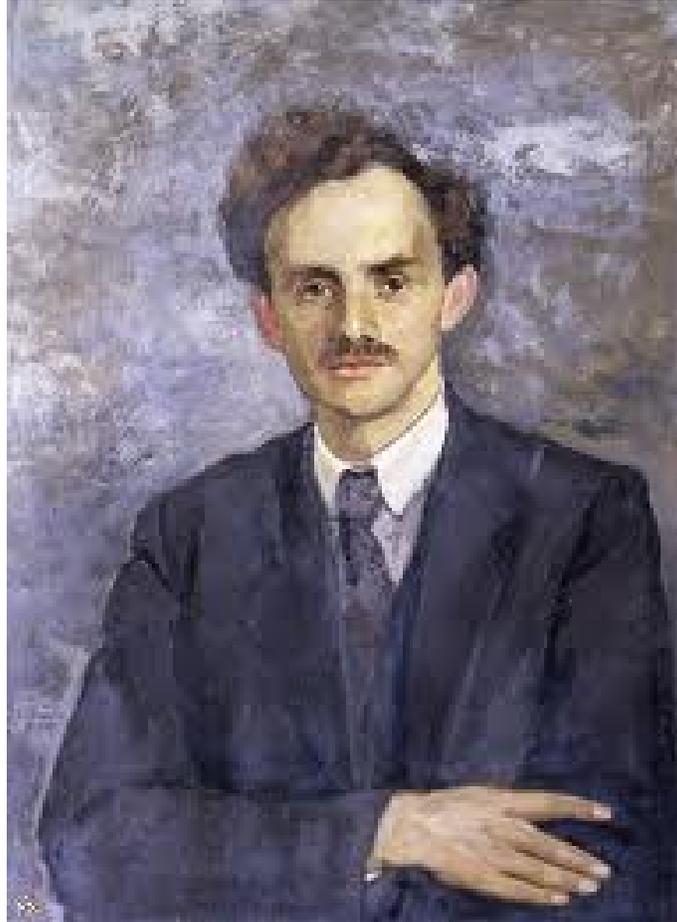


Figure 3.5: P.A.M. Dirac, the greatest British physicist of the 20th century. A memorial inscribed with his relativistic wave equation stands in Westminster Cathedral, near to the statue of Newton.



Figure 3.6: Niels Bohr, Werner Heisenberg and Wolfgang Pauli, c. 1935.



Figure 3.7: Peter Debye, (1884-1966).

3.4 Some equations

The relativistic relationship between energy and momentum

$$E^2 - p^2 c^2 = m^2 c^4 \quad (3.1)$$

Here E stands for energy, p for momentum, m for mass, and c for the velocity of light.

The Klein-Gordon equation

$$\left(-\frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} + \hbar^2 \nabla^2 \right) \psi = m^2 c^2 \psi \quad (3.2)$$

The Klein-Gordon equation can be derived from equation 3.1 by making the substitutions

$$\begin{aligned} E &\rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x_4} & x_4 &\equiv i c t \\ p_j &\rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x_j} & j &= 1, 2, 3 \end{aligned} \quad (3.3)$$

where \hbar is Planck's constant divided by 2π .

Schrödinger's non-relativistic wave equation

The non-relativistic relationship between energy and momentum is given by

$$E = c\sqrt{p^2 + m^2 c^2} + V \approx \frac{p^2}{2m} + V \quad m^2 c^2 \gg p^2 \quad (3.4)$$

Schrödinger's non-relativistic wave equation,

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V \right) \psi = E\psi \quad (3.5)$$

can be derived by making the substitutions

$$p_j \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x_j} \quad j = 1, 2, 3 \quad (3.6)$$

If the wave function ψ has time-dependence of the form

$$\psi(\mathbf{x}, t) = \psi(\mathbf{x}) e^{iEt/\hbar} \quad (3.7)$$

then we can write

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \quad (3.8)$$

where

$$\begin{aligned} H &\equiv \left(-\frac{\hbar^2}{2m} \nabla^2 + V \right) \\ \nabla^2 &\equiv \sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2} \end{aligned} \tag{3.9}$$

Suggestions for further reading

1. Akhlesh Lakhtakia. *Models and Modelers of Hydrogen: Thales, Thomson, Rutherford, Bohr, Sommerfeld, Goudsmit, Heisenberg, Schrödinger, Dirac, Sallhofer*. World Scientific, (1996).
2. O'Connor, John J.; Robertson, Edmund F., *Erwin Schrödinger*, MacTutor History of Mathematics archive, University of St Andrews.
3. Halpern, Paul. *Einstein's Dice and Schrödinger's Cat*. Perseus Books Group, (2015).
4. Heitler, W. *Erwin Schrödinger. 1887-1961*. Biographical Memoirs of Fellows of the Royal Society. 7: 221-226, (1961).
5. Farmelo, Graham. *The Strangest Man: The Hidden Life of Paul Dirac, Quantum Genius*. Faber and Faber, (2009).
6. Dalitz, R. H.; Peierls, R. *Paul Adrien Maurice Dirac. 8 August 1902 - 20 October 1984*. Biographical Memoirs of Fellows of the Royal Society. 32: 137-185, (1986).

Chapter 4

HARMONIC POLYNOMIALS AND SPHERICAL HARMONICS

4.1 Spherical polar coordinates

Spherical harmonics are very useful in many branches of physics and engineering. They are especially used when the problem to be solved has spherical symmetry. In our case, we shall need spherical harmonics to solve the spherically symmetric problem of the Schrödinger equation for hydrogenlike atoms:

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2Z^2}{r}\right)\psi(\mathbf{x}) = E\psi(\mathbf{x}) \quad (4.1)$$

Here m_e is the electron's mass, e is its charge, Z is the atomic number of the atom, r is the distance between the electron and the nucleus, and \hbar is Planck's constant divided by 2π . It is convenient to solve this problem using spherical polar coordinates:

$$\begin{aligned} \sqrt{x^2 + y^2 + z^2} &= r \\ \cos(\theta) &= \frac{z}{r} \\ \tan(\varphi) &= \frac{y}{x} \end{aligned} \quad (4.2)$$

The meaning of r , θ and φ is illustrated in the figure shown below.

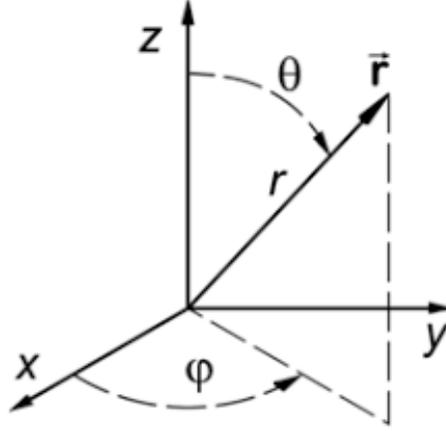


Figure 4.1: This figure illustrates the meaning of spherical polar coordinates. In geography, θ would represent latitude, while φ would represent longitude.

4.2 The Laplacian operator in spherical coordinates

The Laplacian operator, ∇^2 , expressed in terms of spherical polar coordinates, is given by

$$\begin{aligned}\nabla^2 &\equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\ &= \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}\end{aligned}\quad (4.3)$$

The details of the derivation of this equation are given in Appendix B. As we shall see, the spherical harmonics $Y_{l,m}(\theta, \varphi)$ obey the relationship

$$\nabla^2 Y_{l,m}(\theta, \varphi) = -\frac{l(l+1)}{r^2} Y_{l,m}(\theta, \varphi) \quad (4.4)$$

Thus, if we let

$$\psi(\mathbf{x}) = R(r) Y_{l,m}(\theta, \varphi) \quad (4.5)$$

the Schrödinger equation for hydrogenlike (one electron) atoms is separable:

$$\nabla^2 R(r) Y_{l,m}(\theta, \varphi) = \left(\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{l(l+1)}{r^2} \right) R(r) Y_{l,m}(\theta, \varphi) \quad (4.6)$$

Dividing both sides by $Y_{l,m}(\theta, \varphi)$, we find that the radial part of the Schrödinger equation for one-electron atoms must obey the equation

$$\left(-\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{e^2 Z^2}{r} + \frac{l(l+1)}{r^2} \right) R(r) = E R(r) \quad (4.7)$$

4.3 Homogeneous and harmonic polynomials

Let us consider a 3-dimensional space whose Cartesian coordinates are x , y and z . A polynomial in these coordinates is defined to be a sum of terms in which the coordinates are raised to various powers. A polynomial is said to be homogeneous if, in each term, the sum of the powers is the same as it is in all the other terms. A homogeneous polynomial h is harmonic if it also obeys the Laplace equation:

$$\nabla^2 h(x, y, z) \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) h(x, y, z) = 0 \quad (4.8)$$

For example, we can think of the two polynomials

$$\begin{aligned} f_2 &= x^2 + y^2 \\ h_2 &= x^2 - y^2 \end{aligned} \quad (4.9)$$

Both f_2 and h_2 are homogeneous, but only h_2 is harmonic. The reader is invited to invent a few other polynomials that illustrate this difference.

4.4 Harmonic polynomials and spherical harmonics

Harmonic polynomials are very closely related to spherical harmonics. The order l of a harmonic polynomial is the sum of the powers to which the coordinates are raised in each term. Thus, for example, the order of h_2 in the previous equation is 2. For each harmonic polynomial h_l , there is a spherical harmonic related to it by

$$Y_l(\theta, \varphi) = \frac{h_l(x, y, z)}{r^l} \quad (4.10)$$

The quantum number m is absent from this relationship because the spherical harmonic generated in this way is not necessarily an eigenfunction of the operator $\frac{\partial}{\partial \varphi}$. However $Y_l(\theta, \varphi)$ always satisfies the relationship

$$\nabla^2 Y_l(\theta, \varphi) = -\frac{l(l+1)}{r^2} Y_l(\theta, \varphi) \quad (4.11)$$

The proof is as follows: For any homogeneous polynomial f_l of order l ,

$$\nabla^2 (r^\beta f_l) = \beta(\beta + 2l + 1)r^{\beta-2} f_l + r^\beta \nabla^2 f_l \quad (4.12)$$

This relationship can be proved directly by carrying out the differentiation:

$$\begin{aligned} \nabla^2 (r^\beta f_l) &= \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left(\beta r^{\beta-1} \frac{\partial r}{\partial x_j} f_l + r^\beta \frac{\partial f_l}{\partial x_j} \right) \\ &= \sum_{j=1}^3 \left[\beta(\beta - 2)r^{\beta-4} x_j^2 f_l + \beta r^{\beta-2} f_l + 2\beta r^{\beta-2} x_j \frac{\partial f_l}{\partial x_j} + r^\beta \frac{\partial^2 f_l}{\partial x_j^2} \right] \\ &= \beta(\beta + 2l + 1)r^{\beta-2} f_l + r^\beta \nabla^2 f_l \end{aligned} \quad (4.13)$$

If the homogeneous function in is also harmonic, then (since $\nabla^2 h_l = 0$)

$$\nabla^2 (r^\beta h_l) = \beta(\beta + 2l + 1)r^{\beta-2}h_l \quad (4.14)$$

In the special case where $\beta = -l$, this becomes

$$\nabla^2 \frac{h_l}{r^l} = -\frac{l(l+1)}{r^2} \frac{h_l}{r^l} \quad (4.15)$$

or, letting $Y_l = h_l/r^l$,

$$\nabla^2 Y_l = -\frac{l(l+1)}{r^2} Y_l \quad (4.16)$$

which is what we wanted to prove. Hyperspherical harmonics are the d-dimensional analogues of spherical harmonics. Like 3-dimensional spherical harmonics, they can be generated from harmonic polynomials. The 4-dimensional hyperspherical harmonics, a table of which is shown below, are of particular interest because from them, the Fourier transforms of hydrogenlike atomic orbitals can be generated, as we shall see in Chapter 5.

Table 4.1: 3-dimensional spherical harmonics, $u_j \equiv \frac{x_j}{r}$ $j = 1, 2, 3$

l	m	$\sqrt{4\pi} Y_{l,m}(\mathbf{u})$	$\sqrt{4\pi} Y_{l,m}(\theta, \varphi)$
0	0	1	1
1	1	$-\sqrt{\frac{3}{2}}(u_1 + iu_2)$	$-\sqrt{\frac{3}{2}}e^{i\varphi} \sin(\theta)$
1	0	$\sqrt{3}u_3$	$\sqrt{3} \cos(\theta)$
1	-1	$\sqrt{\frac{3}{2}}(u_1 - iu_2)$	$\sqrt{\frac{3}{2}}e^{-i\varphi} \sin(\theta)$
2	2	$\frac{1}{2}\sqrt{\frac{15}{2}}(u_1 + iu_2)^2$	$\frac{1}{2}\sqrt{\frac{15}{2}}e^{2i\varphi} \sin^2(\theta)$
2	1	$\sqrt{\frac{15}{2}}u_3(u_1 + iu_2)$	$-\sqrt{\frac{15}{2}}e^{i\varphi} \sin(\theta) \cos(\theta)$
2	0	$\frac{1}{2}\sqrt{5}(3u_3^2 - 1)$	$\frac{1}{2}\sqrt{5}(3 \cos^2(\theta) - 1)$
2	-1	$\sqrt{\frac{15}{2}}u_3(u_1 - iu_2)$	$\sqrt{\frac{15}{2}}e^{-i\varphi} \sin(\theta) \cos(\theta)$
2	-2	$\frac{1}{2}\sqrt{\frac{15}{2}}(u_1 - iu_2)^2$	$\frac{1}{2}\sqrt{\frac{15}{2}}e^{-2i\varphi} \sin^2(\theta)$

Table 4.2: 4-dimensional hyperspherical harmonics

λ	l	m	$\sqrt{2\pi} Y_{\lambda,l,m}(\mathbf{u})$
0	0	0	1
1	1	1	$-\sqrt{2}(u_1 + iu_2)$
1	1	0	$-2u_3$
1	1	-1	$\sqrt{2}(u_1 - iu_2)$
1	0	0	$2u_4$
2	2	2	$\sqrt{3}(u_1 + iu_2)^2$
2	2	1	$-2\sqrt{3}u_3(u_1 + iu_2)$
2	2	0	$-\sqrt{2}(u_1^2 + u_2^2 - 2u_3^2)$
2	2	-1	$2\sqrt{3}u_3(u_1 - iu_2)$
2	2	-2	$\sqrt{3}(u_1 - iu_2)^2$
2	1	1	$-2\sqrt{3}u_4(u_1 + iu_2)$
2	1	0	$2\sqrt{6}u_4u_3$
2	1	-1	$2\sqrt{3}u_4(u_1 - iu_2)$
2	0	0	$3u_4^2 - u_1^2 - u_2^2 - u_3^2$

4.5 An angular integration theorem

Theorem

Let

$$I(\mathbf{n}) \equiv \int d\Omega \left(\frac{x_1}{r}\right)^{n_1} \left(\frac{x_2}{r}\right)^{n_2} \cdots \left(\frac{x_d}{r}\right)^{n_d} \quad (4.17)$$

where x_1, x_2, \dots, x_d are the Cartesian coordinates of a d -dimensional space, $d\Omega_d$ is the generalized solid angle, r is the hyperradius, and $Re[n_j] > -1$, $j = 1, 2, 3, \dots, d$. Then

$$I(\mathbf{n}) = \frac{2}{\Gamma\left(\frac{d+n}{2}\right)} \prod_{j=1}^d \frac{1}{2} (1 + e^{i\pi n_j}) \Gamma\left(\frac{n_j + 1}{2}\right) \quad (4.18)$$

where

$$n \equiv \sum_{j=1}^d n_j \quad (4.19)$$

Proof

Consider the integral

$$\int_0^\infty dr r^{d-1} e^{-r^2} \int d\Omega x_1^{n_1} x_2^{n_2} \cdots x_d^{n_d} = \prod_{j=1}^d \int_{-\infty}^\infty dx_j x_j^{n_j} e^{-x_j^2} \quad (4.20)$$

The right-hand side of (4.20) can be evaluated and has the form

$$\prod_{j=1}^d \int_{-\infty}^\infty dx_j x_j^{n_j} e^{-x_j^2} = \prod_{j=1}^d \frac{1}{2} (1 + e^{i\pi n_j}) \Gamma\left(\frac{n_j + 1}{2}\right) \quad (4.21)$$

provided that $Re[n_j] > -1$, $j = 1, 2, 3, \dots, d$. The left-hand side of (4.20) can be written as

$$\int_0^\infty dr r^{d+n-1} e^{-r^2} \int d\Omega \left(\frac{x_1}{r}\right)^{n_1} \left(\frac{x_2}{r}\right)^{n_2} \cdots \left(\frac{x_d}{r}\right)^{n_d} = \frac{I(\mathbf{n})}{2} \Gamma\left(\frac{d+n}{2}\right) \quad (4.22)$$

Combining equations (4.20)-(4.22), we obtain (4.18), Q.E.D.

Comments

If all the n_j 's are zero or positive integers, then (4.18) reduces to

$$I(\mathbf{n}) = \begin{cases} \frac{\pi^{d/2}}{2^{(n/2-1)} \Gamma\left(\frac{d+n}{2}\right)} \prod_{j=1}^d (n_j - 1)!! & \text{if all the } n_j\text{'s are even} \\ 0 & \text{otherwise} \end{cases} \quad (4.23)$$

When $d = 3$, (4.23) becomes

$$\int d\Omega_3 \left(\frac{x_1}{r}\right)^{n_1} \left(\frac{x_2}{r}\right)^{n_2} \left(\frac{x_3}{r}\right)^{n_3} = \begin{cases} \frac{4\pi}{(n+1)!!} \prod_{j=1}^3 (n_j - 1)!! & \text{all } n_j\text{'s even} \\ 0 & \text{otherwise} \end{cases} \quad (4.24)$$

while when $d = 4$, (4.23) tells us that

$$\int d\Omega_4 \left(\frac{x_1}{r}\right)^{n_1} \left(\frac{x_2}{r}\right)^{n_2} \left(\frac{x_3}{r}\right)^{n_3} \left(\frac{x_4}{r}\right)^{n_4} = \begin{cases} \frac{4\pi^2}{(n+2)!!} \prod_{j=1}^4 (n_j - 1)!! & \text{all } n_j\text{'s even} \\ 0 & \text{otherwise} \end{cases} \quad (4.25)$$

Let us now consider a general polynomial (not necessarily homogeneous) of the form:

$$P(\mathbf{x}) = \sum_{\mathbf{n}} c_{\mathbf{n}} x_1^{n_1} x_2^{n_2} \cdots x_d^{n_d} \quad (4.26)$$

Then we have

$$\int d\Omega_d P(\mathbf{x}) = \sum_{\mathbf{n}} c_{\mathbf{n}} \int d\Omega x_1^{n_1} x_2^{n_2} \cdots x_d^{n_d} = \sum_{\mathbf{n}} c_{\mathbf{n}} r^n I(\mathbf{n}) \quad (4.27)$$

It can be seen that equation (4.18) can be used to evaluate the generalized angular integral of any polynomial whatever, regardless of whether or not it is homogeneous.

The utility of these relationships is very great indeed. They provide a method for developing the theory of angular momentum and hyperangular momentum which supplements the usual group-theoretical methods. For example, Clebsch-Gordan coefficients and generalized Clebsch-Gordan coefficients may be generated by means of (4.23), rather than through equations derived from group theory, which become quite complicated when they are generalized to d dimensions.

Chapter 5

THE SCHRÖDINGER EQUATION FOR HYDROGEN

5.1 Separation of the equation

The Schrödinger equation for hydrogenlike (1-electron) atoms is

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2 Z^2}{r}\right)\psi(\mathbf{x}) = E\psi(\mathbf{x}) \quad (5.1)$$

If we let

$$\psi(\mathbf{x}) = R(r)Y_{l,m}(\theta, \varphi) \quad (5.2)$$

the Schrödinger equation for hydrogenlike (one electron) atoms is separable:

$$\nabla^2 R(r)Y_{l,m}(\theta, \varphi) = \left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} - \frac{l(l+1)}{r^2}\right)R(r)Y_{l,m}(\theta, \varphi) \quad (5.3)$$

Dividing both sides by $Y_{l,m}(\theta, \varphi)$, we find that the radial part of the Schrödinger equation for one-electron atoms must obey the equation

$$\left(-\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} - \frac{e^2 Z^2}{r} + \frac{l(l+1)}{r^2}\right)R(r) = ER(r) \quad (5.4)$$

where we have made use of the relationship

$$\nabla^2 Y_{l,m}(\theta, \varphi) = -\frac{l(l+1)}{r^2} Y_{l,m}(\theta, \varphi) \quad (5.5)$$

and where we have replaced partial derivatives in the radial equation by ordinary derivatives, since we now have a differential equation in a single variable.

5.2 Solutions to the radial equation

If we let

$$\rho \equiv Zr \quad (5.6)$$

and

$$\epsilon = \frac{2E}{Z^2} \quad (5.7)$$

then the radial equation becomes:

$$\left[\frac{1}{\rho^2} \frac{d}{d\rho} \rho^2 \frac{d}{d\rho} - \frac{l(l+1)}{\rho^2} + \frac{2}{\rho} + \epsilon \right] R_{n,l}(\rho) = 0 \quad (5.8)$$

Equation (5.8) has solutions of the form

$$R_{n,l}(\rho) = N_{n,l} \rho^l e^{-\rho/n} F [l + 1 - n \mid 2l + 2 \mid 2\rho/n] \quad (5.9)$$

where

$$N_{n,l} = \frac{Z^{3/2}}{2(2l+1)!} \left(\frac{(l+n)!}{(n-1-l)!} \right)^{1/2} \left(\frac{2}{n} \right)^{l+2} \quad (5.10)$$

and where $F[a|b|x]$ is a confluent hypergeometric function:

$$F[a|b|x] \equiv 1 + \frac{ax}{b} + \frac{a(a+1)x^2}{b(b+1)2!} + \frac{a(a+1)(a+2)x^3}{b(b+1)(b+2)3!} + \dots \quad (5.11)$$

The confluent hypergeometric series terminates and reduces to a polynomial when a is a negative integer. In our case this means that $l + 1 - n$ must be a negative integer, and thus, for the series to terminate, as is required for finiteness at large values of r , l cannot exceed $n - 1$. A table of the first few radial wave functions for hydrogenlike atoms is shown below:

Table 5.1: Radial wave functions for hydrogenlike atoms

n	l	$R_{nl}(r)$
1	0	$e^{-Zr} Z^{3/2}$
2	0	$\frac{e^{-Zr/2} Z^{3/2} (1 - \frac{Zr}{2})}{\sqrt{2}}$
2	1	$\frac{e^{-Zr/2} Z^{3/2} Zr}{2\sqrt{6}}$
3	0	$\frac{2e^{-Zr/3} Z^{3/2} (\frac{2Zr^2}{27} - \frac{2Zr}{3} + 1)}{3\sqrt{3}}$
3	1	$\frac{4}{27} \sqrt{\frac{2}{3}} e^{-Zr/3} Z^{3/2} (1 - \frac{Zr}{6}) Zr$
3	2	$\frac{2}{81} \sqrt{\frac{2}{15}} e^{-Zr/3} Z^{3/2} Zr^2$

Table 5.2: Atomic orbitals

n	l	m	$\sqrt{4\pi}R_{n,l}(r)Y_{l,m}(\theta, \varphi)$
1	0	0	$e^{-Zr}Z^{3/2}$
2	0	0	$\frac{e^{-Zr/2}Z^{3/2}(1-\frac{Zr}{2})}{\sqrt{2}}$
2	1	1	$-\frac{e^{-Zr/2}Z^{3/2}Zr}{2\sqrt{6}}\sqrt{\frac{3}{2}}e^{i\varphi}\sin(\theta)$
2	1	0	$\frac{e^{-Zr/2}Z^{3/2}Zr}{2\sqrt{6}}\sqrt{3}\cos(\theta)$
2	1	-1	$\frac{e^{-Zr/2}Z^{3/2}Zr}{2\sqrt{6}}\sqrt{\frac{3}{2}}e^{-i\varphi}\sin(\theta)$
3	0	0	$\frac{2e^{-Zr/3}Z^{3/2}(\frac{2Zr^2}{27}-\frac{2Zr}{3}+1)}{3\sqrt{3}}$
3	1	1	$-\frac{4}{27}\sqrt{\frac{2}{3}}e^{-Zr/3}Z^{3/2}(1-\frac{Zr}{6})Zr\sqrt{\frac{3}{2}}e^{i\varphi}\sin(\theta)$
3	1	0	$\frac{4}{27}\sqrt{\frac{2}{3}}e^{-Zr/3}Z^{3/2}(1-\frac{Zr}{6})Zr\sqrt{3}\cos(\theta)$
3	1	-1	$\frac{4}{27}\sqrt{\frac{2}{3}}e^{-Zr/3}Z^{3/2}(1-\frac{Zr}{6})Zr\sqrt{\frac{3}{2}}e^{-i\varphi}\sin(\theta)$

Table 5.3: Atomic orbitals (continued)

n	l	m	$\sqrt{4\pi}R_{n,l}(r)Y_{l,m}(\theta, \varphi)$
3	2	2	$\frac{2}{81}\sqrt{\frac{2}{15}}e^{-Zr/3}Z^{3/2}Zr^2\frac{1}{2}\sqrt{\frac{15}{2}}e^{2i\varphi}\sin^2(\theta)$
3	2	1	$-\frac{2}{81}\sqrt{\frac{2}{15}}e^{-Zr/3}Z^{3/2}Zr^2\sqrt{\frac{15}{2}}e^{i\varphi}\sin(\theta)\cos(\theta)$
3	2	0	$\frac{2}{81}\sqrt{\frac{2}{15}}e^{-Zr/3}Z^{3/2}Zr^2\frac{1}{2}\sqrt{5}(3\cos^2(\theta) - 1)$
3	2	-1	$\frac{2}{81}\sqrt{\frac{2}{15}}e^{-Zr/3}Z^{3/2}Zr^2\sqrt{\frac{15}{2}}e^{-i\varphi}\sin(\theta)\cos(\theta)$
3	2	-2	$\frac{2}{81}\sqrt{\frac{2}{15}}e^{-Zr/3}Z^{3/2}Zr^2\frac{1}{2}\sqrt{\frac{15}{2}}e^{-2i\varphi}\sin^2(\theta)$

5.3 Fock's momentum-space treatment of hydrogen

In a brilliant 1935 paper, the Russian physicist V. Fock was able to show that a relationship exists between 4-dimensional hyperspherical harmonics and the solutions to the Fourier transformed Schrödinger for hydrogenlike (1-electron) atoms. In direct space, the Schrödinger equation (in atomic units) for an electron moving in the potential $V(\mathbf{x})$ is

$$\left[-\frac{1}{2}\nabla^2 + V(\mathbf{x}) \right] \psi(\mathbf{x}) = E\psi(\mathbf{x}) \quad (5.12)$$

We can let

$$\psi(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int d^3p \, e^{i\mathbf{p}\cdot\mathbf{x}} \psi^t(\mathbf{p}) \quad (5.13)$$

where

$$\psi^t(\mathbf{p}) = \frac{1}{(2\pi)^{3/2}} \int d^3x \, e^{-i\mathbf{p}\cdot\mathbf{x}} \psi(\mathbf{x}) \quad (5.14)$$

Substituting (5.13) into (5.12), we have

$$\frac{1}{(2\pi)^{3/2}} \int d^3p \, \left[\frac{p^2}{2} + V(\mathbf{x}) - E \right] e^{i\mathbf{p}\cdot\mathbf{x}} \psi^t(\mathbf{p}) = 0 \quad (5.15)$$

We now multiply on the left by $e^{-i\mathbf{p}'\cdot\mathbf{x}}$ and integrate over d^3x . This gives:

$$\left[\frac{p'^2}{2} - E \right] \psi^t(\mathbf{p}') = \frac{-1}{(2\pi)^{3/2}} \int d^3p \, V^t(\mathbf{p}' - \mathbf{p}) \psi^t(\mathbf{p}) \quad (5.16)$$

which is the 1-particle Schrödinger equation in reciprocal space. For hydrogenlike atoms,

$$V(\mathbf{x}) = -\frac{Z}{r} \quad (5.17)$$

so that from (??),

$$V^t(\mathbf{p}) = -\sqrt{\frac{2}{\pi}} \frac{Z}{p^2} \quad (5.18)$$

Letting

$$-2E = k^2 \quad (5.19)$$

and combining (5.16), (5.17) and (5.18), we obtain

$$[p'^2 + k^2] \psi^t(\mathbf{p}') = \frac{Z}{\pi^2} \int d^3p \, \frac{1}{|\mathbf{p}' - \mathbf{p}|^2} \psi^t(\mathbf{p}) \quad (5.20)$$

Fock then made the transformation:

$$\begin{aligned}
 u_1 &= \frac{2kp_1}{k^2 + p^2} \equiv \sin \chi \sin \theta \cos \varphi \\
 u_2 &= \frac{2kp_2}{k^2 + p^2} \equiv \sin \chi \sin \theta \sin \varphi \\
 u_3 &= \frac{2kp_3}{k^2 + p^2} \equiv \sin \chi \cos \theta \\
 u_4 &= \frac{k^2 - p^2}{k^2 + p^2} \equiv \cos \chi
 \end{aligned} \tag{5.21}$$

Here θ and φ are the polar angles of the vector \mathbf{p} :

$$\begin{aligned}
 p_1 &= p \sin \theta \cos \varphi \\
 p_2 &= p \sin \theta \sin \varphi \\
 p_3 &= p \cos \theta
 \end{aligned} \tag{5.22}$$

while

$$\chi \equiv \cos^{-1} \left(\frac{k^2 - p^2}{k^2 + p^2} \right) = \sin^{-1} \left(\frac{2kp}{k^2 + p^2} \right) \tag{5.23}$$

is an angle introduced by Fock in order to transform the integral d^3p into an integral over solid angle in a 4-dimensional space. Fock's transformation maps the 3-dimensional p -space onto the surface of a unit sphere in a 4-dimensional space. It is easy to verify from (5.21) that

$$u_1^2 + u_2^2 + u_3^2 + u_4^2 = 1 \tag{5.24}$$

From the Jacobian of the transformation from Cartesian coordinates to 4-dimensional hyperspherical coordinates, one finds that the element of solid angle in the 4-dimensional space is given by

$$\begin{aligned}
 d\Omega &= \sin^2 \chi \sin \theta d\chi d\theta d\varphi \\
 &= \left(\frac{2kp}{k^2 + p^2} \right)^2 \sin \theta d\chi d\theta d\varphi
 \end{aligned} \tag{5.25}$$

Comparing this with

$$d^3p = p^2 dp \sin \theta d\theta d\varphi \tag{5.26}$$

and making use of the fact that

$$\frac{d\chi}{dp} = \frac{2k}{k^2 + p^2} \tag{5.27}$$

we have

$$\begin{aligned} d\Omega &= \left(\frac{2k}{k^2 + p^2} \right)^3 d^3p \\ d^3p &= \left(\frac{k^2 + p^2}{2k} \right)^3 d\Omega \end{aligned} \quad (5.28)$$

Also, from (5.21), we have:

$$\begin{aligned} \mathbf{p} \cdot \mathbf{p}' &= \frac{4k^2}{(k^2 + p^2)(k^2 + p'^2)(\mathbf{u} \cdot \mathbf{u}' - u_4 u'_4)} \\ \frac{1}{|\mathbf{p} - \mathbf{p}'|^2} &= \frac{4k^2}{(k^2 + p^2)(k^2 + p'^2)} \frac{1}{|\mathbf{u} - \mathbf{u}'|^2} \end{aligned} \quad (5.29)$$

Inserting (5.28) and (5.29) into (5.20), we obtain:

$$[p'^2 + k^2]^2 \psi^t(\mathbf{p}') = \frac{Z}{2k\pi^2} \int d\Omega \frac{(k^2 + p^2)^2}{|\mathbf{u}' - \mathbf{u}|^2} \psi^t(\mathbf{p}) \quad (5.30)$$

We now let

$$\psi^t(\mathbf{p}) = \frac{4k^{5/2}}{(k^2 + p^2)^2} \varphi(\Omega) \quad (5.31)$$

(As shown in Section 5.3 below, the factor $4k^{5/2}$ in the numerator is needed to normalize $\psi^t(\mathbf{p})$). Equation (5.30) then takes on the simple form

$$\varphi(\Omega') = \frac{Z}{2k\pi^2} \int d\Omega \frac{1}{|\mathbf{u}' - \mathbf{u}|^2} \varphi(\Omega) \quad (5.32)$$

From equation (??), with $d = 4$ and $\alpha = d/2 - 1 = 1$, we have

$$\frac{1}{|\mathbf{u}' - \mathbf{u}|^2} = \sum_{\lambda=0}^{\infty} C_{\lambda}^1(\mathbf{u} \cdot \mathbf{u}') \quad (5.33)$$

so that (5.32) becomes

$$\varphi(\Omega') = \frac{Z}{2k\pi^2} \sum_{\lambda=0}^{\infty} \int d\Omega C_{\lambda}^1(\mathbf{u} \cdot \mathbf{u}') \varphi(\Omega) \quad (5.34)$$

Remembering equation (??) we can rewrite this in the form

$$\varphi(\Omega') = \frac{Z}{2k\pi^2} \sum_{\lambda=0}^{\infty} K_{\lambda} O_{\lambda}[\varphi(\Omega)] \quad (5.35)$$

For $d = 4$,

$$K_\lambda = \frac{I(0)}{\lambda + 1} = \frac{2\pi^2}{\lambda + 1} \quad (5.36)$$

so that equation (5.35) becomes:

$$\varphi(\Omega') = \frac{Z}{k} \sum_{\lambda=0}^{\infty} \frac{1}{\lambda + 1} O_\lambda[\varphi(\Omega')] \quad (5.37)$$

If $\varphi(\Omega)$ is an eigenfunction of Λ^2 . so that

$$O_{\lambda'}[\varphi(\Omega')] = \delta_{\lambda'\lambda} \varphi(\Omega) \quad (5.38)$$

then (5.37) will be satisfied provided that

$$\frac{Z}{k(\lambda + 1)} = 1 \quad (5.39)$$

or, from (5.19),

$$E = -\frac{k^2}{2} = -\frac{Z^2}{2(\lambda + 1)^2} = -\frac{Z^2}{2n^2} \quad \lambda = 0, 1, 2, \dots \quad n = 1, 2, 3, \dots \quad (5.40)$$

where we have made the identification $\lambda + 1 = n$. We can see that Fock's treatment gives the usual energy levels for hydrogenlike atoms. For the transformed wave function $\varphi(\Omega)$, any 4-dimensional hyperspherical harmonic will do, but for most applications, it is convenient to use hyperspherical harmonics of the type shown in Table 2.1. Thus we obtain the Fourier transformed hydrogenlike orbitals:

$$\begin{aligned} \psi_{n,l,m}^t(\mathbf{p}) &= \frac{4k^{5/2}}{(k^2 + p^2)^2} Y_{n-1,l,m}(\Omega_4) \equiv M(p) Y_{n-1,l,m}(\Omega_4) \\ M(p) &\equiv \frac{4k^{5/2}}{(k^2 + p^2)^2} \end{aligned} \quad (5.41)$$

For the first few values of n , l and m , 5.41 yields:

$$\begin{aligned} \psi_{1,0,0}^t(\mathbf{p}) &= \frac{2\sqrt{2} k^{5/2}}{(k^2 + p^2)^2 \pi} \\ \psi_{2,0,0}^t(\mathbf{p}) &= \frac{4\sqrt{2} k^{5/2} (k^2 - p^2)}{(k^2 + p^2)^3 \pi} \\ \psi_{2,1,-1}^t(\mathbf{p}) &= -\frac{8ik^{7/2} (p_1 - ip_2)}{(k^2 + p^2)^3 \pi} \\ \psi_{2,1,0}^t(\mathbf{p}) &= -\frac{8i\sqrt{2} k^{7/2}}{(k^2 + p^2) \pi} \\ \psi_{2,1,1}^t(\mathbf{p}) &= \frac{8ik^{7/2} (p_1 + ip_2)}{(k^2 + p^2)^3 \pi} \\ \vdots & \quad \quad \quad \vdots \end{aligned} \quad (5.42)$$

magnetic field. The “spin” and magnetic moment of electrons were completely explained in 1928 by P.A.M. Dirac’s relativistic wave equation.

Meanwhile, the Austrian physicist Wolfgang Pauli proposed his famous *exclusion principle*, which explained the periodic table and the chemical properties of the elements. According to the Pauli exclusion principle, in the lowest energy state of an atom, the electrons fill the atomic orbitals in the order (1s), (2p), (3d), ... Two electrons are allowed in each linearly independent orbital, one with spin up and the other with spin down. This leads to the following electron configurations for the elements:

- Hydrogen; (1s); very active metal; valence=1
- Helium; (1s)²; noble (inert) gas; valence=0
- Lithium; (1s)²(2p)¹; very active metal; valence=1
- Beryllium; (1s)²(2p)²; metal; valence=2
- Boron; (1s)²(2p)³; less active metal; valence=3
- Carbon; (1s)²(2p)⁴; intermediate; valence=4
- Nitrogen; (1s)²(2p)⁵; less active nonmetal; valence=5
- Oxygen; (1s)²(2p)⁶; nonmetal; valence=6
- Fluorine; (1s)²(2p)⁷; very active nonmetal; valence=7
- Neon; (1s)²(2p)⁸; noble gas; valence=0
- Sodium; (1s)²(2p)⁸(2s)¹; very active metal; valence=1

In chemical reactions, the metals tend to give away their outer-shell electrons, while the non-metals tend to accept electrons. The most active metals, hydrogen, lithium, sodium, potassium, rubidium and cesium, all have a single electron in their outer shell, and they tend to give this electron away. The most active nonmetals, fluorine, chlorine, bromine and iodine, all are missing a single electron to complete their outer shell. We can notice that common table salt, is a cubic crystal structure formed from Na⁺ ions and Cl⁻ ions. When it is dissolved in water, the sodium-chloride crystal dissociates into Na⁺ ions, complexed with water molecules and Cl⁻ ions, also forming complexes with water. We see here the strong tendency of very active metals to give up their outer shell electron and to form positive ions, while very active nonmetals have an equally strong tendency to form negative ions. Helium, neon, argon, krypton, and radon, all with completely filled outer shell, are unreactive noble gases, with no tendency at all to give away or accept electrons or to form ions.

The Hartree-Fock equations

The application of the Schrödinger equation to our understanding of chemical reactivity and the periodic table was made quantitative through the work of Douglas Hartree (1897-1958) and Vladimir A. Fock (1898-1974).

Douglas Hartree was born in Cambridge, England, where his father was a professor of engineering at Cambridge University and his mother was the mayor of the city. In his work on the electronic structure of atoms, Hartree visualized the electrons moving

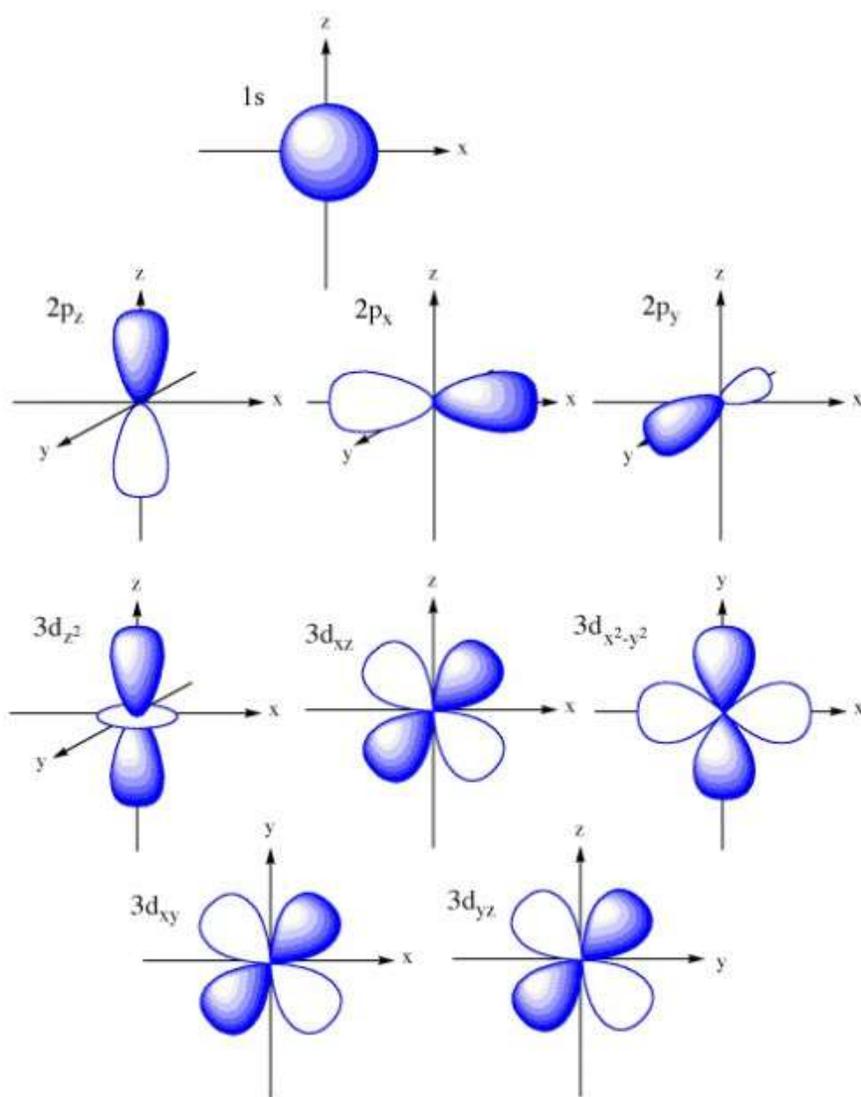


Figure 5.1: Atomic orbitals.

1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	*	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	**	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Uut	114 Fl	115 Uup	116 Lv	117 Uus	118 Uuo
119 Uun																	
			57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
			89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr
Alkali metals		Alkaline earth metals		Lanthanides				Actinides				Transition metals					
Poor metals		Metalloids		Nonmetals				Halogens				Noble gases					

red
blue

Figure 5.2: The periodic table of the elements.

in both the attractive field of the atomic nucleus and in a repulsive potential produced collectively by all the electrons. Hartree's method for treating this problem was to make an initial guess of the size of the atomic orbitals (Figure 11.6) occupied by the electrons. He then calculated the repulsive potential that would result, and combined it with the nuclear attraction potential. Solving the Schrödinger equation for the an electron moving in this new potential, he obtained a set of improved atomic orbitals, and from these he could calculate an improved total potential. He continued to iterate this process until the change resulting from successive iterations became very small, at which point he described the electrical field in which the electrons moved as being *self-consistent*. Hartree called his procedure the Self-Consistent-Field (or SCF) Method. He published his first results in 1927, only a year after Schrödinger's discovery of his wave equation.

The Russian physicist Vladimir A. Fock was able to refine Hartree's method by postulating that the total electronic wave function of an atom or molecule had to be antisymmetric with respect to the exchange of the coordinates of any two electrons in the system. When spin was included in the wave function, this requirement led in a natural way to the exclusion principle postulated by Wolfgang Pauli. When combined with Hartree's SCF method, Fock's antisymmetry requirement led to more accurate results and better agreement between theory and experiment. However, the Hartree-Fock SCF equations were much more difficult to solve. Later Clemens C.J. Roothaan (1918-2019) converted the Hartree-Fock equations into a matrix form suitable for solution by digital computers. The method in use today is thus known as the Hartree-Fock-Roothaan SCF Method. When applied to molecules, it is called the Hartree-Fock-Roothaan LCAO SCF Method. The LCAO in the name stands for the fact that molecular orbitals are represented as Linear Combinations of Atomic Orbitals.

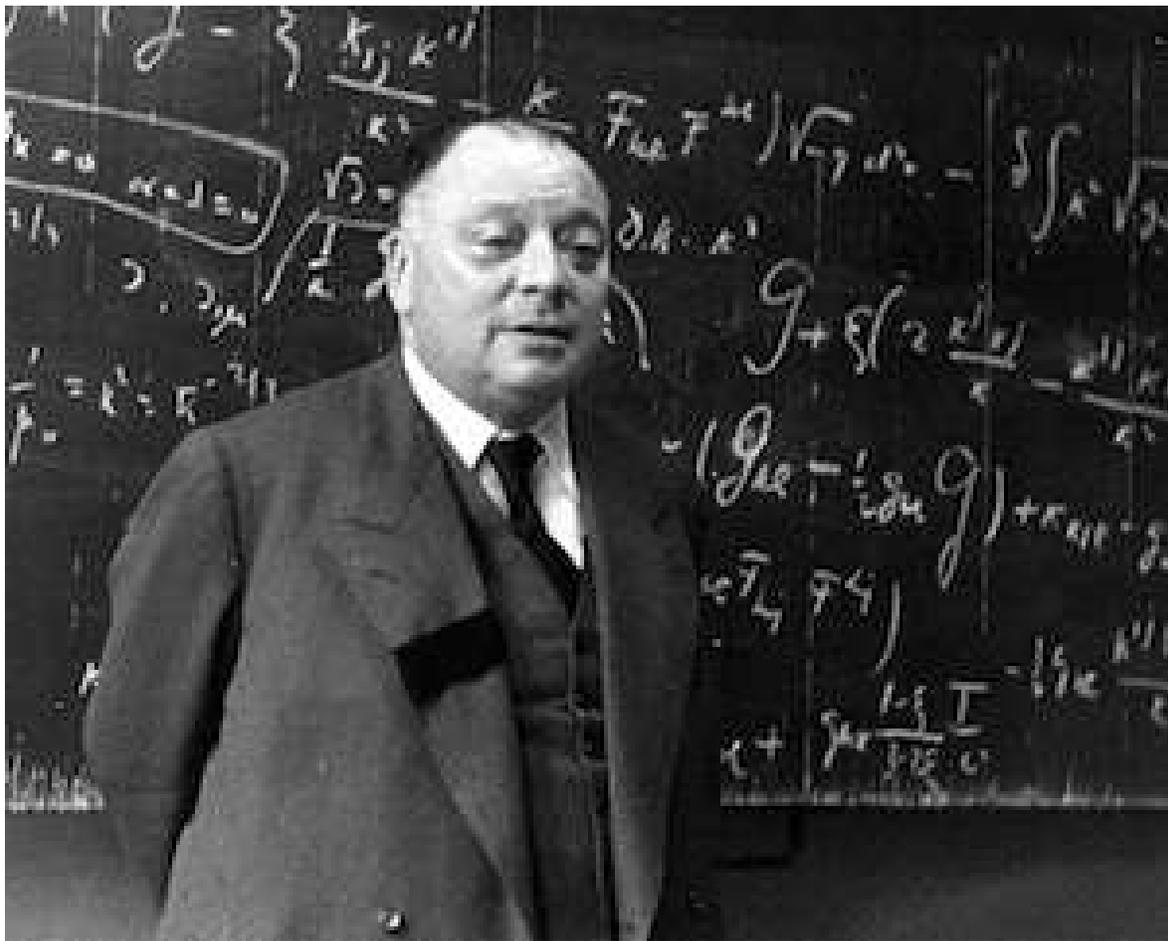


Figure 5.3: Wolfgang Pauli (1900-1958).



Figure 5.4: Douglas Hartree (1897-1958).



Figure 5.5: Vladimir A. Fock (1898-1974).

5.5 Valence bond theory

Lewis structures

In 1916, G.N. Lewis of the University of California, Berkeley, proposed a theory of chemical bonding in which a pair of electrons, one donated by each of the bonding atoms, together form a covalent bond. For example, in the hydrogen molecule, H_2 , one electron is contributed by each of the two hydrogen atoms. Another example of a Lewis structure is the NO_2^- ion shown in Figure 15.2. The outer-shell electrons that do not contribute to bonding are represented by pairs of dots and are called *lone pairs*.

Heitler-London theory

The next step in valence bond theory was taken in 1927 by Walter Heitler and Fritz London, who used Erwin Schrödinger's wave equation and Wolfgang Pauli's exclusion principle to study the covalent bonding of the hydrogen molecule.

Linus Pauling's contributions

Linus Pauling developed these ideas further by introducing the key concepts of resonance and orbital hybridization. Pauling's two famous books, *Introduction to Quantum Mechanics, With Applications to Chemistry* (with E. Bright Wilson, 1935), and *The Nature of the Chemical Bond* (1939) were extremely important and influential, as was Charles Coulson's *Valence* (1952).

5.6 Molecular orbital theory

In molecular orbital theory, atomic orbitals shown in in Figure 11.6 are used to build up a representation of the orbital of an electron moving in a molecule. For example, Figure 15.3 shows a schematic diagram of the LCAO (Linear Combination of Atomic Orbitals) model of bonding in the H_2 molecule. When two hydrogen atoms approach each other, the two electrons move in an attractive potential produced by both nuclei. Thus the range of their motion is enlarged. According to the Pauli exclusion principle, both electrons are allowed to move in this enlarged region, if they have opposite spins. The electronic wave function in the enlarged region of motion is called a *molecular orbital*.

In the LCAO approximation, molecular orbitals are built up from atomic orbitals centered on the individual atoms of a molecule. For example, suppose that in the H_2 molecule we denote the positions of the two nuclei by \mathbf{X}_1 and \mathbf{X}_2 . Then we can approximate the molecular orbitals $\phi_\sigma(\mathbf{x})$ and $\phi_{\sigma^*}(\mathbf{x})$ by symmetric and antisymmetric combinations of the two atomic orbitals, $\chi_{1s}(\mathbf{x} - \mathbf{X}_1)$ and $\chi_{1s}(\mathbf{x} - \mathbf{X}_2)$:

$$\begin{aligned}\phi_\sigma(\mathbf{x}) &= N [\chi_{1s}(\mathbf{x} - \mathbf{X}_1) + \chi_{1s}(\mathbf{x} - \mathbf{X}_2)] \\ \phi_{\sigma^*}(\mathbf{x}) &= N' [\chi_{1s}(\mathbf{x} - \mathbf{X}_1) - \chi_{1s}(\mathbf{x} - \mathbf{X}_2)]\end{aligned}\tag{5.45}$$

where N and N' are normalizing constants. The symmetric combination is called a *bonding orbital*, and in the hydrogen molecule ground state it is doubly occupied, the two electrons having opposite spin quantum numbers. In the ground state of H_2 , the antibonding orbital, $\phi_{\sigma^*}(\mathbf{x})$, is unoccupied, as is illustrated schematically in Figure 15.3.

Quantum biochemistry

Erich Hückel devised an extremely simple semiempirical approximation for treating the highest filled and lowest empty molecular orbitals of flat organic molecules such as benzene, naphthalene, pyridine, guanine, cytosine, etc. Solution to Hückel's equations only required the diagonalization of a small matrix, and this could be done even before the advent of electronic computers. During World War II, the French scientist Alberte Pullman sat in a basement room in Paris diagonalizing Hückel matrices with a desk-top mechanical calculator, while her husband Bernard drove a tank for the Free French forces in North Africa. After the war Alberte and Bernard Pullman published a pioneering book entitled *Quantum Biochemistry* (1963), which opened a new field of research. Meanwhile, the great Hungarian-American biochemist and physiologist Albert Szent-Györgyi, published a book entitled *Introduction to a Submolecular Biology* (1960) in which he explored the biochemical roles of charge donors and charge acceptors. Charge donors are molecules whose highest filled molecular orbitals are relatively high in energy while charge acceptors are molecules are relatively low in energy.

A typical programmable minicomputer or "microprocessor", manufactured in the 1970's, could have 30,000 circuit elements, all of which were contained on a single chip. By 1989, more than a million transistors were being placed on a single chip; and by 2000, the number reached 42,000,000.

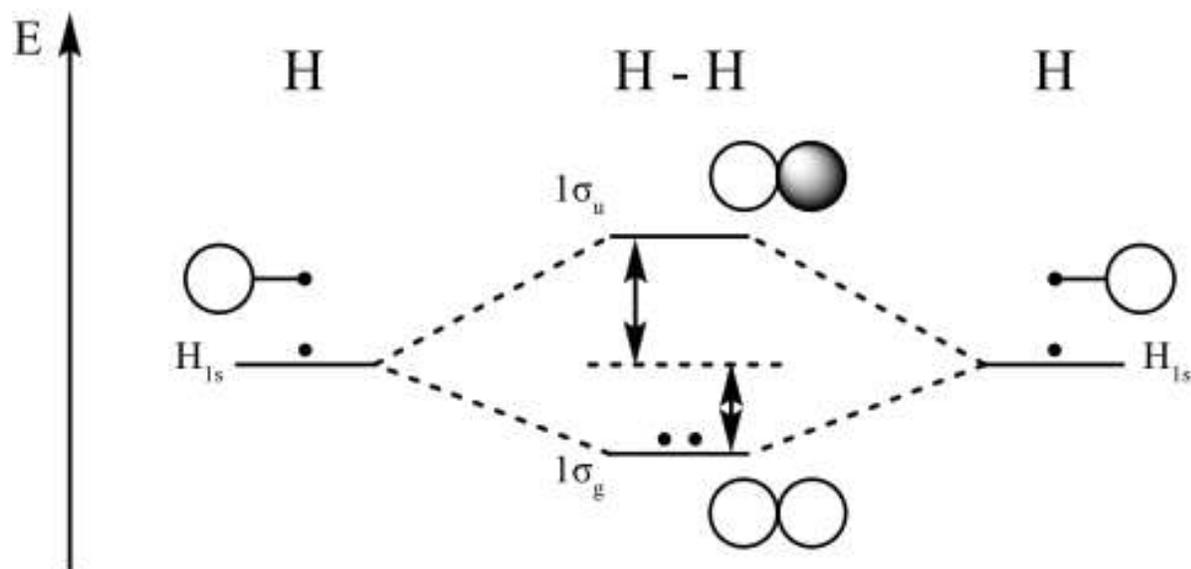


Figure 5.8: A schematic diagram of the LCAO (Linear Combination of Atomic Orbitals) model of bonding in the H_2 molecule.

As a result of miniaturization and parallelization, the speed of computers rose exponentially. In 1960, the fastest computers could perform a hundred thousand elementary operations in a second. By 1970, the fastest computers took less than a second to perform a million such operations. In 1987, a massively parallel computer, with 566 parallel processors, called GF11 was designed to perform 11 billion floating-point operations per second (flops). By 2002 the fastest computer performed 40 at teraflops, making use of 5120 parallel CPU's.

Computer disk storage has also undergone a remarkable development. In 1987, the magnetic disk storage being produced could store 20 million bits of information per square inch; and even higher densities could be achieved by optical storage devices. Storage density has until followed a law similar to Moore's law.

In the 1970's and 1980's, computer networks were set up linking machines in various parts of the world. It became possible (for example) for a scientist in Europe to perform a calculation interactively on a computer in the United States just as though the distant machine were in the same room; and two or more computers could be linked for performing large calculations. It also became possible to exchange programs, data, letters and manuscripts very rapidly through the computer networks.

The exchange of large quantities of information through computer networks was made easier by the introduction of fiber optics cables. By 1986, 250,000 miles of such cables had been installed in the United States. If a ray of light, propagating in a medium with a large refractive index, strikes the surface of the medium at a grazing angle, then the ray undergoes total internal reflection. This phenomenon is utilized in fiber optics: A light signal can propagate through a long, hairlike glass fiber, following the bends of the fiber without



Figure 5.9: **Robert Mulliken (1896-1986)**. His contributions to molecular orbital theory won him the 1966 Nobel Prize in Chemistry.



Figure 5.10: Erich Hückel (1896-1980). He developed a very simple semiempirical approximation for treating the highest filled and lowest empty molecular orbitals of flat organic molecules.



Figure 5.11: Alberte Pullman (1920-2011).



Figure 5.12: Alberte Pullman with her husband, Bernard.



Figure 5.13: Linus Pauling (1901-1994). The New Scientist called him one of the 20 most important scientists in history. He was awarded the Nobel Prize in Chemistry in 1954 and the Nobel Peace Prize in 1962.

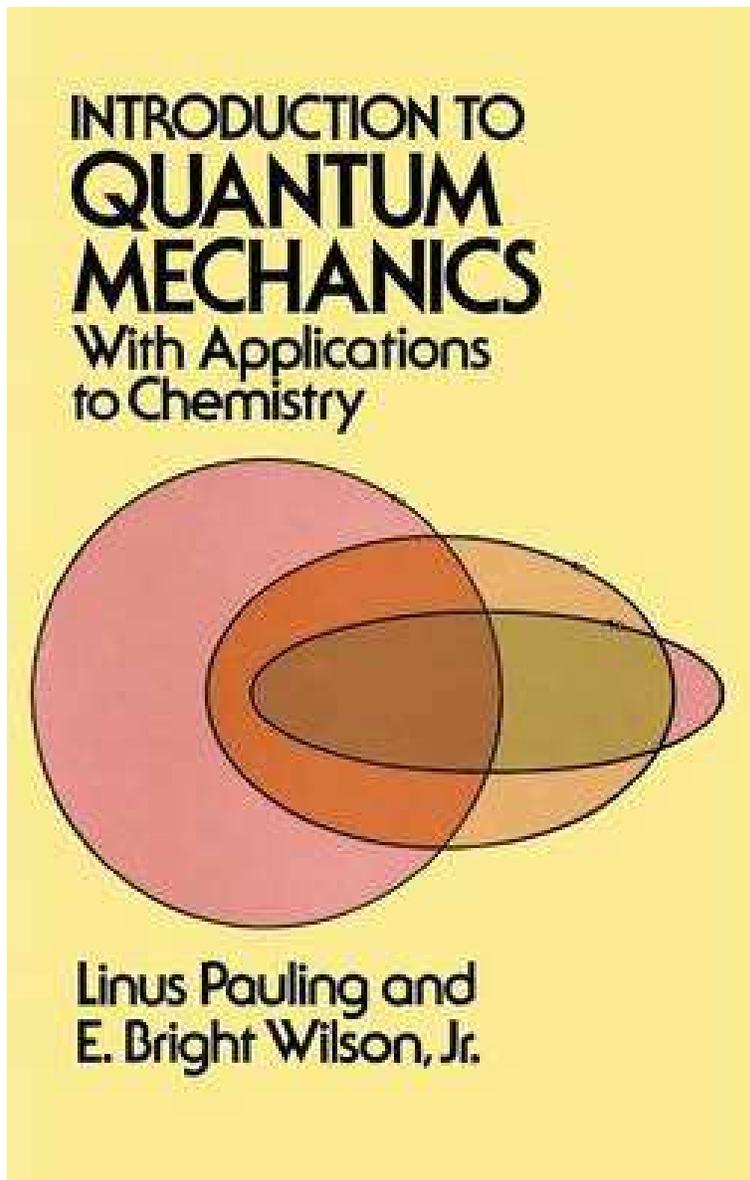


Figure 5.14: Two-time Nobel Prize winner Linus Pauling, Research Professor at the Linus Pauling Institute of Science and Medicine, Palo Alto, California, and E. Bright Wilson, Jr., Professor Emeritus of Chemistry at Harvard University, provide a readily understandable study of “wave mechanics,” discussing the Schrödinger wave equation and the problems which can be solved with it. The book was first published in 1935, and it is still in use today.

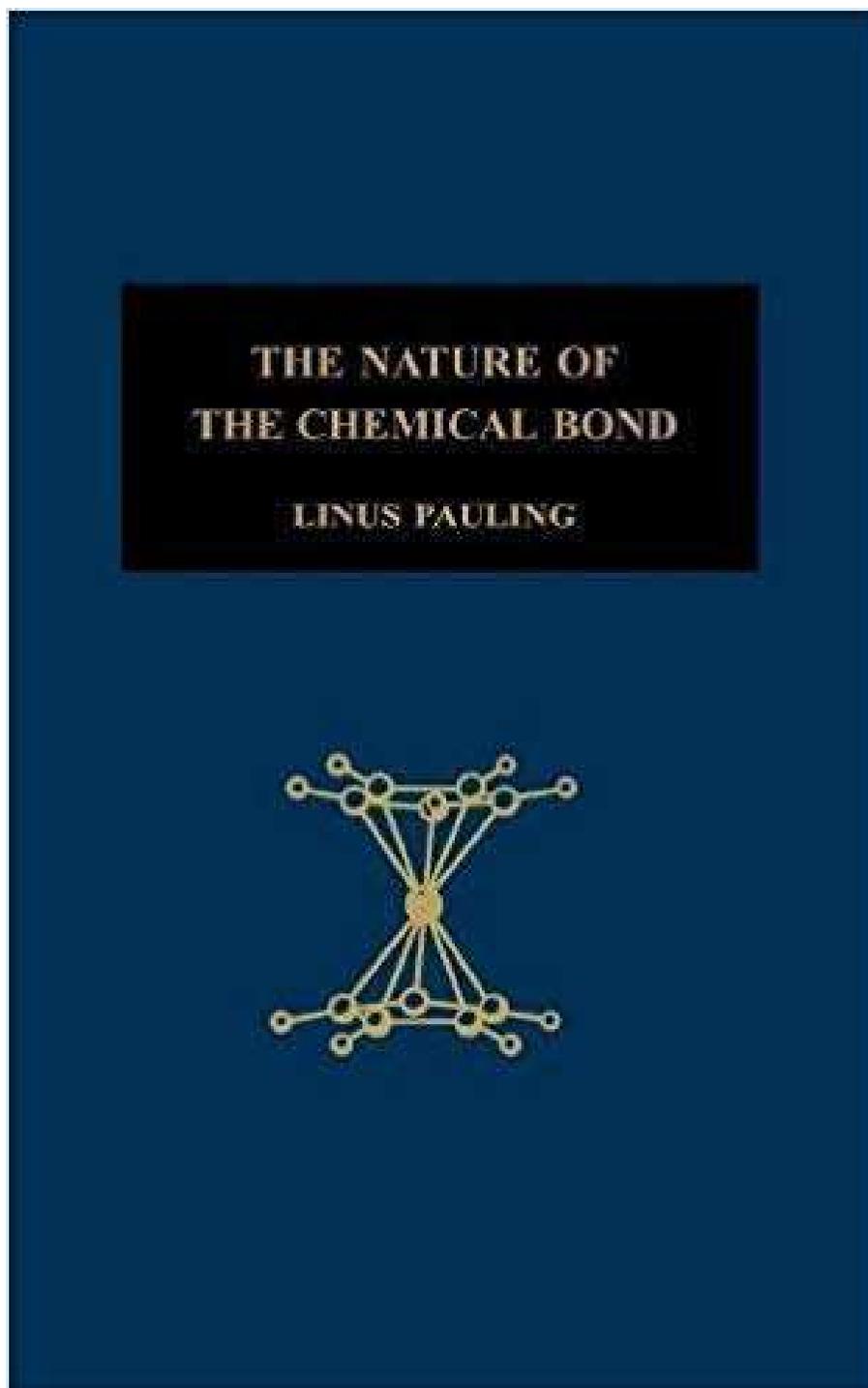


Figure 5.15: Linus Pauling's famous book, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals*. Cornell University Press, 1939.

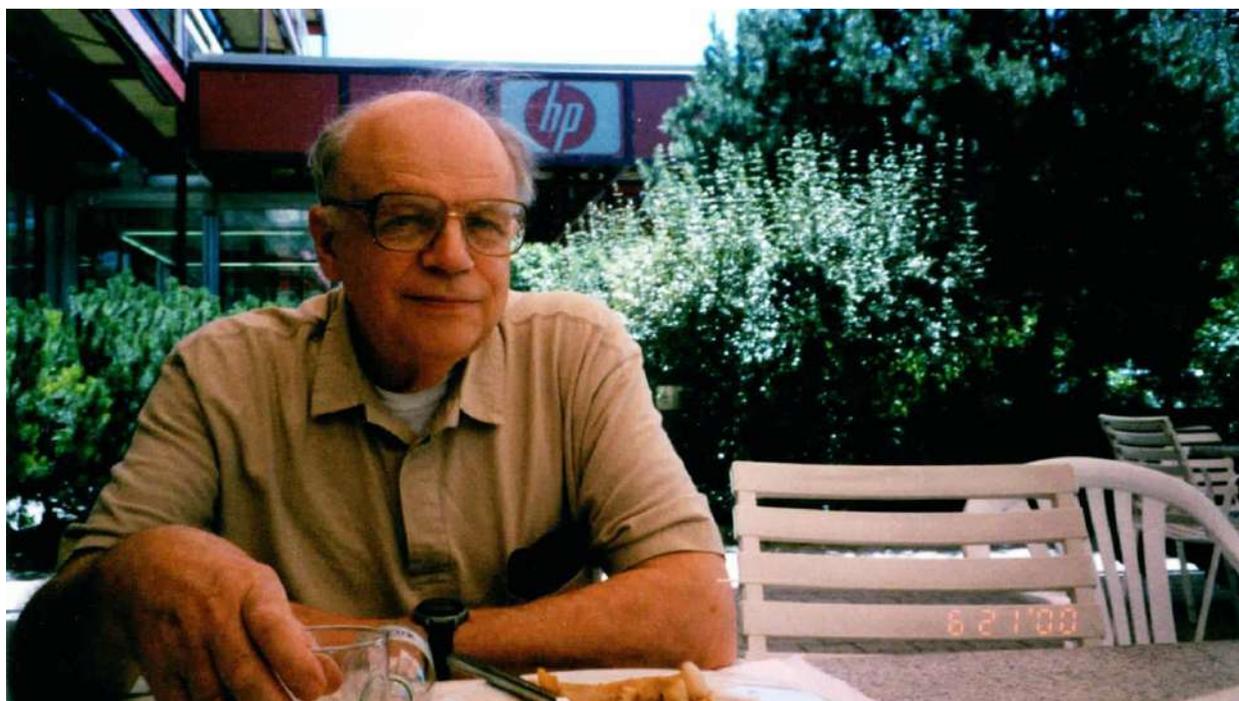


Figure 5.16: Clemens C.J. Roothaan (1918-2019). His thesis supervisor, Robert Mulliken, said of him, “I tried to induce Roothaan to do his Ph.D. thesis on Hückel-type calculations on substituted benzenes. But after carrying out some very good calculations on these he revolted against the Hückel method, threw his excellent calculations out the window, and for his thesis developed entirely independently his now well known all-electron LCAO SCF self-consistent-field method for the calculation of atomic and molecular wave functions, now appropriately referred to, I believe, as the Hartree-Fock-Roothaan method”.

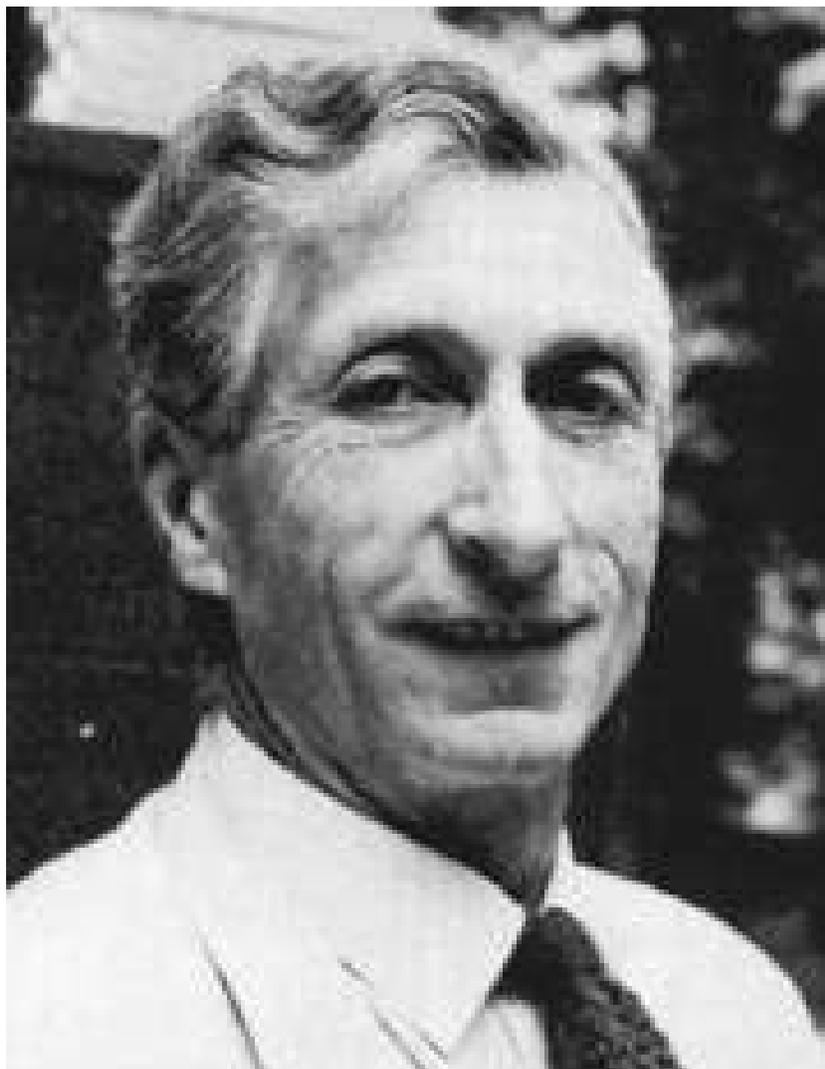


Figure 5.17: Charles Coulson (1910-1974). He was the leader of important quantum chemistry groups, first at Kings College London, and then later at Oxford University. He was a very tall man, and also a religious one, with a strong sense of duty. On the day of his death in 1974, he worked until the last minute. His last words were, “I think I can’t do any more”. His book, *Valence*, greatly influenced the development of quantum chemistry.

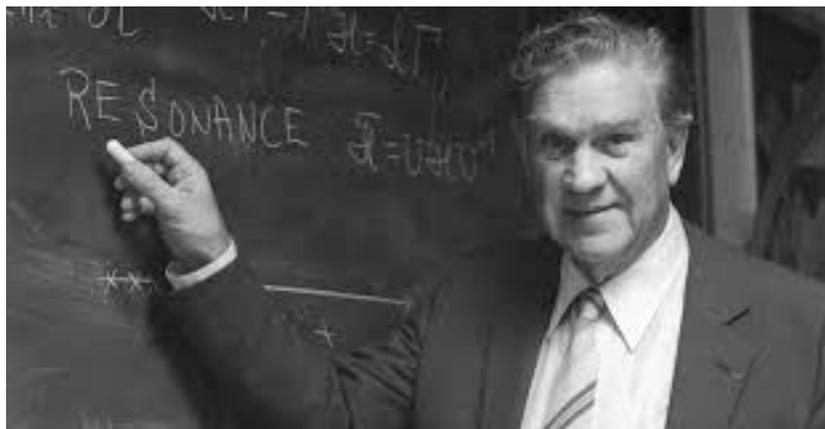


Figure 5.18: Per-Olov Löwdin (1916-2000). Prof. Löwdin founded two important quantum chemistry groups, one at the University of Uppsala in Sweden, and another at the University of Florida in Gainesville. Many scientists were introduced to quantum chemistry at his famous Uppsala summer schools, which often included mountain-climbing expeditions. Prof. Löwdin also founded two important journals, *The International Journal of Quantum Chemistry* and *Advances in Quantum Chemistry*. He invented a number of mathematical techniques much used in the field, for example his method for symmetrical orthogonalization of basis functions.

losing intensity because of total internal reflection. However, before fiber optics could be used for information transmission over long distances, a technological breakthrough in glass manufacture was needed, since the clearest glass available in 1940 was opaque in lengths more than 10 m. Through studies of the microscopic properties of glasses, the problem of absorption was overcome. By 1987, devices were being manufactured commercially that were capable of transmitting information through fiber-optic cables at the rate of 1.7 billion bits per second.

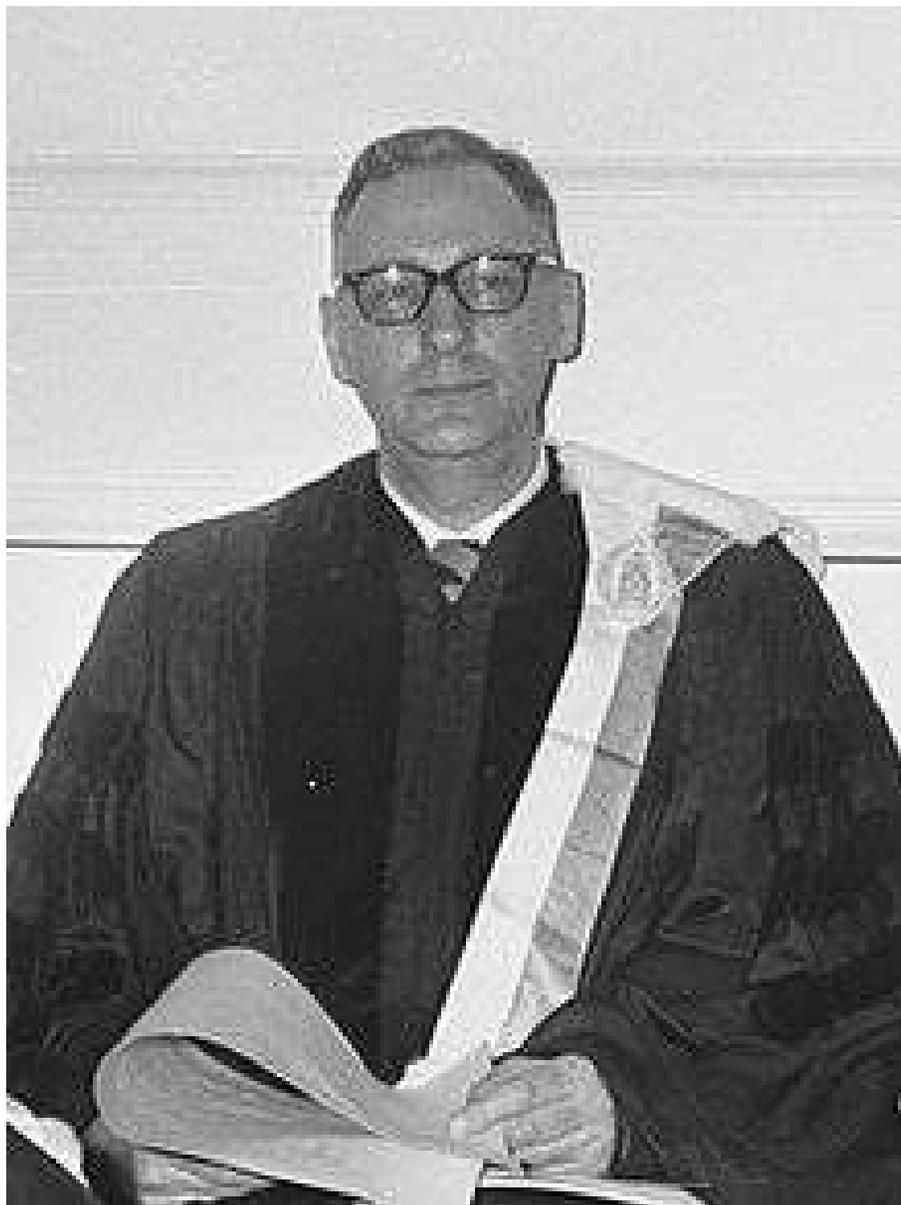


Figure 5.19: Tjalling C. Koopmans (1910-1985). Born in the Netherlands he studied mathematics and physics at the University of Utrecht. After publishing his famous theorem in quantum physics, he switched his interest to applications of mathematics in economics. In 1975, he shared the Nobel Memorial Prize in Economics for his work in that field. Koopmans' theorem states that if we neglect the readjustment of orbitals resulting from ionization, the energy needed to remove an electron from the k th molecular orbital is just the eigenvalue of the Fock operator corresponding to that orbital.



Figure 5.20: Prof. Clemens C. J. Roothaan in the University of Chicago Computation Center (ca. 1964). As computers gained more and more speed and memory capacity, quantum chemists like Prof. Roothaan have become able to use progressively more accurate calculational methods. Prof. Roothaan was my teacher in group theory when I studied at the University of Chicago. I later had the privilege of knowing him quite well as the result of meeting him at many quantum chemistry conferences. He and his brother were both in the Dutch resistance movement against the Nazis during World War II. Although his brother was killed, Prof. Roothaan survived, and lived an extremely long and productive life. He lived to be 101. I remember talking with him when he was a very old man. He told me about a new and powerful computer chip that he was developing.

5.7 The Hartree-Fock-Roothaan equations

In the Hartree-Fock-Roothaan method, the wave function is built up from a linear combination of atomic spin-orbitals χ_a , located on the various atoms in the molecule:

$$\Phi_i = \sum_{a=1}^M \chi_a C_{a,i} \quad (5.46)$$

The Fock operator is defined by the equation

$$F(1) \equiv H^c(1) + \sum_{s=1}^N \int d\tau_2 \Phi_s^*(2) \frac{e^2}{r_{1,2}} (1 - P_{1,2}) \Phi_s(2) \quad (5.47)$$

Here $\int d\tau_2$ indicates integration over the space and spin coordinates of electron 2, while $H^c(1)$ is the *core Hamiltonian* as a function of the space and spin coordinates of electron 1:

$$H^c(1) = -\frac{\hbar^2}{2m_e} \nabla_1^2 + V(1) \quad (5.48)$$

The operator $P_{1,2}$ exchanges the space and spin coordinates of electrons 1 and 2:

$$P_{1,2} \Phi_s(2) \Phi_i(1) \equiv \Phi_s(1) \Phi_i(2) \quad (5.49)$$

When the Fock operator acts on a filled molecular spin-orbital, it generates a linear combination of filled molecular spin-orbitals:

$$F(1) \Phi_j(1) = \sum_{i=1}^N \Phi_i(1) \lambda_{i,j} \quad (5.50)$$

where N is the number of electrons in the molecule, and where

$$\int d\tau_1 \Phi_i(1) F(1) \Phi_j(1) = \lambda_{i,j} \quad (5.51)$$

is a matrix representation of the Fock operator based on the filled molecular spin-orbitals. We can bring the Fock operator into a diagonal form by means of a unitary transformation. Letting ϵ_k , $k = 1, 2, 3, \dots, N$ be the diagonal elements, we then have:

$$\{F(1) - \epsilon_k\} \Phi_k(1) = 0 \quad k = 1, 2, 3, \dots, N \quad (5.52)$$

These simultaneous equations are called the *Hartree-Fock equations*.

The Fock operator $F(1)$ can be expressed in the form

$$F(1) = H^c(1) + \sum_{s=1}^N [J_s(1) - K_s(1)] \quad (5.53)$$

where

$$\begin{aligned}
H^c(1) &\equiv -\frac{\hbar^2}{2m_e}\nabla_1^2 + V(1) \\
J_s(1)\Phi_k(2) &\equiv \int d\tau_2 \Phi_s^*(2) \frac{e^2}{r_{1,2}} \Phi_s(2)\Phi_k(1) \\
K_s(1)\Phi_k(2) &\equiv \int d\tau_2 \Phi_s^*(2) \frac{e^2}{r_{1,2}} \Phi_s(1)\Phi_k(2)
\end{aligned} \tag{5.54}$$

Roothaan's equations

In his famous Ph.D. thesis, Clemens Roothaan applied the Hartree-Fock equations to the case where molecular spin-orbitals are represented as linear combinations of atomic spin-orbitals:

$$\Phi_k(1) = \sum_{b=1}^M \chi_a(1)C_{b,k} \tag{5.55}$$

Inserting this into the Hartree-Fock equations, multiplying from the left by $\chi_a^*(1)$ and integrating over the space and spin coordinates of electron 1, we obtain

$$\sum_{b=1}^M \int d\tau_1 \chi_a^*(1) [F(1) - \epsilon_k] \chi_b(1)C_{a,k} = 0 \tag{5.56}$$

or in matrix form

$$\sum_{b=1}^M [F_{a,b} - \epsilon_k S_{a,b}] C_{b,k} = 0 \tag{5.57}$$

where

$$\begin{aligned}
F_{a,b} &\equiv \int d\tau_1 \chi_a^*(1) F(1) \chi_b(1) \\
S_{a,b} &\equiv \int d\tau_1 \chi_a^*(1) \chi_b(1)
\end{aligned} \tag{5.58}$$

The matrix $S_{a,b}$ is a matrix of overlap integrals, and it must be taken into account because the atomic spin-orbitals are not necessarily orthonormal. Roothaan expressed the Fock matrix $F_{a,b}$ in the form

$$F_{a,b} = H_{a,b}^c + \sum_{c,d} P_{c,d} \Gamma_{(ab)cd} \tag{5.59}$$

where

$$\begin{aligned}
\Gamma_{(ab)cd} &\equiv \int d\tau_1 \int d\tau_2 \chi_a^*(1) \chi_c^*(2) \frac{e^2}{r_{1,2}} \chi_d(2) \chi_b(1) \\
&\quad - \int d\tau_1 \int d\tau_2 \chi_a^*(1) \chi_c^*(2) \frac{e^2}{r_{1,2}} \chi_d(1) \chi_b(2)
\end{aligned} \tag{5.60}$$

The density matrix $P_{c,d}$ is defined by the relationship

$$P_{c,d} \equiv \sum_{i=1}^M \nu_i C_{c,i}^* C_{d,i} \quad (5.61)$$

where

$$\nu_i \equiv \begin{cases} 1 & \text{filled spin orbitals} \\ 0 & \text{empty spin orbitals} \end{cases} \quad (5.62)$$

In Roothaan's method, one begins by evaluating and storing the the $M \times M \times M \times M$ dimensional matrix of integrals $\Gamma_{(ab)cd}$. An initial guess of the coefficients $C_{a,i}$ is then used to evaluate the core Hamiltonian, overlap and Fock matrices $H_{a,b}^c$, $S_{a,b}$ and $F_{a,b}$. By solving the matrix form of the Hartree-Fock equations, new and improved values of the coefficients are obtained. The process is repeated many times until it (hopefully) converges to an accurate self-consistent solution.

5.8 Koopmans' theorem

A Hartree-Fock-Roothaan calculation gives us a ground state N -electron wave function Δ_0 which is totally antisymmetric with respect to the exchange of the space and spin coordinates of any two electrons. The physicist J.C. Slater introduced a way of writing such a totally antisymmetric wave function as a determinant in the atomic spin-orbitals.

$$\Delta_0 = |\chi_\mu \chi_{\mu'} \chi_{\mu''} \cdots| \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_\mu(\mathbf{x}_1) & \chi_{\mu'}(\mathbf{x}_1) & \chi_{\mu''}(\mathbf{x}_1) & \cdots \\ \chi_\mu(\mathbf{x}_2) & \chi_{\mu'}(\mathbf{x}_2) & \chi_{\mu''}(\mathbf{x}_2) & \cdots \\ \chi_\mu(\mathbf{x}_3) & \chi_{\mu'}(\mathbf{x}_3) & \chi_{\mu''}(\mathbf{x}_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \quad (5.63)$$

involving atomic orbitals of the form

$$\chi_\mu(\mathbf{x}_i) \equiv \chi_{n,l,m,m_s}(\mathbf{x}_i) \equiv R_{n,l}(r_i) Y_{l,m}(\theta_i, \phi_i) \begin{cases} \alpha_i & m_s = 1/2 \\ \beta_i & m_s = -1/2 \end{cases} \quad (5.64)$$

In fact, the Hartree-Fock equations can be derived by assuming that the ground-state wave function has this form, and then using a variational principle to minimize the energy of the ground state (see, for example, ([Avery, 1976])). The ground state energy then has the

form:

$$\begin{aligned}
E_0 &= \int d\tau_1 \int d\tau_2 \cdots \int d\tau_N \Delta_0^* H \Delta_0 \\
&= \sum_{s=1}^N \int d\tau_1 \Phi_s^*(1) H^c(1) \Phi_s(1) \\
&+ \frac{1}{2} \sum_{s=1}^N \sum_{t=1}^N \int d\tau_1 \int d\tau_2 \Phi_s^*(1) \Phi_t^*(2) \frac{e^2}{r_{1,2}} \Phi_t(2) \Phi_s(1) \\
&- \frac{1}{2} \sum_{s=1}^N \sum_{t=1}^N \int d\tau_1 \int d\tau_2 \Phi_s^*(1) \Phi_t^*(2) \frac{e^2}{r_{1,2}} \Phi_t(1) \Phi_s(2) \tag{5.65}
\end{aligned}$$

The ground-state energy can also be expressed in terms of the core Hamiltonian matrix, the Fock matrix, and the density matrix, based on atomic orbitals:

$$E_0 = \frac{1}{2} \sum_{a=1}^M \sum_{b=1}^M P_{a,b} (H_{a,b}^c + F_{a,b}) \tag{5.66}$$

Since M , the number of basis functions, is larger than N , the number of electrons, solutions to the Hartree-Fock-Roothaan equations give us a number of states greater than the number of electrons. The lowest N of the resulting molecular orbitals are interpreted as being filled, while those of higher energy are seen as empty or “virtual” orbitals. If we lift an electron from a filled orbital to a virtual orbital, we change the occupation numbers, and hence we also change the density matrix. The resulting change in energy is given by

$$(\Delta E)_{\nu_i \rightarrow \nu'_i} = \sum_{a,b} \left[(P'_{a,b} - P_{a,b}) H_{a,b}^c + \frac{1}{2} \sum_{c,d} (P'_{a,b} P'_{c,d} - P_{a,b} P_{c,d}) \Gamma_{(ab)cd} \right] \tag{5.67}$$

In the special case where

$$\nu'_i = \nu_i - \delta_{i,k} \tag{5.68}$$

we are removing an electron from the molecule, in other words, turning it into a positive ion with a single electronic charge. The approximate change in energy is then

$$\Delta E_k = -\epsilon_k \tag{5.69}$$

where

$$\begin{aligned}
\epsilon_k &\equiv \int d\tau_1 \Phi_k^*(1) H^c(1) \Phi_k(1) \\
&+ \frac{1}{2} \sum_{t=1}^N \int d\tau_1 \int d\tau_2 \Phi_k^*(1) \Phi_t^*(2) \frac{e^2}{r_{1,2}} \Phi_t(2) \Phi_s(1) \\
&- \frac{1}{2} \sum_{t=1}^N \int d\tau_1 \int d\tau_2 \Phi_k^*(1) \Phi_t^*(2) \frac{e^2}{r_{1,2}} \Phi_t(1) \Phi_s(2) \tag{5.70}
\end{aligned}$$

is the eigenvalue of the Fock matrix belonging to the k th spin-orbital. This result is known as Koopmans' theorem.

5.9 Electron creation and annihilation operators

We shall now consider an alternative way of representing many-electron wave functions. We define a set of “electron creation operators”, $b_1^\dagger, b_2^\dagger, b_3^\dagger, \dots$ corresponding to the 1-electron spin-orbitals $\chi_1, \chi_2, \chi_3, \dots$. In this new notation, the Slater determinant

$$\Delta_\nu = |\chi_\mu \chi_{\mu'} \chi_{\mu''} \dots| \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_\mu(\mathbf{x}_1) & \chi_{\mu'}(\mathbf{x}_1) & \chi_{\mu''}(\mathbf{x}_1) & \dots \\ \chi_\mu(\mathbf{x}_2) & \chi_{\mu'}(\mathbf{x}_2) & \chi_{\mu''}(\mathbf{x}_2) & \dots \\ \chi_\mu(\mathbf{x}_3) & \chi_{\mu'}(\mathbf{x}_3) & \chi_{\mu''}(\mathbf{x}_3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \quad (5.71)$$

is represented by

$$|\Delta_\nu\rangle = b_\mu^\dagger b_{\mu'}^\dagger b_{\mu''}^\dagger \dots |0\rangle \quad (5.72)$$

We also introduce a set of “electron annihilation operators”, b_1, b_2, b_3, \dots corresponding to the same set of 1-electron spin-orbitals. When one of these operators acts on an N -electron state, it produces an $(N - 1)$ -electron state. We also impose the condition that when any electron annihilation operator acts on the “vacuum state” $|0\rangle$, it gives zero.

$$b_\mu |0\rangle = 0 \quad (5.73)$$

for all values of μ . The adjoint of this equation has the form

$$\langle 0|b_\mu^\dagger = 0 \quad (5.74)$$

The many-electron wave functions expressed in terms of electron creation and annihilation operators will be totally antisymmetric with respect to the exchange of the space and spin coordinates of any two electrons provided that these operators obey the following “anticommutation relations”:

$$\begin{aligned} b_i^\dagger b_j + b_j b_i^\dagger &= \delta_{i,j} \\ b_i^\dagger b_j^\dagger + b_j^\dagger b_i^\dagger &= 0 \\ b_i b_j + b_j b_i &= 0 \end{aligned} \quad (5.75)$$

The vacuum state is assumed to be normalized, so that

$$\langle 0|0\rangle = 1 \quad (5.76)$$

The adjoint of a many-electron state is represented by

$$\langle \Delta_\nu| = \langle 0|\dots, b_{\mu''}, b_{\mu'}, b_\mu \quad (5.77)$$

Then, using the anticommutation relations to move the annihilation operators over to the right, where they annihilate the vacuum state, we find that

$$\begin{aligned}
 \langle \Delta_\nu | \Delta_\nu \rangle &= \langle 0 | \cdots, b_{\mu''}, b_{\mu'}, b_\mu, b_\mu^\dagger, b_{\mu'}^\dagger, b_{\mu''}^\dagger, \cdots | 0 \rangle \\
 &= \langle 0 | \cdots, b_{\mu''}, b_{\mu'}, b_{\mu'}^\dagger, b_{\mu''}^\dagger, \cdots | 0 \rangle \\
 &= \langle 0 | \cdots, b_{\mu''}, b_{\mu''}^\dagger, \cdots | 0 \rangle \\
 &\vdots \\
 &= \langle 0 | 0 \rangle = 1
 \end{aligned} \tag{5.78}$$

so that the many-electron state is properly normalized.

1-electron and 2-electron operators

In our new notation, 1-electron are represented by

$$\mathcal{F} = \sum_{i,j} f_{i,j} b_i^\dagger b_j \tag{5.79}$$

where

$$f_{i,j} \equiv \int d\tau_1 \chi_i^*(1) f(1) \chi_j(1) \tag{5.80}$$

Examples of 1-electron operators are the core Hamiltonian or externally applied electric or magnetic fields.

By contrast, 2-electron operators involve the interactions of electrons with each other, an example being the Coulomb repulsion between electrons. In our new notation, 2-electron operators have the form

$$\mathcal{G} = \frac{1}{2} \sum_{s,t,u,v} g_{st|uv} b_s^\dagger b_t^\dagger b_u b_v \tag{5.81}$$

where

$$g_{st|uv} = \int d\tau_1 \int d\tau_2 \chi_s^*(1) \chi_t^*(2) g(1,2) \chi_u(2) \chi_v(1) \tag{5.82}$$

5.10 Quantum chemistry and the development of computers

Some personal memories of early computers

I hope that readers will forgive me if I tell them of my own personal memories of early computers:

When I arrived at Imperial College (then part of the University of London) in 1962, I worked with a crystallographic group that using the Mercury computer at University College to do the calculations needed to arrive at molecular structures. This gave me the chance to use Mercury to do quantum chemical calculations. I always went over to University College with the crystallographers at night, because time on the computer was so expensive that we could only afford to use it at night. I would make a bed for myself out of three rather hard chairs in a row and would try to sleep. At 3 AM or 4 AM they would wake me up and would say “Now it’s your turn”.

Mercury was as big as a house, but could do far less than a modern laptop. It had 50,000 or so vacuum tubes which required cooling. The cooling system sometimes broke down, and one or another of the vacuum tubes sometimes failed, so one had to be grateful for the periods when Mercury was working. Our programs were written on punched tape in a language called CHLF3. (The letters stood for Cambridge, London, Harwell and Farnborough, the four places that had Mercurys). After we had read the paper tape into the computer, the program was converted into a magnetic form on a rapidly rotating drum, and then checked against the original input. If it did not check, we had a so-called “drum parity”, which meant that we had to stop the computer and restart it by hand, using a bewildering array of manual controls.

After finishing the work on Mercury at 6 AM or so, I would walk home, passing through the almost-deserted streets of Soho, and seeing pale-faced teenagers who had been up all night, high on amphetamines. They were sitting on the pavement near an underground station, waiting for it to open.

After we had used Mercury for two years or so, IBM gave Imperial College one of their early computers. Using this was much better. Programs for the IBM machine were written on punched cards. We just went over to the machine with our punched cards and stood in line to have them read into the computer. Then a few minutes later we were handed a printout of the output.

The IBM was much better than the machines that were available in eastern Europe, and for this reason I was contacted by Janos Ladik and his group at the Hungarian Academy of Science, who proposed a collaboration. We worked together for several years, calculating the electronic structure of a number of polypeptides and polynucleotides.

In 1965, Janos Ladik invited me to attend a meeting of quantum theorists and computer scientists from both East and West, held at a town on the Hungarian Puszta, the great Hungarian plain east of Budapest. Both Charles Coulson and Per-Olov Löwdin were there, as well as many scientists from the eastern side of the Iron Curtain. At the meeting, Enrico Clementi spoke about computer programs that he had developed for performing *ab-initio*¹ calculation of the electronic structure of molecules. Clementi was an important IBM scientist, and he had his own laboratory with a large computer which he could use as he liked. The programs that he described to us took hundreds of hours to complete an electronic structure calculation on a single molecule.

¹*ab-initio* is a Latin expression meaning “from the beginning”. Such programs are completely free of input parameters based on experiments.



Figure 5.21: **Enrico Clementi (born 1931) explained to us that microminiaturization would soon make computers hundreds of times faster, smaller and less expensive. He was completely right.**

In the question period after Clementi’s lecture, someone from the audience said: “It’s all right for you, Clementi. You can use hundreds of hours on a single calculation if you want to, because you are sitting at IBM with your own dedicated computer. But what about the rest of us? What good are these programs to us?”

Clementi answered: “In a few years, computers will be hundreds of times faster, and they will also be cheaper.” The audience asked: “And how will this happen?”. Clementi answered: “Through microminiaturization.” He was completely right. That was exactly what happened.

The invention of transistors; Microelectronics

The problem of unreliable vacuum tubes was solved in 1948 by John Bardeen, William Shockley and Walter Brattain of the Bell Telephone Laboratories. Application of quantum theory to solids had led to an understanding of the electrical properties of crystals. Like atoms, crystals were found to have allowed and forbidden energy levels.

The allowed energy levels for an electron in a crystal were known to form bands, i.e., some energy ranges with many allowed states (allowed bands), and other energy ranges with none (forbidden bands). The lowest allowed bands were occupied by electrons, while higher bands were empty. The highest filled band was called the “valence band”, and the lowest empty band was called the “conduction band”.

According to quantum theory, whenever the valence band of a crystal is only partly filled, the crystal is a conductor of electricity; but if the valence band is completely filled with electrons, the crystal is an electrical insulator. (A completely filled band is analogous to a room so packed with people that none of them can move.)

In addition to conductors and insulators, quantum theory predicted the existence of “semiconductors” - crystals where the valence band is completely filled with electrons, but where the energy gap between the conduction band and the valence band is very small. For example, crystals of the elements silicon and germanium are semiconductors. For such a crystal, thermal energy is sometimes enough to lift an electron from the valence band to the conduction band.

Bardeen, Shockley and Brattain found ways to control the conductivity of germanium crystals by injecting electrons into the conduction band, or alternatively by removing electrons from the valence band. They could do this by “doping” the crystals with appropriate impurities, or by injecting electrons with a special electrode. The semiconducting crystals whose conductivity was controlled in this way could be used as electronic valves, in place of vacuum tubes.

By the 1960’s, replacement of vacuum tubes by transistors in electronic computers had led not only to an enormous increase in reliability and a great reduction in cost, but also to an enormous increase in speed. It was found that the limiting factor in computer speed was the time needed for an electrical signal to propagate from one part of the central processing unit to another. Since electrical impulses propagate with the speed of light, this time is extremely small; but nevertheless, it is the limiting factor in the speed of electronic computers.

Integrated circuits

In order to reduce the propagation time, computer designers tried to make the central processing units very small; and the result was the development of integrated circuits and microelectronics. (Another motive for miniaturization of electronics came from the requirements of space exploration.)

Integrated circuits were developed in which single circuit elements were not manufactured separately. Instead, the whole circuit was made at one time. An integrated circuit is a sandwich-like structure, with conducting, resisting and insulating layers interspersed with layers of germanium or silicon, “doped ” with appropriate impurities. At the start of the manufacturing process, an engineer makes a large drawing of each layer. For example, the drawing of a conducting layer would contain pathways which fill the role played by wires in a conventional circuit, while the remainder of the layer would consist of areas destined to be etched away by acid.

The next step is to reduce the size of the drawing and to multiply it photographically. The pattern of the layer is thus repeated many times, like the design on a piece of wallpaper. The multiplied and reduced drawing is then focused through a reversed microscope onto the surface to be etched.

Successive layers are built up by evaporating or depositing thin films of the appropriate substances onto the surface of a silicon or germanium wafer. If the layer being made is to be conducting, the surface would consist of an extremely thin layer of copper, covered with a photosensitive layer called a “photoresist”. On those portions of the surface receiving light from the pattern, the photoresist becomes insoluble, while on those areas not receiving

light, the photoresist can be washed away.

The surface is then etched with acid, which removes the copper from those areas not protected by photoresist. Each successive layer of a wafer is made in this way, and finally the wafer is cut into tiny “chips”, each of which corresponds to one unit of the wallpaper-like pattern.

Although the area of a chip may be much smaller than a square centimeter, the chip can contain an extremely complex circuit. A typical programmable minicomputer or “microprocessor”, manufactured during the 1970’s, could have 30,000 circuit elements, all of which were contained on a single chip. By 1986, more than a million transistors were being placed on a single chip.

As a result of miniaturization, the speed of computers rose steadily. In 1960, the fastest computers could perform a hundred thousand elementary operations in a second. By 1970, the fastest computers took less than a second to perform a million such operations. In 1987, a computer called GF11 was designed to perform 11 billion floating-point operations (flops) per second.

GF11 (Gigaflop 11) is a scientific parallel-processing machine constructed by IBM. Approximately ten floating-point operations are needed for each machine instruction. Thus GF11 runs at the rate of approximately a thousand million instructions per second (1,100 MIPS). The high speed achieved by parallel-processing machines results from dividing a job into many sub-jobs on which a large number of processing units can work simultaneously.

Computer memories have also undergone a remarkable development. In 1987, the magnetic disc memories being produced could store 20 million bits of information per square inch; and even higher densities could be achieved by optical storage devices. (A “bit” is the unit of information. For example, the number 25, written in the binary system, is 11001. To specify this 5-digit binary number requires 5 bits of information. To specify an n-digit binary number requires n bits of information. Eight bits make a “byte”.)

In the 1970’s and 1980’s, computer networks were set up linking machines in various parts of the world. It became possible (for example) for a scientist in Europe to perform a calculation interactively on a computer in the United States just as though the distant machine were in the same room; and two or more computers could be linked for performing large calculations. It also became possible to exchange programs, data, letters and manuscripts very rapidly through the computer networks.

Suggestions for further reading

1. Akhlesh Lakhtakia. *Models and Modelers of Hydrogen: Thales, Thomson, Rutherford, Bohr, Sommerfeld, Goudsmit, Heisenberg, Schrödinger, Dirac, Sallhofer*. World Scientific, (1996).
2. O’Connor, John J.; Robertson, Edmund F., *Erwin Schrödinger*, MacTutor History of Mathematics archive, University of St Andrews.
3. Halpern, Paul. *Einstein’s Dice and Schrödinger’s Cat*. Perseus Books Group, (2015).
4. Heitler, W. *Erwin Schrödinger. 1887-1961*. Biographical Memoirs of Fellows of the Royal Society. 7: 221-226, (1961).

5. Farmelo, Graham. *The Strangest Man: The Hidden Life of Paul Dirac, Quantum Genius*. Faber and Faber, (2009).
6. Dalitz, R. H.; Peierls, R. *Paul Adrien Maurice Dirac. 8 August 1902 - 20 October 1984*. Biographical Memoirs of Fellows of the Royal Society. 32: 137-185, (1986).
7. Douglas, Sandy, *Some Memories of EDSAC 1: 1950-1952*, Ann. Hist. Comp., Vol. 1, No. 2, 1980, pp. 98-99, 208.
8. Medwick, Paul A., *Douglas Hartree and Early Computations in Quantum Mechanics*, Ann. Hist. Comp., Vol. 10, No. 2, 1988, pp. 105-112.
9. Hartree, D. R., *Calculating Machines-Recent and Prospective Developments*, Cambridge Univ. Press, Cambridge, UK, 1947.
10. Hartree, D. R., *Automatic Calculating Machines*, Mathematical Gazette, Vol. 34, Dec. 1950, pp. 241-252.
11. Graham, L. (1982). *The reception of Einstein's ideas: Two examples from contrasting political cultures*. In Holton, G. and Elkana, Y. (Eds.) *Albert Einstein: Historical and cultural perspectives*. Princeton, NJ: Princeton UP, pp. 107-136
12. Fock, V. A. (1964). *The Theory of Space, Time and Gravitation*. Macmillan.
13. Lewis, G. N. (1916), *The Atom and the Molecule*, J. Am. Chem. Soc., 38 (4): 762-85,
14. Miburo, Barnabe B. (1993), *Simplified Lewis Structure Drawing for Non-science Majors*, J. Chem. Educ., 75 (3): 317
15. Coffey, Patrick (2008). *Cathedrals of Science: The Personalities and Rivalries That Made Modern Chemistry*. Oxford University Press.
16. Davenport, Derek A. (1996). *The Many Lives of Linus Pauling: A Review of Reviews*. Journal of Chemical Education. 73 (9): A210.
17. Hargittai, István (2000). Hargittai, Magdolna (ed.). *Candid science: conversations with famous chemists (Reprinted ed.)*. London: Imperial College Press.
18. Marinacci, Barbara, ed. (1995). *Linus Pauling: In His Own Words; Selected Writings, Speeches, and Interviews*. Introduction by Linus Pauling. New York: Simon & Schuster.
19. Sturchio, Jeffrey L. (April 6, 1987). *Linus C. Pauling, Transcript of an Interview Conducted by Jeffrey L. Sturchio in Denver, Colorado on 6 April 1987 (PDF)*. Philadelphia, PA: Chemical Heritage Foundation.
20. Avery, John, *The Quantum Theory of Atoms, Molecules and Photons*, McGraw-Hill, (1972).
21. Avery, John, *Creation and Annihilation Operators*, McGraw-Hill, (1976).
22. Avery, John, Dahl, J.P., and Hansen, A.E., editors, *Understanding Molecular Properties*, Reidel, Dordrecht, (1987).
23. Avery, John, *Hyperspherical Harmonics; Applications in Quantum Theory*, Kluwer Academic Publishers, Dordrecht, (1989).
24. Avery, John, editor, *The Chemical Bond*, Special Issue of the International Journal of Quantum Chemistry, Volume 58, (1996).
25. Avery, John, *Hyperspherical Harmonics and Generalized Sturmians*, Kluwer Academic Publishers, Dordrecht, Netherlands, (2000).

26. Avery, James and Avery, John, *Generalized Sturmians and Atomic Spectra*, World Scientific, (2000).
27. Avery, John, Rettrup, Sten and Avery, James, *Symmetry Adapted Basis Functions*. World Scientific, (2012).
28. Avery, James and Avery, John, *Hyperspherical Harmonics and Their Physical Applications*, World Scientific, (2017).
29. H. Babbage, *Babbages Calculating Engines: A Collection of Papers by Henry Prevost Babbage*, MIT Press, (1984).
30. A.M. Turing, *The Enigma of Intelligence*, Burnett, London (1983).
31. Ft. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford University Press, (1989).
32. S. Wolfram, *A New Kind of Science*, Wolfram Media, Champaign IL, (2002).
33. A.M. Turing, *On computable numbers, with an application to the Entscheidungsproblem*, Proc. Lond. Math. Soc. Ser 2, **42**, (1937). Reprinted in M. David Ed., *The Undecidable*, Raven Press, Hewlett N.Y., (1965).
34. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
35. J. Shurkin, *Engines of the Mind: A History of Computers*, W.W. Norton, (1984).
36. J. Palfreman and D. Swade, *The Dream Machine: Exploring the Computer Age*, BBC Press (UK), (1991).
37. T.J. Watson, Jr. and P. Petre, *Father, Son, and Co.*, Bantam Books, New York, (1991).
38. A. Hodges, Alan Turing: *The Enigma*, Simon and Schuster, (1983).
39. H.H. Goldstein, *The Computer from Pascal to Von Neumann*, Princeton University Press, (1972).
40. C.J. Bashe, L.R. Johnson, J.H. Palmer, and E.W. Pugh, *IBM's Early Computers*, Vol. 3 in the History of Computing Series, MIT Press, (1986).
41. K.D. Fishman, *The Computer Establishment*, McGraw-Hill, (1982).
42. S. Levy, *Hackers*, Doubleday, (1984).
43. S. Franklin, *Artificial Minds*, MIT Press, (1997).
44. P. Freiberger and M. Swaine, *Fire in the Valley: The Making of the Personal Computer*, Osborne/McGraw-Hill, (1984).
45. R.X. Cringely, *Accidental Empires*, Addison-Wesley, (1992).
46. R. Randell editor, *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
47. H. Lukoff, *From Dits to Bits*, Robotics Press, (1979).
48. D.E. Lundstrom, *A Few Good Men from Univac*, MIT Press, (1987).
49. D. Rutland, *Why Computers Are Computers (The SWAC and the PC)*, Wren Publishers, (1995).
50. P.E. Ceruzzi, *Reckoners: The Prehistory of the Digital Computer, from Relays to the Stored Program Concept, 1935-1945*, Greenwood Press, Westport, (1983)
51. S.G. Nash, *A History of Scientific Computing*, Adison-Wesley, Reading Mass., (1990).

52. P.E. Ceruzzi, *Crossing the divide: Architectural issues and the emergence of stored programme computers, 1935-1953*, IEEE Annals of the History of Computing, **19**, 5-12, January-March (1997).
53. P.E. Ceruzzi, *A History of Modern Computing*, MIT Press, Cambridge MA, (1998).
54. K. Zuse, *Some remarks on the history of computing in Germany*, in *A History of Computing in the 20th Century*, N. Metropolis et al. editors, 611-627, Academic Press, New York, (1980).
55. A.R. Mackintosh, *The First Electronic Computer*, Physics Today, March, (1987).
56. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
57. A. Hodges, *Alan Turing: The Enigma*, Simon and Schuster, New York, (1983).
58. A. Turing, *On computable numbers with reference to the Entscheidungsproblem*, Journal of the London Mathematical Society, **II**, **2**, **42**, 230-265 (1937).
59. J. von Neumann, *The Computer and the Brain*, Yale University Press, (1958).
60. I.E. Sutherland, *Microelectronics and computer science*, Scientific American, 210-228, September (1977).
61. W. Aspray, *John von Neumann and the Origins of Modern Computing*, M.I.T. Press, Cambridge MA, (1990, 2nd ed. 1992).
62. W. Aspray, *The history of computing within the history of information technology*, History and Technology, **11**, 7-19 (1994).
63. G.F. Luger, *Computation and Intelligence: Collected Readings*, MIT Press, (1995).
64. Z.W. Pylyshyn, *Computation and Cognition: Towards a Foundation for Cognitive Science*, MIT Press, (1986).
65. D.E. Shasha and C. Lazere, *Out of Their Minds: The Creators of Computer Science*, Copernicus, New York, (1995).
66. W. Aspray, *An annotated bibliography of secondary sources on the history of software*, Annals of the History of Computing **9**, 291-243 (1988).
67. R. Kurzweil, *The Age of Intelligent Machines*, MIT Press, (1992).
68. S.L. Garfinkel and H. Abelson, eds., *Architects of the Information Society: Thirty-Five Years of the Laboratory for Computer Sciences at MIT*, MIT Press, (1999).
69. J. Haugeland, *Artificial Intelligence: The Very Idea*, MIT Press, (1989).
70. M.A. Boden, *Artificial Intelligence in Psychology: Interdisciplinary Essays*, MIT Press, (1989).
71. J.W. Cortada, *A Bibliographic Guide to the History of Computer Applications, 1950-1990*, Greenwood Press, Westport Conn., (1996).
72. M. Campbell-Kelly and W. Aspry, *Computer: A History of the Information Machine*, Basic Books, New York, (1996).
73. B.I. Blum and K. Duncan, editors, *A History of Medical Informatics*, ACM Press, New York, (1990).
74. J.-C. Guedon, La Planete Cyber, *Internet et Cyberspace*, Gallimard, (1996).
75. S. Augarten, *Bit by Bit: An Illustrated History of Computers*, Unwin, London, (1985).
76. N. Wiener, *Cybernetics; or Control and Communication in the Animal and the Machine*, The Technology Press, John Wiley and Sons, New York, (1948).
77. W.R. Ashby, *An Introduction to Cybernetics*, Chapman and Hall, London, (1956).

78. M.A. Arbib, *A partial survey of cybernetics in eastern Europe and the Soviet Union*, Behavioral Sci., **11**, 193-216, (1966).
79. A. Rosenblueth, N. Wiener and J. Bigelow, *Behavior, purpose and teleology*, Phil. Soc. **10** (1), 18-24 (1943).
80. N. Wiener and A. Rosenblueth, *Conduction of impulses in cardiac muscle*, Arch. Inst. Cardiol. Mex., **16**, 205-265 (1946).
81. H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems. Transactions of sixth-tenth conferences*, Josiah J. Macy Jr. Foundation, New York, (1950- 1954).
82. W.S. McCulloch and W. Pitts, *A logical calculus of ideas immanent in nervous activity*, Bull. Math. Biophys., **5**, 115-133 (1943).
83. W.S. McCulloch, *An Account of the First Three Conferences on Teleological Mechanisms*, Josiah Macy Jr. Foundation, (1947).
84. G.A. Miller, *Languages and Communication*, McGraw-Hill, New York, (1951).
85. G.A. Miller, *Statistical behavioristics and sequences of responses*, Psychol. Rev. **56**, 6 (1949).
86. G. Bateson, *Bali - the value system of a steady state*, in M. Fortes, editor, *Social Structure Studies Presented to A.R. Radcliffe-Brown*, Clarendon Press, Oxford, (1949).
87. G. Bateson, *Communication, the Social Matrix of Psychiatry*, Norton, (1951).
88. G. Bateson, *Steps to an Ecology of Mind*, Chandler, San Francisco, (1972).
89. G. Bateson, *Communication et Societe*, Seuil, Paris, (1988).
90. S. Heims, *Gregory Bateson and the mathematicians: From interdisciplinary interactions to societal functions*, J. History Behavioral Sci., **13**, 141-159 (1977).
91. S. Heims, *John von Neumann and Norbert Wiener. From Mathematics to the Technology of Life and Death*, MIT Press, Cambridge MA, (1980).
92. S. Heims, *The Cybernetics Group*, MIT Press, Cambridge MA, (1991).
93. G. van de Vijver, *New Perspectives on Cybernetics (Self-Organization, Autonomy and Connectionism)*, Kluwer, Dordrecht, (1992).
94. A. Bavelas, *A mathematical model for group structures*, Appl. Anthropol. **7** (3), 16 (1948).
95. P. de Latil, *La Pensee Artificielle - Introduction a la Cybernetique*, Gallimard, Paris, (1953).
96. L.K. Frank, G.E. Hutchinson, W.K. Livingston, W.S. McCulloch and N. Wiener, *Teleological Mechanisms*, Ann. N.Y. Acad. Sci. **50**, 187- 277 (1948).
97. H. von Foerster, *Quantum theory of memory*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems. Transactions of the sixth conferences*, Josiah J. Macy Jr. Foundation, New York, (1950).
98. H. von Foerster, *Observing Systems*, Intersystems Publications, California, (1984).
99. H. von Foerster, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer, New York, (2002).
100. M. Newborn, *Kasparov vs. Deep Blue: Computer Chess Comes of age*, Springer Verlag, (1996).

101. K.M. Colby, *Artificial Paranoia: A Computer Simulation of the Paranoid Process*, Pergamon Press, New York, (1975).
102. J.Z. Young, *Discrimination and learning in the octopus*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the ninth conference, Josiah J. Macy Jr. Foundation, New York, (1953).
103. M.J. Apter and L. Wolpert, *Cybernetics and development*. I. Information theory, *J. Theor. Biol.* **8**, 244-257 (1965).
104. H. Atlan, *L'Organization Biologique et la Theorie de l'Information*, Hermann, Paris, (1972).
105. H. Atlan, *On a formal definition of organization*, *J. Theor. Biol.* **45**, 295-304 (1974).
106. H. Atlan, *Organization du vivant, information et auto-organization*, in Volume Symposium 1986 de l'Encyclopediea Universalis, pp. 355-361, Paris, (1986).
107. E.R. Kandel, Nerve cells and behavior, *Scientific American*, **223**, 57-70, July, (1970).
108. E.R. Kandel, *Small systems of neurons*, *Scientific American*, **241 no.3**, 66-76 (1979).
109. A.K. Katchalsky et al., *Dynamic patterns of brain cell assemblies*, *Neurosciences Res. Prog. Bull.*, **12 no.1**, (1974).
110. G.E. Moore, *Cramming more components onto integrated circuits*, *Electronics*, April 19, (1965).
111. P. Gelsinger, P. Gargini, G. Parker and A. Yu, *Microprocessors circa 2000*, *IEEE Spectrum*, October, (1989).
112. P. Baron, *On distributed communications networks*, *IEEE Trans. Comm. Systems*, March (1964).
113. V.G. Cerf and R.E. Khan, *A protocol for packet network intercommunication*, *Trans. Comm. Tech.* **COM-22**, **V5**, 627-641, May (1974).
114. L. Kleinrock, *Communication Nets: Stochastic Message Flow and Delay*, McGraw-Hill, New York, (1964).
115. L. Kleinrock, *Queueing Systems: Vol. II, Computer Applications*, Wiley, New York, (1976).
116. R. Kahn, editor, *Special Issue on Packet Communication Networks*, *Proc. IEEE*, **66**, November, (1978).
117. L.G. Roberts, *The evolution of packet switching*, *Proc. of the IEEE* **66**, 1307-13, (1978).
118. J. Abbate, *The electrical century: Inventing the web*, *Proc. IEEE* **87**, November, (1999).
119. J. Abbate, *Inventing the Internet*, MIT Press, Cambridge MA, (1999).
120. J.C. McDonald, editor, *Fundamentals of Digital Switching, 2nd Edition*, Plenum, New York, (1990).
121. B. Metcalfe, *Packet Communication, Peer-to-Peer Communication*, San Jose Calif, (1996).
122. T. Berners-Lee, *The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*, Harper San Francisco, (1999).

123. J. Clark, *Netscape Time: The Making of the Billion-Dollar Start-Up That Took On Microsoft*, St. Martin's Press, New York, (1999).
124. J. Wallace, *Overdrive: Bill Gates and the Race to Control Cyberspace*, Wiley, New York, (1997).
125. P. Cunningham and F. Froschl, *The Electronic Business Revolution*, Springer Verlag, New York, (1999).
126. J.L. McKenny, *Waves of Change: Business Evolution Through Information Technology*, Harvard Business School Press, (1995).
127. M.A. Cosumano, *Competing on Internet Time: Lessons From Netscape and Its Battle with Microsoft*, Free Press, New York, (1998).
128. F.J. Dyson, *The Sun, the Genome and the Internet: Tools of Scientific Revolutions*, Oxford University Press, (1999).
129. L. Bruno, *Fiber Optimism: Nortel, Lucent and Cisco are battling to win the high-stakes fiber-optics game*, Red Herring, June (2000).
130. N. Cochrane, *We're insatiable: Now it's 20 million million bytes a day*, Melbourne Age, January 15, (2001).
131. K.G. Coffman and A.N. Odlyzko, The size and growth rate of the Internet, First Monday, October, (1998).
132. C.A. Eldering, M.L. Sylla, and J.A. Eisenach, *Is there a Moore's law for bandwidth?*, IEEE Comm. Mag., 2-7, October, (1999).
133. G. Gilder, *Fiber keeps its promise: Get ready, bandwidth will triple each year for the next 25 years*, Forbes, April 7, (1997).
134. A.M. Noll, *Does data traffic exceed voice traffic?*, Comm. ACM, 121- 124, June, (1999).
135. B. St. Arnaud, J. Coulter, J. Fitchett, and S. Mokbel, *Architectural and engineering issues for building an optical Internet*, Proc. Soc. Optical Eng. (1998).
136. M. Weisner, *The computer for the 21st century*, Scientific American, September, (1991).
137. R. Wright, *Three Scientists and Their Gods*, Time Books, (1988).
138. S. Nora and A. Mine, *The Computerization of Society*, MIT Press, (1981).
139. T. Forester, *Computers in the Human Context: Information Theory, Productivity, and People*, MIT Press, (1989).
140. J. Hoffmeyer, *Some semiotic aspects of the psycho-physical relation: the endo-exosemiotic boundary*, in Biosemiotics. The Semiotic Web, T.A. Sebeok and J. Umiker-Sebeok, editors, Mouton de Gruyter, Berlin/New York, (1991).
141. J. Hoffmeyer, *The swarming cyberspace of the body*, Cybernetics and Human Knowing, **3(1)**, 1-10 (1995).
142. J. Hoffmeyer, *Signs of Meaning in the Universe*, Indiana University Press, Bloomington IN, (1996).
143. J. Hoffmeyer, *Biosemiotics: Towards a new synthesis in biology*, European J. Semiotic Stud. **9(2)**, 355-376 (1997).

144. J. Hoffmeyer and C. Emmeche, *Code-duality and the semiotics of nature*, in *On Semiotic Modeling*, M. Anderson and F. Merrell, editors, Mouton de Gruyter, New York, (1991).
145. C. Emmeche and J. Hoffmeyer, *From language to nature - The semiotic metaphor in biology*, *Semiotica*, **84**, 1-42 (1991).
146. C. Emmeche, *The biosemiotics of emergent properties in a pluralist ontology*, in *Semiosis, Evolution, Energy: Towards a Reconceptualization of the Sign*, E. Taborsky, editor, Shaker Verlag, Aachen, (1999).
147. S. Brier, *Information and consciousness: A critique of the mechanistic concept of information*, in *Cybernetics and Human Knowing*, **1(2/3)**, 71-94 (1992).
148. S. Brier, *Ciber-Semiotics: Second-order cybernetics and the semiotics of C.S. Peirce*, Proceedings from the Second European Congress on System Science, Prague, October 5-8, 1993, AFCET, (1993).
149. S. Brier, *A cybernetic and semiotic view on a Galilean theory of psychology*, *Cybernetics and Human Knowing*, **2 (2)**, 31-46 (1993).
150. S. Brier, *Cybersemiotics: A suggestion for a transdisciplinary framework for description of observing, anticipatory, and meaning producing systems*, in D.M. Dubois, editor, *Computing Anticipatory Systems, CASYS - First International Conference*, Liege, Belgium 1997, AIP Conference Proceedings no. 437, (1997).
151. S. Oyama, *The Ontogeny of Information*, Cambridge University Press, (1985).
152. J. Hoffmeyer, *The swarming cyberspace of the body*, *Cybernetics and Human Knowing*, **3(1)**, 1-10 (1995).
153. J.L. Casti and A. Karlqvist, editors, *Complexity, Language, and Life: Mathematical Approaches*, Springer, Berlin, (1985).
154. H. Maturana and F. Varla, *Autopoiesis and Cognition: The Realization of the Living*, Reidel, London, (1980).
155. J. Mingers, *Self-Producing Systems: Implications and Application of Autopoiesis*, Plenum Press, New York, (1995).
156. J. Buchler, editor, *Philosophical Writings of Peirce: Selected and Edited with an Introduction by Justus Buchler*, Dover Publications, New York, (1955).
157. T.L. Short, *Peirce's semiotic theory of the self*, *Semiotica*, **91 (1/2)**, 109-131 (1992).
158. J. von Uexküll, *Umwelt und Innenwelt der Tiere. 2. verm. und verb. Aufl.*, Springer, Berlin, (1921).
159. J. von Uexküll, *The theory of meaning*, *Semiotica*, **42(1)**, 25-87 (1982 [1940]).
160. T. von Uexküll, *Introduction: Meaning and science in Jacob von Uexküll's concept of biology*, *Semiotica*, **42**, 1-24 (1982).
161. T. von Uexküll, *Medicine and semiotics*, *Semiotica*, **61**, 201-217 (1986).
162. G. Bateson, *Form, substance, and difference. Nineteenth Annual Korzybski Memorial Lecture*, (1970). Reprinted in G. Bateson, *Steps to an Ecology of Mind*, Balentine Books, New York, (1972), pp. 448-464.
163. G. Bateson, *Mind and Nature: A Necessary Unity*, Bantam Books, New York, (1980).
164. G. Bateson, *Sacred Unity: Further Steps to an Ecology of Mind*, Harper Collins, New York, (1991).

165. J. Ruesch and G. Bateson, *Communication*, Norton, New York, (1987).
166. E.F. Yates, *Semiotics as a bridge between information (biology) and dynamics (physics)*, *Recherches Semiotiques/Semiotic Inquiry* **5**, 347- 360 (1985).
167. T.A. Sebeok, *Communication in animals and men*, *Language*, **39**, 448-466 (1963).
168. T.A. Sebeok, *The Sign and its Masters*, University of Texas Press, (1979).
169. P. Bouissac, *Ecology of semiotic space: Competition, exploitation, and the evolution of arbitrary signs*, *Am. J. Semiotics*, **10**, 145-166 (1972).
170. F. Varla, *Autopoiesis: A Theory of Living Organization*, North Holland, New York, (1986).
171. R. Posner, K. Robins and T.A. Sebeok, editors, *Semiotics: A Handbook of the Sign-Theoretic Foundations of Nature and Culture*, Walter de Gruyter, Berlin, (1992).
172. R. Paton, *The ecologies of hereditary information*, *Cybernetics and Human Knowing*, **5(4)**, 31-44 (1998).
173. T. Stonier, *Information and the Internal Structure of the Universe*, Springer, Berlin, (1990).
174. T. Stonier, *Information and Meaning: An Evolutionary Perspective*, Springer, Berlin, (1997).
175. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
176. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
177. Alan Turing, *The Enigma of Intelligence*, Burnett, London (1983).
178. R. Randell (editor), *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
179. Allan R. Mackintosh, *The First Electronic Computer*, *Physics Today*, March, (1987).
180. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
181. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
182. R. Randell (editor), *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
183. Allan R. Mackintosh, *The First Electronic Computer*, *Physics Today*, March, (1987).
184. H. Babbage, *Babbages Calculating Engines: A Collection of Papers by Henry Prevost Babbage*, MIT Press, (1984).
185. A.M. Turing, *The Enigma of Intelligence*, Burnett, London (1983).
186. Ft. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford University Press, (1989).
187. S. Wolfram, *A New Kind of Science*, Wolfram Media, Champaign IL, (2002).
188. A.M. Turing, *On computable numbers, with an application to the Entscheidungsproblem*, *Proc. Lond. Math. Soc. Ser 2*, **42**, (1937). Reprinted in M. David Ed., *The Undecidable*, Raven Press, Hewlett N.Y., (1965).
189. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
190. J. Shurkin, *Engines of the Mind: A History of Computers*, W.W. Norton, (1984).

191. J. Palfreman and D. Swade, *The Dream Machine: Exploring the Computer Age*, BBC Press (UK), (1991).
192. T.J. Watson, Jr. and P. Petre, *Father, Son, and Co.*, Bantam Books, New York, (1991).
193. A. Hodges, Alan Turing: *The Enigma*, Simon and Schuster, (1983).
194. H.H. Goldstein, *The Computer from Pascal to Von Neumann*, Princeton University Press, (1972).
195. C.J. Bashe, L.R. Johnson, J.H. Palmer, and E.W. Pugh, *IBM's Early Computers*, Vol. 3 in the History of Computing Series, MIT Press, (1986).
196. K.D. Fishman, *The Computer Establishment*, McGraw-Hill, (1982).
197. S. Levy, *Hackers*, Doubleday, (1984).
198. S. Franklin, *Artificial Minds*, MIT Press, (1997).
199. P. Freiberger and M. Swaine, *Fire in the Valley: The Making of the Personal Computer*, Osborne/McGraw-Hill, (1984).
200. R.X. Cringely, *Accidental Empires*, Addison-Wesley, (1992).
201. R. Randell editor, *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
202. H. Lukoff, *From Dits to Bits*, Robotics Press, (1979).
203. D.E. Lundstrom, *A Few Good Men from Univac*, MIT Press, (1987).
204. D. Rutland, *Why Computers Are Computers (The SWAC and the PC)*, Wren Publishers, (1995).
205. P.E. Ceruzzi, *Reckoners: The Prehistory of the Digital Computer, from Relays to the Stored Program Concept, 1935-1945*, Greenwood Press, Westport, (1983)
206. S.G. Nash, *A History of Scientific Computing*, Adison-Wesley, Reading Mass., (1990).
207. P.E. Ceruzzi, *Crossing the divide: Architectural issues and the emergence of stored programme computers, 1935-1953*, IEEE Annals of the History of Computing, **19**, 5-12, January-March (1997).
208. P.E. Ceruzzi, *A History of Modern Computing*, MIT Press, Cambridge MA, (1998).
209. K. Zuse, *Some remarks on the history of computing in Germany*, in *A History of Computing in the 20th Century*, N. Metropolis et al. editors, 611-627, Academic Press, New York, (1980).
210. A.R. Mackintosh, *The First Electronic Computer*, Physics Today, March, (1987).
211. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
212. A. Hodges, *Alan Turing: The Enigma*, Simon and Schuster, New York, (1983).
213. A. Turing, *On computable numbers with reference to the Entscheidungsproblem*, Journal of the London Mathematical Society, **II**, **2**, **42**, 230-265 (1937).
214. J. von Neumann, *The Computer and the Brain*, Yale University Press, (1958).
215. I.E. Sutherland, *Microelectronics and computer science*, Scientific American, 210-228, September (1977).
216. W. Aspray, *John von Neumann and the Origins of Modern Computing*, M.I.T. Press, Cambridge MA, (1990, 2nd ed. 1992).
217. W. Aspray, *The history of computing within the history of information technology*, History and Technology, **11**, 7-19 (1994).

218. G.F. Luger, *Computation and Intelligence: Collected Readings*, MIT Press, (1995).
219. Z.W. Pylyshyn, *Computation and Cognition: Towards a Foundation for Cognitive Science*, MIT Press, (1986).
220. D.E. Shasha and C. Lazere, *Out of Their Minds: The Creators of Computer Science*, Copernicus, New York, (1995).
221. W. Aspray, *An annotated bibliography of secondary sources on the history of software*, *Annals of the History of Computing* **9**, 291-243 (1988).
222. R. Kurzweil, *The Age of Intelligent Machines*, MIT Press, (1992).
223. S.L. Garfinkel and H. Abelson, eds., *Architects of the Information Society: Thirty-Five Years of the Laboratory for Computer Sciences at MIT*, MIT Press, (1999).
224. J. Haugeland, *Artificial Intelligence: The Very Idea*, MIT Press, (1989).
225. M.A. Boden, *Artificial Intelligence in Psychology: Interdisciplinary Essays*, MIT Press, (1989).
226. J.W. Cortada, *A Bibliographic Guide to the History of Computer Applications, 1950-1990*, Greenwood Press, Westport Conn., (1996).
227. M. Campbell-Kelly and W. Aspry, *Computer: A History of the Information Machine*, Basic Books, New York, (1996).
228. B.I. Blum and K. Duncan, editors, *A History of Medical Informatics*, ACM Press, New York, (1990).
229. J.-C. Guedon, La Planete Cyber, *Internet et Cyberspace*, Gallimard, (1996).
230. S. Augarten, *Bit by Bit: An Illustrated History of Computers*, Unwin, London, (1985).
231. N. Wiener, *Cybernetics; or Control and Communication in the Animal and the Machine*, The Technology Press, John Wiley and Sons, New York, (1948).
232. W.R. Ashby, *An Introduction to Cybernetics*, Chapman and Hall, London, (1956).
233. M.A. Arbib, *A partial survey of cybernetics in eastern Europe and the Soviet Union*, *Behavioral Sci.*, **11**, 193-216, (1966).
234. A. Rosenblueth, N. Weiner and J. Bigelow, *Behavior, purpose and teleology*, *Phil. Soc.* **10** (1), 18-24 (1943).
235. N. Weiner and A. Rosenblueth, *Conduction of impulses in cardiac muscle*, *Arch. Inst. Cardiol. Mex.*, **16**, 205-265 (1946).
236. H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems. Transactions of sixth-tenth conferences*, Josiah J. Macy Jr. Foundation, New York, (1950- 1954).
237. W.S. McCulloch and W. Pitts, *A logical calculus of ideas immanent in nervous activity*, *Bull. Math. Biophys.*, **5**, 115-133 (1943).
238. W.S. McCulloch, *An Account of the First Three Conferences on Teleological Mechanisms*, Josiah Macy Jr. Foundation, (1947).
239. G.A. Miller, *Languages and Communication*, McGraw-Hill, New York, (1951).
240. G.A. Miller, *Statistical behavioristics and sequences of responses*, *Psychol. Rev.* **56**, 6 (1949).
241. G. Bateson, *Bali - the value system of a steady state*, in M. Fortes, editor, *Social Structure Studies Presented to A.R. Radcliffe-Brown*, Clarendon Press, Oxford, (1949).
242. G. Bateson, *Communication, the Social Matrix of Psychiatry*, Norton, (1951).

243. G. Bateson, *Steps to an Ecology of Mind*, Chandler, San Francisco, (1972).
244. G. Bateson, *Communication et Societe*, Seuil, Paris, (1988).
245. S. Heims, *Gregory Bateson and the mathematicians: From interdisciplinary interactions to societal functions*, J. History Behavioral Sci., **13**, 141-159 (1977).
246. S. Heims, *John von Neumann and Norbert Wiener. From Mathematics to the Technology of Life and Death*, MIT Press, Cambridge MA, (1980).
247. S. Heims, *The Cybernetics Group*, MIT Press, Cambridge MA, (1991).
248. G. van de Vijver, *New Perspectives on Cybernetics (Self-Organization, Autonomy and Connectionism)*, Kluwer, Dordrecht, (1992).
249. A. Bavelas, *A mathematical model for group structures*, Appl. Anthropol. **7 (3)**, 16 (1948).
250. P. de Latil, *La Pensee Artificielle - Introduction a la Cybernetique*, Gallimard, Paris, (1953).
251. L.K. Frank, G.E. Hutchinson, W.K. Livingston, W.S. McCulloch and N. Wiener, *Teleological Mechanisms*, Ann. N.Y. Acad. Sci. **50**, 187- 277 (1948).
252. H. von Foerster, *Quantum theory of memory*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the sixth conferences, Josiah J. Macy Jr. Foundation, New York, (1950).
253. H. von Foerster, *Observing Systems*, Intersystems Publications, California, (1984).
254. H. von Foerster, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer, New York, (2002).
255. M. Newborn, *Kasparov vs. Deep Blue: Computer Chess Comes of age*, Springer Verlag, (1996).
256. K.M. Colby, *Artificial Paranoia: A Computer Simulation of the Paranoid Process*, Pergamon Press, New York, (1975).
257. J.Z. Young, *Discrimination and learning in the octopus*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the ninth conference, Josiah J. Macy Jr. Foundation, New York, (1953).
258. M.J. Apter and L. Wolpert, *Cybernetics and development*. I. Information theory, J. Theor. Biol. **8**, 244-257 (1965).
259. H. Atlan, *L'Organisation Biologique et la Theorie de l'Information*, Hermann, Paris, (1972).
260. H. Atlan, *On a formal definition of organization*, J. Theor. Biol. **45**, 295-304 (1974).
261. H. Atlan, *Organization du vivant, information et auto-organization*, in Volume Symposium 1986 de l'Encyclopediea Universalis, pp. 355-361, Paris, (1986).
262. E.R. Kandel, *Nerve cells and behavior*, Scientific American, **223**, 57-70, July, (1970).
263. E.R. Kandel, *Small systems of neurons*, Scientific American, **241 no.3**, 66-76 (1979).
264. A.K. Katchalsky et al., *Dynamic patterns of brain cell assemblies*, Neurosciences Res. Prog. Bull., **12 no.1**, (1974).
265. G.E. Moore, *Cramming more components onto integrated circuits*, Electronics, April 19, (1965).

266. P. Gelsinger, P. Gargini, G. Parker and A. Yu, *Microprocessors circa 2000*, IEEE Spectrum, October, (1989).
267. P. Baron, *On distributed communications networks*, IEEE Trans. Comm. Systems, March (1964).
268. V.G. Cerf and R.E. Khan, *A protocol for packet network intercommunication*, Trans. Comm. Tech. **COM-22**, **V5**, 627-641, May (1974).
269. L. Kleinrock, *Communication Nets: Stochastic Message Flow and Delay*, McGraw-Hill, New York, (1964).
270. L. Kleinrock, *Queueing Systems: Vol. II, Computer Applications*, Wiley, New York, (1976).
271. R. Kahn, editor, *Special Issue on Packet Communication Networks*, Proc. IEEE, **66**, November, (1978).
272. L.G. Roberts, *The evolution of packet switching*, Proc. of the IEEE **66**, 1307-13, (1978).
273. J. Abbate, *The electrical century: Inventing the web*, Proc. IEEE **87**, November, (1999).
274. J. Abbate, *Inventing the Internet*, MIT Press, Cambridge MA, (1999).
275. J.C. McDonald, editor, *Fundamentals of Digital Switching, 2nd Edition*, Plenum, New York, (1990).
276. B. Metcalfe, *Packet Communication, Peer-to-Peer Communication*, San Jose Calif, (1996).
277. T. Berners-Lee, *The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*, Harper San Francisco, (1999).
278. J. Clark, *Netscape Time: The Making of the Billion-Dollar Start-Up That Took On Microsoft*, St. Martin's Press, New York, (1999).
279. J. Wallace, *Overdrive: Bill Gates and the Race to Control Cyberspace*, Wiley, New York, (1997).
280. P. Cunningham and F. Froschl, *The Electronic Business Revolution*, Springer Verlag, New York, (1999).
281. J.L. McKenny, *Waves of Change: Business Evolution Through Information Technology*, Harvard Business School Press, (1995).
282. M.A. Cosumano, *Competing on Internet Time: Lessons From Netscape and Its Battle with Microsoft*, Free Press, New York, (1998).
283. F.J. Dyson, *The Sun, the Genome and the Internet: Tools of Scientific Revolutions*, Oxford University Press, (1999).
284. L. Bruno, *Fiber Optimism: Nortel, Lucent and Cisco are battling to win the high-stakes fiber-optics game*, Red Herring, June (2000).
285. N. Cochrane, *We're insatiable: Now it's 20 million million bytes a day*, Melbourne Age, January 15, (2001).
286. K.G. Coffman and A.N. Odlyzko, *The size and growth rate of the Internet*, First Monday, October, (1998).
287. C.A. Eldering, M.L. Sylla, and J.A. Eisenach, *Is there a Moore's law for bandwidth?*, IEEE Comm. Mag., 2-7, October, (1999).

288. G. Gilder, *Fiber keeps its promise: Get ready, bandwidth will triple each year for the next 25 years*, Forbes, April 7, (1997).
289. A.M. Noll, *Does data traffic exceed voice traffic?*, Comm. ACM, 121- 124, June, (1999).
290. B. St. Arnaud, J. Coulter, J. Fitchett, and S. Mokbel, *Architectural and engineering issues for building an optical Internet*, Proc. Soc. Optical Eng. (1998).
291. M. Weisner, *The computer for the 21st century*, Scientific American, September, (1991).
292. R. Wright, *Three Scientists and Their Gods*, Time Books, (1988).
293. S. Nora and A. Mine, *The Computerization of Society*, MIT Press, (1981).
294. T. Forester, *Computers in the Human Context: Information Theory, Productivity, and People*, MIT Press, (1989).

Chapter 6

PERIODIC SYSTEMS

6.1 The discovery of X-rays

In 1895, while the work leading to the discovery of the electron was still going on, a second revolutionary discovery was made. In the autumn of that year, Wilhelm Konrad Roentgen (1845-1923), the head of the department of physics at the University of Würzburg in Bavaria, was working with a discharge tube, repeating some of the experiments of Crookes.

Roentgen was especially interested in the luminescence of certain materials when they were struck by cathode rays. He darkened the room, and turned on the high voltage. As the current surged across the tube, a flash of light came from an entirely different part of the room! To Roentgen's astonishment, he found that a piece of paper which he had coated with barium platinocyanide was glowing brightly, even though it was so far away from the discharge tube that the cathode rays could not possibly reach it!

Roentgen turned off the tube, and the light from the coated paper disappeared. He turned on the tube again, and the bright glow on the screen reappeared. He carried the coated screen into the next room. Still it glowed! Again he turned off the tube, and again the screen stopped glowing. Roentgen realized that he had discovered something completely strange and new. Radiation of some kind was coming from his discharge tube, but the new kind of radiation could penetrate opaque matter!

Years later, when someone asked Roentgen what he thought when he discovered X-rays, he replied: "I didn't think. I experimented!" During the next seven weeks he experimented like a madman; and when he finally announced his discovery in December, 1895, he was able to report all of the most important properties of X-rays, including their ability to ionize gases and the fact that they cannot be deflected by electric or magnetic fields. Roentgen correctly believed X-rays to be electromagnetic waves, just like light waves, but with very much shorter wavelength.

It turned out that X-rays were produced by electrons from the cathode of the discharge tube. These electrons were accelerated by the strong electric field as they passed across the tube from the cathode (the negative terminal) to the anode (the positive terminal).



Dr. W. C. Roentgen

Figure 6.1: Wilhelm Konrad Roentgen (1845-1923). Wellcome Images.



Figure 6.2: X-ray photograph by W.K. Roentgen. Wellcome Images.

They struck the platinum anode with very high velocity, knocking electrons out of the inner parts of the platinum atoms. As the outer electrons fell inward to replace these lost inner electrons, electromagnetic waves of very high frequency were emitted.

On January 23, 1896, Roentgen gave the first public lecture on X-rays; and in this lecture he demonstrated to his audience that X-ray photographs could be used for medical diagnosis. When Roentgen called for a volunteer from the audience, the 79 year old physiologist, Rudolf von Kölliker stepped up to the platform, and an X-ray photograph was taken of the old man's hand. The photograph, still in existence, shows the bones beautifully.

Wild enthusiasm for Roentgen's discovery swept across Europe and America, and soon many laboratories were experimenting with X-rays. The excitement about X-rays led indirectly to a third revolutionary discovery - radioactivity.

6.2 Bragg father and son

Sir William Henry Bragg (1862-1942) graduated from Trinity College, Cambridge University. with first class honors in mathematics in 1885. In that year, at the age of 25, he was appointed Professor of Mathematics and Experimental Physics at the University of Adelaide in Australia. In 1908, Bragg returned to England after 23 years in Australia.



Figure 6.3: Max von Laue (1879-1960).

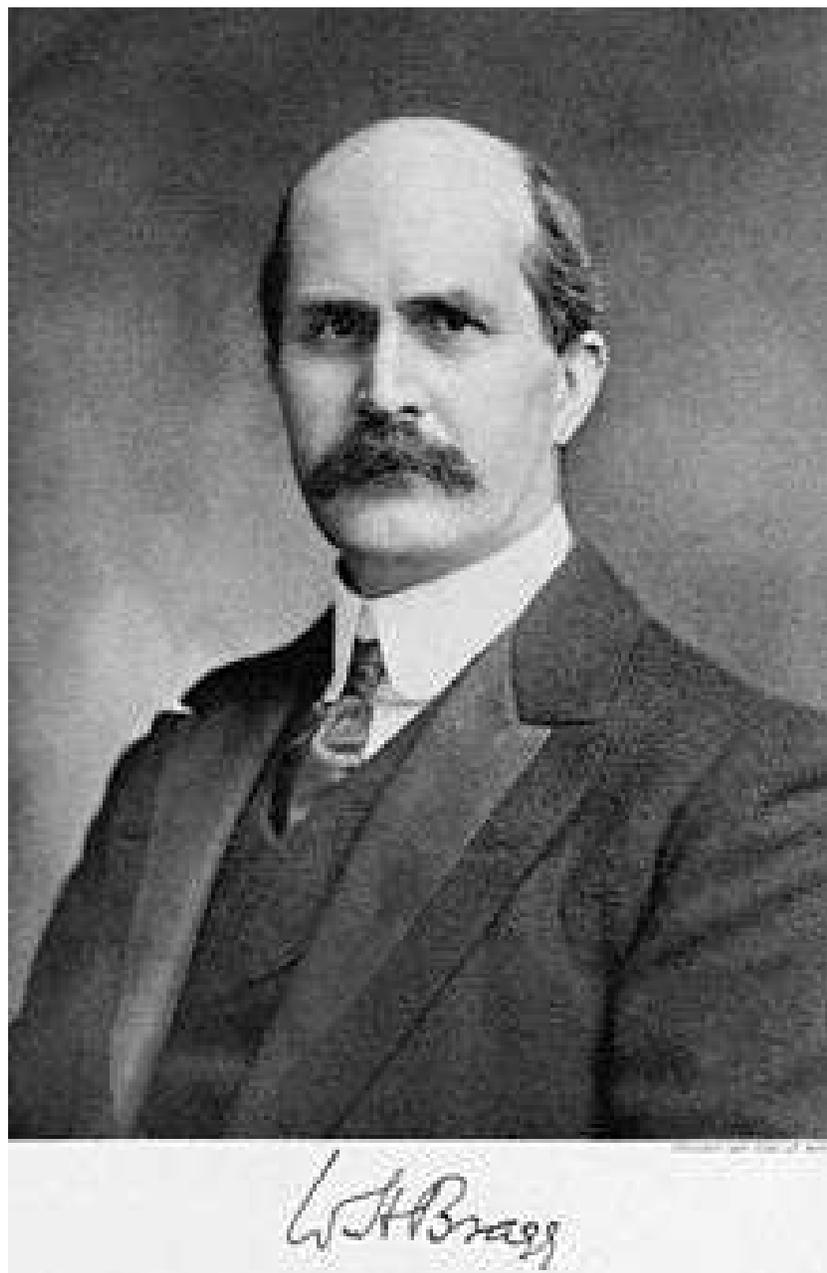


Figure 6.4: Sir William Henry Bragg (1862-1942). He and his son, Lawrence Bragg shared the 1915 Nobel Prize in Physics “for their services to the analysis of crystal structure by means of X-rays”. He studied with J.J. Thomson at Cambridge University after having won a scholarship to Trinity College in 1885. X-ray crystallography, pioneered by Bragg and his son, has proved to be enormously important both in chemistry and in biology. It has allowed us to understand the structure of both organic and inorganic molecules, and initiated the science of molecular biology.

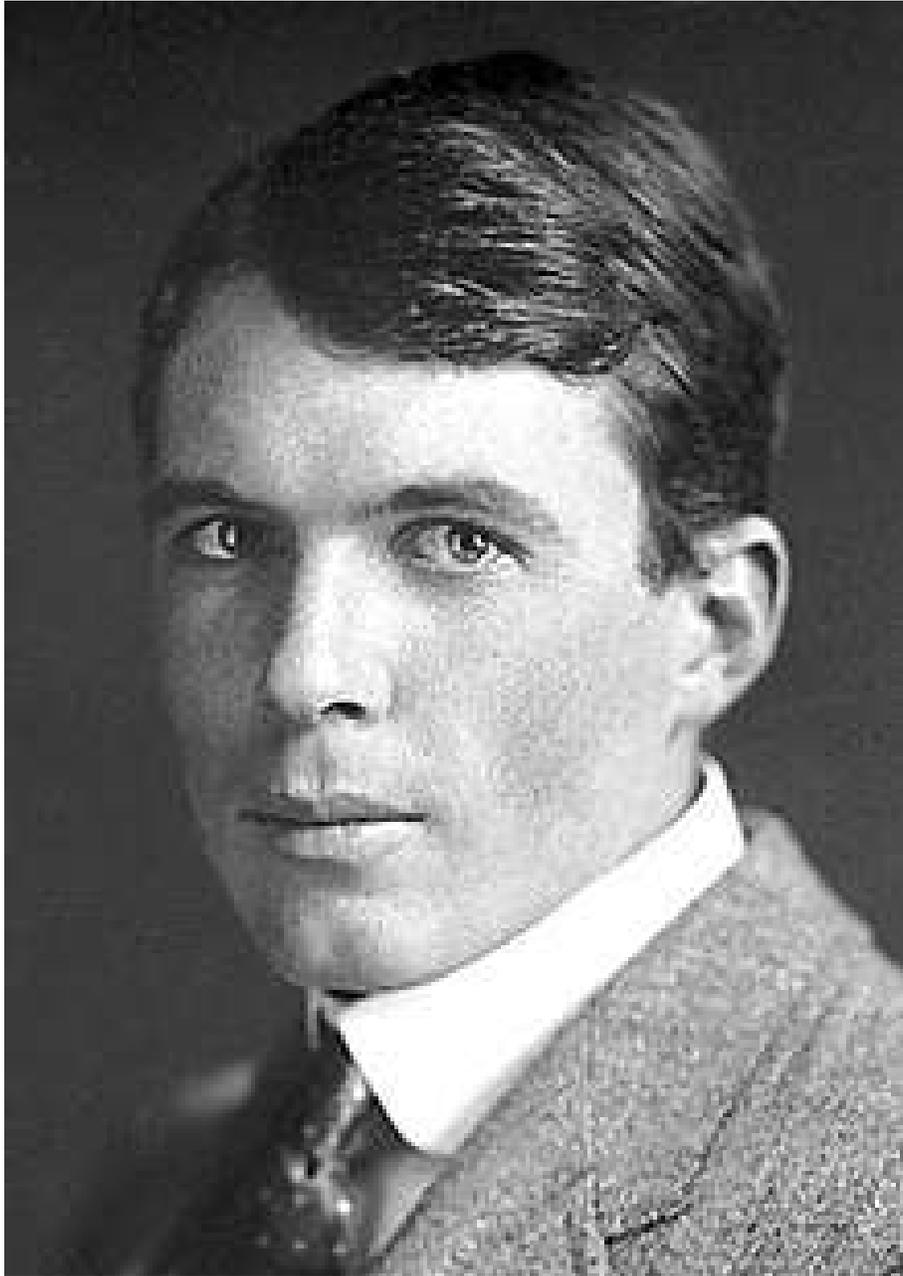


Figure 6.5: Sir William Lawrence Bragg (1890-1971).

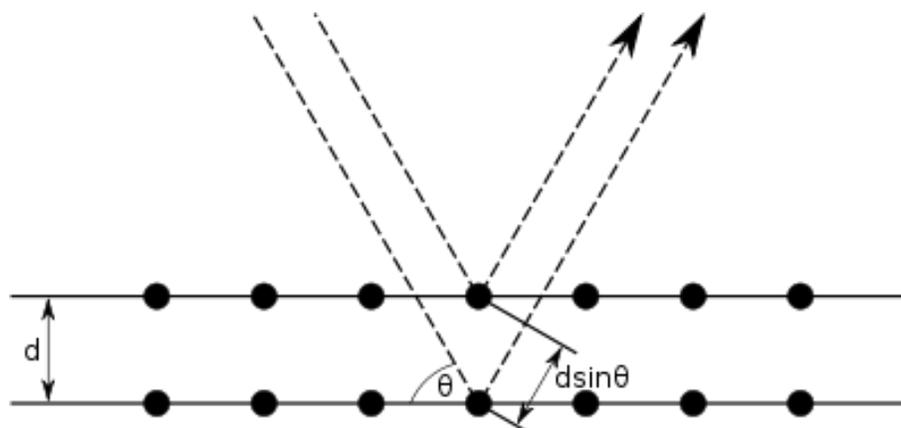


Figure 6.6: **Bragg's law of diffraction.** Two beams with identical wavelength and phase approach a crystalline solid and are scattered off two different atoms within it. The lower beam traverses an extra length of $2d \sin \theta$. Constructive interference occurs when this length is equal to an integer multiple of the wavelength of the radiation.

6.3 J.D. Bernal and Dorothy Crowfoot Hodgkin

In England, J.D. Bernal and Dorothy Crowfoot Hodgkin pioneered the application of X-ray diffraction methods to the study of complex biological molecules. In 1949, Hodgkin determined the structure of penicillin; and in 1955, she followed this with the structure of vitamin B12. In 1960, Max Perutz and John C. Kendrew obtained the structures of the blood proteins myoglobin and hemoglobin. This was an impressive achievement for the Cambridge crystallographers, since the hemoglobin molecule contains roughly 12,000 atoms.

The structure obtained by Perutz and Kendrew showed that hemoglobin is a long chain of amino acids, folded into a globular shape, like a small, crumpled ball of yarn. They found that the amino acids with an affinity for water were on the outside of the globular molecule; while the amino acids for which contact with water was energetically unfavorable were hidden on the inside. Perutz and Kendrew deduced that the conformation of the protein - the way in which the chain of amino acids folded into a 3-dimensional structure - was determined by the sequence of amino acids in the chain.

In 1966, D.C. Phillips and his co-workers at the Royal Institution in London found the crystallographic structure of the enzyme lysozyme (an egg-white protein which breaks down the cell walls of certain bacteria). Again, the structure showed a long chain of amino acids, folded into a roughly globular shape. The amino acids with hydrophilic groups were on the outside, in contact with water, while those with hydrophobic groups were on the inside. The structure of lysozyme exhibited clearly an active site, where sugar molecules of bacterial cell walls were drawn into a mouth-like opening and stressed by electrostatic forces, so that bonds between the sugars could easily be broken.

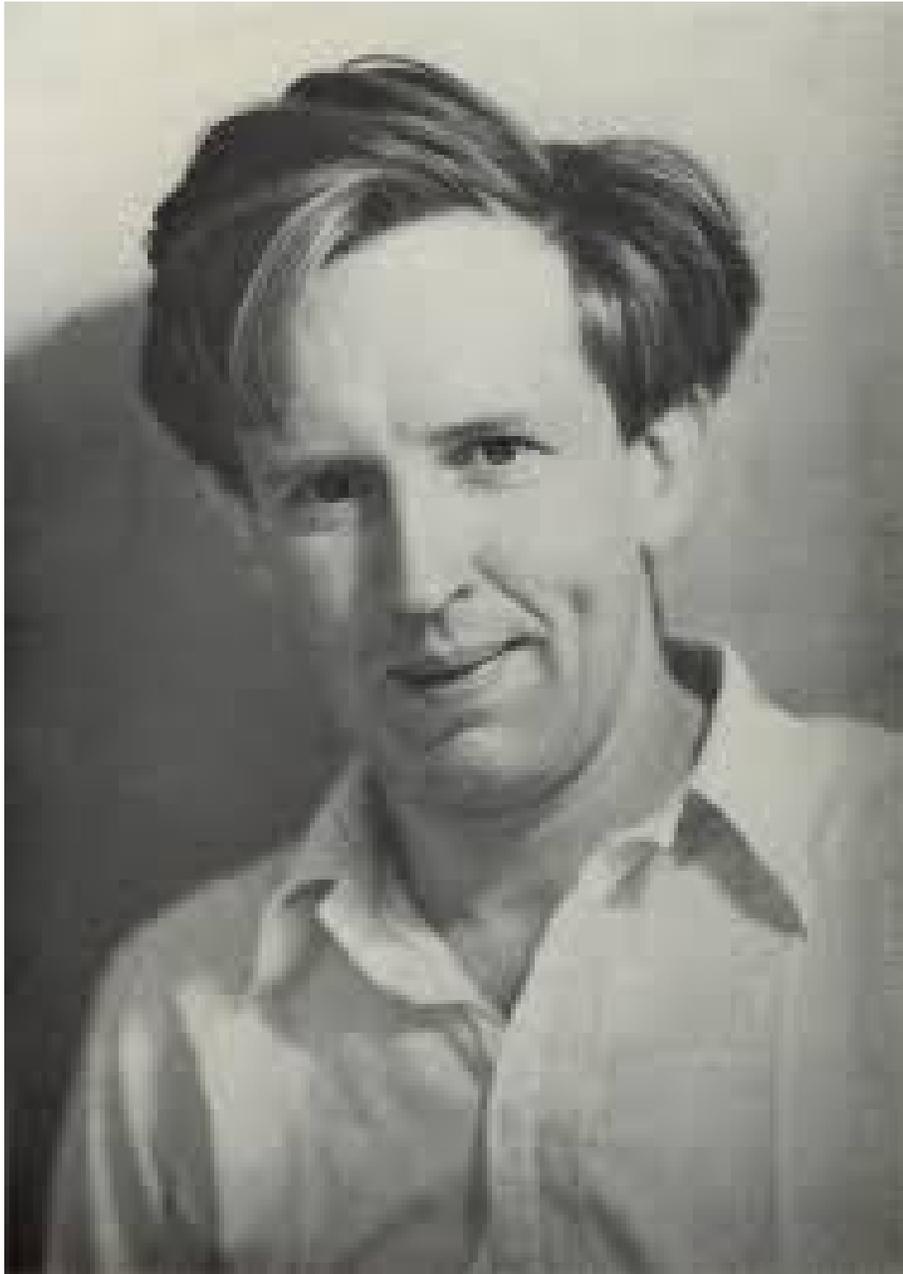


Figure 6.7: J.D. Bernal, (1901-1971).



Figure 6.8: Dorothy Crowfoot Hodgkin (1910-1994). She and her mentor J.D Bernal were great pioneers in the application of X-ray crystallography to determination of the structure of biological molecules, such as proteins. She was awarded the Nobel Prize in Chemistry in 1964.



Figure 6.9: **Sir John C. Kendrew (1917-1997) with an unfinished model of myoglobin. He shared the 1962 Nobel Prize in Chemistry with Peruz, Crick, Watson and Wilkins.**

6.4 The structure of DNA: Molecular biology

The discovery of the molecular structure of DNA was an event of enormous importance for genetics, and for biology in general. The structure was a revelation! The giant, helical DNA molecule was like a twisted ladder: Two long, twisted sugar-phosphate backbones formed the outside of the ladder, while the rungs were formed by the base pairs, A, T, G and C. The base adenine (A) could only be paired with thymine (T), while guanine (G) fit only with cytosine (C). Each base pair was weakly joined in the center by hydrogen bonds - in other words, there was a weak point in the center of each rung of the ladder - but the bases were strongly attached to the sugar-phosphate backbone. In their 1953 paper, Crick and Watson wrote:

”It has not escaped our notice that the specific pairing we have postulated suggests a possible copying mechanism for genetic material”. Indeed, a sudden blaze of understanding illuminated the inner workings of heredity, and of life itself.

If the weak hydrogen bonds in the center of each rung were broken, the ladderlike DNA macromolecule could split down the center and divide into two single strands. Each single strand would then become a template for the formation of a new double-stranded molecule.

Because of the specific pairing of the bases in the Watson-Crick model of DNA, the two strands had to be complementary. T had to be paired with A, and G with C. Therefore, if the sequence of bases on one strand was (for example) TTTGCTAAAGGTGAACCA... , then the other strand necessarily had to have the sequence AAACGATTTCCACTTGGT... The Watson-Crick model of DNA made it seem certain that all the genetic information

needed for producing a new individual is coded into the long, thin, double-stranded DNA molecule of the cell nucleus, written in a four-letter language whose letters are the bases, adenine, thymine, guanine and cytosine.

In 1956, George Emil Palade of the Rockefeller Institute used electron microscopy to study subcellular particles rich in RNA (ribosomes). Ribosomes were found to consist of two subunits - a smaller subunit, with a molecular weight one million times the weight of a hydrogen atom, and a larger subunit with twice this weight.

It was shown by means of radioactive tracers that a newly synthesized protein molecule is attached temporarily to a ribosome, but neither of the two subunits of the ribosome seemed to act as a template for protein synthesis. Instead, Palade and his coworkers found that genetic information is carried from DNA to the ribosome by a messenger RNA molecule (mRNA). Electron microscopy revealed that mRNA passes through the ribosome like a punched computer tape passing through a tape-reader. It was found that the adapter molecules, whose existence Crick had postulated, were smaller molecules of RNA; and these were given the name "transfer RNA" (tRNA). It was shown that, as an mRNA molecule passes through a ribosome, amino acids attached to complementary tRNA adaptor molecules are added to the growing protein chain.

The solution of the DNA structure in 1953 initiated a new kind of biology - molecular biology. This new discipline made use of recently-discovered physical techniques - X-ray diffraction, electron microscopy, electrophoresis, chromatography, ultracentrifugation, radioactive tracer techniques, autoradiography, electron spin resonance, nuclear magnetic resonance and ultraviolet spectroscopy. In the 1960's and 1970's, molecular biology became the most exciting and rapidly-growing branch of science.

Since DNA was known to carry the genetic message, coded into the sequence of the four nucleotide bases, A, T, G and C, and since proteins were known to be composed of specific sequences of the twenty amino acids, it was logical to suppose that the amino acid sequence in a protein was determined by the base sequence of DNA. The information somehow had to be read from the DNA and used in the biosynthesis of the protein.

It was known that, in addition to DNA, cells also contain a similar, but not quite identical, polynucleotide called ribonucleic acid (RNA). The sugar-phosphate backbone of RNA was known to differ slightly from that of DNA; and in RNA, the nucleotide thymine (T) was replaced by a chemically similar nucleotide, uracil (U). Furthermore, while DNA was found only in cell nuclei, RNA was found both in cell nuclei and in the cytoplasm of cells, where protein synthesis takes place. Evidence accumulated indicating that genetic information is first transcribed from DNA to RNA, and afterwards translated from RNA into the amino acid sequence of proteins.

The crystallographic determination of the structures of DNA and RNA and proteins opened the way for a new field of science, Molecular Biology, now one of the most rapidly developing of all research fields.

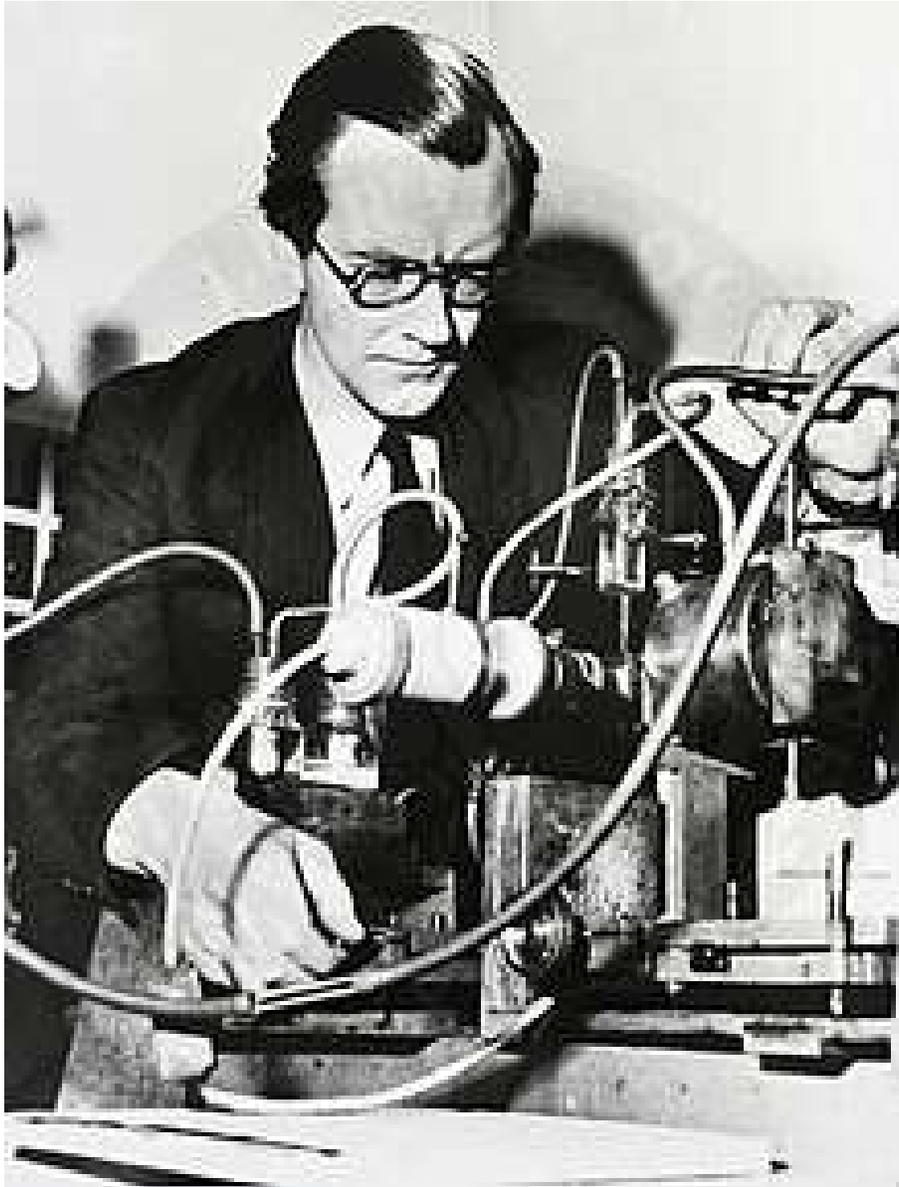


Figure 6.10: Maurice Wilkins (1916-2004). He applied to DNA the X-ray diffraction methods pioneered by Dorothy Hodgkin. It was his work, and that of Rosalind Franklin, together with Linus Pauling's model-building methods, that enabled Crick and Watson to correctly solve the structure of DNA. He shared the 1962 Nobel Prize in Physiology or Medicine with them.



Figure 6.11: Rosalind Franklin (1920-1958). It was one of her high-quality diffraction photographs, taken in Maurice Wilkins' laboratory, that proved to be critical for the DNA structure. She might have shared the Nobel Prize with Wilkins, Crick and Watson, but before this could be considered by the committee, she died of ovarian cancer.

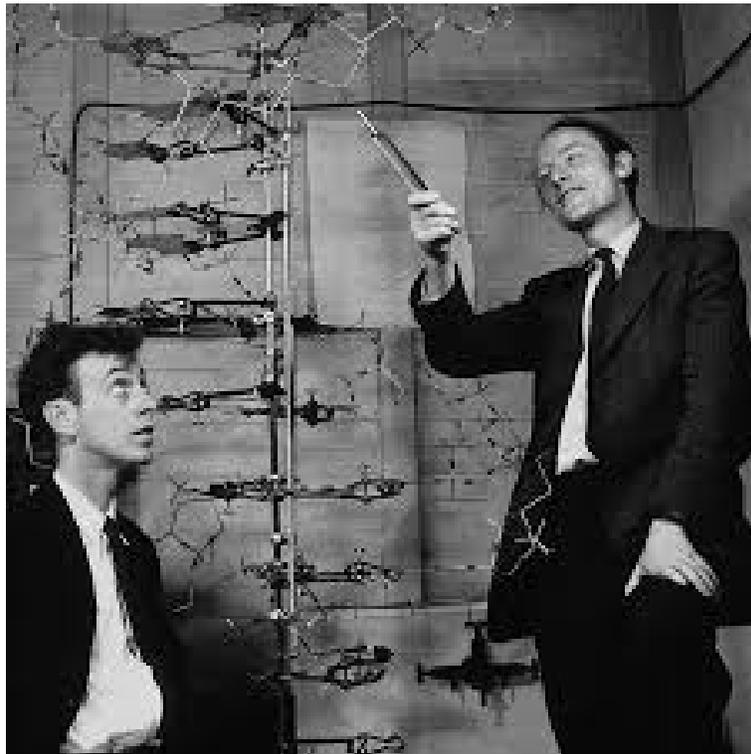


Figure 6.12: Francis Crick (1916-2004) and James Dewey Watson (born 1928) at the Cavendish Laboratory with their model of DNA. After their discovery of the structure of DNA, it became clear that it was this molecule that carried genetic information between generations.

6.5 Direct and reciprocal lattice vectors

In order to understand crystal structure, X-ray diffraction, and the quantum mechanical treatment of crystals, we can begin by considering a 2-dimensional lattice points in space defined by the relationship

$$\mathbf{R}_n = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 \quad n_1 = 0, \pm 1, \pm 2, \pm 3, \dots \quad n_2 = 0, \pm 1, \pm 2, \pm 3, \dots \quad (6.1)$$

Here \mathbf{a}_1 and \mathbf{a}_2 are vectors defining the unit cell of the 2-dimensional system in direct space. For example, the substance graphite, from which pencil leads are made, consists of planar arrays in which carbon atoms are tightly bound together in linked hexagonal arrays. These planes are much more loosely bound together, so it makes sense to treat each graphite plane by itself. In the case of the 2-dimensional lattice of a graphite plane, the vectors \mathbf{a}_1 and \mathbf{a}_2 are given by

$$\begin{aligned} \mathbf{a}_1 &= \left\{ \frac{3}{2}d, -\frac{\sqrt{3}}{2}d \right\} \\ \mathbf{a}_2 &= \left\{ 0, \sqrt{3}d \right\} \end{aligned} \quad (6.2)$$

Here d is a small length defining the size of the unit cell. We can also define a reciprocal lattice, whose points in momentum-space are given by the relationships

$$\mathbf{G}_m = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 \quad m_1 = 0, \pm 1, \pm 2, \pm 3, \dots \quad m_2 = 0, \pm 1, \pm 2, \pm 3, \dots \quad (6.3)$$

We would like the function $f(\mathbf{r})$ defined by the Fourier series

$$f(\mathbf{r}) = \sum_m f_m e^{i\mathbf{G}_m \cdot \mathbf{r}} \quad (6.4)$$

to have the periodicity of the direct lattice. In other words, we would like this Fourier series to have the property:

$$f(\mathbf{r} + \mathbf{R}_n) = \sum_m f_m e^{i\mathbf{G}_m \cdot (\mathbf{r} + \mathbf{R}_n)} = f(\mathbf{r}) \quad (6.5)$$

regardless of the values of the coefficients f_m . For this to be the case, we require that

$$e^{i\mathbf{G}_m \cdot \mathbf{R}_n} = 1 \quad (6.6)$$

a relationship that holds only when

$$\mathbf{G}_m \cdot \mathbf{R}_n = 2\pi N \quad N = 0, \pm 1, \pm 2, \pm 3, \dots \quad (6.7)$$

Remembering the definitions of \mathbf{G}_m and \mathbf{R}_n , we have

$$(m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2) \cdot (n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2) = 2\pi N \quad N = 0, \pm 1, \pm 2, \pm 3, \dots \quad (6.8)$$

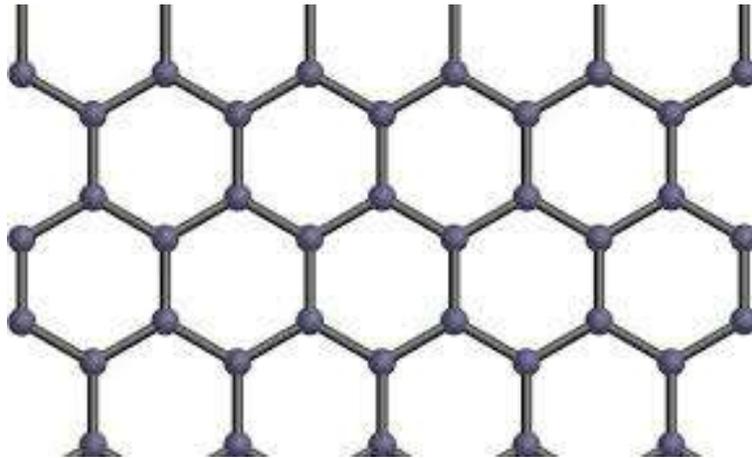


Figure 6.13: The arrangement of carbon atoms, tightly bound together in a graphite plane, like a giant planar molecule.

If

$$\mathbf{b}_i \cdot \mathbf{a}_j = \delta_{i,j} \quad (6.9)$$

then the desired relationship will be fulfilled, since any product of integers is also an integer. We are now able to determine the vectors \mathbf{b}_1 and \mathbf{b}_2 . In the example of the graphite plane, the conditions

$$\begin{aligned} \mathbf{b}_1 \cdot \mathbf{a}_1 &= 1 \\ \mathbf{b}_1 \cdot \mathbf{a}_2 &= 0 \\ \mathbf{b}_2 \cdot \mathbf{a}_2 &= 1 \\ \mathbf{b}_2 \cdot \mathbf{a}_1 &= 0 \end{aligned} \quad (6.10)$$

can be solved to yield

$$\begin{aligned} \mathbf{b}_1 &= \left\{ \frac{2}{3d}, 0 \right\} \\ \mathbf{b}_2 &= \left\{ \frac{1}{3d}, \frac{1}{\sqrt{3}d} \right\} \end{aligned} \quad (6.11)$$

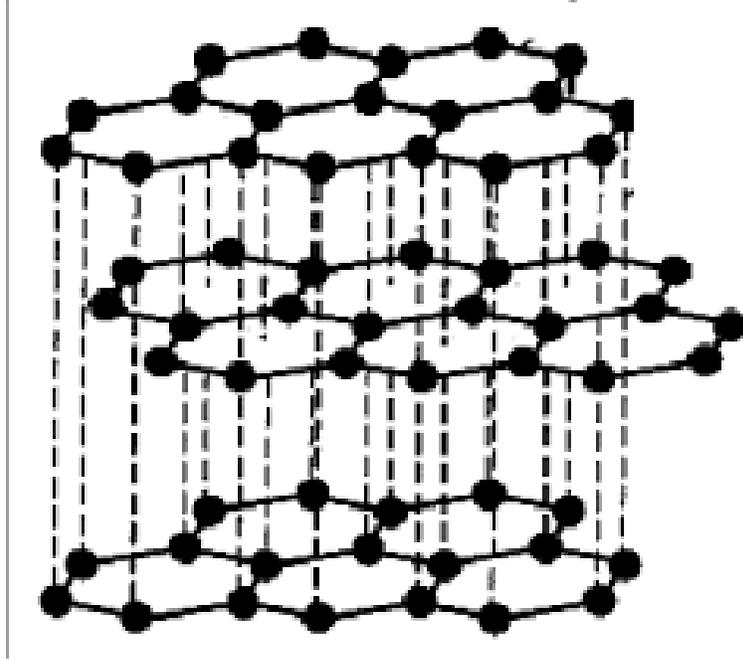


Figure 6.14: In a crystal of graphite, the planes interact with each other very weakly, so it makes sense to treat each plane by itself.

6.6 A Hückel calculation for a graphite plane

We can introduce a basis functions with three indices, $|n_1, n_2, j\rangle$. Here n_1 and n_2 are the integers that specify a particular direct lattice vector, $\mathbf{R}_n = n_1\mathbf{a}_1 + n_2\mathbf{a}_2$ that indicates the position of a particular unit cell. Each unit cell of a graphite plane contains 2 carbon atoms and the index $j = 1, 2$ indicates on which of these the basis function is located. In the Hückel approximation, the matrix elements of the Hamiltonian are given by

$$\langle n'_1, n'_2, j' | H | n_1, n_2, j \rangle = \begin{cases} \alpha & \text{diagonal elements} \\ \beta & \text{nearest neighbors} \\ 0 & \text{otherwise} \end{cases} \quad (6.12)$$

Then

$$\begin{aligned} H_{j',j}(\mathbf{k}) &= \sum_{n_1, n_2} \langle 0, 0, j' | H | n_1, n_2, j \rangle e^{i\mathbf{k} \cdot (n_1\mathbf{a}_1 + n_2\mathbf{a}_2)} \\ &= \begin{pmatrix} \alpha & \beta \{1 + e^{i\mathbf{k} \cdot \mathbf{a}_1} + e^{i\mathbf{k} \cdot (\mathbf{a}_1 + \mathbf{a}_2)}\} \\ \beta \{1 + e^{i\mathbf{k} \cdot \mathbf{a}_1} + e^{i\mathbf{k} \cdot (\mathbf{a}_1 + \mathbf{a}_2)}\} & \alpha \end{pmatrix} \end{aligned} \quad (6.13)$$

Solving the quadratic equation

$$\det |H_{j',j}(\mathbf{k}) - E_b(\mathbf{k})\delta_{j',j}| = 0 \quad (6.14)$$

we find two bands of allowed energy levels:

$$E_{\pm}(\mathbf{k}) = \alpha \pm |\beta| \{3 + 2 \cos(\mathbf{k} \cdot \mathbf{a}_1) + 2 \cos(\mathbf{k} \cdot \mathbf{a}_2) + 2 \cos(\mathbf{k} \cdot (\mathbf{a}_1 + \mathbf{a}_2))\} \quad (6.15)$$

The allowed values of \mathbf{k} can be determined by imposing periodic boundary conditions. Suppose that we impose periodic boundary conditions at the edges of a parallelogram in the graphite plane, and that the edges are determined by the vector $n\mathbf{a}_1$ and $n\mathbf{a}_2$. Then the parallelogram will contain $N = n^2$ unit cells. The periodic boundary conditions require that

$$\begin{aligned} e^{i\mathbf{k} \cdot \mathbf{a}_1} &= 1 \\ e^{i\mathbf{k} \cdot \mathbf{a}_2} &= 1 \end{aligned} \quad (6.16)$$

These boundary conditions will be satisfied if

$$\begin{aligned} \mathbf{k} &= \frac{2\pi}{n} \left\{ \left(m_1 - \frac{\mathbf{b}_1}{2} \right) + \left(m_2 - \frac{\mathbf{b}_2}{2} \right) \right\} \\ m_1 &= 1, 2, 3, \dots, n \\ m_2 &= 1, 2, 3, \dots, n \end{aligned} \quad (6.17)$$

Other values of m_1 and m_2 are possible, but they do not lead to any new independent solutions.

6.7 3-dimensional crystal lattices

The discussion given above can be extended to 3-dimensional crystal lattices with only a slight increase in complexity. For 3-dimensional lattices, the direct lattice vectors are given by

$$\mathbf{R}_n = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad n_j = 0, \pm 1, \pm 2, \pm 3, \dots \quad j = 1, 2, 3 \quad (6.18)$$

While the reciprocal lattice vectors are

$$\mathbf{G}_m = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3 \quad m_j = 0, \pm 1, \pm 2, \pm 3, \dots \quad j = 1, 2, 3 \quad (6.19)$$

The relationship

$$\mathbf{b}_i \cdot \mathbf{a}_j = \delta_{i,j} \quad (6.20)$$

will be fulfilled if

$$\begin{aligned} \mathbf{b}_1 &= 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \\ \mathbf{b}_2 &= 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \\ \mathbf{b}_3 &= 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \end{aligned} \quad (6.21)$$

where the vector product of two vectors \mathbf{A} and \mathbf{B} is defined by the relationship:

$$\mathbf{A} \times \mathbf{B} \equiv \{A_2B_3 - A_3B_2, A_3B_1 - A_1B_3, A_1B_2 - A_2B_1\} \quad (6.22)$$

The sum of any two direct lattice vectors \mathbf{R}_n and $\mathbf{R}_{n'}$ is a direct lattice vector, and the difference between them is also a direct lattice vector. Similarly, if we take the sum or difference between two reciprocal lattice vectors, \mathbf{G}_m and $\mathbf{G}_{m'}$, we obtain a reciprocal lattice vector. These relationships follow from the fact that sums of integers and differences between integers are always integers.

6.8 Quantum treatment of electrons in crystals

The electrons in a crystal move in a periodic potential of the form

$$V(\mathbf{r}) = \sum_m V_m e^{i\mathbf{G}_m \cdot \mathbf{r}} \quad (6.23)$$

where the Fourier coefficients of the potential, V_m depend on the nature of the crystal, and where the vectors \mathbf{G}_m are the appropriate reciprocal lattice vectors. We now introduce a set of basis functions of the form:

$$\psi_k(\mathbf{r}) = u_k(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} \quad (6.24)$$

where $u_k(\mathbf{r})$ has the periodicity of the crystal lattice, and can therefore be represented by the Fourier series:

$$u_k(\mathbf{r}) = \sum_{m'} u_{m'} e^{i\mathbf{G}_{m'} \cdot \mathbf{r}} \quad (6.25)$$

6.9 The nearly-free electron approximation

In this approximation we let the basis functions be plane waves, so that

$$\begin{aligned} u_k(\mathbf{r}) &= \frac{1}{\sqrt{\mathcal{V}}} \\ \psi_k(\mathbf{r}) &= \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k} \cdot \mathbf{r}} \end{aligned} \quad (6.26)$$

where \mathcal{V} is the volume of the crystal. Then the matrix elements of the potential between basis functions are given by

$$\begin{aligned} &\int d^3x \psi_{k'}^*(\mathbf{r}) V(\mathbf{r}) \psi_k(\mathbf{r}) \equiv \langle k' | V | k \rangle \\ &= \frac{1}{\mathcal{V}} \int d^3x e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}} V(\mathbf{r}) \\ &= \sum_m \frac{V_m}{\mathcal{V}} \int d^3x e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{G}_m) \cdot \mathbf{r}} \end{aligned} \quad (6.27)$$

Non-zero matrix elements occur only when $\mathbf{k} - \mathbf{k}'$ is exactly equal to a reciprocal lattice vector. We can use perturbation theory to treat the electrons in a crystal in the nearly-free electron approximation. Letting

$$\psi_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}^0(\mathbf{r}) + \psi_{\mathbf{k}}^1(\mathbf{r}) + \dots \quad (6.28)$$

where

$$\psi_{\mathbf{k}}^0(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (6.29)$$

If the perturbation produced by the potential is small, then the zeroth-order energy is given by

$$E_{\mathbf{k}}^0 = \frac{\hbar^2}{2m_e} \mathbf{k} \cdot \mathbf{k} = \frac{\hbar^2 k^2}{2m_e} \quad (6.30)$$

just as would be the case for completely free electrons. If we take only the first-order correction into account, the wave function is given approximately by

$$\begin{aligned} \psi_{\mathbf{k}}(\mathbf{r}) &= \psi_{\mathbf{k}}^0(\mathbf{r}) + \psi_{\mathbf{k}}^1(\mathbf{r}) \\ &= \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}\cdot\mathbf{r}} + \sum_{\mathbf{k}' \neq \mathbf{k}} \frac{\langle \mathbf{k}' | V | \mathbf{k} \rangle}{k'^2 - k^2} \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}'\cdot\mathbf{r}} \end{aligned} \quad (6.31)$$

while the energy becomes

$$\begin{aligned} E_{\mathbf{k}} &= E_{\mathbf{k}}^0 + E_{\mathbf{k}}^1 + E_{\mathbf{k}}^2 \\ &= \frac{\hbar^2 k^2}{2m_e} + \langle \mathbf{k} | V | \mathbf{k} \rangle + \frac{2m_e}{\hbar^2} \sum_{\mathbf{k}' \neq \mathbf{k}} \frac{\langle \mathbf{k} | V | \mathbf{k}' \rangle \langle \mathbf{k}' | V | \mathbf{k} \rangle}{k'^2 - k^2} \end{aligned} \quad (6.32)$$

Using the Fourier series expansion of $V(\mathbf{r})$ and remembering that $e^{i\mathbf{k}'\cdot\mathbf{r}}$ is required to obey periodic boundary conditions, we can evaluate matrix elements of the potential:

$$\begin{aligned} \langle \mathbf{k}' | V | \mathbf{k} \rangle &= \frac{1}{\mathcal{V}} \sum_m V_m \int d^3x e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{G}_m)\cdot\mathbf{r}} \\ &= \sum_m V_m \delta_{\mathbf{k}'-\mathbf{k}, \mathbf{G}_m} \end{aligned} \quad (6.33)$$

The matrix elements are zero unless $\mathbf{k}' - \mathbf{k}$ is exactly equal to a reciprocal lattice vector. Substituting this result into our expression for the energies, we obtain:

$$\begin{aligned} E_{\mathbf{k}} &= E_{\mathbf{k}}^0 + E_{\mathbf{k}}^1 + E_{\mathbf{k}}^2 \\ &= \frac{\hbar^2 k^2}{2m_e} + V_0 + \frac{2m_e}{\hbar^2} \sum_{\mathbf{G}_m \neq 0} \frac{|V_m|^2}{(\mathbf{k} - \mathbf{G}_m) \cdot (\mathbf{k} - \mathbf{G}_m) - \mathbf{k} \cdot \mathbf{k}} \end{aligned} \quad (6.34)$$

The denominator of the second term vanishes when

$$(\mathbf{k} - \mathbf{G}_m) \cdot (\mathbf{k} - \mathbf{G}_m) - \mathbf{k} \cdot \mathbf{k} = 0 \quad (6.35)$$

This condition will be fulfilled whenever a vector \mathbf{k} touches a plane perpendicularly bisecting a reciprocal lattice vector. Such a plane is defined to be the boundary of a Brillion zone.

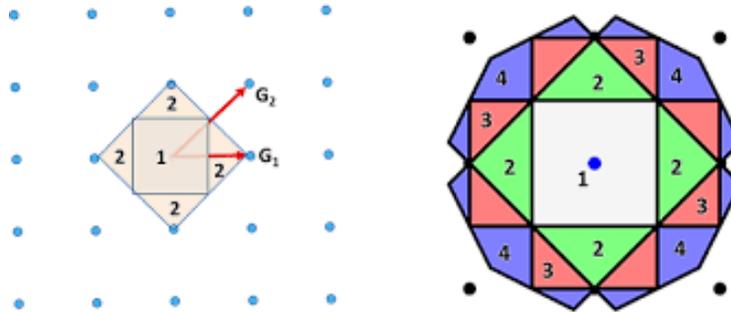


Figure 6.15: The first, second, third and fourth Brillouin zones in a plane of a simple cubic reciprocal lattice.

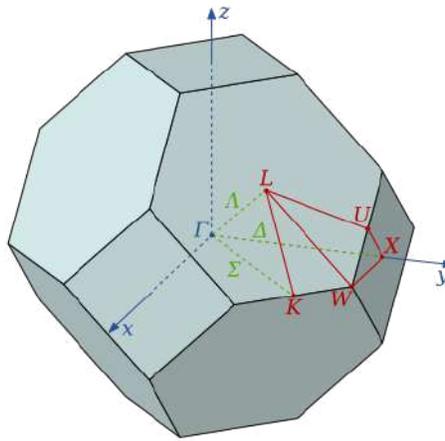


Figure 6.16: The first Brillouin zone in the reciprocal lattice of a face-centered crystal.

The band gap at the edge of a Brillouin zone

As we approach the boundary of a Brillouin zone (i.e. as we approach a plane perpendicularly bisecting a reciprocal lattice vector), the effect of the periodic potential $V(\mathbf{r})$ becomes more and more pronounced. Finally, very near to the boundary, the perturbation series fails to converge. Near to the boundary, we can approximate the wave function by a linear combination of two nearly degenerate zeroth-order wave functions:

$$\psi_{k,m}(\mathbf{r}) = C_1 \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}\cdot\mathbf{r}} + C_2 \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}\cdot(\mathbf{r}+\mathbf{G}_m)} \quad (6.36)$$

Then

$$\left(-\frac{\hbar^2}{2m_e} \nabla^2 + V(\mathbf{r}) - E_{k,m} \right) \left(C_1 \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}\cdot\mathbf{r}} + C_2 \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}\cdot(\mathbf{r}+\mathbf{G}_m)} \right) = 0 \quad (6.37)$$

Multiplying from the left by $e^{-i\mathbf{k}\cdot\mathbf{r}}$ and integrating over the volume of the crystal, and then doing the same with $e^{i\mathbf{k}\cdot(\mathbf{r}+\mathbf{G}_m)}$, we obtain two simultaneous equations which will have a

solution provided that the secular determinant vanishes

$$\det \begin{vmatrix} \frac{\hbar^2 k^2}{2m_e} + V_0 - E_{k,m} & V_m \\ V_m & \frac{\hbar^2 k^2}{2m_e} + V_0 - E_{k,m} \end{vmatrix} = 0 \quad (6.38)$$

Thus, finally, we obtain the energy at the Brillouin zone boundary

$$E_{k,m} = \frac{\hbar^2 k^2}{2m_e} + V_0 \pm V_m \quad (6.39)$$

and we see that the band gap is $2V_m$.

6.10 Molecular crystals

A molecular crystal is a weakly-bound periodic arrangement of tightly-bound constituent molecules. Examples are crystals of benzene or naphthalene. The low melting point of benzene reflects the small amount of energy required to separate the benzene molecules from one another. A much larger amount of energy would be needed to break the covalent bonds of all the benzene rings.

The wave function of the electrons in a molecular crystal can be built from products of the wave functions of the individual molecules. Suppose that we have found wave functions representing the ground state and an excited state of an isolated constituent molecule. We might do this by means of the Hartree-Fock SCF method, in which case the wave function of the isolated molecule would be antisymmetrized. In molecular crystals, overlap of the wave functions of neighboring molecules can be neglected, so antisymmetrization over the entire crystal is unnecessary. The basis functions used to build up the electronic wave function of a molecular crystal in the Frenkel exciton picture can be simple products of wave functions of the constituent molecules.

We can separate the Hamiltonian of the crystal into two parts, \mathbf{H}_0 and \mathbf{H}' , where \mathbf{H}_0 represents the Hamiltonian of the individual molecules, while \mathbf{H}' represents the interaction between them. Our basis set will consist of eigenfunctions of \mathbf{H}_0 . Let

$$|0\rangle = \Delta_0(1)\Delta_0(2)\cdots\Delta_0(N) \quad (6.40)$$

represent the eigenfunction of \mathbf{H}_0 with all N subunits in their ground state, and let

$$|n\rangle = \Delta_0(1)\Delta_0(2)\cdots\Delta_\zeta(n)\cdots\Delta_0(N) \quad (6.41)$$

represent the eigenfunction of \mathbf{H}_0 with all the subunits in their ground states except the n th subunit, which is in its ζ th excited state. In the simplest approximation, we can say

that the effect of \mathbf{H}' is to hybridize the N degenerate states of the form shown in equation (6.41). We can try to build up eigenfunctions of the complete Hamiltonian, $\mathbf{H} = \mathbf{H}_0 + \mathbf{H}'$ from linear combinations of these basis functions:

$$|\mathbf{k}\rangle = \sum_{n=1}^N |n\rangle \langle n|\mathbf{k}\rangle \quad (6.42)$$

In order that $|\mathbf{k}\rangle$ should be an eigenfunction of \mathbf{H} , we require that

$$\sum_{n=1}^N \langle n'|\mathbf{H}|n\rangle \langle n|\mathbf{k}\rangle = E_{\mathbf{k}} \langle n'|\mathbf{k}\rangle \quad (6.43)$$

Inserting the trial solution

$$\langle n|\mathbf{k}\rangle = \frac{1}{\sqrt{N}} e^{i\mathbf{k}\cdot\mathbf{a}_n} \quad (6.44)$$

into equation (6.43), we obtain

$$E_{\mathbf{k}} = \sum_{n=1}^N \langle n'|\mathbf{H}|n\rangle e^{i\mathbf{k}\cdot(\mathbf{a}_n - \mathbf{a}_{n'})} \quad (6.45)$$

because of the translational symmetry of the crystal, the lattice sum shown in equation (6.45) depends only on the relative distance, $\mathbf{a}_{n'} - \mathbf{a}_n$, except at the surfaces of the crystal, where special boundary conditions must be imposed. Written out in full, the wave function of the ζ -band is

$$|\mathbf{k}, \zeta\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^N \Delta_0(1)\Delta_0(2)\cdots\Delta_{\zeta}(n)\cdots\Delta_0(N) e^{i\mathbf{k}\cdot\mathbf{a}_n} \quad (6.46)$$

Here \mathbf{k} is called the “exciton wave number”, while ζ is called the “band index”. States of this kind were introduced by J. Frenkel and R. Peierls to explain the optical properties of molecular crystals, and they have become known as “Frenkel exciton states”. A Frenkel exciton state is an excited state of the entire crystal. All of the electrons are involved, and the excited state is delocalized, although each electron is localized on its own monomeric subunit.

The Frenkel exciton picture can be used to describe collective excited states of polymers, provided that they have translational symmetry, and also to describe arrays of atoms, if the atoms are sufficiently separated so that orbital overlap between them can be neglected.

6.11 Periodic boundary conditions

In solid state theory, periodic boundary conditions (sometimes called “Born-von Kármán boundary conditions”) are often used. These boundary conditions restrict the allowed

values of the wave number \mathbf{k} . Suppose, for example, we are using the Frenkel exciton picture to describe the collective excited states of a polymer with N subunits. Then the wave number \mathbf{k} is not a vector, but reduces simply to a number, k . If we require that

$$e^{ika_1} = e^{ika_{N+1}} \quad (6.47)$$

This is the same as requiring that

$$\frac{e^{ika_{N+1}}}{e^{ika_1}} = e^{ik(a_{N+1}-a_1)} = e^{ikL} = 1 \quad (6.48)$$

where L is the length of the polymer. This condition will be fulfilled if

$$k = \frac{2\pi j}{L} \quad j = 1, 2, \dots, N \quad (6.49)$$

No new linearly independent states are produced if higher values of the integer j are used. For a cubic crystal consisting of N^3 molecules, periodic boundary conditions require that

$$\begin{aligned} k_x &= \frac{2\pi j_x}{L} & j_x &= 1, 2, \dots, N \\ k_y &= \frac{2\pi j_y}{L} & j_y &= 1, 2, \dots, N \\ k_z &= \frac{2\pi j_z}{L} & j_z &= 1, 2, \dots, N \end{aligned} \quad (6.50)$$

assuming that the crystal is also cubic in shape, so that each side has length L

$$(6.51)$$

6.12 Homogeneous boundary conditions

For cases where only interactions between nearest neighbors are important, another type of boundary condition (called “homogeneous”) is often used. Homogeneous boundary conditions are especially appropriate for small systems, such as small polymers. For an oligomer, where only nearest-neighbor interactions are important, the set of secular equations to be solved have the form

$$\begin{aligned} (\alpha - E_k)\langle 1|k\rangle + \beta\langle 2|k\rangle &= 0 \\ \beta\langle 1|k\rangle + (\alpha - E_k)\langle 2|k\rangle + \beta\langle 3|k\rangle &= 0 \\ \beta\langle 2|k\rangle + (\alpha - E_k)\langle 3|k\rangle + \beta\langle 4|k\rangle &= 0 \\ \beta\langle 3|k\rangle + (\alpha - E_k)\langle 4|k\rangle + \beta\langle 5|k\rangle &= 0 \\ \beta\langle 4|k\rangle + (\alpha - E_k)\langle 5|k\rangle + \beta\langle 6|k\rangle &= 0 \\ &\vdots &&\vdots \\ \beta\langle N-1|k\rangle + (\alpha - E_k)\langle N|k\rangle &= 0 \end{aligned} \quad (6.52)$$

where β represents the nearest-neighbor interaction energy, while α is the energy of an isolated monomer. The symmetry of this set of secular equations is spoiled by the lack of coupling to anything at the two ends of the polymer. We can, however, restore the symmetry by adding a fictitious monomer at each end so that the equations become

$$\begin{aligned}
\beta\langle 0|k\rangle + (\alpha - E_k)\langle 1|k\rangle + \beta\langle 2|k\rangle &= 0 \\
\beta\langle 1|k\rangle + (\alpha - E_k)\langle 2|k\rangle + \beta\langle 3|k\rangle &= 0 \\
\beta\langle 2|k\rangle + (\alpha - E_k)\langle 3|k\rangle + \beta\langle 4|k\rangle &= 0 \\
\beta\langle 3|k\rangle + (\alpha - E_k)\langle 4|k\rangle + \beta\langle 5|k\rangle &= 0 \\
\beta\langle 4|k\rangle + (\alpha - E_k)\langle 5|k\rangle + \beta\langle 6|k\rangle &= 0 \\
&\vdots & \vdots & \vdots \\
\beta\langle N-1|k\rangle + (\alpha - E_k)\langle N|k\rangle + \beta\langle N+1|k\rangle &= 0
\end{aligned} \tag{6.53}$$

and requiring that

$$\langle 0|k\rangle = \langle N+1|k\rangle = 0 \tag{6.54}$$

The set of secular equations (6.53), together with the boundary conditions (6.54), have normalized solutions of the form

$$\langle n|k\rangle = \sqrt{\frac{1}{N+1}} \sin(nkd) \tag{6.55}$$

where

$$d = a_{n+1} - a_n \tag{6.56}$$

is the lattice spacing. The homogeneous boundary conditions require that

$$(N+1)kd = \pi \tag{6.57}$$

so that we obtain N linearly-independent states, with

$$kd = \frac{\pi}{N+1}, \frac{2\pi}{N+1}, \dots, \frac{N\pi}{N+1} \tag{6.58}$$

Notice that regardless of whether we use periodic boundary conditions or homogeneous boundary conditions, we obtain the correct number of linearly independent eigenstates of the total Hamiltonian. We start with a set of N -fold degenerate basis functions, and the inter-monomeric interactions split these into a band of N exciton states. For a polymer treated with homogeneous boundary conditions, the exciton energies are

$$E_k = \alpha + 2\beta \cos(kd) \tag{6.59}$$

This expression also holds for a polymer treated with periodic boundary conditions, but the allowed values of k are slightly different, being given by (6.49) rather than (6.58). Similar results are obtained in the case of 3-dimensional crystals.

6.13 Taylor series expansion of the inter-monomer interaction

The unperturbed Hamiltonian of a molecular crystal or polymer can be written in the form

$$\mathbf{H}_0 = \sum_{n=1}^N \mathbf{H}(n) \quad (6.60)$$

where

$$\begin{aligned} \mathbf{H}(n)\Delta_0(n) &= E_0\Delta_0(n) & n = 1, 2, \dots, N \\ \mathbf{H}(n)\Delta_\zeta(n) &= E_\zeta\Delta_\zeta(n) & n = 1, 2, \dots, N \end{aligned} \quad (6.61)$$

These eigenstates of the unperturbed Hamiltonian are orthonormal, and so we have the relationships

$$\begin{aligned} \int d\tau_n \Delta_0^*(n)\Delta_0(n) &= 1 \\ \int d\tau_n \Delta_\zeta^*(n)\Delta_\zeta(n) &= 1 \\ \int d\tau_n \Delta_\zeta^*(n)\Delta_0(n) &= 0 \end{aligned} \quad (6.62)$$

We now introduce the 6-fold Taylor series expansion

$$\begin{aligned} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} &= \left(1 + (\mathbf{x}_i - \mathbf{a}_n) \frac{\partial}{\partial \mathbf{a}_n} + \dots\right) \left(1 + (\mathbf{x}_j - \mathbf{a}_{n'}) \frac{\partial}{\partial \mathbf{a}_{n'}} + \dots\right) \frac{1}{|\mathbf{a}_n - \mathbf{a}_{n'}|} \\ &= \end{aligned} \quad (6.63)$$

We then obtain, as the leading nonzero term in the inter-monomer interaction:

$$\langle n' | \mathbf{H}' | n \rangle = \frac{\mathbf{D}_n \cdot \mathbf{D}_{n'}}{R^3} - \frac{3(\mathbf{D}_n \cdot \mathbf{R})(\mathbf{D}_{n'} \cdot \mathbf{R})}{R^5} \quad (6.64)$$

where

$$\mathbf{R} \equiv \mathbf{a}_n - \mathbf{a}_{n'} \quad (6.65)$$

and

$$\mathbf{D}_n \equiv e \sum_j \int d\tau_n \Delta_\zeta^*(n)(\mathbf{x}_j - \mathbf{a}_n)\Delta_0(n) \quad (6.66)$$

the sum being taken over all the electrons in the n th monomer. The quantity \mathbf{D}_n is called the “transition dipole moment”.

6.14 X-ray diffraction experiments

It can be shown that if the electron density $\rho(\mathbf{r})$ in a crystal is represented by the Fourier series:

$$\rho(\mathbf{r}) = \sum_m F_m e^{i\mathbf{G}_m \cdot \mathbf{r}} \quad (6.67)$$

then the intensity of an X-ray with wave number \mathbf{k} scattered into a state with wave number \mathbf{k}' is given by

$$I(\mathbf{k} - \mathbf{k}') \sim |F_m|^2 \quad (6.68)$$

The phase problem

In diffraction experiments, F_m is not measured, but only a quantity proportional to $|F_m|^2$, and thus F_m is determined only up to a constant of proportionality and a phase factor. A common method for overcoming this difficulty is to guess the atomic positions in a crystal, and then assume that the atoms are relatively unaffected by the presence of their neighbors. In this way, a trial electron density distribution $\rho(\mathbf{r})$ can be calculated, and hence the Fourier coefficients F_m . The calculated diffraction intensities thus generated can be compared with those experimentally observed. The atomic positions in the model are then changed slightly, and the comparison process is repeated and refined until a good fit is obtained. Very similar considerations also hold for electron diffraction experiments and neutron diffraction experiments.

Suggestions for further reading

1. Nitske, Robert W., *The Life of W. C. Röntgen, Discoverer of the X-Ray*, University of Arizona Press, 1971.
2. Glasser, Otto (1933). *Wilhelm Conrad Röntgen and the Early History of the Roentgen Rays*. London: John Bale, Sons and Danielsson, Ltd
3. Neil W. Ashcroft and N. David Mermin, *Solid State Physics* (Harcourt: Orlando, 1976).
4. Bragg W (1913). *The Diffraction of Short Electromagnetic Waves by a Crystal*. Proceedings of the Cambridge Philosophical Society. 17: 43-57.
5. H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology, 4th Edition*, W.H. Freeman, New York, (2000).
6. Lily Kay, *Who Wrote the Book of Life? A History of the Genetic Code*, Stanford University Press, Stanford CA, (2000).
7. Sahotra Sarkar (editor), *The Philosophy and History of Molecular Biology*, Kluwer Academic Publishers, Boston, (1996).
8. James D. Watson et al. *Molecular Biology of the Gene, 4th Edition*, Benjamin-Cummings, (1988).

9. J.S. Fruton, *Proteins, Enzymes, and Genes*, Yale University Press, New Haven, (1999).
10. S.E. Lauria, *Life, the Unfinished Experiment*, Charles Scribner's Sons, New York (1973).
11. A. Lwoff, *Biological Order*, MIT Press, Cambridge MA, (1962).
12. James D. Watson, *The Double Helix*, Athenium, New York (1968).
13. F. Crick, *The genetic code*, Scientific American, **202**, 66-74 (1962).
14. F. Crick, *Central dogma of molecular biology*, Nature, **227**, 561-563 (1970).
15. David Freifelder (editor), *Recombinant DNA, Readings from the Scientific American*, W.H. Freeman and Co. (1978).
16. James D. Watson, John Tooze and David T. Kurtz, *Recombinant DNA, A Short Course*, W.H. Freeman, New York (1983).

Chapter 7

HARMONIC OSCILLATORS

7.1 Normal modes

Let us consider the small vibrations of a classical system of particles about the equilibrium positions. Suppose that the kinetic energy of the system is given by

$$T = \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d m_i \delta_{i,j} \frac{dx^i}{dt} \frac{dx^j}{dt} \quad (7.1)$$

while the leading term in a Taylor series expansion of the potential energy has the form

$$V = \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d V_{i,j} x^i x^j \quad (7.2)$$

The coordinates x^1, x^2, \dots, x^d , which represent small displacements from the equilibrium positions of the particles, are by no means the most convenient ones for solving the equations of motion of the system. We can bring the kinetic energy into a more convenient form by going over to the mass-weighted coordinates defined by

$$X^i \equiv \sqrt{m_i} x^i \quad i = 1, 2, \dots, d \quad (7.3)$$

In terms of these coordinates, the kinetic energy has the form

$$T = \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \delta_{i,j} \frac{dX^i}{dt} \frac{dX^j}{dt} \quad (7.4)$$

while the potential energy becomes

$$V = \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \frac{V_{i,j}}{\sqrt{m_i m_j}} X^i X^j \quad (7.5)$$

The mass-weighted coordinates are still not the most convenient ones that we can find, since the potential energy matrix $V_{i,j}$ may contain off-diagonal terms. and we would like to get rid of these. We can find a unitary transformation which diagonalizes $V_{i,j}/\sqrt{m_i m_j}$ by solving the secular equations

$$\sum_{j=1}^d \left(\frac{V_{i,j}}{\sqrt{m_i m_j}} - \mathcal{V}(k) \delta_{i,j} \right) U_{j,k} = 0 \quad (7.6)$$

Having performed the diagonalization, we can express the potential energy and the kinetic energy of the system in terms of the *normal coordinates* defined by

$$q^k = \sum_{i=1}^d X^i U_{i,k} = \sum_{i=1}^d \sqrt{m_i} x^i U_{i,k} \quad (7.7)$$

When we do this, the kinetic energy retains its diagonal form because of the unitarity of $U_{j,k}$:

$$T = \frac{1}{2} \sum_{k=1}^d \left(\frac{dq^k}{dt} \right)^2 \quad (7.8)$$

but the off-diagonal terms in the potential energy disappear:

$$V = \frac{1}{2} \sum_{k=1}^d \mathcal{V}(k) (q^k)^2 \quad (7.9)$$

From (7.8) and (7.9) we can see that the Lagrangian of the system can be written in the form

$$L = T - V = \sum_{k=1}^d L_k \quad (7.10)$$

where

$$L_k = \frac{1}{2} \left[\left(\frac{dq^k}{dt} \right)^2 - \mathcal{V}(k) (q^k)^2 \right] \quad (7.11)$$

The canonically conjugate momentum paired with the coordinate q^k is defined in mechanics to be

$$p_k = \frac{\partial L}{\partial \dot{q}^k} = \frac{dq^k}{dt} \quad (7.12)$$

the Hamiltonian of the system can be written in the form

$$H = T + V = \sum_{k=1}^d H_k \quad (7.13)$$

where

$$H_k = \frac{1}{2} (p_k^2 + \omega_k^2 q_k^2) \quad (7.14)$$

and

$$\omega_k = \sqrt{\mathcal{V}(k)} \quad (7.15)$$

In other words, when the Hamiltonian which represents small vibrations of a classical system is expressed in terms of the normal coordinates (or normal modes), it reduces to a sum of simple harmonic oscillator Hamiltonians. The normal coordinates are found by diagonalizing the mass-weighted potential energy matrix. The harmonic oscillator frequency of each is found by taking the square root of the corresponding eigenvalue of the mass-weighted potential energy matrix.

To illustrate this procedure, we can think of a system, whose Lagrangian is given by

$$\frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \left(m \delta_{i,j} \frac{dx^i}{dt} \frac{dx^j}{dt} - V_{i,j} x^i x^j \right) \quad (7.16)$$

where

$$V_{i,j} = \begin{cases} 2\kappa & i = j \\ -\kappa & i = j \pm 1 \end{cases} \quad (7.17)$$

This Lagrangian corresponds to a linear system of point masses, each joined elastically to the next. Then the secular equations (7.6) have the form

$$-\kappa U_{k-1,k} + [2\kappa - \mathcal{V}(k)] U_{k,k} - \kappa U_{k+1,k} = 0 \quad k = 2, \dots, d-1 \quad (7.18)$$

The trial solution

$$U_{j,k} = \sqrt{\frac{2}{d+1}} \sin(jka) \quad (7.19)$$

makes all of the secular equations redundant, All of them redundantly require that

$$\mathcal{V}(k) = \kappa [1 - \cos(ka)] \quad (7.20)$$

Imposing homogeneous boundary conditions (i.e clamping the two ends of the line) restricts the allowed values of k , and we must have

$$k = \frac{\pi}{(d+1)a}, \frac{2\pi}{(d+1)a}, \dots, \frac{\pi d}{(d+1)a} \quad (7.21)$$

where $(d+1)a$ is the length of the chain. The frequency spectrum of the normal modes is given by

$$\omega_k = \sqrt{\frac{\mathcal{V}(k)}{m}} = \sqrt{\frac{2\kappa [1 - \cos(2ka)]}{m}} \quad (7.22)$$

In terms of the normal mode coordinates and their time derivatives, the Lagrangian of the system becomes

$$L = \frac{1}{2} \sum_k \left[\left(\frac{dq_k}{dt} \right)^2 - (\omega_k q_k)^2 \right] \quad (7.23)$$

which can be recognized as a sum of harmonic oscillator Lagrangians.

7.2 Molecular vibrations and rotations

In the simplest possible approximation, we can regard a molecule (or a cluster in a non-melted state) as a collection of point masses held together by springlike bonds. When we calculate the normal modes of such a system, we always find that there are six zero-frequency modes. Three of these correspond to the degrees of freedom associated with translation of the whole system, and three with rotation. Let us use the symbols \mathbf{R}_s to represent the equilibrium position of the atom s , and \mathbf{x}_s to represent the displacement of the atom from its equilibrium position. Then in our simple model, the classical potential energy of the molecule can be written in the form

$$V = \frac{1}{2} \sum_{t>s}^N \sum_{s=1}^N k_{st} (|\mathbf{x}_s + \mathbf{R}_s - \mathbf{x}_t - \mathbf{R}_t| - |\mathbf{R}_s - \mathbf{R}_t|)^2 \quad (7.24)$$

Here k_{st} represents the force constant of the “spring” which connects atom s with atom t . Let us also introduce the notation

$$\begin{aligned} \mathbf{R}_{st} &\equiv \mathbf{R}_s - \mathbf{R}_t \\ \mathbf{x}_{st} &\equiv \mathbf{x}_s - \mathbf{x}_t \end{aligned} \quad (7.25)$$

Then, if we assume that $|\mathbf{x}_{st}| \ll |\mathbf{R}_{st}|$ and expand V in a Taylor series, we obtain the leading term

$$\begin{aligned} V &\approx \sum_{t>s}^N \sum_{s=1}^N \frac{k_{st}}{|\mathbf{R}_{st}|^2} (\mathbf{R}_{st} \cdot \mathbf{x}_{st})^2 \\ &\equiv \frac{1}{2} \sum_{t>s}^N \sum_{s=1}^N \sum_{\mu=1}^3 \sum_{\nu=1}^3 V_{s,\mu;t,\nu} x_{s,\mu} x_{t,\nu} \end{aligned} \quad (7.26)$$

where

$$\mathbf{x}_s \equiv (x_{s1}, x_{s2}, x_{s3}) \quad (7.27)$$

By diagonalizing the mass-weighted potential energy matrix

$$\frac{V_{s,\mu;t,\nu}}{\sqrt{m_s m_t}} \quad (7.28)$$

we can find the normal modes of the system, and as mentioned, six of them will be zero-frequency modes corresponding to translations and rotations of the entire system.

7.3 Commutation relations

Let us consider a simple harmonic oscillator whose Hamiltonian is given by

$$H = \frac{1}{2} (p^2 + \omega^2 q^2) \quad (7.29)$$

We can go over to a quantum treatment of the harmonic oscillator by letting let

$$p = -i\hbar \frac{\partial}{\partial q} \quad (7.30)$$

so that

$$[p, q] \equiv pq - qp = -i\hbar \quad (7.31)$$

Thus

$$[H, p] = \frac{\omega^2}{2} [q^2, p] = \frac{\omega^2}{2} (q [q, p] + [q, p] q) = i\hbar\omega^2 q \quad (7.32)$$

and similarly

$$[H, q] = \frac{1}{2} [p^2, q] = \frac{1}{2} (p [p, q] + [p, q] q) = -i\hbar p \quad (7.33)$$

Now suppose that $|n\rangle$ is an eigenfunction of H , so that

$$H|n\rangle = E_n|n\rangle \quad (7.34)$$

If we act on such an eigenfunction with an operator of the form $p \pm i\omega q$, then we can show by means of the commutation relations that the result is also an eigenfunction of H :

$$\begin{aligned} H(p + i\omega q)|n\rangle &= \{[H, p] + i\omega[H, q] + (p + i\omega q)H\} |n\rangle \\ &= \{i\hbar\omega^2 q + i\omega(-i\hbar p) + (p + i\omega q)E_n\} |n\rangle \\ &= \{(E_n + \hbar\omega)(p + i\omega q)\} |n\rangle \end{aligned} \quad (7.35)$$

and

$$\begin{aligned}
 H(p - i\omega q)|n\rangle &= \{[H, p] - i\omega[H, q] + (p - i\omega q)H\} |n\rangle \\
 &= \{i\hbar\omega^2 q - i\omega(-i\hbar p) + (p - i\omega q)E_n\} |n\rangle \\
 &= \{(E_n - \hbar\omega)(p - i\omega q)\} |n\rangle
 \end{aligned} \tag{7.36}$$

These equations show that $(p \pm i\omega q)|n\rangle$ is an eigenfunction of H corresponding to the raised or lowered energy $E_n \pm \hbar\omega$. The operator $p + i\omega q$ is thus a “raising operator” (sometimes called a “creation operator”). When it acts on an eigenfunction of the harmonic oscillator Hamiltonian H , it produces another eigenfunction of H , whose energy is raised by an amount $\hbar\omega$. Similarly, $p - i\omega q$ can be thought of as a “lowering operator”, or “annihilation operator”. Acting on an eigenfunction of H , it produces another eigenfunction, whose eigenvalue is lowered by an amount $\hbar\omega$.

If we continue to act on an eigenfunction of H with the lowering operator $p - i\omega q$, we must eventually reach the ground state, a state whose energy cannot be lowered further. Since it cannot lower the energy of the ground state, the lowering operator must give zero when it acts on it. Thus we have the relationship:

$$(p - i\omega q)|0\rangle = 0 \tag{7.37}$$

where $|0\rangle$ represents the ground state. If we next act on this equation from the left with the raising operator, we obtain:

$$\begin{aligned}
 (p + i\omega q)(p - i\omega q)|0\rangle &= (p^2 + \omega^2 q^2 + i\omega[p, q]) |0\rangle \\
 &= (2H - \hbar\omega) |0\rangle \\
 &= (2E_0 - \hbar\omega) |0\rangle = 0
 \end{aligned} \tag{7.38}$$

and thus

$$E_0 = \frac{1}{2}\hbar\omega \tag{7.39}$$

Combining this with our previous results, we obtain the entire energy spectrum of a quantum harmonic oscillator:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad n = 0, 1, 2, 3, \dots \tag{7.40}$$

7.4 Phonon creation and annihilation operators

It is convenient to define a normalized raising operator, which we shall call a “creation operator”:

$$a^\dagger \equiv -i\mathbb{N}(p + i\omega q) \tag{7.41}$$

and a normalized lowering operator, which we shall call an “annihilation operator”:

$$a \equiv i\mathbb{N}(p - i\omega q) \quad (7.42)$$

The normalizing constant, \mathbb{N} , is chosen in such a way that

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (7.43)$$

and

$$a|n+1\rangle = \sqrt{n+1}|n\rangle \quad (7.44)$$

Then

$$\begin{aligned} aa^\dagger|n\rangle &= \sqrt{n+1}a|n+1\rangle = (n+1)|n\rangle \\ &= \mathbb{N}^2(p - i\omega q)(p + i\omega q)|n\rangle \\ &= \mathbb{N}^2\{2H + i\omega[p, q]\}|n\rangle \\ &= \mathbb{N}^2(n+1)2\hbar\omega|n\rangle \end{aligned} \quad (7.45)$$

Solving for the normalization constant, we obtain

$$\mathbb{N} = \frac{1}{\sqrt{2\hbar\omega}} \quad (7.46)$$

Thus

$$\begin{aligned} a^\dagger &= \frac{-i}{\sqrt{2\hbar\omega}}(p + i\omega q) \\ a &= \frac{i}{\sqrt{2\hbar\omega}}(p - i\omega q) \end{aligned} \quad (7.47)$$

We can also express p and q in terms of the creation and annihilation operators:

$$\begin{aligned} p &= i\sqrt{\frac{\hbar\omega}{2}}(a^\dagger - a) \\ q &= -i\sqrt{\frac{\hbar\omega}{2}}(a^\dagger + a) \end{aligned} \quad (7.48)$$

We can see from these relationships that the creation and annihilation operators obey the commutation rules:

$$[a, a^\dagger] = 1 \quad (7.49)$$

$$[a, a] = 0 \quad (7.50)$$

$$[a^\dagger, a^\dagger] = 0 \quad (7.51)$$

The operator $a^\dagger a$ is called the “number operator”, since it is easy to show from the relationships given above that

$$a^\dagger a |n\rangle = n |n\rangle \quad (7.52)$$

We can also note that the harmonic oscillator Hamiltonian can be written in the form:

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \quad (7.53)$$

7.5 Collections of harmonic oscillators

Let us now consider a system whose Hamiltonian is a sum of simple harmonic oscillator Hamiltonians:

$$H = \sum_k H_k = \sum_k (p_k^2 + \omega_k^2 q_k^2) = \sum_k \hbar\omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right) \quad (7.54)$$

The commutation relationships for the momenta and coordinates of the system are

$$[p_k, q_{k'}] = -i\hbar\delta_{k,k'} \quad (7.55)$$

$$[p_k, p_{k'}] = 0 \quad (7.56)$$

$$[q_k, q_{k'}] = 0 \quad (7.57)$$

while those of the creation and annihilation operators are

$$[a_k, a_{k'}^\dagger] = \delta_{k,k'} \quad (7.58)$$

$$[a_k^\dagger, a_{k'}^\dagger] = 0 \quad (7.59)$$

$$[a_k, a_{k'}] = 0 \quad (7.60)$$

The eigenfunctions of the total Hamiltonian are products of simple harmonic oscillator eigenfunctions. If we use the symbol $|n_1, n_2, n_3, \dots\rangle$ to represent such an eigenfunction, then we can write:

$$\begin{aligned} H|n_1, n_2, n_3, \dots\rangle &= E_{n_1, n_2, n_3, \dots} |n_1, n_2, n_3, \dots\rangle \\ &= \sum_k \hbar\omega_k \left(n_k + \frac{1}{2} \right) |n_1, n_2, n_3, \dots\rangle \end{aligned} \quad (7.61)$$

Chapter 8

THE DIRAC EQUATION

8.1 Lorentz invariance and 4-vectors

Albert Einstein's special theory of relativity was built on the negative result of the Michaelson-Morley experiment, an experiment that attempted to measure the absolute velocity of the earth through space. Einstein boldly postulated that no experiment whatever can measure absolute motion, that is to say, according to his postulate it is impossible for an observer to know whether he is in a state of rest or in a state of uniform motion. All inertial frames are equivalent. Einstein's postulate has been amply confirmed by experiment, and today it is one of the basic principles of modern physics.

The equivalence of all inertial frames can be expressed in another way: Every fundamental physical law must exhibit symmetry between the space and time coordinates in such a way that ict enters on the same footing as the Cartesian coordinates x , y and z . (Here $i \equiv \sqrt{-1}$, while c is the velocity of light, and t is the time.) In relativistic theory, space and time combine to form a pseudo-Euclidean space-time continuum (Minkowski space). A transformation from one inertial frame to another (a Lorentz transformation) corresponds to a rotation in this space, and such a transformation must leave all fundamental physical laws invariant in form.

Every physical quantity that is represented by a 3-component vector in non-relativistic theory has a 4th component in the relativistic 4-dimensional space-time continuum. Thus, for example, the position vector $\mathbf{x} = (x, y, z)$ in 3-dimensional space has a 4th component in relativistic theory:

$$x_\lambda = (x, y, z, ict) = (\mathbf{x}, ict) \quad (8.1)$$

while the vector potential $\mathbf{A} = (A_x, A_y, A_z)$ in electromagnetic theory is the space component of a 4-vector, whose 4th component is i multiplied by the electrostatic potential ϕ :

$$A_\lambda = (A_x, A_y, A_z, i\phi) = (\mathbf{A}, i\phi) \quad (8.2)$$

Similarly, the current density vector $\mathbf{j} = (j_x, j_y, j_z)$ is the space-component of a 4-vector

$$j_\lambda = (j_x, j_y, j_z, ic\rho) = (\mathbf{j}, ic\rho) \quad (8.3)$$

whose time-component is ic multiplied by the charge density ρ . (Throughout this chapter we will represent 3-vectors by writing them in bold-face letters. Thus $j_\lambda = (\mathbf{j}, ic\rho)$ means that the first three components of the 4-vector j_λ are given by $\mathbf{j} = (j_x, j_y, j_z)$, while the 4th component is $ic\rho$.) The gradient operator $\boldsymbol{\partial}$ also becomes the space-component of a 4-vector in relativistic theory:

$$\partial_\lambda \equiv \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_4} \right) = \left(\boldsymbol{\partial}, -\frac{i}{c} \frac{\partial}{\partial t} \right) \quad (8.4)$$

while the Laplacian operator is replaced by the d'Alembertian operator:

$$\square \equiv \sum_{\lambda=1}^4 \partial_\lambda^2 \equiv \sum_{\lambda=1}^4 \frac{\partial^2}{\partial x_\lambda^2} \quad (8.5)$$

an operator which exhibits the required space-time symmetry, so that its form is the same in all inertial frames. In relativistic electrodynamics, the electric field vector $\boldsymbol{\mathcal{E}}$ and the magnetic field vector $\boldsymbol{\mathcal{H}}$ are components of an antisymmetric tensor $F_{\lambda',\lambda}$, which is related to A_λ by

$$F_{\lambda',\lambda} \equiv \partial_{\lambda'} A_\lambda - \partial_\lambda A_{\lambda'} = \begin{pmatrix} 0 & \mathcal{H}_z & -\mathcal{H}_y & -i\mathcal{E}_x \\ -\mathcal{H}_z & 0 & \mathcal{H}_x & -i\mathcal{E}_y \\ \mathcal{H}_y & -\mathcal{H}_x & 0 & -i\mathcal{E}_z \\ i\mathcal{E}_x & i\mathcal{E}_y & i\mathcal{E}_z & 0 \end{pmatrix} \quad (8.6)$$

The 4-vector A_λ , which represents the electromagnetic potential, is related to the 4-vector representing current density by

$$\square A_\lambda = -\frac{4\pi}{c} j_\lambda \quad (8.7)$$

When both the current density j_λ and the electromagnetic potential 4-vector A_λ are independent of time, equation (8.7) reduces to:

$$\nabla_1^2 A_\lambda(\mathbf{x}_1) = -\frac{4\pi}{c} j_\lambda(\mathbf{x}_1) \quad (8.8)$$

which has the Green's function solution

$$A_\lambda(\mathbf{x}_1) = \frac{1}{c} \int d^3x_2 \frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|} j_\lambda(\mathbf{x}_2) \quad (8.9)$$

We can see that (8.9) is a solution to (8.8) because

$$\nabla_1^2 \frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|} = -4\pi \delta^3(\mathbf{x}_1 - \mathbf{x}_2) \equiv -4\pi \delta(x_1 - x_2) \delta(y_1 - y_2) \delta(z_1 - z_2) \quad (8.10)$$

8.2. THE DIRAC EQUATION FOR AN ELECTRON IN AN EXTERNAL ELECTROMAGNETIC POTENTIAL

and therefore

$$\begin{aligned}
 \nabla_1^2 A_\lambda(\mathbf{x}_1) &= \frac{1}{c} \int d^3x_2 \nabla_1^2 \frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|} j_\lambda(\mathbf{x}_2) \\
 &= -\frac{4\pi}{c} \int d^3x_2 \delta^3(\mathbf{x}_1 - \mathbf{x}_2) j_\lambda(\mathbf{x}_2) \\
 &= -\frac{4\pi}{c} j_\lambda(\mathbf{x}_1)
 \end{aligned} \tag{8.11}$$

The subscript 1 on the Laplacian operator means that the operator is acting on the coordinates of the field-point \mathbf{x}_1 rather than on the source-point, \mathbf{x}_2 .

Because of charge conservation, the current density 4-vector obeys the condition

$$\sum_{\lambda=1}^4 \partial_\lambda j_\lambda = 0 \tag{8.12}$$

Since the current density is related to the electromagnetic potential 4-vector through (8.7), it is natural to work in the Lorentz gauge, where a similar condition is imposed on A_λ :

$$\sum_{\lambda=1}^4 \partial_\lambda A_\lambda = 0 \tag{8.13}$$

Equations (8.7) and (8.13) are Maxwell's equations in a vacuum, written in a form that makes the space-time symmetry apparent.

8.2 The Dirac equation for an electron in an external electromagnetic potential

P.A.M. Dirac's relativistic wave equation for an electron moving in an external potential A_λ can be written in the form:

$$\left[\sum_{\lambda=1}^4 \gamma_\lambda \left(\partial_\lambda - \frac{i}{c} A_\lambda \right) + c \right] \chi_\mu = 0 \tag{8.14}$$

where atomic units are used and where the γ_λ 's are 4×4 matrices:

$$\gamma_1 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad \gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{8.15}$$

$$\gamma_3 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} \quad \gamma_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (8.16)$$

In atomic units, the electron rest-mass is equal to 1, and Planck's constant divided by 2π is also equal to 1, while the velocity of light has a value equal to the reciprocal of the fine structure constant:

$$m_0 = 1 \quad \hbar = 1 \quad c = 137.036 \quad (8.17)$$

From the definitions of the γ_λ 's, it follows that they anticommute:

$$\gamma_{\lambda'}\gamma_\lambda + \gamma_\lambda\gamma_{\lambda'} = 2I\delta_{\lambda',\lambda} \quad (8.18)$$

In equation (8.18), I is a 4×4 unit matrix. Solutions to the 1-electron Dirac equation are 4-component spinors.

8.3 Time-independent problems

In the special case where the external electromagnetic potential 4-vector A_λ is independent of time, it is convenient to write the Dirac equation (8.14) in a different form, where we introduce the notation

$$\boldsymbol{\alpha} = i\gamma_0\boldsymbol{\gamma} \quad \gamma_0 \equiv \gamma_4 \quad (8.19)$$

From equations (8.15), (8.16) and (8.19) it follows that the components of the 3-vector $\boldsymbol{\alpha}$ can be written in block form as

$$\alpha_j = \left(\begin{array}{c|c} 0 & \sigma_j \\ \hline \sigma_j & 0 \end{array} \right) \quad j = 1, 2, 3 \quad (8.20)$$

where, in the off-diagonal blocks, σ_j , $j = 1, 2, 3$ are the 2×2 Pauli spin matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (8.21)$$

For time-independent problems, the Dirac equation for a single electron can then be written in the form:

$$[H - \epsilon_\mu] \chi_\mu(\mathbf{x}) = 0 \quad (8.22)$$

where

$$H = -ic\boldsymbol{\alpha} \cdot \left(\boldsymbol{\partial} - \frac{i}{c}\mathbf{A}(\mathbf{x}) \right) + I\phi(\mathbf{x}) + \gamma_0c^2 \quad (8.23)$$

is the Dirac Hamiltonian of an electron moving in a constant external electromagnetic potential, ϵ_μ is the 1-electron energy, and $\chi_\mu(\mathbf{x})$ is the 4-component time-independent

spinor of the electron. The kinetic energy term in the Dirac Hamiltonian is given by

$$-ic\boldsymbol{\alpha} \cdot \boldsymbol{\partial} = -ic \begin{pmatrix} 0 & 0 & \partial_3 & \partial_- \\ 0 & 0 & \partial_+ & -\partial_3 \\ \partial_3 & \partial_- & 0 & 0 \\ \partial_+ & -\partial_3 & 0 & 0 \end{pmatrix} \quad (8.24)$$

where

$$\partial_{\pm} \equiv \partial_1 \pm i\partial_2 \quad (8.25)$$

Similarly, the part of the Dirac Hamiltonian involving potentials is

$$-\boldsymbol{\alpha} \cdot \mathbf{A} + I\phi = \begin{pmatrix} \phi & 0 & -A_3 & -A_- \\ 0 & \phi & -A_+ & A_3 \\ -A_3 & -A_- & \phi & 0 \\ -A_+ & A_3 & 0 & \phi \end{pmatrix} \quad (8.26)$$

where

$$A_{\pm} \equiv A_1 \pm iA_2 \quad (8.27)$$

8.4 The Dirac equation for an electron in the field of a nucleus

When $\mathbf{A}(\mathbf{x}) = 0$, and $\phi(\mathbf{x}) = -Z/r$, equation (8.22) reduces to

$$\left[-ic\boldsymbol{\alpha} \cdot \boldsymbol{\partial} - \frac{Z}{r} + \gamma_0 c^2 - \epsilon_{\mu} \right] \chi_{\mu}(\mathbf{x}) = 0 \quad (8.28)$$

which is the Dirac equation for an electron moving in the attractive electrostatic potential of a nucleus with charge Z . Equation (8.28) can be solved exactly, and the solutions have the form

$$\chi_{\mu}(\mathbf{x}) = \chi_{njlM}(\mathbf{x}) = \begin{pmatrix} ig_{njl}(r)\Omega_{j,l,M}(\theta, \varphi) \\ -f_{njl}(r)\Omega_{j,2j-l,M}(\theta, \varphi) \end{pmatrix} \quad (8.29)$$

Examples are shown in equations (??) and (??). In equation (8.29), the angular function $\Omega_{j,l,M}(\theta, \varphi)$ is a two-component ‘‘spherical spinor’’, which is an eigenfunction of orbital angular momentum corresponding to the quantum number l , total angular momentum (orbital plus spin) with quantum number j , and the z -component of total angular momentum, with quantum number M . The spherical spinors are built up from spherical harmonics and 2-component spinors by combining them with the appropriate Clebsch-Gordan coefficients in such a way as to produce eigenfunctions of total angular momentum. The Clebsch-Gordan coefficients that enter are different, depending on whether $j = l + \frac{1}{2}$ or $j = l - \frac{1}{2}$.

When $j = l + \frac{1}{2}$,

$$\Omega_{j,l,M}(\theta, \varphi) = \begin{pmatrix} \sqrt{\frac{l+M+\frac{1}{2}}{2l+1}} Y_{l,M-\frac{1}{2}}(\theta, \varphi) \\ \sqrt{\frac{l-M+\frac{1}{2}}{2l+1}} Y_{l,M+\frac{1}{2}}(\theta, \varphi) \end{pmatrix} \quad (8.30)$$

while when $j = l - \frac{1}{2}$,

$$\Omega_{j,l,M}(\theta, \varphi) = \begin{pmatrix} -\sqrt{\frac{l-M+\frac{1}{2}}{2l+1}} Y_{l,M-\frac{1}{2}}(\theta, \varphi) \\ \sqrt{\frac{l+M+\frac{1}{2}}{2l+1}} Y_{l,M+\frac{1}{2}}(\theta, \varphi) \end{pmatrix} \quad (8.31)$$

The radial function $g_{njl}(r)$ is much larger than $f_{njl}(r)$. The large and small radial functions are defined respectively by

$$g_{njl}(r) = \mathcal{N} r^{\gamma-1} e^{-Zr/\bar{n}} (W_1(r) - W_2(r)) \quad (8.32)$$

and

$$f_{njl}(r) = \mathcal{N} \sqrt{\frac{c^2 - \epsilon_{nj}}{c^2 + \epsilon_{nj}}} r^{\gamma-1} e^{-Zr/\bar{n}} (W_1(r) + W_2(r)) \quad (8.33)$$

where

$$\begin{aligned} W_1(r) &\equiv n_r F \left(j + \frac{1}{2} - n + 1 \middle| 2\gamma + 1 \middle| \frac{2Zr}{\bar{n}} \right) \\ W_2(r) &\equiv (\bar{n} - \kappa) F \left(j + \frac{1}{2} - n \middle| 2\gamma + 1 \middle| \frac{2Zr}{\bar{n}} \right) \end{aligned} \quad (8.34)$$

with

$$\kappa \equiv \begin{cases} -(j + \frac{1}{2}) & j = l + \frac{1}{2} \\ j + \frac{1}{2} & j = l - \frac{1}{2} \end{cases} \quad (8.35)$$

$$\gamma \equiv \sqrt{\left(j + \frac{1}{2} \right)^2 - \left(\frac{Z}{c} \right)^2} \quad (8.36)$$

$$n_r \equiv n - j - \frac{1}{2} \quad (8.37)$$

and

$$\bar{n} \equiv \sqrt{n^2 - 2n_r(j + \frac{1}{2} - \gamma)} \quad (8.38)$$

Just as in the definition of the non-relativistic hydrogenlike orbitals, $F(a|b|\zeta)$ is a confluent hypergeometric function:

$$F(a|b|\zeta) \equiv 1 + \frac{a}{b}\zeta + \frac{a(a+1)}{b(b+1)2!}\zeta^2 + \dots \quad (8.39)$$

When $Z \ll 137$, the 1-electron energies

$$\epsilon_{nj} = \frac{c^2}{\sqrt{1 + \left(\frac{Z}{c(\gamma+n_r)}\right)^2}} \quad (8.40)$$

Chapter 9

INTERACTION BETWEEN MATTER AND RADIATION

9.1 Lagrangian densities for fields

Let us now turn our attention to the problem of how to treat continuous systems, or fields, using the Lagrangian formalism. We shall begin by considering a 3-dimensional space, with coordinates x , y and z , and later we will generalize to spaces of higher dimension and also to spaces where the metric is non-Euclidean. We have just been discussing normal modes for small vibrations of a classical system of particles about the equilibrium positions. We might also wish to discuss the normal modes of a violin string or a drumhead, or the harmonics of sound waves inside a closed container. In these examples, the medium can be regarded as continuous as long as we are considering low frequencies, but at high frequencies, where the wavelength is comparable to the distance between the particles of which the system is composed, this picture breaks down. But at low frequencies, the displacement from equilibrium can be treated as a field.

There are other fields that one might wish to consider, such as electromagnetic fields. What we now wish to discuss is the problem of how to start with a wave equation in classical mechanics, and how to pass over to a quantum mechanical treatment of the system. In order to do this, it is useful to introduce the concept of a Lagrangian density, \mathcal{L} . This is sometimes defined as the kinetic energy per unit volume of space, minus the density of potential energy. However, in some cases, for example when the Lagrangian must be invariant under a Lorentz transformation, this definition cannot be used, and we must instead define the Lagrangian density as that which gives the correct wave equations. The usual Lagrangian of the system is the Lagrangian density integrated over the whole volume of the system.

$$L = \int \int \int \mathcal{L} \, dx \, dy \, dz \quad (9.1)$$

The variational principle from which the equations of motion can be derived can be written

in the form

$$\int \delta L dt = \int \int \int \int \delta \mathcal{L} dx dy dz dt = 0 \quad (9.2)$$

Suppose that the continuous system is described by a number of fields, $\phi_1, \phi_2, \dots, \phi_n$, which are continuous functions of the space and time coordinates x, y, z and t . For example, ϕ_1, ϕ_2, \dots might represent the components of the electromagnetic field. Suppose also that the Lagrangian density can be expressed in terms of the fields and their first derivatives with respect to the space and time coordinates, so that

$$\mathcal{L} = \mathcal{L} \left(\phi_j, \frac{\partial \phi_j}{\partial x^\mu} \right) \quad \begin{cases} j = 1, 2, \dots, n \\ \mu = 1, 2, 3, 4 \\ x^4 = ict \end{cases} \quad (9.3)$$

If we change one of the fields by a small amount, the resulting change in the Lagrangian density is given by

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \phi_i} \delta \phi_i + \sum_{\mu=1}^4 \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x^\mu)} \frac{\partial (\delta \phi_i)}{\partial x^\mu} \quad (9.4)$$

The variational principle then becomes

$$\int \delta L dt = \int \int \int \int \left[\frac{\partial \mathcal{L}}{\partial \phi_i} + \sum_{\mu=1}^4 \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x^\mu)} \frac{\partial (\delta \phi_i)}{\partial x^\mu} \right] dx dy dz dt \quad (9.5)$$

The second term on the right-hand side of (9.5) can be integrated by parts in a manner analogous to equations (??)-(??). Thus, for example, we can write

$$\begin{aligned} & \int \int \int \int \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x)} \frac{\partial (\delta \phi_i)}{\partial x} dx dy dz dt \\ &= \left[\int \int \int \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x)} \delta \phi_i \right]_{x_a}^{x_b} dy dz dt \\ &- \int \int \int \int \frac{\partial}{\partial x} \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x)} \delta \phi_i dx dy dz dt \end{aligned} \quad (9.6)$$

The first term to the right of the = sign is required to vanish because the variation must be zero at the end points. In this way, we can rewrite the variational principle (9.2) in the form:

$$\int \delta L dt = \int \int \int \int \left[\frac{\partial \mathcal{L}}{\partial \phi_i} - \sum_{\mu=1}^4 \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x^\mu)} \right] \delta \phi_i dx dy dz dt = 0 \quad (9.7)$$

Since the variation $\delta \phi_i$ is arbitrary, (9.7) can be fulfilled only if

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \sum_{\mu=1}^4 \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial \phi_i / \partial x^\mu)} = 0 \quad (9.8)$$

Thus if we are able to write down the Lagrangian density of a continuous system, the variational principle leads us to the equations of motion (9.8). Conversely, knowledge of the equations of motion of a continuous system can be used to infer the Lagrangian density from which they are derived. The canonical conjugate of the field ϕ^i is called the *momentum density*, and is defined as

$$\pi_i \equiv \frac{\partial \mathcal{L}}{\partial(\partial\phi_i/\partial t)} \quad (9.9)$$

and the Hamiltonian density \mathcal{H} , corresponding to the Lagrangian density \mathcal{L} , is given by

$$\mathcal{H} = \sum_{i=1}^n \pi_i \frac{\partial\phi^i}{\partial t} - \mathcal{L} \quad (9.10)$$

Notice that equation (9.8) is invariant under a Lorentz transformation. but equation (9.10) is not, since the time coordinate is singled out for special treatment. This corresponds to the fact that energy, by itself, is not a Lorentz-invariant quantity, but only the 4th component of a 4-vector. It is possible to perform approximately relativistic calculations using the Hamiltonian formulation, but the Lagrangian formalism should be used rather than the Hamiltonian in cases where Lorentz invariance is important. Having found the equations of motion (9.8), we can next solve them, subject to the boundary conditions, and find solutions corresponding to pure harmonic oscillations. The fields are then expressed as sums of these harmonic modes. When the sums are substituted into the Lagrangian $L = \int \int \int \mathcal{L} \, dx \, dy \, dz$, then L can be recognized as a sum of simple harmonic oscillator Lagrangians. The quantization of these can proceed in the standard way.

9.2 Electromagnetic potentials

Let us now attempt to apply these methods to the electromagnetic potential 4-vector, A_λ . This is a case where the definition of the Lagrangian density as the density of kinetic energy minus the density of potential energy cannot be used. We must instead use the concept of the Lagrangian density as a quantity which leads, through the variational principle, to the correct wave equations. In the case of the electromagnetic potentials, we know that they obey the wave equation

$$\square^2 A_\lambda = -\frac{4\pi}{c} j_\lambda \quad j_\lambda = 1, 2, 3, 4 \quad (9.11)$$

where

$$\square^2 \equiv \sum_{\mu=1}^4 \frac{\partial^2}{\partial x^{\mu 2}} = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \quad x^4 = ict \quad (9.12)$$

We must invent a Lagrangian density that leads to the wave equation (9.11). The way to do this is to try all possibilities, starting with the simplest ones, always remembering

that the forms must be Lorentz invariant. The Lagrangian density which actually leads to equation (9.11) is

$$\mathcal{L} = -\frac{1}{2} \sum_{\lambda=1}^4 \sum_{\mu=1}^4 \left(\frac{\partial A_\lambda}{\partial x^\mu} \right)^2 + \frac{4\pi}{c} \sum_{\lambda=1}^4 A_\lambda j_\lambda \quad (9.13)$$

From (9.8) we have:

$$\frac{\partial \mathcal{L}}{\partial A_\lambda} - \sum_{\mu=1}^4 \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial A_\lambda / \partial x^\mu)} = 0 \quad (9.14)$$

With the trial Lagrangian density of equation (9.13), we have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial A_\lambda} &= \frac{4\pi}{c} j_\lambda \\ - \sum_{\mu=1}^4 \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial A_\lambda / \partial x^\mu)} &= \sum_{\mu=1}^4 \frac{\partial}{\partial x^\mu} \frac{\partial A_\lambda}{\partial x^\mu} = \square^2 A_\lambda \end{aligned} \quad (9.15)$$

Substituting these results into the Euler-Lagrange equations (9.14) gives

$$\frac{4\pi}{c} j_\lambda + \square^2 A_\lambda = 0 \quad (9.16)$$

which is just the wave equation that we wanted to reproduce. The next step is to solve the wave equation and find the harmonic normal modes of the system. However, we can see that if we include the term involving the current density, the solutions will not necessarily be harmonic. Therefore we omit the interaction term, hoping to include it later by means of perturbation theory. This leaves us with the homogeneous wave equation:

$$\square^2 A_\lambda = 0 \quad (9.17)$$

which has harmonic solutions of the form

$$A_{\mathbf{k},\lambda} = \mathcal{N} \{ a_{\mathbf{k},\lambda}(t) e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k},\lambda}^*(t) e^{-i\mathbf{k}\cdot\mathbf{x}} \} \quad (9.18)$$

Here the values of \mathbf{k} are restricted by periodic boundary conditions imposed at the faces of a box of normalization of volume V . \mathcal{N} is a constant which we will choose later in a convenient way. $a_{\mathbf{k},\lambda}$ and its complex conjugate $a_{\mathbf{k},\lambda}^*$ are Fourier coefficients which are harmonic functions of time. Substituting (9.18) into (9.17), we can see that $A_{\mathbf{k},\lambda}$ will be a solution, provided that

$$\begin{aligned} \frac{\partial^2 a_{\mathbf{k},\lambda}}{\partial t^2} + \omega_k^2 a_{\mathbf{k},\lambda} &= 0 \\ \frac{\partial^2 a_{\mathbf{k},\lambda}^*}{\partial t^2} + \omega_k^2 a_{\mathbf{k},\lambda}^* &= 0 \end{aligned} \quad (9.19)$$

where $\omega_k = |\mathbf{k}|/c$. Having found harmonic solutions to the wave equation, we must expand the potentials in terms of these, and substitute them back into the Lagrangian. We then hope to be able to recognize the Lagrangian as a sum of simple harmonic oscillator Lagrangians. Thus, expanding the potentials in terms of their normal modes, we write

$$A_\lambda(\mathbf{x}, t) = \mathcal{N} \sum'_{\mathbf{k}} (a_{\mathbf{k},\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k},\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{x}}) \quad (9.20)$$

In (9.20), the sum runs over only half of k -space for the following reason: By writing the harmonic solutions in the form (9.18), (i.e as a function plus its complex conjugate), we ensure that they shall be real. However, we do not want the solution corresponding to positive \mathbf{k} to repeat that corresponding to $-\mathbf{k}$. Therefore we restrict the sum to half of k -space, and we denote this with a prime on the summation. From (9.20), we have

$$\frac{\partial A_\lambda}{\partial x^\mu} = \mathcal{N} \sum'_{\mathbf{k}} (ik_\mu a_{\mathbf{k},\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} - ik_\mu a_{\mathbf{k},\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{x}}) \quad \mu = 1, 2, 3 \quad (9.21)$$

and

$$\frac{\partial A_\lambda}{\partial x^4} = \frac{\mathcal{N}}{ic} \sum'_{\mathbf{k}} \left(\frac{\partial a_{\mathbf{k},\lambda}}{\partial t} e^{i\mathbf{k}\cdot\mathbf{x}} + \frac{\partial a_{\mathbf{k},\lambda}^*}{\partial t} e^{-i\mathbf{k}\cdot\mathbf{x}} \right) \quad (9.22)$$

Let us now calculate the Hamiltonian density \mathcal{H} associated with the electromagnetic potentials. According to our earlier discussion, the Hamiltonian density is given by

$$\begin{aligned} \mathcal{H} &= \sum_{\lambda=1}^4 \frac{\partial A_\lambda}{\partial t} \left\{ \frac{\partial \mathcal{L}}{\partial(\partial A_\lambda/\partial t)} \right\} - \mathcal{L} \\ &= \frac{1}{2} \sum_{\lambda=1}^4 \left[\sum_{\mu=1}^3 \left(\frac{\partial A_\lambda}{\partial x^\mu} \right)^2 - \left(\frac{\partial A_\lambda}{\partial x^4} \right)^2 \right] - \frac{4\pi}{c} \sum_{\lambda=1}^4 j_\lambda A_\lambda \\ &= \mathcal{H}_{\text{radiation}} + \mathcal{H}_{\text{interaction}} \equiv \mathcal{H}_r + \mathcal{H}_i \end{aligned} \quad (9.23)$$

Now we know that the Fourier coefficients $a_{\mathbf{k},\lambda}$ depend harmonically on time, so that

$$\begin{aligned} \frac{\partial a_{\mathbf{k},\lambda}}{\partial t} &= \pm ikc a_{\mathbf{k},\lambda} \\ \frac{\partial a_{\mathbf{k},\lambda}^*}{\partial t} &= \mp ikc a_{\mathbf{k},\lambda}^* \end{aligned} \quad (9.24)$$

Also we know that the box-normalized functions $e^{i\mathbf{k}\cdot\mathbf{x}}/\sqrt{V}$ obey the orthonormality relations

$$\frac{1}{V} \int d^3x e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} = \delta_{\mathbf{k},\mathbf{k}'} \quad (9.25)$$

Substituting (9.25) and (9.24) into (9.23), we obtain the Hamiltonian for the free electromagnetic field in the form

$$H_r = \int d^3x \mathcal{H}_r = V \sum_{\mathbf{k}} \sum_{\lambda=1}^4 \mathcal{N}^2 (a_{\mathbf{k},\lambda} a_{\mathbf{k},\lambda}^* + a_{\mathbf{k},\lambda}^* a_{\mathbf{k},\lambda}) \quad (9.26)$$

Notice that the terms proportional to $a_{\mathbf{k},\lambda} a_{\mathbf{k},\lambda}$ and $a_{\mathbf{k},\lambda}^* a_{\mathbf{k},\lambda}^*$ do not occur in H_r because the sum over \mathbf{k} is restricted to one half of k -space. With the substitutions

$$\begin{aligned} a_{\mathbf{k},\lambda} &= \frac{1}{\sqrt{2kc}} (-ip_{\mathbf{k},\lambda} + kcq_{\mathbf{k},\lambda}) \\ a_{\mathbf{k},\lambda}^* &= \frac{1}{\sqrt{2kc}} (ip_{\mathbf{k},\lambda} + kcq_{\mathbf{k},\lambda}) \end{aligned} \quad (9.27)$$

H_r becomes

$$H_r = \frac{V}{c} \sum_{\mathbf{k}} \sum_{\lambda=1}^4 \mathcal{N}^2 k (p_{\mathbf{k},\lambda}^2 + k^2 c^2 q_{\mathbf{k},\lambda}^2) \quad (9.28)$$

This can be identified as the Hamiltonian of a collection of harmonic oscillators, provided that

$$\mathcal{N} = \sqrt{\frac{c}{2kV}} \quad (9.29)$$

If we make the identification $a_{\mathbf{k},\lambda}^* \rightarrow a_{\mathbf{k},\lambda}^\dagger$, and choose the value of \mathcal{N} given by (9.29), then H_r takes on the form

$$H_r = \frac{1}{2} \sum_{\mathbf{k}} \sum_{\lambda=1}^4 kc (a_{\mathbf{k},\lambda} a_{\mathbf{k},\lambda}^\dagger + a_{\mathbf{k},\lambda}^\dagger a_{\mathbf{k},\lambda}) \quad (9.30)$$

This can be recognized as the Hamiltonian of a collection of harmonic oscillators with frequencies $\omega_k = kc$. In the quantum treatment of such a system, $a_{\mathbf{k},\lambda}^\dagger$ and $a_{\mathbf{k},\lambda}$ correspond to creation and annihilation operators, and they obey the commutation relations

$$\begin{aligned} [a_{\mathbf{k},\lambda}, a_{\mathbf{k},\lambda}^\dagger] &= \delta_{\mathbf{k},\mathbf{k}'} \delta_{\lambda,\lambda'} \\ [a_{\mathbf{k},\lambda}^\dagger, a_{\mathbf{k},\lambda}^\dagger] &= 0 \\ [a_{\mathbf{k},\lambda}, a_{\mathbf{k},\lambda}] &= 0 \end{aligned} \quad (9.31)$$

(See equations (??)-(??)). Expanding the potentials in terms of the normal modes of the system, we have

$$A_\lambda = \sqrt{\frac{c}{2kV}} \sum_{\mathbf{k}} (a_{\mathbf{k},\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k},\lambda}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}}) \quad (9.32)$$

The interaction Hamiltonian, expressed in terms of the photon creation and annihilation operators, then becomes:

$$\begin{aligned} H_i &= -\frac{4\pi}{c} \sum_{\lambda=1}^4 \int d^3x A_\lambda j_\lambda \\ &= -\sum_{\mathbf{k}} \sum_{\lambda=1}^4 \frac{4\pi}{\sqrt{2\omega_k V}} \left[a_{\mathbf{k},\lambda} \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} j_\lambda + a_{\mathbf{k},\lambda}^\dagger \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} j_\lambda \right] \end{aligned} \quad (9.33)$$

The potential A_λ is a 4-vector

$$A_\lambda = \{\mathbf{A}, i\varphi\} \quad (9.34)$$

of which the first three (spacelike) components are the vector potential, which is related to the magnetic field, while the last (timelike) component is the scalar or electrostatic potential. The current density j_λ is also a 4-vector:

$$j_\lambda = \{\mathbf{j}, i\rho\} \quad (9.35)$$

is also a 4-vector. Its first three components form the current density vector in 3-dimensional space, while the 4th component represents the charge density. For an electron in the state ψ satisfying the Dirac equation, its contribution to the current density is given by

$$\mathbf{j} = ec\psi^* \boldsymbol{\alpha} \psi \quad (9.36)$$

where the three components of the vector $\boldsymbol{\alpha}$ are the Dirac matrices representing the velocity operator:

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (9.37)$$

$$\alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad (9.38)$$

$$\alpha_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad (9.39)$$

In the non-relativistic limit, this reduces approximately to

$$\mathbf{j}_{N.R.} \rightarrow \frac{e}{m_e} \psi^* \mathbf{p} \psi \quad (9.40)$$

For an electron in the state ψ , its contribution to the charge density ρ is given by

$$\rho = e\psi^* \psi \quad (9.41)$$

9.3 Separation of the longitudinal and transverse potentials

The approximate Hamiltonian of the electron-photon system

The sum of the Hamiltonian that represents the radiation field by itself and the Hamiltonian representing the interaction between matter and radiation can be written in the form:

$$H_r + H_i = \sum_{\lambda=1}^4 H_{\lambda} \quad (9.42)$$

where

$$H_{\lambda} = \int d^3x \left[\frac{1}{8\pi} \left\{ \sum_{\mu=1}^3 \left(\frac{\partial A_{\lambda}}{\partial x_{\mu}} \right)^2 - \left(\frac{\partial A_{\lambda}}{\partial x_4} \right)^2 \right\} - \frac{1}{c} j_{\lambda} A_{\lambda} \right] \quad (9.43)$$

Thus the Hamiltonian can be split up into four parts, each of which depends on only one of the potentials. This means that the potentials would be independent of each other, except that the components of the 4-current j_{λ} are not independent, but obey a conservation law. For most physical and chemical applications, it is convenient to treat the vector potential \mathbf{A} in the way that we have outlined above, but to treat the scalar potential φ in a different way. The classical wave equation for the scalar potential is

$$\square^2 A_4 = \frac{4\pi}{c} j_4 \quad (9.44)$$

or

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \varphi = -4\pi\rho = -4\pi e\psi^*\psi \quad (9.45)$$

If we are in the near neighborhood of a charge distribution, the time-dependence of φ can usually be neglected in comparison with $\nabla^2\varphi$, and we have the approximate relationship

$$\nabla^2\varphi = -4\pi\rho \quad (9.46)$$

which has the solution

$$\varphi(\mathbf{x}) = \int d^3x' \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \quad (9.47)$$

We now introduce an approximation which is useful for practical calculations in atomic or molecular physics. We divide the charged particles in the universe into two parts: (1) the small system that we intend to study, and (2) all the rest. We assume that in some approximation, the vector potential in our small system can be neglected in comparison with the scalar potential. Conversely, we assume that for the interaction of our small system with

the distant remainder of the universe, the scalar potential φ can be neglected in comparison with the vector potential \mathbf{A} . In this approximation, the Hamiltonian representing the interaction of the particles in our small system with other charges, both inside and outside it, is

$$H_i = -\frac{1}{c} \int d^3x \mathbf{j} \cdot \mathbf{A} + \int d^3x \rho \varphi \quad (9.48)$$

where \mathbf{j} , ρ and φ belong to the small system that we are studying, while the vector potential \mathbf{A} belongs to the universe outside. We can write the approximate interaction Hamiltonian H_i in the form

$$\begin{aligned} H_i = & - \sum_{\mathbf{k}} \sum_{\lambda=1}^3 \sqrt{\frac{2\pi\hbar c}{V k}} \left[a_{\mathbf{k},\lambda} \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} \psi^* \alpha_\lambda \psi + a_{\mathbf{k},\lambda}^\dagger \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \psi^* \alpha_\lambda \psi \right] \\ & + e^2 \int d^3x \int d^3x' \frac{\psi^*(\mathbf{x}) \psi(\mathbf{x}) \psi^*(\mathbf{x}') \psi(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \end{aligned} \quad (9.49)$$

9.4 Linear polarization

The equations

$$\square^2 A_\lambda = -\frac{4\pi}{c} j_\lambda \quad \lambda = 1, 2, 3, 4 \quad (9.50)$$

are equivalent to Maxwell's equations for free space, provided that we impose the subsidiary condition

$$\sum_{\lambda=1}^4 \frac{\partial A_\lambda}{\partial x_\lambda} = 0 \quad x_4 \equiv ict \quad (9.51)$$

This is called the ‘‘Lorenz condition’’, and it defines the ‘‘gauge’’ in which we will be working. The Lorenz condition is a natural one to impose, since the potentials A_λ arise from 4-currents which obey the conservation law:

$$\sum_{\lambda=1}^4 \frac{\partial j_\lambda}{\partial x_\lambda} = 0 \quad (9.52)$$

When $A_4 = 0$, the Lorenz condition reduces to

$$\nabla \cdot \mathbf{A} = 0 \quad (9.53)$$

which requires the divergence of the vector potential to vanish. Let us consider a case where $A_4 = 0$, and where \mathbf{A} represents a plane wave propagating through space with wave number \mathbf{k} . Then we can write \mathbf{A} in the form

$$\mathbf{A} = A_0 \operatorname{Re} \left\{ \mathbf{u} e^{i(\mathbf{k}\cdot\mathbf{x} - kct)} \right\} \quad (9.54)$$

where A_0 is a constant and \mathbf{u} is a unit vector in the direction of \mathbf{A} . Then the Lorenz condition requires that

$$\nabla \cdot \mathbf{A} = A_0 \text{Re} \{ \mathbf{k} \cdot \mathbf{u} e^{i(\mathbf{k} \cdot \mathbf{x} - kct)} \} = 0 \quad (9.55)$$

which will be satisfied for all values of \mathbf{x} if

$$\mathbf{k} \cdot \mathbf{u} = 0 \quad (9.56)$$

In other words, the unit vector \mathbf{u} must be perpendicular to the direction of propagation of the wave. The plane monochromatic wave described by these equations is said to be linearly polarized, with polarization vector \mathbf{u} . For a fixed value of \mathbf{k} , there are two linearly independent polarization vectors, \mathbf{u}_1 and \mathbf{u}_2 , which are perpendicular to each other, and also perpendicular to \mathbf{k} . For example, if the z -axis is chosen to be in the direction of \mathbf{k} , then \mathbf{u}_1 and \mathbf{u}_2 can be chosen to be in the directions of the x -axis and y -axis respectively.

9.5 Spontaneous photon emission

Using the electron creation and annihilation operator notation introduced in Section 5.9 of this book, we can write the Hamiltonians of the electrons in the small system that we are studying and the Hamiltonian of the radiation field by itself in the form

$$\begin{aligned} H_0 &= H_r + H_e \\ &= \frac{1}{2} \sum_{\mathbf{k}, \lambda} hck \left(a_{\mathbf{k}, \lambda}^\dagger a_{\mathbf{k}, \lambda} + \frac{1}{2} \right) + \sum_{r, s} H_{r, s}^c b_r^\dagger b_s + \frac{1}{2} \sum_{r, s, t, u} g_{r, s|t, u} b_r^\dagger b_s^\dagger b_t b_u \end{aligned} \quad (9.57)$$

We can also introduce a number of states which are eigenfunctions of H_0 :

$$\begin{aligned} H_0|A\rangle &= E_A|A\rangle \\ H_0|B\rangle &= E_B|B\rangle \\ H_0|C\rangle &= E_C|C\rangle \\ &\vdots \quad \vdots \quad \vdots \end{aligned} \quad (9.58)$$

We would like to calculate the probability per unit time that the interaction Hamiltonian will induce a transition between two of these eigenstates, $|A\rangle$ and $|B\rangle$, accompanied by the spontaneous emission of a photon. To do so, we start by representing $|B\rangle$ in the form:

$$|B\rangle = \frac{1}{\sqrt{n_{\mathbf{k}, \lambda} + 1}} b_f^\dagger b_i a_{\mathbf{k}, \lambda}^\dagger |A\rangle \quad (9.59)$$

In other words, $|B\rangle$ differs from $|A\rangle$ by having an extra photon in the mode \mathbf{k}, λ . and by the transition $\chi_i \rightarrow \chi_f$. We can check that $|B\rangle$ is properly normalized by calculating its scalar product with its adjunct state

$$\langle B| = \frac{1}{\sqrt{n_{\mathbf{k}, \lambda} + 1}} \langle A| a_{\mathbf{k}, \lambda} b_i^\dagger b_f \quad (9.60)$$

Making use of the relations

$$a_{\mathbf{k},\lambda} a_{\mathbf{k},\lambda}^\dagger |A\rangle = (1 + a_{\mathbf{k},\lambda}^\dagger a_{\mathbf{k},\lambda}) |A\rangle = (1 + n_{\mathbf{k},\lambda}) |A\rangle \quad (9.61)$$

and

$$\begin{aligned} b_i^\dagger b_i |A\rangle &= \nu_i |A\rangle = |A\rangle \\ b_f b_f^\dagger |A\rangle &= (1 - \nu_f) |A\rangle = |A\rangle \end{aligned} \quad (9.62)$$

(where ν_i and ν_f are occupation numbers, equal to 1 or 0 depending on whether the spin-orbital is occupied or unoccupied) we obtain

$$\langle B|B\rangle = \frac{1 + n_{\mathbf{k},\lambda}}{(\sqrt{1 + n_{\mathbf{k},\lambda}})^2} = 1 \quad (9.63)$$

so that $|B\rangle$ is properly normalized. Here $n_{\mathbf{k},\lambda}$ is the number of photons in the mode \mathbf{k},λ in the initial state $|A\rangle$, and since we are treating spontaneous photon emission, this number is zero. Letting

$$M_{r,s}(\mathbf{k}, \lambda) \equiv -e \sqrt{\frac{2\pi\hbar c}{V k}} \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} \chi_r^*(\mathbf{x}) \alpha_\lambda \chi_s(\mathbf{x}) \quad (9.64)$$

the transition matrix element of the interaction Hamiltonian then becomes:

$$\begin{aligned} \langle A|H_i|B\rangle &= \sum_{r,s} \sum_{\mathbf{k}',\lambda'} \langle A|b_r^\dagger b_s \{ a_{\mathbf{k}',\lambda'} M_{r,s}(\mathbf{k}', \lambda') + a_{\mathbf{k}',\lambda'}^\dagger M_{r,s}^\dagger(\mathbf{k}', \lambda') \} b_f^\dagger b_i a_{\mathbf{k},\lambda}^\dagger |A\rangle \\ &= \sum_{r,s} \langle A|b_r^\dagger b_s b_f^\dagger b_i |A\rangle M_{r,s}(\mathbf{k}, \lambda) \\ &= M_{i,f}(\mathbf{k}, \lambda) \end{aligned} \quad (9.65)$$

For a full derivation of the rate of spontaneous photon emission, see [Avery, 1976]. For cases where the transition from the excited state to the ground state is dipole-allowed, i.e. when

$$\int d^3x \chi_i^*(\mathbf{x}) \mathbf{x} \chi_f(\mathbf{x}) \neq 0 \quad (9.66)$$

one finds excited state lifetimes of the order of 10^{-8} seconds, the decay of the excited state being due to spontaneous photon emission.

9.6 Photon absorption

Suppose that the initial state $|A\rangle$ of the electron-photon system contains $n_{\mathbf{k},\lambda}$ photons in the mode \mathbf{k},λ , so that

$$a_{\mathbf{k},\lambda}^\dagger a_{\mathbf{k},\lambda} |A\rangle = n_{\mathbf{k},\lambda} |A\rangle \quad (9.67)$$

Now let us consider a transition from this state to the state

$$|B\rangle = \frac{1}{\sqrt{n_{\mathbf{k},\lambda}}} b_f^\dagger b_i a_{\mathbf{k},\lambda} |A\rangle \quad (9.68)$$

which differs from $|A\rangle$ by having one fewer photon in the mode $\mathbf{k}\lambda$, and by the transition $i \rightarrow f$. The factor $1/\sqrt{n_{\mathbf{k},\lambda}}$ is needed to make $|B\rangle$ properly normalized. The matrix element of the interaction Hamiltonian between the two states is given by

$$\begin{aligned} \langle A|H_i|B\rangle &= \frac{1}{\sqrt{n_{\mathbf{k},\lambda}}} \langle A|H_i|b_f^\dagger b_i a_{\mathbf{k},\lambda}|A\rangle \\ &= \frac{1}{\sqrt{n_{\mathbf{k},\lambda}}} \langle A|\sum_{r,s} b_r^\dagger b_s \sum_{\mathbf{k}',\lambda'} \left\{ a_{\mathbf{k}',\lambda'} M_{r,s}(\mathbf{k}',\lambda') + a_{\mathbf{k}',\lambda'}^\dagger M_{r,s}^\dagger(\mathbf{k}',\lambda') \right\} b_f^\dagger b_i a_{\mathbf{k},\lambda}|A\rangle \\ &= \frac{1}{\sqrt{n_{\mathbf{k},\lambda}}} \langle A|a_{\mathbf{k},\lambda}^\dagger a_{\mathbf{k},\lambda}|A\rangle M_{i,f}^\dagger(\mathbf{k},\lambda) \\ &= \sqrt{n_{\mathbf{k},\lambda}} M_{i,f}^\dagger(\mathbf{k},\lambda) \end{aligned} \quad (9.69)$$

where

$$M_{r,s}^\dagger(\mathbf{k},\lambda) \equiv -e\sqrt{\frac{2\pi\hbar c}{Vk}} \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \chi_r^*(\mathbf{x}) \alpha_\lambda \chi_s(\mathbf{x}) \quad (9.70)$$

Using Fermi's "Golden Rule Number Two" from perturbation theory, we obtain the probability per unit time for the transition $|A\rangle \rightarrow |B\rangle$:

$$\frac{1}{\tau} = \frac{2\pi}{\hbar} n_{\mathbf{k},\lambda} |M_{i,f}^\dagger(\mathbf{k},\lambda)|^2 \rho(B) \quad (9.71)$$

where $\rho(B)$ is the density of final states. The number of photons in the mode $\mathbf{k}\lambda$ is $n_{\mathbf{k},\lambda}/V$, where V is the volume of normalization. Therefore, since the photons move with velocity c , the flux of photons in this mode is $n_{\mathbf{k},\lambda}c/V$. The cross-section σ for the absorption of a photon is defined to be the probability per unit time for absorption, divided by the flux. Therefore, in the non-relativistic limit, we obtain the cross-section

$$\sigma = (2\pi)^2 \left(\frac{e^2}{\hbar c}\right) \left(\frac{\hbar}{m_e c}\right)^2 \frac{\hbar c}{k} \left| \int d\tau e^{-i\mathbf{k}\cdot\mathbf{x}} \chi_i^* \left(\mathbf{u} \cdot \frac{\partial}{\partial \mathbf{x}} \right) \chi_f \right|^2 \rho(B) \quad (9.72)$$

Here σ represents the cross-section for the absorption of a photon with wave number \mathbf{k} and polarization vector \mathbf{u} , accompanied by the 1-electron transition $i \rightarrow f$. Similarly, the cross-section for absorption of a left- or right-circularly polarized photon is given by

$$\sigma_\pm = (2\pi)^2 \left(\frac{e^2}{\hbar c}\right) \left(\frac{\hbar}{m_e c}\right)^2 \frac{\hbar c}{2k} \left| \int d\tau e^{-i\mathbf{k}\cdot\mathbf{x}} \chi_i^* \left((\mathbf{u}_1 \pm i\mathbf{u}_2) \cdot \frac{\partial}{\partial \mathbf{x}} \right) \chi_f \right|^2 \rho(B) \quad (9.73)$$

If χ_f corresponds to a bound state with a finite lifetime Γ , the density of final states can be represented by the function

$$\rho(B) = \frac{\hbar\Gamma}{\pi \{(E_A - E_B)^2 + (\hbar\Gamma)^2\}} \quad (9.74)$$

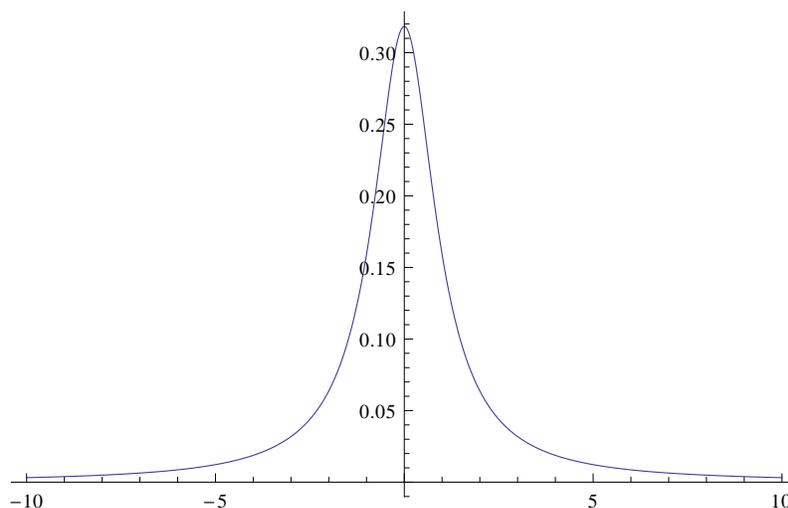


Figure 9.1: In this figure, the density of final states $\rho(B)$ is plotted as a function of $\hbar\Gamma$ on a scale where $|E_A - E_B| = 1$.

9.7 Problems with field theories

Quantum electrodynamics was pioneered very largely by the great British physicist P.A.M. Dirac, and yet Dirac was dissatisfied with the theory as it stood because of logical inconsistencies which had to be overcome by “renormalization”, a procedure that Dirac regarded as a “stop-gap measure”. There is a remarkable contrast between the enormous calculational success of quantum electrodynamics, and the logical inconsistencies that worried Dirac.

One serious problem with quantum field theories is the infinite self-energies that they predict. Another difficulty is connected with Haag’s theorem. In an article entitled “Haag’s Theorem as a Reason to Reconsider Direct-Action Theories” [Kastner, 2015a], R.E. Kastner writes: “Haag showed that the interacting field demands an inequivalent representation from that of the free field; the vacuum states of the two fields cannot be defined in the same representation. This result presents a serious problem for the basic mathematical consistency of quantum field theories, and has led to much discussion... Haag presented a more general and formal result in which the infinite degrees of freedom of the quantized field can be seen as the actual source of the problem.[203]”.

In their famous 1949 paper, J.A. Wheeler and R.P. Feynman say “...action at a distance must have a close connection with field theory. But never does it consider the action of a charge on itself. The theory of direct interparticle action is equivalent, not to the usual field theory, but to a modified or *adjunct field theory*..” In this adjunct field theory, which incorporates the concept of direct interparticle interaction, the logical difficulties of the usual theory are removed, for example the difficulties connected with giving the electromagnetic field an infinite number of degrees of freedom, but the calculational power of the usual theory is retained.



Figure 9.2: Paul Adrien Maurice Dirac (1902-1984). Although he was the principal founder of quantum electrodynamics, he was dissatisfied with the theory as it stood. Dirac once said that renormalization is just a stop gap procedure, and there had to occur a fundamental change in our ideas.



Figure 9.3: Rudolf Haag (1922-2016). He was the author of a theorem criticizing all field theories. In one form of Haag's theorem, the root of the problem is seen to be the fact that field theories lead to an infinite number of degrees of freedom. Haag was awarded the Max Planck medal in 1970 and the Henri Poincaré Prize of the International Association of Mathematical Physics in 1997.

Chapter 10

RESONANCE ENERGY TRANSFER

10.1 Introduction

The phenomenon of resonance energy transfer or “sensitized fluorescence” was discovered in 1922 by G. Cario and J. Franck [Cario and Franck, 1923]. They exposed a mixture of mercury and thallium vapors to a frequency of light which could only be absorbed by the mercury, and observed in the fluorescence a frequency which only the thallium could emit. An initial theoretical explanation of this phenomenon on the basis of classical coupled oscillator theory was put forward by J. Perrin in 1926. Later F. Perrin, Kallman and London, developed quantum mechanical theories to explain sensitized fluorescence. These theories were improved by Th. Förster [Scholes, 2003]. Much experimental work has also been done in this field, and interest is high today because resonance energy transfer has proved to be a very useful tool in nonoscillation [Chou and Dennis, 2015], [Wegner, 2015], [Kusku and Akan, 2011] and biophysics [Avery, Bay and Szent-Györgyi, 1961] [Pearlstein, 2002].

In the present chapter, we extend the early theories by considering relativistic effects. We replace the Coulomb interaction by a relativistic interaction, as was done in early papers by one of us [Avery, 1966], [Avery, 1984]. Similar results were obtained by Gomberoff and Power [Gomberoff and Power, 1966]. More recently, very important theoretical work applying relativistic theory to resonance energy transfer was carried out by D.L. Andrews and his co-workers [Andrews and Demidov, 1999] [Juzeliunas and Andrews, 2000] [Anderson, 2008] [Daniels et al., 2003] [Bradshaw and Andrews, 2008].

The theory presented here is a quantum mechanical direct-interaction theory. In other words it does not involve photons, and thus it avoids problems such as the infinities associated self-interaction energies, and the consequences of Haag’s theorem that appear when photons are regarded as an infinite assembly of harmonic oscillators. In 1949, J.A. Wheeler and R.P. Feynman published a famous paper [Wheeler and Feynman, 1949], in which they showed that the classical behavior of a system of charged particles can be correctly explained by a direct interparticle interaction principle. They were motivated by the desire

to avoid worrying infinities and other problems in the usual formulation of quantum electrodynamics, [Kastner, 2015a], [Kastner, 2015b], [Grundler, pdf], [Lazarovici, preprint, pdf], [Klaczynski, 2015]. The Wheeler-Feynman action principle is time-symmetric. In other words, it makes use of both retarded and advanced potentials in a symmetrical way.

10.2 Review of the Perrin-Förster theory

In the Perrin-Förster theory of resonance energy transfer (also called sensitized fluorescence), one considers two atoms or molecules, located respectively at the points \mathbf{a}_1 and \mathbf{a}_2 so that they are separated by the distance $R \equiv \mathbf{a}_1 - \mathbf{a}_2$. This distance is assumed to be large enough so that overlap between the electronic wave functions of the two systems can be neglected. If it were not for the perturbation due to the interaction of the electrons of the first system with those of the second, the product of the electronic wave functions of the two systems, $\psi_r(1)\psi_s(2)$, would be a stationary state, i.e. an eigenfunction of the total Hamiltonian. However, there is a perturbation term due to the Coulomb interaction of the electrons on one system with those on the other, and this perturbation can induce a transition to another state, which we can denote by $\psi_t(1)\psi_u(2)$. The perturbation matrix element is given by:

$$H'_{rs|tu} = \int d\tau_1 \int d\tau_2 \psi_r^\dagger(1)\psi_s^\dagger(2) \sum_{i,j} \frac{e^2}{|\mathbf{x}_i - \mathbf{x}_j|} \psi_t(1)\psi_u(2) \quad (10.1)$$

Here \mathbf{x}_i indicates the position of electron i on system 1, while \mathbf{x}_j is the position of electron j in system 2. Expanding $1/|\mathbf{x}_i - \mathbf{x}_j|$ in a six-fold Taylor series about the centers of the two systems, \mathbf{a}_1 and \mathbf{a}_2 , one obtains for the first non-zero term the interaction between the transition dipole moments of the two systems.

$$H'_{rs|tu} \approx e^2 \left[\frac{\mathbf{d}_1^{r,t} \cdot \mathbf{d}_2^{s,u}}{R^3} - \frac{3(\mathbf{d}_1^{r,t} \cdot \mathbf{R})(\mathbf{d}_2^{s,u} \cdot \mathbf{R})}{R^5} \right] \quad (10.2)$$

where

$$\begin{aligned} \mathbf{d}_1^{r,t} &\equiv \int d\tau_1 \psi_r^\dagger(1)\mathbf{r}_1\psi_t(1) \\ \mathbf{d}_2^{s,u} &\equiv \int d\tau_2 \psi_s^\dagger(2)\mathbf{r}_2\psi_u(2) \end{aligned} \quad (10.3)$$

and

$$\begin{aligned} \mathbf{r}_1 &\equiv \mathbf{x}_i - \mathbf{a}_1 \\ \mathbf{r}_2 &\equiv \mathbf{x}_j - \mathbf{a}_2 \end{aligned} \quad (10.4)$$

From equation (10.2), one can see that the perturbation inducing the transition is proportional to $1/R^3$. When first-order perturbation theory is used to calculate the transition

probability, the perturbation matrix element is squared, and it is thus proportional to $1/R^6$. In Förster's version of the theory, a distance R_0 is defined to be the distance at which the probability for resonance transfer of electronic excitation energy between two systems is the same as the probability, $1/\tau$, of spontaneous photon emission by the initially excited system. Thus, Förster wrote:

$$\frac{1}{\tau_{s \rightarrow a}} = \frac{1}{\tau} \left(\frac{R_0}{R} \right)^6 \quad (10.5)$$

Typical values of R_0 are of the order of magnitude 50 Ångströms.

10.3 A relativistic interaction

It is interesting to try to extend the Perrin-Förster theory by replacing the Coulomb interaction by its relativistic counterpart.[Avery, 1966], [Avery, 1984], [Andrews and Demidov, 1999]. We start by recalling the definition of the contravariant anticommuting Dirac matrices:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (10.6)$$

$$\gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \quad (10.7)$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad (10.8)$$

$$\gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0-1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad (10.9)$$

We next introduce the second-order electron-electron S-matrix[204] defined by

$$S_{rs|tu} \equiv \sum_{\mu=0}^3 e^2 \int d^4x_1 \int d^4x_2 \bar{\psi}_r(1) \gamma^\mu(1) \psi_s(1) G(1,2) \bar{\psi}_t(2) \gamma^\mu(2) \psi_u(2) \quad (10.10)$$

where $d^4x = d^3x dt$ and $\bar{\psi} \equiv \psi^\dagger \gamma^0$, and where $G(1,2)$ is the Green's function for the d'Alembertian operator:

$$G(1,2) \equiv \frac{1}{4\pi^3} \int d^4k \frac{e^{i\mathbf{k} \cdot (\mathbf{x}_1 - \mathbf{x}_2) - i\omega(t_1 - t_2)}}{\mathbf{k} \cdot \mathbf{k} - (\omega/c)^2 - i\epsilon} \quad (10.11)$$

and $d^4k = d^3k d\omega$. We now let

$$\begin{aligned} ec\bar{\psi}_r(1)\gamma^\mu(1)\psi_t(1) &\equiv j_\mu^{r,t}(\mathbf{x}_1)e^{i(\omega_r-\omega_t)t_1} \\ ec\bar{\psi}_s(2)\gamma^\mu(2)\psi_u(2) &\equiv j_\mu^{s,u}(\mathbf{x}_2)e^{i(\omega_s-\omega_u)t_2} \end{aligned} \quad (10.12)$$

Then [204]

$$\begin{aligned} S_{rs|tu} &= \sum_{\mu=0}^3 \int d^3x_1 dt_1 \int d^3x_2 dt_2 j_\mu^{r,t}(\mathbf{x}_1) e^{i(\omega_r-\omega_t)t_1} G(1,2) j_\mu^{s,u}(\mathbf{x}_2) e^{i(\omega_s-\omega_u)t_2} \\ &= -2\pi\delta(\omega_r - \omega_t + \omega_s - \omega_u) U_{rs|tu} \end{aligned} \quad (10.13)$$

where [204]

$$\begin{aligned} U_{rs|tu} &= -\frac{1}{c^2} \sum_{\mu=0}^3 \int d^3x_1 \int d^3x_2 j_\mu^{rt}(\mathbf{x}_1) \frac{e^{ik|\mathbf{x}_1-\mathbf{x}_2|}}{|\mathbf{x}_1-\mathbf{x}_2|} j_\mu^{su}(\mathbf{x}_2) \\ &\equiv -\frac{1}{c^2} \sum_{\mu=0}^3 \int d^3x_1 \int d^3x_2 j_\mu^{rt}(\mathbf{x}_1) \mathcal{G}(1,2) j_\mu^{su}(\mathbf{x}_2) \end{aligned} \quad (10.14)$$

and

$$k \equiv \frac{\omega_{rt}}{c} \quad (10.15)$$

10.4 The Green's function of the Helmholtz equation

The function

$$\mathcal{G}(1,2) \equiv \frac{1}{2} \left[\frac{e^{ik|\mathbf{x}_1-\mathbf{x}_2|}}{|\mathbf{x}_1-\mathbf{x}_2|} + \frac{e^{-ik|\mathbf{x}_1-\mathbf{x}_2|}}{|\mathbf{x}_1-\mathbf{x}_2|} \right] \quad (10.16)$$

is the Green's function of the Helmholtz equation and it obeys the relationships

$$[\nabla_1^2 + k^2]\mathcal{G}(1,2) = [\nabla_2^2 + k^2]\mathcal{G}(1,2) = 4\pi\delta(\mathbf{x}_1 - \mathbf{x}_2) \quad (10.17)$$

The significance of this function is as follows: Suppose that the current 4-vector and the electromagnetic potential 4-vector are simple harmonic functions of time:

$$\begin{aligned} j_\mu(\mathbf{x}, ict) &= j_\mu(\mathbf{x}) e^{ikct} & \mu = 1, 2, 3, 4 \\ A_\mu(\mathbf{x}, ict) &= A_\mu(\mathbf{x}) e^{ikct} \end{aligned} \quad (10.18)$$

In that case, the wave equation relating the potentials to the source current becomes

$$[\nabla^2 + k^2] A_\mu(\mathbf{x}, ict) = \frac{1}{4\pi} j_\mu(\mathbf{x}, ict) \quad \mu = 1, 2, 3, 4 \quad (10.19)$$

Making use of the 3-dimensional symmetric form of Green's theorem and equation (10.17), we have

$$A_\mu(1) = \int d^3x_2 \mathcal{G}(1,2) j_\mu(2) \quad (10.20)$$

10.5 Matrix elements

We wish to evaluate matrix elements of the form

$$\int d^3x_1 \bar{\psi}_r(1) \gamma^\mu(1) \psi_t(1) = \int d^3x_1 \psi_r^\dagger(1) \gamma^0(1) \gamma^\mu(1) \psi_t(1) \quad (10.21)$$

where

$$\begin{aligned} \gamma^0 \gamma^0 &= I \\ \gamma^0 \gamma^1 &= \alpha_1 \\ \gamma^0 \gamma^2 &= \alpha_2 \\ \gamma^0 \gamma^3 &= \alpha_3 \end{aligned} \quad (10.22)$$

where

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (10.23)$$

is the identity matrix, and where the components of the 3-vector $\boldsymbol{\alpha}$ are given by

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (10.24)$$

$$\alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad (10.25)$$

$$\alpha_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad (10.26)$$

The matrix elements of the Dirac operator $\boldsymbol{\alpha}$ can be related to the transition dipole moments by the commutation relation:

$$[\mathbf{H}_j, \mathbf{d}_j] = \frac{\hbar c}{i} \boldsymbol{\alpha}(j) \quad j = 1, 2 \quad (10.27)$$

where \mathbf{H}_j is the Dirac Hamiltonian operator of the j th electron.

$$\mathbf{H}_j = \boldsymbol{\alpha}(j) \cdot \left(\frac{\hbar c}{i} \frac{\partial}{\partial \mathbf{x}_j} - e \mathbf{A}(\mathbf{x}_j) \right) - e \phi(\mathbf{x}_j) + \beta_j m c^2 \quad (10.28)$$

Then

$$\begin{aligned}
\int d^3x_1 \psi_r^\dagger(1) \boldsymbol{\alpha}(1) \psi_t(1) &= \int d\tau_1 \psi_r^\dagger(1) (\mathbf{H}_1 \mathbf{r}_1 - \mathbf{r}_1 \mathbf{H}_1) \psi_t(1) \\
&= \frac{i(E_r - E_t)}{\hbar c} \int d\tau_1 \psi_r^\dagger(1) \mathbf{r}_1 \psi_t(1) \\
&= i(\omega_r - \omega_t) \mathbf{d}_1^{r,t}
\end{aligned} \tag{10.29}$$

with a similar expression for $j = 2$. Finally, with the help of equation (E.43), we can obtain the relations:

$$\begin{aligned}
\int d\tau_1 \psi_r^\dagger(1) \boldsymbol{\alpha}(1) \psi_t(1) &= ik \mathbf{d}_1^{r,t} \\
\int d\tau_2 \psi_s^\dagger(2) \boldsymbol{\alpha}(2) \psi_u(2) &= -ik \mathbf{d}_2^{s,u}
\end{aligned} \tag{10.30}$$

where

$$k \equiv \frac{E_r - E_t}{\hbar c} \approx \frac{E_s - E_u}{\hbar c} \tag{10.31}$$

The sign difference between the first and second parts of equation (10.30) comes from the fact that one system is losing excitation energy, while the other is gaining it.

10.6 Transition probability at macroscopic separations

When the sensitizer and the acceptor are separated by a macroscopic distance, the non-relativistic probability of resonance energy transfer between them becomes vanishingly small. However, in a relativistic calculation, the terms involving $\boldsymbol{\alpha}$ lead to a transition probability which falls off only as $1/R^2$. At macroscopic sensitizer-acceptor separations, we can make the approximation

$$\mathbf{x}_1 - \mathbf{x}_2 \approx \mathbf{a}_1 - \mathbf{a}_2 \equiv \mathbf{R} \tag{10.32}$$

so that the Green's function of the Helmholtz equation becomes

$$\mathcal{G}(1, 2) \approx \frac{1}{2} \frac{(e^{ikR} + e^{-ikR})}{R} \equiv \frac{\cos(kR)}{R} \tag{10.33}$$

Then equation (10.14) becomes

$$U_{rs|tu} = -\frac{1}{c^2} \frac{\cos(kR)}{R} \sum_{\mu=0}^3 \left(\int d^3x_1 j_\mu^{rt}(\mathbf{x}_1) \right) \left(\int d^3x_2 j_\mu^{su}(\mathbf{x}_2) \right) \tag{10.34}$$

From equations (10.12) and (10.22), we have

$$\begin{aligned}
\int d^3x_1 j_0^{rt}(\mathbf{x}_1) &= ec \int d^3x_1 \psi_r^\dagger(\mathbf{x}_1) I(1) \psi_t(\mathbf{x}_1) = 0 \\
\int d^3x_1 j_1^{rt}(\mathbf{x}_1) &= ec \int d^3x_1 \psi_r^\dagger(\mathbf{x}_1) \alpha_1(1) \psi_t(\mathbf{x}_1) \\
\int d^3x_1 j_2^{rt}(\mathbf{x}_1) &= ec \int d^3x_1 \psi_r^\dagger(\mathbf{x}_1) \alpha_2(1) \psi_t(\mathbf{x}_1) \\
\int d^3x_1 j_3^{rt}(\mathbf{x}_1) &= ec \int d^3x_1 \psi_r^\dagger(\mathbf{x}_1) \alpha_3(1) \psi_t(\mathbf{x}_1)
\end{aligned} \tag{10.35}$$

With the help of equation (10.30) we then have

$$\begin{aligned}
\int d^3x_1 j_\mu^{r,t}(\mathbf{x}_1) &= ec \int d^3x_1 \psi_r^\dagger(\mathbf{x}_1) \alpha_\mu(1) \psi_t(\mathbf{x}_1) \\
&= ikec d_{1,\mu}^{r,t} \quad \mu = 1, 2, 3
\end{aligned} \tag{10.36}$$

The expressions for $j_\mu^{s,u}(\mathbf{x}_2)$ are similar. Thus, at macroscopic sensitizer-acceptor separations, we have

$$\begin{aligned}
U_{rs|tu} &= -\frac{1}{c^2} \frac{\cos(kR)}{R} \sum_{\mu=0}^3 \left(\int d^3x_1 j_\mu^{rt}(\mathbf{x}_1) \right) \left(\int d^3x_2 j_\mu^{su}(\mathbf{x}_2) \right) \\
&\approx -e^2 k^2 \mathbf{d}_1^{r,t} \cdot \mathbf{d}_2^{s,u} \frac{\cos(kR)}{R}
\end{aligned} \tag{10.37}$$

Then

$$|U_{rs|tu}|^2 \approx \frac{(e^2 k^2 \mathbf{d}_1^{r,t} \cdot \mathbf{d}_2^{s,u} \cos(kR))^2}{R^2} \tag{10.38}$$

falls off only as $1/R^2$. Now suppose that the sensitizer is at the center of a large sphere of radius R , on the surface of which acceptors are uniformly distributed. Then the probability of resonance energy transfer from the electronically excited sensitizer to one or another of the acceptors will be independent of the radius of the sphere. This shows that the sensitizer can lose its excitation energy to acceptors over macroscopic distances. A much more detailed calculation leading to the prediction of resonance energy transfer over macroscopic distances has been carried through by G. Juzeliunas and D.L. Andrews [Andrews and Demidov, 1999].

10.7 Comparison with spontaneous photon emission

Is long-range resonance energy transfer a process that competes with spontaneous photon emission? Or is it another way of describing the same process? In the usual theory of spontaneous photon emission, the presence of acceptors is not needed. The spontaneous

photon emission would take place even if the remainder of space were entirely empty. At the end of this article, we propose an experimental test. The presence of acceptors may also be involved in the recently observed violation of Planck's law ([Thompson et. al., 2018]). We are working on a longer paper which will explore these questions.

10.8 The Perrin-Förster region

We have, until now, been discussing the relativistic calculation of resonance energy transfer probabilities at very large sensitizer-acceptor separations, where $\mathbf{x}_i - \mathbf{x}_j \approx \mathbf{R}$ and $kR \gg 1$. Let us now turn to the region of Förster's R_0 , where we have $kR \ll 1$ and $k|\mathbf{x}_i - \mathbf{x}_j| \ll 1$. We must ask whether the relativistic calculation in the short-range limit gives us the Perrin-Förster's result, which has been confirmed by so many experiments. In this region,

$$\begin{aligned} & \sum_{i,j} \frac{e^2(1 - \boldsymbol{\alpha}(i) \cdot \boldsymbol{\alpha}(j))e^{ik|\mathbf{x}_i - \mathbf{x}_j|}}{|\mathbf{x}_i - \mathbf{x}_j|} \rightarrow \sum_{i,j} \frac{e^2(1 - \boldsymbol{\alpha}(i) \cdot \boldsymbol{\alpha}(j))}{|\mathbf{x}_i - \mathbf{x}_j|} \\ = & \sum_{i,j} \frac{e^2}{|\mathbf{x}_i - \mathbf{x}_j|} - \sum_{i,j} \frac{e^2 \boldsymbol{\alpha}(i) \cdot \boldsymbol{\alpha}(j)}{|\mathbf{x}_i - \mathbf{x}_j|} \end{aligned} \quad (10.39)$$

The first term in the last line of equation (E.46) is that which gives rise to the Perrin-Förster results. Our task now is to find matrix elements of the perturbation Hamiltonian due to the second term, and to show that when $kR \ll 1$, the contribution of this term is small compared to the contribution of the first term. Let

$$\mathbf{H}_{1,2}'' \equiv -e^2 \int d\tau_1 \int d\tau_2 \psi_r^\dagger(1) \psi_s^\dagger(2) \frac{(\boldsymbol{\alpha}(1) \cdot \boldsymbol{\alpha}(2))}{|\mathbf{x}_i - \mathbf{x}_j|} \psi_t(1) \psi_u(2) \quad (10.40)$$

Making use of equation (E.41), we obtain for the leading non-zero term in the Taylor series expansion of $\mathbf{H}_{1,2}''$ in terms of $1/|\mathbf{x}_i - \mathbf{x}_j|$:

$$\begin{aligned} \mathbf{H}_{1,2}'' & \approx -\frac{e^2(ik\mathbf{d}_1^{r,t}) \cdot (-ik\mathbf{d}_2^{s,u})}{R} + \dots \\ & = -e^2 k^2 \frac{(\mathbf{d}_1^{r,t}) \cdot (\mathbf{d}_2^{s,u})}{R} + \dots \end{aligned} \quad (10.41)$$

The ratio between this term and the leading Perrin-Förster term in equation (10.2) is

$$\begin{aligned} \frac{\mathbf{H}_{1,2}''}{\mathbf{H}_{1,2}'} & \approx \left(-e^2 k^2 \frac{\mathbf{d}_1^{r,t} \cdot \mathbf{d}_2^{s,u}}{R} \right) / \left(e^2 \frac{\mathbf{d}_1^{r,t} \cdot \mathbf{d}_2^{s,u}}{R^3} \right) \\ & = \left(-k^2 \frac{R^3}{R} \right) = -k^2 R^2 \end{aligned} \quad (10.42)$$

so that when $kR \ll 1$, the term $\mathbf{H}_{1,2}''$ can be neglected in comparison with the Perrin-Förster term. Thus, in this region, the relativistic treatment closely approximates the well-verified non-relativistic Perrin-Förster result shown in equations (10.2) and (10.5).

10.9 A proposed experiment

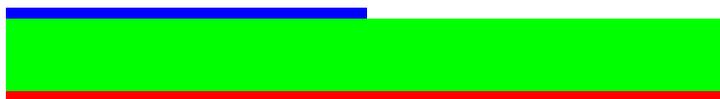


Figure 10.1: We propose the experiment illustrated in this figure for detecting resonance energy transfer over macroscopic distances. The central rectangle represents a glass slide, whose thickness (for example 0.2 cm.) is a macroscopic distance. On one side, the slide is coated with sensitizer molecules, while on the other side it is partially coated with acceptor molecules. If we illuminate the slide from the sensitizer side with a frequency of light which only the sensitizer molecules can absorb, and observe the fluorescence of the sensitizer from the same side, we can answer the question of whether resonance energy transfer can take place over macroscopic distances. If the transfer takes place, the presence of the acceptors on the opposite side of the slide will reduce the observed fluorescence of the sensitizer.

In the proposed experiment, sensitizer and acceptor molecules are chosen in such a way that the emission maximum of the sensitizers corresponds to the absorption maximum of the acceptors. Let ω_1 be the frequency of light corresponding to the absorption maximum of the sensitizer molecules, while $\omega_2 \approx \omega_3$ is the frequency corresponding both to the emission maximum of the sensitizers and the absorption maximum of the acceptors. The slide is illuminated from the sensitizer side with light of frequency ω_1 , and radiation at the frequency ω_2 is observed from the sensitizer side, both on portions of the slide which are coated with acceptors and portions where the acceptors are not present. If resonance energy transfer over macroscopic distances is taking place, the fluorescence of the sensitizer will be damped by the presence of acceptors on the opposite side of the slide.

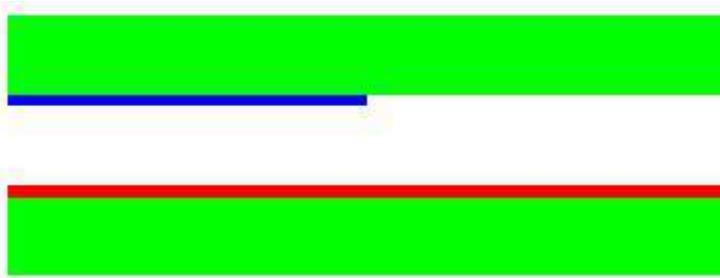


Figure 10.2: This figure shows an alternative arrangement, where the 2mm. gap is empty. Above the gap is a thin glass slide partially coated with acceptor molecules. Below the gap is another thin glass slide, completely coated with sensitizer molecules.

Experiments on long-distance resonance energy transfer, for example [205] [207] [208], [209], [210], [211], have detected departures from the $1/R^6$ distance dependence predicted Perrin-Förster theory, but until now only deviations in the nanometer range have been detected. We propose the following set of measurements, which aim at the detection of resonance energy transfer over a macroscopic distance:

A glass slide roughly 2 mm. in thickness (a macroscopic distance) is uniformly coated on one side with sensitizer molecules, as shown in Figure 1, while the opposite side of the slide is partially coated with acceptor molecules. The sensitizer and acceptor molecules are chosen in such a way that their absorption and emission bands lie in the visual range of the spectrum, a range in which the glass slide is transparent. Furthermore, the sensitizer and acceptor molecules should be chosen in such a way the the Stokes-shifted emission band of the sensitizers overlaps maximally with the absorption band of the acceptors.

The slide is illuminated from the sensitizer side with a frequency of light, ω_1 , which only the sensitizers can absorb, and, from the same side, the emission is observed at a frequency, ω_2 , at which the sensitizer molecules fluoresce. If this fluorescence is damped on the portion of the slide coated with acceptors, resonance energy transfer over macroscopic distances will have been demonstrated. The trivial process, where a photon emitted by the sensitizer, is absorbed by the acceptor, Stokes-shifted, and then fluoresced, could explain an enhancement of intensity in the frequency region near the acceptor fluorescence maximum, but not the loss of intensity in the frequency region near the fluorescence maximum of of the sensitizer. No explanation besides resonance energy transfer can be devised to explain such a loss.

The arrangement shown in Figure 2

The measurements should be repeated with the experimental arrangement shown in Figure 2. In this case, it is possible to perform the following initial experiments for the purpose of calibration:

1. The following calibration experiments are performed on the slide with the acceptor molecules by itself: The slide is illuminated from the acceptor side with light of constant intensity I_1 , with frequencies varying over the entire relevant range. The intensity of emitted light, I_2 , is measured at various points along the length of the slide from the coated side, and the ratio I_1/I_2 is recorded, again with frequencies varying over the entire range.
2. Another calibration experiment can be performed on the slide coated with sensitizer molecules by itself: The slide is illuminated with a constant intensity, I_0 , from the glass side, with frequencies running over the entire range. The corresponding fluorescence intensities, I_2 , are measured from the sensitizer side, and the ratio I_0/I_2 is recorded, again with frequencies varying over the entire range. Finally, the fluorescence intensity, I_3 , is measured from the glass side, and the ratio, I_0/I_3 . is recorded at all frequency combinations.

Having performed these initial calibration experiments, we are ready for the sensitized fluorescence experiment: The two slides are placed together with a 2 mm. gap, as shown in Figure 2. We illuminate the slide with constant intensity, I_0 , from the sensitizer side, with frequencies running over the entire range, and we record the fluorescence intensity, I_3 , from the sensitizer side, at various points on the slide, at frequencies combinations running over the entire range. If we see a loss of intensity of I_3 at the frequency range corresponding to the fluorescence maximum of the sensitizer, localized at the points of the side which are nearest to the acceptor molecules, this can be interpreted as evidence of resonance energy transfer to the acceptors.

If desired, the experiment can be simplified. The initial calibration experiments can be omitted if the absorption maximum of the sensitizers, ω_1 , is known. The two slides, arranged as shown in Figure 2 can be illuminated from the sensitizer side with intensity, I_0 , and frequency ω_1 . The fluorescence intensity, I_3 , can then be measured at all frequencies, at various points, from the sensitizer side. Loss of intensity in the frequency range corresponding to the fluorescence maximum of the sensitizer, at points on the slide where the sensitizers and acceptors are close, can still be interpreted as evidence of resonance energy transfer. To explain this loss of sensitizer fluorescence intensity, one has to imagine that before the sensitizer could fluoresce, it transferred its excitation energy to an acceptor. As mentioned above, the trivial process might be invoked to explain an enhancement in intensity in the frequency region near to the fluorescence maximum of the absorber, but it cannot explain a loss of intensity in the frequency region near the fluorescence maximum of the sensitizers in those portions of the slide where the sensitizers and acceptors are close together. If observed, such a loss could only be explained by resonance energy transfer.

10.10 Discussion: Direct interparticle interaction

Despite the great successes of quantum field theories, they contain some logical inconsistencies. P.A.M. Dirac himself, the principal founder of quantum electrodynamics, was dissatisfied with the theory as it stood because of these problems. Dirac once said that renormalization is just a stop gap procedure, and there had to occur a fundamental change in our ideas.

One serious problem with quantum field theories is the infinite self-energies that they predict. Another difficulty is connected with Haag's theorem. In an article entitled "Haag's Theorem as a Reason to Reconsider Direct-Action Theories" [Kastner, 2015a], R.E. Kastner writes: "Haag showed that the interacting field demands an inequivalent representation from that of the free field; the vacuum states of the two fields cannot be defined in the same representation. This result presents a serious problem for the basic mathematical consistency of quantum field theories, and has led to much discussion... Haag presented a more general and formal result in which the infinite degrees of freedom of the quantized field can be seen as the actual source of the problem.[203]".

In their famous 1949 paper, J.A. Wheeler and R.P. Feynman say "...action at a distance must have a close connection with field theory. But never does it consider the action

of a charge on itself. The theory of direct interparticle action is equivalent, not to the usual field theory, but to a modified or *adjunct field theory*.” In this adjunct field theory, which incorporates the concept of direct interparticle interaction, the logical difficulties of the usual theory are removed, for example the difficulties connected with giving the electromagnetic field an infinite number of degrees of freedom, but the calculational power of the usual theory is retained.

10.11 Feynman’s first seminar at Princeton

After graduating from MIT with a B.Sc. in physics (having published two papers while still an undergraduate), Richard Feynman entered Princeton University as a graduate student. Here is Wikipedia’s account of the first seminar that he gave there:

“Attendees at Feynman’s first seminar, which was on the classical version of the Wheeler-Feynman absorber theory, included Albert Einstein, Wolfgang Pauli, and John von Neumann. Pauli made the prescient comment that the theory would be extremely difficult to quantize, and Einstein said that one might try to apply this method to gravity in general relativity, which Sir Fred Hoyle and Jayant Narlikar did much later as the Hoyle-Narlikar theory of gravity. Feynman received a Ph.D. from Princeton in 1942; his thesis advisor was John Archibald Wheeler. His doctoral thesis was titled “The Principle of Least Action in Quantum Mechanics”. Feynman had applied the principle of stationary action to problems of quantum mechanics, inspired by a desire to quantize the Wheeler-Feynman absorber theory of electrodynamics, and laid the groundwork for the path integral formulation and Feynman diagrams. A key insight was that positrons behaved like electrons moving backwards in time.”

James Gleick wrote: “This was Richard Feynman nearing the crest of his powers. At twenty-three ... there may now have been no physicist on earth who could match his exuberant command over the native materials of theoretical science. It was not just a facility at mathematics (though it had become clear ... that the mathematical machinery emerging in the Wheeler-Feynman collaboration was beyond Wheeler’s own ability). Feynman seemed to possess a frightening ease with the substance behind the equations, like Einstein at the same age, like the Soviet physicist Lev Landau - but few others.”



Figure 10.3: **Richard Feynman (1918-1988)**. Feynman shared the 1965 Nobel Prize in Physics for his contributions to the development of quantum electrodynamics. He was an inspiring and popular lecturer, much loved by his students. Many of his lectures are available on the Internet.

10.12 The absorber theory of radiation

Wheeler and Feynman's 1949 paper

Here are some quotations from this famous paper:

“We... propose here to go back to the great basic problem of classical physics the motion of a system of charged particles under the forces and to inquire influence of electromagnetic what description of the interactions and motions is possible’which is at the same time (1) well defined (2) economical in postulates and (3) in agreement with experience.

“We conclude that these requirements are satisfied by the theory of action at a distance of Schwarzschild, Tetrode, and Fokker. In this description of nature no direct use is made of the notion of field. Each particle moves in compliance with the principle of stationary action:

$$J = - \sum_a m_a c \int (-da_\mu da^\mu)^{1/2} + \sum_{a < b} \left(\frac{e_a e_b}{c} \right) \int \int \delta(ab_\mu ab^\mu) (da_\nu db^\nu) = \text{extremum} \quad (10.43)$$

“However unfamiliar this direct interparticle treatment compared to the electrodynamics of Maxwell and Lorentz, it deals with the same problems, talks about the same charges, considers the interaction of the same current elements, obtains the same capacities, predicts the same inductances and yields the same physical conclusions. Consequently action at a distance must have a close connection with field theory. But never does it consider the action of a charge on itself. The theory of direct interparticle action is equivalent, not to the usual field theory, but to a modified or adjunct field theory...”

10.13 Karl Schwarzschild's contributions to physics

Here are some quotations from the Wikipedia article about Karl Schwarzschild:

“Schwarzschild provided the first exact solution to the Einstein field equations of general relativity, for the limited case of a single spherical non-rotating mass, which he accomplished in 1915, the same year that Einstein first introduced general relativity. The Schwarzschild solution, which makes use of Schwarzschild coordinates and the Schwarzschild metric, leads to a derivation of the Schwarzschild radius, which is the size of the event horizon of a non-rotating black hole...”

“Thousands of dissertations, articles, and books have since been devoted to the study of Schwarzschild's solutions to the Einstein field equations. However, although Schwarzschild's best known work lies in the area of general

relativity, his research interests were extremely broad, including work in celestial mechanics, observational stellar photometry, quantum mechanics, instrumental astronomy, stellar structure, stellar statistics, Halley's comet, and spectroscopy...

“He also introduced a field free variational formulation of electrodynamics (also known as “action at distance” or “direct interparticle action”) based only on the world line of particles as

$$S = \sum_i m_i \int_{C_i} ds_i + \frac{1}{2} \int \int_{C_i C_j} q_i q_j \delta(\|P_i P_j\|) ds_i ds_j \quad (10.44)$$

where C_α are the world lines of the particle, ds_a the (vectorial) arc element along the world line. Two points on two world lines contribute to the Lagrangian (are coupled) only if they are a zero Minkowskian distance (connected by a light ray), hence the term $\delta(\|P_i P_j\|)$. The idea was further developed by Tetrode and Fokker in the 1920s and Wheeler and Feynman in the 1940s and constitutes an alternative/equivalent formulation of electrodynamics.



Figure 10.4: **Karl Schwarzschild (1873-1916)**. He showed his exceptional abilities early by publishing two papers on binary orbits in astronomy at the age of 16.

Bibliography

- [Cario and Franck, 1923] G. Cario and J. Frank, *Z. Phys.* **17**, 202, (1923)
- [Avery, Bay and Szent-Györgyi, 1961] J.S. Avery, Z. Bay and A. Szent-Györgyi, *On the Energy Transfer in Biological Systems*, *Proc. Nat. Acad. Sci. (US)*, **47**, 1742-1744, (1961)
- [Avery, 1966] J.S. Avery, *Resonance energy transfer and spontaneous photon emission*, *Proc. Phys. Soc. (London)* **88**, Part 1, (1966)
- [Avery, 1984] J.S. Avery, *Use of the S-Matrix in the Relativistic Treatment of Resonance Energy Reansfer*. *Int. J. Quantum Chemistry*, **XXV**, 79-96, (1984)
- [Gomberoff and Power, 1966] L. Gomberoff and E.A. Power, *The resonance transfer of excitation*, *Proc. Pys. Soc.* **88**, 281. (1966)
- [Scholes, 2003] G.D. Scholes, *Long-Range Resonance Energy Transfer in Molecular Systems*, *Annu. Rev. Phys. Chem.* **54**, 57-87, (2003)
- [Chou and Dennis, 2015] K.F. Chou and A.M. Dennis, *Förster Resonance Energy Transfer Between Quantum Dot Donors and Quantum Dot Acceptors*, *Sensors*, **2015**, 15, 13288-13325, (2015)
- [Wegner, 2015] K-D. Wegner, *Förster Resonance Energy Transfer From Terbium Complexes to Quantum Dots for Multiplexed Immunoassays and Molecular Rulers*, Doctoral Thesis, Université Paris-Sud, (2015)
- [Kusku and Akan, 2011] M. Kusku and O.B. Akan, *A Nanoscale Communication Channel with Fluorescence Resonance Energy Transfer (FRET)*, 1st IEEE International Workshop on Molecular and Nano Scale Communication (MoNaCom), Published in: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), (2011)
- [Pearlstein, 2002] R.M. Pearlstein, *Photosynthetic Exciton Transfer in the 1960's*, *Photosynthesis Research* **73**, 119-126, (2002)
- [Andrews and Demidov, 1999] D.L. Andrews and A.A. Demidov, *Resonance Energy Transfer*, (Wiley, 1999)

- [Juzeliunas and Andrews, 2000] G. Juzeliunas and D.L. Andrews, *Quantum Electrodynamics of Resonance Energy Transfer*, Adv. Chem. Physics, **112**, (2000)
- [Anderson, 2008] D.L. Anderson, Editor, *Encyclopedia of Applied Spectroscopy*, Wiley, (2008)
- [Daniels et al., 2003] G.J. Daniels, R.G. Jenkins, D.S. Bradshaw, and D.L. Andrews, *Resonance Energy Transfer: The Unified Theory Revisited*, J. Chem. Phys. **119**, 2264-2274, (2003)
- [Bradshaw and Andrews, 2008] D.S. Bradshaw and D.L. Andrews, *Optically controlled resonance energy transfer: Mechanisms and configuration for all-optical switching*, J. Chem. Phys. **128**, 144506, (2008)
- [Wheeler and Feynman, 1949] J.A. Wheeler and R.P. Feynman, *Classical Electrodynamics in Terms of Direct Interparticle Interaction*, Rev. Mod. Phys. **21**, 3, 425-433, (1949)
- [Kastner, 2015a] R.E. Kastner, *Haag's Theorem as a Reason to Reconsider Direct-Action Theories*, International Journal of Quantum Foundations, **1**, (2), 56-64, (2015) (<https://arxiv.org/pdf/1502.03814.pdf>)
- [Kastner, 2015b] R.E. Kastner, *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility* (Cambridge University Press, 2012)
- [Grundler, pdf] G. Gründler, *Remarks on Wheeler-Feynman Absorber Theory* (<https://arxiv.org/pdf/1501.03516.pdf>)
- [Lazarovici, preprint, pdf] D. Lazarovici, *Against Fields*, (preprint, pdf available online),
- [Klaczynski, 2015] L. Klaczynski, *Haag's Theorem in Renormalisable Quantum Field Theories*, Doctoral Thesis at Humboldt-Universität zu Berlin, (2015), (available on line at <http://www2.mathematik.hu-berlin.de/kreimer/wp-content/uploads/PhDLutz.pdf>)
- [Thompson et al., 2018] D. Thompson et al. *Hundred-fold enhancement in far-field radiative heat transfer over the blackbody limit*, Nature, 3 September, 2018.
- [1] R. Haag, *On quantum field theories*, Matematisk-fysiske Meddelelser, **29**, 12, 1955.
- [2] A.I. Akhiezer and V.B. Berestetski, *Quantum Electrodynamics*, Interscience, New York, 1965.
- [3] R.S. Swathi and K.L. Sebastian, *Distance dependence of fluorescence resonance energy transfer*, J. Chem. Soc. **121**, 777-787, (2009)
- [4] R.S. Swathi and K.L. Sebastian, *Long range resonance energy transfer from a dye molecule to graphene has (distance)⁻⁴ dependence*, J. Chem Phys. **130**, 086101, (2009).

- [5] L.I. Rangel-Zamudio, D.S. Rushforth, and R.P. Van Duyne, *Long-Range Resonance Energy Transfer from Aromatic Hydrocarbons to the Anion Radical of Tetracyanoquinodimethane*, J. Phys. Chem. **90**, 807, (1986)
- [6] R. Deshmukh et al., *Long-Range Resonant Energy Transfer Using Optical Topological Transitions in Metamaterials*, ACS Photonics, **5**, 2737-2741, (2018)
- [7] K. Akulov et al., *Long-Distance Resonant Energy Transfer Mediated by Hybrid Plasmonic-Photonic Modes*, J. Phys. Chem. C., just-accepted manuscript, 18 June, (2018).
- [8] D. Martin-Cano et al., *Resonance Energy Transfer and Superradiance Mediated by Plasmonic Nanowaveguides*, Nano Lett. **10** (8), 3129-3134, (2010).
- [9] City College of New York, Team study breaks Forster resonant energy transfer (FRET) distance limit, Physics, Optics and Photonics, Oct. 25, (2018).

Suggestions for further reading

1. G. Cario and J. Frank, Z. Phys. **17**, 202, (1923)
2. J.S. Avery, Z. Bay and A. Szent-Györgyi, *On the Energy Transfer in Biological Systems*, Proc. Nat. Acad. Sci. (US), **47**, 1742-1744, (1961)
3. J.S. Avery, *Resonance energy transfer and spontaneous photon emission*, Proc. Phys. Soc. (London) **88**, Part 1, (1966)
4. J.S. Avery, *Use of the S-Matrix in the Relativistic Treatment of Resonance Energy Transfer*. Int. J. Quantum Chemistry, **XXV**, 79-96, (1984)
5. L. Gomberoff and E.A. Power, *The resonance transfer of excitation*, Proc. Pys. Soc. **88**, 281. (1966)
6. G.D. Scholes, *Long-Range Resonance Energy Transfer in Molecular Systems*, Annu. Rev. Phys. Chem. **54**, 57-87, (2003)
7. K.F. Chou and A.M. Dennis, *Förster Resonance Energy Transfer Between Quantum Dot Donors and Quantum Dot Acceptors*, Sensors, **2015**, 15, 13288-13325, (2015)
8. K-D. Wegner, *Förster Resonance Energy Transfer From Terbium Complexes to Quantum Dots for Multiplexed Immunoassays and Molecular Rulers*, Doctoral Thesis, Université Paris-Sud, (2015)
9. M. Kusku and O.B. Akan, *A Nanoscale Communication Channel with Fluorescence Resonance Energy Transfer (FRET)*, 1st IEEE International Workshop on Molecular and Nano Scale Communication (MoNaCom), Published in: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), (2011)
10. R.M. Pearlstein, *Photosynthetic Exciton Transfer in the 1960's*, Photosynthesis Research **73**, 119-126, (2002)
11. D.L. Andrews and A.A. Demidov, *Resonance Energy Transfer*, (Wiley, 1999)
12. G. Juzeliunas and D.L. Andrews, *Quantum Electrodynamics of Resonance Energy Transfer*, Adv. Chem. Physics, **112**, (2000)

13. D.L. Anderson, Editor, *Encyclopedia of Applied Spectroscopy*, Wiley, (2008)
14. G.J. Daniels, R.G. Jenkins, D.S. Bradshaw, and D.L. Andrews, *Resonance Energy Transfer: The Unified Theory Revisited*, J. Chem. Phys. **119**, 2264-2274, (2003)
15. D.S. Bradshaw and D.L. Andrews, *Optically controlled resonance energy transfer: Mechanisms and configuration for all-optical switching*, J. Chem. Phys. **128**, 144506, (2008)
16. J.A. Wheeler and R.P. Feynman, *Classical Electrodynamics in Terms of Direct Interparticle Interaction*, Rev. Mod. Phys. **21**, 3, 425-433, (1949)
17. R.E. Kastner, *Haag's Theorem as a Reason to Reconsider Direct-Action Theories*, International Journal of Quantum Foundations, **1**, (2), 56-64, (2015) (<https://arxiv.org/pdf/1502.03814>)
18. R.E. Kastner, *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility* (Cambridge University Press, 2012)
19. G. Gründler, *Remarks on Wheeler-Feynman Absorber Theory* (<https://arxiv.org/pdf/1501.03516.pdf>)
20. D. Lazarovici, *Against Fields*, (preprint, pdf available online),
21. L. Klaczynski, *Haag's Theorem in Renormalisable Quantum Field Theories*, Doctoral Thesis at Humboldt-Universität zu Berlin, (2015), (available on line at <http://www2.mathematik.hu-berlin.de/kreimer/wp-content/uploads/PhDLutz.pdf>)
22. D. Thompsen et al. *Hundred-fold enhancement in far-field radiative heat transfer over the blackbody limit*, Nature, 3 September, 2018.
23. R. Haag, *On quantum field theories*, Matematisk-fysiske Meddelelser, **29**, 12, 1955.
24. A.I. Akhiezer and V.B. Berestetski, *Quantum Electrodynamics*, Interscience, New York, 1965.
25. R.S. Swathi and K.L. Sebastian, *Distance dependance of fluorescence resonance energy transfer*, J. Chem. Soc. **121**, 777-787, (2009)
26. R.S. Swathi and K.L. Sebastian, *Long range resonance energy transfer from a dye molecule to graphene has (distance)⁻⁴ dependence*, J. Chem Phys. **130**, 086101, (2009).
27. L.I. Rangel-Zamudio, D.S. Rushforth, and R.P. Van Duyne, *Long-Range Resonance Energy Transfer from Aromatic Hydrocarbons to the Anion Radical of Tetracyanoquinodimethane*, J. Phys. Chem. **90**, 807, (1986)
28. R. Deshmukh et al., *Long-Range Resonant Energy Transfer Using Optical Topological Transitions in Metamaterials*, ACS Photonics, **5**, 2737-2741, (2018)
29. K. Akulov et al., *Long-Distance Resonant Energy Transfer Mediated by Hybrid Plasmonic-Photonic Modes*, J. Phys. Chem. C., just-accepted manuscript, 18 June, (2018).
30. D. Martin-Cano et al., *Resonance Energy Transfer and Superradiance Mediated by Plasmonic Nanowaveguides*, Nano Lett. **10** (8), 3129-3134, (2010).
31. City College of New York, Team study breaks Forster resonant energy transfer (FRET) distance limit, Physics, Optics and Photonics, Oct. 25, (2018).

Appendix A

THE HISTORY OF COMPUTERS

A.1 Pascal and Leibniz

If civilization survives, historians in the distant future will undoubtedly regard the invention of computers as one of the most important steps in human cultural evolution - as important as the invention of writing or the invention of printing. The possibilities of artificial intelligence have barely begun to be explored, but already the impact of computers on society is enormous.

The invention of transistors was a crucial step in the history of computers, and this invention in turn depended on the development of quantum theory. Thus quantum theory, despite its highly abstract nature, has had an enormous impact on the modern world.

The first programmable universal computers were completed in the mid-1940's; but they had their roots in the much earlier ideas of Blaise Pascal (1623-1662), Gottfried Wilhelm Leibniz (1646-1716), Joseph Marie Jacquard (1752-1834) and Charles Babbage (1791-1871).

In 1642, the distinguished French mathematician and philosopher Blaise Pascal completed a working model of a machine for adding and subtracting. According to tradition, the idea for his "calculating box" came to Pascal when, as a young man of 17, he sat thinking of ways to help his father (who was a tax collector). In describing his machine, Pascal wrote: "I submit to the public a small machine of my own invention, by means of which you alone may, without any effort, perform all the operations of arithmetic, and may be relieved of the work which has often times fatigued your spirit when you have worked with the counters or with the pen."

Pascal's machine worked by means of toothed wheels. It was much improved by Leibniz, who constructed a mechanical calculator which, besides adding and subtracting, could also multiply and divide. His first machine was completed in 1671; and Leibniz' description of it, written in Latin, is preserved in the Royal Library at Hanover: "There are two parts of the machine, one designed for addition (and subtraction), and the other designed for multiplication (and division); and they should fit together. The adding (and subtracting) machine coincides completely with the calculating box of Pascal. Something, however,



Figure A.1: **Blaise Pascal (1623-1662)** was a French mathematician, physicist, writer, inventor and theologian. Pascal, a child prodigy, was educated by his father, who was a tax-collector. He invented his calculating box to make his father's work less tedious.



Figure A.2: The German mathematician, philosopher and universal genius Gottfried Wilhelm von Leibniz (1646-1716) was a contemporary of Isaac Newton. He invented differential and integral calculus independently, just as Newton had done many years earlier. However, Newton had not published his work on calculus, and thus a bitter controversy over priority was precipitated. When his patron, the Elector of Hanover moved to England to become George I, Leibniz was left behind because the Elector feared that the controversy would alienate the English. Leibniz extended Pascal's calculating box so that it could perform multiplication and division. Calculators of his design were still being used in the 1960's.

must be added for the sake of multiplication..."

"The wheels which represent the multiplicand are all of the same size, equal to that of the wheels of addition, and are also provided with ten teeth which, however, are movable so that at one time there should protrude 5, at another 6 teeth, etc., according to whether the multiplicand is to be represented five times or six times, etc."

"For example, the multiplicand 365 consists of three digits, 3, 6, and 5. Hence the same number of wheels is to be used. On these wheels, the multiplicand will be set if from the right wheel there protrude 5 teeth, from the middle wheel 6, and from the left wheel 3."

A.2 Jacquard and Babbage

By 1810, calculating machines based on Leibniz' design were being manufactured commercially; and mechanical calculators of a similar (if much improved) design could be found in laboratories and offices until the 1960's. The idea of a programmable universal computer is due to the English mathematician, Charles Babbage, who was the Lucasian Professor of

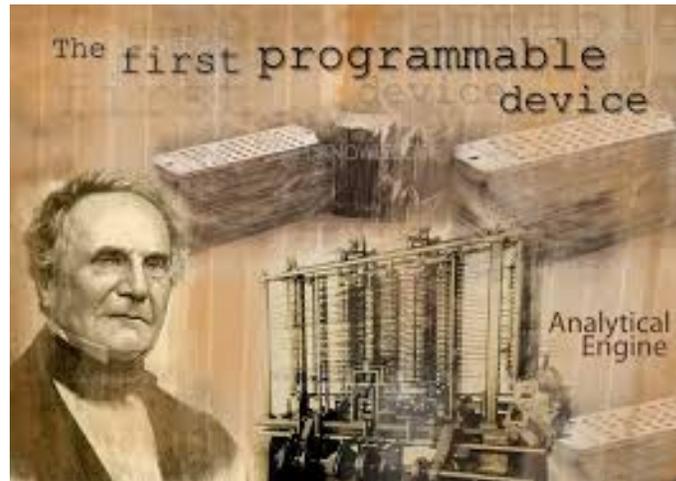


Figure A.3: Charles Babbage (1791-1871) and his analytical engine.

Mathematics at Cambridge University. (In the 17th century, Isaac Newton held this post, and in the 20th century, P.A.M. Dirac and Stephen Hawking also held it.)

In 1812, Babbage conceived the idea of constructing a machine which could automatically produce tables of functions, provided that the functions could be approximated by polynomials. He constructed a small machine, which was able to calculate tables of quadratic functions to eight decimal places, and in 1832 he demonstrated this machine to the Royal Society and to representatives of the British government.

The demonstration was so successful that Babbage secured financial support for the construction of a large machine which would tabulate sixth-order polynomials to twenty decimal places. The large machine was never completed, and twenty years later, after having spent seventeen thousand pounds on the project, the British government withdrew its support. The reason why Babbage's large machine was never finished can be understood from the following account by Lord Moulton of a visit to the mathematician's laboratory:

"One of the sad memories of my life is a visit to the celebrated mathematician and inventor, Mr. Babbage. He was far advanced in age, but his mind was still as vigorous as ever. He took me through his workrooms."

"In the first room I saw the parts of the original Calculating Machine, which had been shown in an incomplete state many years before, and had even been put to some use. I asked him about its present form. 'I have not finished it, because in working at it, I came on the idea of my Analytical Machine, which would do all that it was capable of doing, and much more. Indeed, the idea was so much simpler that it would have taken more work to complete the Calculating Machine than to design and construct the other in its entirety; so I turned my attention to the Analytical Machine.'"

"After a few minutes talk, we went into the next workroom, where he showed me the working of the elements of the Analytical Machine. I asked if I could see it. 'I have never completed it,' he said, 'because I hit upon the idea of doing the same thing by a different and far more effective method, and this rendered it useless to proceed on the old lines.'"



Figure A.4: **Joseph Marie Jacquard (1752-1834) invented a loom which could be programed to produce any design by means of punched cards. News of his invention inspired Babbage to invent a universal programmable computing machine.**

“Then we went into a third room. There lay scattered bits of mechanism, but I saw no trace of any working machine. Very cautiously I approached the subject, and received the dreaded answer: ‘It is not constructed yet, but I am working at it, and will take less time to construct it altogether than it would have taken to complete the Analytical Machine from the stage in which I left it.’ I took leave of the old man with a heavy heart.”

Babbage’s first calculating machine was a special-purpose mechanical computer, designed to tabulate polynomial functions; and he abandoned this design because he had hit on the idea of a universal programmable computer. Several years earlier, the French inventor Joseph Marie Jacquard had constructed an automatic loom in which large wooden “punched cards” were used to control the warp threads. Inspired by Jacquard’s invention, Babbage planned to use punched cards to program his universal computer. (Jacquard’s looms could be programmed to weave extremely complex patterns: A portrait of the inventor, woven on one of his looms in Lyon, hung in Babbage’s drawing room.)

One of Babbage’s frequent visitors was Augusta Ada¹, Countess of Lovelace (1815-1852), the daughter of Lord and Lady Byron. She was a mathematician of considerable ability, and it is through her lucid descriptions that we know how Babbage’s never-completed Analytical Machine was to have worked.

¹ The programming language ADA is named after her.



Figure A.5: **Jacquard's loom.**



Figure A.6: Lord Byron's daughter, Augusta Ada, Countess of Lovelace (1815-1852) was an accomplished mathematician and a frequent visitor to Babbage's workshop. It is through her lucid description of his ideas that we know how Babbage's universal calculating machine was to have worked. The programming language ADA is named after her.

A.3 Harvard's sequence-controlled calculator

The next step towards modern computers was taken by Herman Hollerith, a statistician working for the United States Bureau of the Census. He invented electromechanical machines for reading and sorting data punched onto cards. Hollerith's machines were used to analyze the data from the 1890 United States Census. Because the Census Bureau was a very limited market, Hollerith branched out and began to manufacture similar machines for use in business and administration. His company was later bought out by Thomas J. Watson, who changed its name to International Business Machines.

In 1937, Howard Aiken, of Harvard University, became interested in combining Babbage's ideas with some of the techniques which had developed from Hollerith's punched card machines. He approached the International Business Machine Corporation, the largest manufacturer of punched card equipment, with a proposal for the construction of a large, automatic, programmable calculating machine.

Aiken's machine, the Automatic Sequence Controlled Calculator (ASCC), was completed in 1944 and presented to Harvard University. Based on geared wheels, in the Pascal-Leibniz-Babbage tradition, ASCC had more than three quarters of a million parts and used 500 miles of wire. ASCC was unbelievably slow by modern standards - it took three-tenths of a second to perform an addition - but it was one of the first programmable general-purpose digital computers ever completed. It remained in continuous use, day and night, for fifteen years.



Figure A.7: The Automatic Sequence-Controlled Calculator ASCC can still be seen by visitors at Harvard's science building and cafeteria.

A.4 The first electronic computers

In the ASCC, binary numbers were represented by relays, which could be either on or off. The on position represented 1, while the off position represented 0, these being the only two digits required to represent numbers in the binary (base 2) system. Electromechanical calculators similar to ASCC were developed independently by Konrad Zuse in Germany and by George R. Stibitz at the Bell Telephone Laboratory.

Electronic digital computers

In 1937, the English mathematician A.M. Turing published an important article in the Proceedings of the London Mathematical Society in which envisioned a type of calculating machine consisting of a long row of cells (the “tape”), a reading and writing head, and a set of instructions specifying the way in which the head should move the tape and modify the state and “color” of the cells on the tape. According to a hypothesis which came to be known as the “Church-Turing hypothesis”, the type of computer proposed by Turing was capable of performing every possible type of calculation. In other words, the Turing machine could function as a universal computer.

In 1943, a group of English engineers, inspired by the ideas of Alan Turing and those of the mathematician M.H.A. Newman, completed the electronic digital computer Colossus. Colossus was the first large-scale electronic computer. It was used to break the German Enigma code; and it thus affected the course of World War II.

In 1946, ENIAC (Electronic Numerical Integrator and Calculator) became operational. This general-purpose computer, designed by J.P. Eckert and J.W. Mauchley of the University of Pennsylvania, contained 18,000 vacuum tubes, one or another of which was often out of order. However, during the periods when all its vacuum tubes were working, an electronic computer like Colossus or ENIAC could shoot ahead of an electromechanical machine (such as ASCC) like a hare outdistancing a tortoise.

During the summer of 1946, a course on “The Theory and Techniques of Electronic Digital Computers” was given at the University of Pennsylvania. The ideas put forward in this course had been worked out by a group of mathematicians and engineers headed by J.P. Eckert, J.W. Mauchley and John von Neumann, and these ideas very much influenced all subsequent computer design.

Cybernetics

The word “Cybernetics”, was coined by the American mathematician Norbert Wiener (1894-1964) and his colleagues, who defined it as “the entire field of control and communication theory, whether in the machine or in the animal”. Wiener derived the word from the Greek term for “steersman”.

Norbert Wiener began life as a child prodigy: He entered Tufts University at the age of 11 and received his Ph.D. from Harvard at 19. He later became a professor of mathematics at the Massachusetts Institute of Technology. In 1940, with war on the horizon,



Figure A.8: Alan Turing (1912-1954). He is considered to be the father of theoretical computer science. During World War II, Turing's work allowed the allies to crack the German's code. This appreciably shortened the length of the war in Europe, and saved many lives.



Figure A.9: John von Neumann (1903-1957, right) with J. Robert Oppenheimer. In the background is an electronic digital computer.

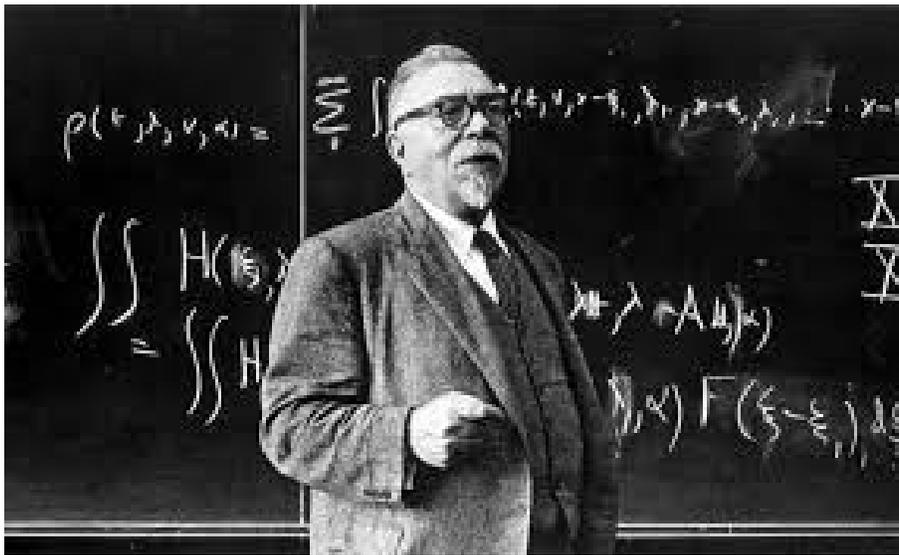


Figure A.10: MIT's Norbert Wiener (1894-1964) coined the word "Cybernetics", derived from a Greek word meaning "steersman". Wiener was one of the principle organizers of the Macy Conferences.

Wiener sent a memorandum to Vannevar Bush, another MIT professor who had done pioneering work with analogue computers, and had afterwards become the chairman of the U.S. National Defense Research Committee. Wiener's memorandum urged the American government to support the design and construction of electronic digital computers, which would make use of binary numbers, vacuum tubes, and rapid memories. In such machines, the memorandum emphasized, no human intervention should be required except when data was to be read into or out of the machine.

Like Leo Szilard, John von Neumann, Claude Shannon and Erwin Schrödinger, Norbert Wiener was aware of the relation between information and entropy. In his 1948 book *Cybernetics* he wrote: "...we had to develop a statistical theory of the amount of information, in which the unit amount of information was that transmitted by a single decision between equally probable alternatives. This idea occurred at about the same time to several writers, among them the statistician R.A. Fisher, Dr. Shannon of Bell Telephone Laboratories, and the author. Fisher's motive in studying this subject is to be found in classical statistical theory; that of Shannon in the problem of coding information; and that of the author in the problem of noise and message in electrical filters... The notion of the amount of information attaches itself very naturally to a classical notion in statistical mechanics: that of entropy. Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization; and the one is simply the negative of the other."

During World War II, Norbert Wiener developed automatic systems for control of anti-aircraft guns. His systems made use of feedback loops closely analogous to those with which animals coordinate their movements. In the early 1940's, he was invited to attend a



Figure A.11: Margaret Mead (1901-1978) and Gregory Bateson (1904-1980). They used the feedback loops studied by Wiener to explain many aspects of human behavior. Bateson is considered to be one of the main founders of the discipline Biosemiotics, which considers information to be the central feature of living organisms.

series of monthly dinner parties organized by Arturo Rosenbluth, a professor of physiology at Harvard University. The purpose of these dinners was to promote discussions and collaborations between scientists belonging to different disciplines. The discussions which took place at these dinners made both Wiener and Rosenbluth aware of the relatedness of a set of problems that included homeostasis and feedback in biology, communication and control mechanisms in neurophysiology, social communication among animals (or humans), and control and communication involving machines.

Wiener and Rosenbluth therefore tried to bring together workers in the relevant fields to try to develop common terminology and methods. Among the many people whom they contacted were the anthropologists Gregory Bateson and Margaret Mead, Howard Aiken (the designer of the Automatic Sequence Controlled Calculator), and the mathematician John von Neumann. The Josiah Macy Jr. Foundation sponsored a series of ten yearly

meetings, which continued until 1949 and which established cybernetics as a new research discipline. It united areas of mathematics, engineering, biology, and sociology which had previously been considered unrelated. Among the most important participants (in addition to Wiener, Rosenbluth, Bateson, Mead, and von Neumann) were Heinz von Foerster, Kurt Lewin, Warren McCulloch and Walter Pitts. The Macy conferences were small and informal, with an emphasis on discussion as opposed to the presentation of formal papers. A stenographic record of the last five conferences has been published, edited by von Foerster. Transcripts of the discussions give a vivid picture of the enthusiastic and creative atmosphere of the meetings. The participants at the Macy Conferences perceived Cybernetics as a much-needed bridge between the natural sciences and the humanities. Hence their enthusiasm. Wiener's feedback loops and von Neumann's theory of games were used by anthropologists Mead and Bateson to explain many aspects of human behavior.

A.5 Biosemiotics

The Oxford Dictionary of Biochemistry and Molecular Biology (Oxford University Press, 1997) defines Biosemiotics as "the study of signs, of communication, and of information in living organisms". The biologists Claus Emmeche and K. Kull offer another definition of Biosemiotics: "biology that interprets living systems as sign systems".

The American philosopher Charles Sanders Peirce (1839-1914) is considered to be one of the founders of Semiotics (and hence also of Biosemiotics). Peirce studied philosophy and chemistry at Harvard, where his father was a professor of mathematics and astronomy. He wrote extensively on philosophical subjects, and developed a theory of signs and meaning which anticipated many of the principles of modern Semiotics. Peirce built his theory on a triad: (1) the sign, which represents (2) something to (3) somebody. For example, the sign might be a broken stick, which represents a trail to a hunter, it might be the arched back of a cat, which represents an aggressive attitude to another cat, it might be the waggle-dance of a honey bee, which represents the coordinates of a source of food to her hive-mates, or it might be a molecule of trans-10-cis-hexadecadienol, which represents irresistible sexual temptation to a male moth of the species *Bombyx mori*. The sign might be a sequence of nucleotide bases which represents an amino acid to the ribosome-transfer-RNA system, or it might be a cell-surface antigen which represents self or non-self to the immune system. In information technology, the sign might be the presence or absence of a pulse of voltage, which represents a binary digit to a computer. Semiotics draws our attention to the sign and to its function, and places much less emphasis on the physical object which forms the sign. This characteristic of the semiotic viewpoint has been expressed by the Danish biologist Jesper Hoffmeyer in the following words: "The sign, rather than the molecule, is the basic unit for studying life."

A second important founder of Biosemiotics was Jakob von Uexküll (1864-1944). He was born in Estonia, and studied zoology at the University of Tartu. After graduation, he worked at the Institute of Physiology at the University of Heidelberg, and later at the Zoological Station in Naples. In 1907, he was given an honorary doctorate by Heidelberg



Figure A.12: Charles Sanders Pearce (1839-1914).



Figure A.13: Jakob Johann Baron von Uexküll (1864-1944). Together with Pearce and Bateson, he is one of the principle founders of Biosemiotics.

for his studies of the physiology of muscles. Among his discoveries in this field was the first recognized instance of negative feedback in an organism. Von Uexküll's later work was concerned with the way in which animals experience the world around them. To describe the animal's subjective perception of its environment he introduced the word *Umwelt*; and in 1926 he founded the Institut für Umweltforschung at the University of Heidelberg. Von Uexküll visualized an animal - for example a mouse - as being surrounded by a world of its own - the world conveyed by its own special senses organs, and processed by its own interpretative systems. Obviously, the *Umwelt* will differ greatly depending on the organism. For example, bees are able to see polarized light and ultraviolet light; electric eels are able to sense their environment through their electric organs; many insects are extraordinarily sensitive to pheromones; and a dog's *Umwelt* far richer in smells than that of most other animals. The *Umwelt* of a jellyfish is very simple, but nevertheless it exists.² Von Uexküll's *Umwelt* concept can even extend to one-celled organisms, which receive chemical and tactile signals from their environment, and which are often sensitive to light. The ideas and research of Jakob von Uexküll inspired the later work of the Nobel Laureate ethologist Konrad Lorenz, and thus von Uexküll can be thought of as one of the founders of ethology as well as of Biosemiotics. Indeed, ethology and Biosemiotics are closely related.

Biosemiotics also values the ideas of the American anthropologist Gregory Bateson (1904-1980), who was mentioned in Chapter 7 in connection with cybernetics and with the Macy Conferences. He was married to another celebrated anthropologist, Margaret Mead, and together they applied Norbert Wiener's insights concerning feedback mechanisms to sociology, psychology and anthropology. Bateson was the originator of a famous epigrammatic definition of information: "...a difference which makes a difference". This definition occurs in Chapter 3 of Bateson's book, *Mind and Nature: A Necessary Unity*, Bantam, (1980), and its context is as follows: "To produce news of a difference, i.e. information", Bateson wrote, "there must be two entities... such that news of their difference can be represented as a difference inside some information-processing entity, such as a brain or, perhaps, a computer. There is a profound and unanswerable question about the nature of these two entities that between them generate the difference which becomes information by making a difference. Clearly each alone is - for the mind and perception - a non-entity, a non-being... the sound of one hand clapping. The stuff of sensation, then, is a pair of values of some variable, presented over time to a sense organ, whose response depends on the ratio between the members of the pair."

Microelectronics

The problem of unreliable vacuum tubes was solved in 1948 by John Bardeen, William Shockley and Walter Brattain of the Bell Telephone Laboratories. Application of quantum theory to solids had led to an understanding of the electrical properties of crystals. Like atoms, crystals were found to have allowed and forbidden energy levels.

² It is interesting to ask to what extent the concept of *Umwelt* can be equated to that of consciousness. To the extent that these two concepts can be equated, von Uexküll's *Umweltforschung* offers us the opportunity to explore the phylogenetic evolution of the phenomenon of consciousness.

The allowed energy levels for an electron in a crystal were known to form bands, i.e., some energy ranges with many allowed states (allowed bands), and other energy ranges with none (forbidden bands). The lowest allowed bands were occupied by electrons, while higher bands were empty. The highest filled band was called the “valence band”, and the lowest empty band was called the “conduction band”.

According to quantum theory, whenever the valence band of a crystal is only partly filled, the crystal is a conductor of electricity; but if the valence band is completely filled with electrons, the crystal is an electrical insulator. (A completely filled band is analogous to a room so packed with people that none of them can move.)

In addition to conductors and insulators, quantum theory predicted the existence of “semiconductors” - crystals where the valence band is completely filled with electrons, but where the energy gap between the conduction band and the valence band is very small. For example, crystals of the elements silicon and germanium are semiconductors. For such a crystal, thermal energy is sometimes enough to lift an electron from the valence band to the conduction band.

Bardeen, Shockley and Brattain found ways to control the conductivity of germanium crystals by injecting electrons into the conduction band, or alternatively by removing electrons from the valence band. They could do this by “doping” the crystals with appropriate impurities, or by injecting electrons with a special electrode. The semiconducting crystals whose conductivity was controlled in this way could be used as electronic valves, in place of vacuum tubes.

By the 1960’s, replacement of vacuum tubes by transistors in electronic computers had led not only to an enormous increase in reliability and a great reduction in cost, but also to an enormous increase in speed. It was found that the limiting factor in computer speed was the time needed for an electrical signal to propagate from one part of the central processing unit to another. Since electrical impulses propagate with the speed of light, this time is extremely small; but nevertheless, it is the limiting factor in the speed of electronic computers.

A.6 The Traitorous Eight

According to the Wikipedia article on Shockley,

“In 1956 Shockley moved from New Jersey to Mountain View, California to start Shockley Semiconductor Laboratory to live closer to his ailing mother in Palo Alto, California. The company, a division of Beckman Instruments, Inc., was the first establishment working on silicon semiconductor devices in what came to be known as Silicon Valley.

“His way [of leading the group] could generally be summed up as domineering and increasingly paranoid. In one well-known incident, he claimed that a secretary’s cut thumb was the result of a malicious act and he demanded lie detector tests to find the culprit, when in reality, the secretary had simply grabbed at a door handle that happened to have an exposed tack on it for the purpose of hanging paper notes on. After he received the Nobel Prize in 1956 his demeanor changed, as evidenced in his increasingly autocratic, erratic and



Figure A.14: **William Shockley (1910-1989) shared the 1956 Nobel Prize in Physics with John Bardeen and Walter Brattain.**

hard-to-please management style. In late 1957, eight of Shockley's researchers, who would come to be known as the 'traitorous eight, resigned after Shockley decided not to continue research into silicon-based semiconductors. They went on to form Fairchild Semiconductor, a loss from which Shockley Semiconductor never recovered. Over the course of the next 20 years, more than 65 new enterprises would end up having employee connections back to Fairchild."



Figure A.15: **The Traitorous Eight:** From left to right, Gordon Moore, C. Sheldon Roberts, Eugene Kleiner, Robert Noyce, Victor Grinich, Julius Blank, Jean Hoerni and Jay Last.

A.7 Integrated circuits

In order to reduce the propagation time, computer designers tried to make the central processing units very small; and the result was the development of integrated circuits and microelectronics. (Another motive for miniaturization of electronics came from the requirements of space exploration.)

Integrated circuits were developed in which single circuit elements were not manufactured separately. Instead, the whole circuit was made at one time. An integrated circuit is a sandwich-like structure, with conducting, resisting and insulating layers interspersed with layers of germanium or silicon, “doped ” with appropriate impurities. At the start of the manufacturing process, an engineer makes a large drawing of each layer. For example, the drawing of a conducting layer would contain pathways which fill the role played by wires in a conventional circuit, while the remainder of the layer would consist of areas destined to be etched away by acid.

The next step is to reduce the size of the drawing and to multiply it photographically. The pattern of the layer is thus repeated many times, like the design on a piece of wallpaper. The multiplied and reduced drawing is then focused through a reversed microscope onto the surface to be etched.

Successive layers are built up by evaporating or depositing thin films of the appropriate substances onto the surface of a silicon or germanium wafer. If the layer being made is to be conducting, the surface would consist of an extremely thin layer of copper, covered with a photosensitive layer called a “photoresist”. On those portions of the surface receiving light from the pattern, the photoresist becomes insoluble, while on those areas not receiving light, the photoresist can be washed away.

The surface is then etched with acid, which removes the copper from those areas not protected by photoresist. Each successive layer of a wafer is made in this way, and finally the wafer is cut into tiny “chips”, each of which corresponds to one unit of the wallpaper-like pattern.

Although the area of a chip may be much smaller than a square centimeter, the chip can contain an extremely complex circuit. A typical programmable minicomputer or “microprocessor”, manufactured during the 1970’s, could have 30,000 circuit elements, all of which were contained on a single chip. By 1986, more than a million transistors were being placed on a single chip.

As a result of miniaturization, the speed of computers rose steadily. In 1960, the fastest computers could perform a hundred thousand elementary operations in a second. By 1970, the fastest computers took less than a second to perform a million such operations. In 1987, a computer called GF11 was designed to perform 11 billion floating-point operations (flops) per second.

GF11 (Gigaflop 11) is a scientific parallel-processing machine constructed by IBM. Approximately ten floating-point operations are needed for each machine instruction. Thus GF11 runs at the rate of approximately a thousand million instructions per second (1,100 MIPS). The high speed achieved by parallel-processing machines results from dividing a job into many sub-jobs on which a large number of processing units can work simultaneously.

Computer memories have also undergone a remarkable development. In 1987, the magnetic disc memories being produced could store 20 million bits of information per square inch; and even higher densities could be achieved by optical storage devices. (A “bit” is the unit of information. For example, the number 25, written in the binary system, is 11001. To specify this 5-digit binary number requires 5 bits of information. To specify an n -digit binary number requires n bits of information. Eight bits make a “byte”.)

In the 1970’s and 1980’s, computer networks were set up linking machines in various parts of the world. It became possible (for example) for a scientist in Europe to perform a calculation interactively on a computer in the United States just as though the distant machine were in the same room; and two or more computers could be linked for performing large calculations. It also became possible to exchange programs, data, letters and manuscripts very rapidly through the computer networks.

A.8 Moore’s law

In 1965, only four years after the first integrated circuits had been produced, Dr. Gordon E. Moore, one of the founders of Intel, made a famous prediction which has come to be known as “Moore’s Law”. He predicted that the number of transistors per integrated circuit would double every two years, and that this trend would continue through 1975. In fact, the general trend predicted by Moore has continued for a much longer time. Although the number of transistors per unit area has not continued to double every two years, the logic density (bits per unit area) has done so, and thus a modified version of Moore’s law still holds today. How much longer the trend can continue remains to be seen. Physical limits to miniaturization of transistors of the present type will soon be reached; but there is hope that further miniaturization can be achieved through “quantum dot” technology, molecular switches, and autoassembly.

A typical programmable minicomputer or “microprocessor”, manufactured in the 1970’s, could have 30,000 circuit elements, all of which were contained on a single chip. By 1989, more than a million transistors were being placed on a single chip; and by 2000, the number reached 42,000,000.

As a result of miniaturization and parallelization, the speed of computers rose exponentially. In 1960, the fastest computers could perform a hundred thousand elementary operations in a second. By 1970, the fastest computers took less than a second to perform a million such operations. In 1987, a massively parallel computer, with 566 parallel processors, called GF11 was designed to perform 11 billion floating-point operations per second (flops). By 2002 the fastest computer performed 40 at teraflops, making use of 5120 parallel CPU’s.

Computer disk storage has also undergone a remarkable development. In 1987, the magnetic disk storage being produced could store 20 million bits of information per square inch; and even higher densities could be achieved by optical storage devices. Storage density has until followed a law similar to Moore’s law.

In the 1970’s and 1980’s, computer networks were set up linking machines in various

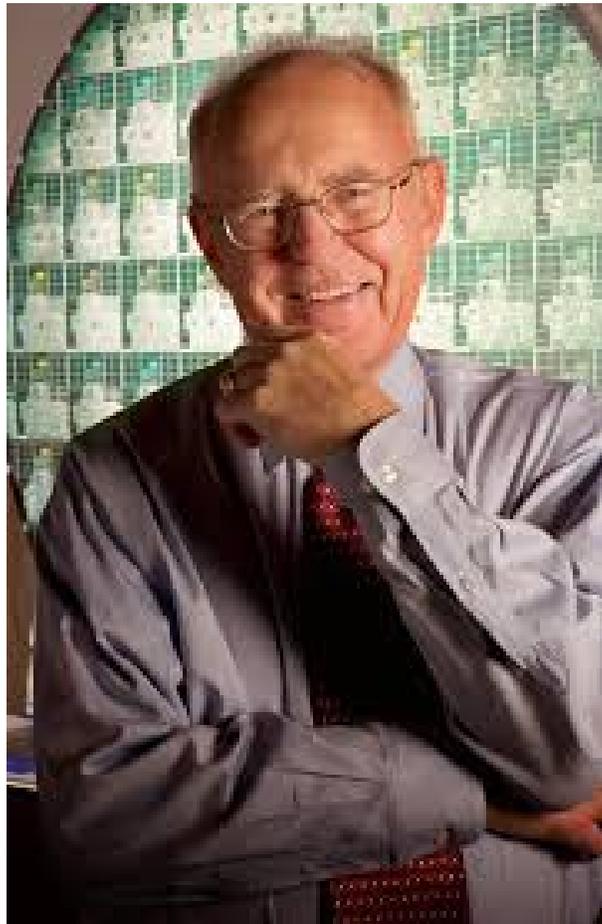


Figure A.16: Gordon E. Moore (born 1929), a founder of Intel and the author of Moore's Law. In 1965 he predicted that the number of components in integrated circuits would double every year for the next 10 years". In 1975 he predicted the this doubling would continue, but revised the doubling rate to "every two years. Astonishingly, Moore's Law has held much longer than he, or anyone else, anticipated.

parts of the world. It became possible (for example) for a scientist in Europe to perform a calculation interactively on a computer in the United States just as though the distant machine were in the same room; and two or more computers could be linked for performing large calculations. It also became possible to exchange programs, data, letters and manuscripts very rapidly through the computer networks.

The exchange of large quantities of information through computer networks was made easier by the introduction of fiber optics cables. By 1986, 250,000 miles of such cables had been installed in the United States. If a ray of light, propagating in a medium with a large refractive index, strikes the surface of the medium at a grazing angle, then the ray undergoes total internal reflection. This phenomenon is utilized in fiber optics: A light signal can propagate through a long, hairlike glass fiber, following the bends of the fiber without losing intensity because of total internal reflection. However, before fiber optics could be used for information transmission over long distances, a technological breakthrough in glass manufacture was needed, since the clearest glass available in 1940 was opaque in lengths more than 10 m. Through studies of the microscopic properties of glasses, the problem of absorption was overcome. By 1987, devices were being manufactured commercially that were capable of transmitting information through fiber-optic cables at the rate of 1.7 billion bits per second.

A.9 Self-reinforcing information accumulation

Humans have been living on the earth for roughly two million years (more or less, depending on where one draws the line between our human and prehuman ancestors, Table 6.1). During almost all of this time, our ancestors lived by hunting and food-gathering. They were not at all numerous, and did not stand out conspicuously from other animals. Then, suddenly, during the brief space of ten thousand years, our species exploded in numbers from a few million to seven billion, populating all parts of the earth, and even setting foot on the moon. This population explosion, which is still going on, has been the result of dramatic cultural changes. Genetically we are almost identical with our hunter-gatherer ancestors, who lived ten thousand years ago, but cultural evolution has changed our way of life beyond recognition.

Beginning with the development of speech, human cultural evolution began to accelerate. It started to move faster with the agricultural revolution, and faster still with the invention of writing and printing. Finally, modern science has accelerated the rate of social and cultural change to a completely unprecedented speed.

The growth of modern science is accelerating because knowledge feeds on itself. A new idea or a new development may lead to several other innovations, which can in turn start an avalanche of change. For example, the quantum theory of atomic structure led to the invention of transistors, which made high-speed digital computers possible. Computers have not only produced further developments in quantum theory; they have also revolutionized many other fields.

The self-reinforcing accumulation of knowledge - the information explosion - which

characterizes modern human society is reflected not only in an explosively-growing global population, but also in the number of scientific articles published, which doubles roughly every ten years. Another example is Moore's law - the doubling of the information density of integrated circuits every two years. Yet another example is the explosive growth of Internet traffic shown in Table 17.1.

The Internet itself is the culmination of a trend towards increasing societal information exchange - the formation of a collective human consciousness. This collective consciousness preserves the observations of millions of eyes, the experiments of millions of hands, the thoughts of millions of brains; and it does not die when the individual dies.

A.10 Automation

During the last three decades, the cost of computing has decreased exponentially by between twenty and thirty percent per year. Meanwhile, the computer industry has grown exponentially by twenty percent per year (faster than any other industry). The astonishing speed of this development has been matched by the speed with which computers have become part of the fabric of science, engineering, industry, commerce, communications, transport, publishing, education and daily life in the industrialized parts of the world.

The speed, power and accuracy of computers has revolutionized many branches of science. For example, before the era of computers, the determination of a simple molecular structure by the analysis of X-ray diffraction data often took years of laborious calculation; and complicated structures were completely out of reach. In 1949, however, Dorothy Crowfoot Hodgkin used an electronic computer to work out the structure of penicillin from X-ray data. This was the first application of a computer to a biochemical problem; and it was followed by the analysis of progressively larger and more complex structures.

Proteins, DNA, and finally even the detailed structures of viruses were studied through the application of computers in crystallography. The enormous amount of data needed for such studies was gathered automatically by computer-controlled diffractometers; and the final results were stored in magnetic-tape data banks, available to users through computer networks.

The application of quantum theory to chemical problems is another field of science which owes its development to computers. When Erwin Schrödinger wrote down his wave equation in 1926, it became possible, in principle, to calculate most of the physical and chemical properties of matter. However, the solutions to the Schrödinger equation for many-particle systems can only be found approximately; and before the advent of computers, even approximate solutions could not be found, except for the simplest systems.

When high-speed electronic digital computers became widely available in the 1960's, it suddenly became possible to obtain solutions to the Schrödinger equation for systems of chemical and even biochemical interest. Quantum chemistry (pioneered by such men as J.C. Slater, R.S. Mullikin, D.R. Hartree, V. Fock, J.H. Van Vleck, L. Pauling, E.B. Wilson, P.O. Löwdin, E. Clementi, C.J. Ballhausen and others) developed into a rapidly-growing field, as did solid state physics. Through the use of computers, it became possible to

design new materials with desired chemical, mechanical, electrical or magnetic properties. Applying computers to the analysis of reactive scattering experiments, D. Herschbach, J. Polanyi and Y. Lee were able to achieve an understanding of the dynamics of chemical reactions.

The successes of quantum chemistry led Albert Szent-Györgyi, A. and B. Pullman, H. Scheraga and others to pioneer the fields of quantum biochemistry and molecular dynamics. Computer programs for drug design were developed, as well as molecular-dynamics programs which allowed the conformations of proteins to be calculated from a knowledge of their amino acid sequences. Studies in quantum biochemistry have yielded insights into the mechanisms of enzyme action, photosynthesis, active transport of ions across membranes, and other biochemical processes.

In medicine, computers began to be used for monitoring the vital signs of critically ill patients, for organizing the information flow within hospitals, for storing patients' records, for literature searches, and even for differential diagnosis of diseases.

The University of Pennsylvania has developed a diagnostic program called INTERNIST-1, with a knowledge of 577 diseases and their interrelations, as well as 4,100 signs, symptoms and patient characteristics. This program was shown to perform almost as well as an academic physician in diagnosing difficult cases. QMR (Quick Medical Reference), a microcomputer adaptation of INTERNIST-1, incorporates the diagnostic functions of the earlier program, and also offers an electronic textbook mode.

Beginning in the 1960's, computers played an increasingly important role in engineering and industry. For example, in the 1960's, Rolls Royce Ltd. began to use computers not only to design the optimal shape of turbine blades for aircraft engines, but also to control the precision milling machines which made the blades. In this type of computer-assisted design and manufacture, no drawings were required. Furthermore, it became possible for an industry requiring a part from a subcontractor to send the machine-control instructions for its fabrication through the computer network to the subcontractor, instead of sending drawings of the part.

In addition to computer-controlled machine tools, robots were also introduced. They were often used for hazardous or monotonous jobs, such as spray-painting automobiles; and they could be programmed by going through the job once manually in the programming mode. By 1987, the population of robots in the United States was between 5,000 and 7,000, while in Japan, the Industrial Robot Association reported a robot population of 80,000.

Chemical industries began to use sophisticated computer programs to control and to optimize the operations of their plants. In such control systems, sensors reported current temperatures, pressures, flow rates, etc. to the computer, which then employed a mathematical model of the plant to calculate the adjustments needed to achieve optimum operating conditions.

Not only industry, but also commerce, felt the effects of computerization during the postwar period. Commerce is an information-intensive activity; and in fact some of the crucial steps in the development of information-handling technology developed because of the demands of commerce: The first writing evolved from records of commercial transactions kept on clay tablets in the Middle East; and automatic business machines, using

punched cards, paved the way for the development of the first programmable computers.

Computerization has affected wholesaling, warehousing, retailing, banking, stockmarket transactions, transportation of goods - in fact, all aspects of commerce. In wholesaling, electronic data is exchanged between companies by means of computer networks, allowing order-processing to be handled automatically; and similarly, electronic data on prices is transmitted to buyers.

The key to automatic order-processing in wholesaling was standardization. In the United States, the Food Marketing Institute, the Grocery Manufacturers of America, and several other trade organizations, established the Uniform Communications System (UCS) for the grocery industry. This system specifies a standard format for data on products, prices and orders.

Automatic warehouse systems were designed as early as 1958. In such systems, the goods to be stored are placed on pallets (portable platforms), which are stacked automatically in aisles of storage cubicles. A computer records the position of each item for later automatic retrieval.

In retailing, just as in wholesaling, standardization proved to be the key requirement for automation. Items sold in supermarkets in most industrialized countries are now labeled with a standard system of machine-readable thick and thin bars known as the Universal Product Code (UPC). The left-hand digits of the code specify the manufacturer or packer of the item, while the right-hand set of digits specify the nature of the item. A final digit is included as a check, to make sure that the others were read correctly. This last digit (called a modulo check digit) is the smallest number which yields a multiple of ten when added to the sum of the previous digits.

When a customer goes through a check-out line, the clerk passes the purchased items over a laser beam and photocell, thus reading the UPC code into a small embedded computer or microprocessor at the checkout counter, which adds the items to the customer's bill. The microprocessor also sends the information to a central computer and inventory data base. When stocks of an item become low, the central computer generates a replacement order. The financial book-keeping for the retailing operation is also carried out automatically by the central computer.

In many places, a customer passing through the checkout counter of a supermarket is able to pay for his or her purchases by means of a plastic card with a magnetic, machine-readable identification number. The amount of the purchase is then transmitted through a computer network and deducted automatically from the customer's bank account. If the customer pays by check, the supermarket clerk may use a special terminal to determine whether a check written by the customer has ever "bounced".

Most checks are identified by a set of numbers written in the Magnetic-Ink Character Recognition (MICR) system. In 1958, standards for the MICR system were established, and by 1963, 85 percent of all checks written in the United States were identified by MICR numbers. By 1968, almost all banks had adopted this system; and thus the administration of checking accounts was automated, as well as the complicated process by which a check, deposited anywhere in the world, returns to the payers bank.

Container ships were introduced in the late 1950's, and since that time, container sys-

tems have increased cargo-handling speeds in ports by at least an order of magnitude. Computer networks contributed greatly to the growth of the container system of transportation by keeping track of the position, ownership and contents of the containers.

In transportation, just as in wholesaling and retailing, standardization proved to be a necessary requirement for automation. Containers of a standard size and shape could be loaded and unloaded at ports by specialized tractors and cranes which required only a very small staff of operators. Standard formats for computerized manifests, control documents, and documents for billing and payment, were instituted by the Transportation Data Coordinating Committee, a non-profit organization supported by dues from shipping firms.

In the industrialized parts of the world, almost every type of work has been made more efficient by computerization and automation. Even artists, musicians, architects and authors find themselves making increasing use of computers: Advanced computing systems, using specialized graphics chips, speed the work of architects and film animators. The author's traditional typewriter has been replaced by a word-processor, the composer's piano by a music synthesizer.

In the Industrial Revolution of the 18th and 19th centuries, muscles were replaced by machines. Computerization represents a Second Industrial Revolution: Machines have begun to perform not only tasks which once required human muscles, but also tasks which formerly required human intelligence.

In industrial societies, the mechanization of agriculture has very much reduced the fraction of the population living on farms. For example, in the United States, between 1820 and 1980, the fraction of workers engaged in agriculture fell from 72 percent to 3.1 percent. There are signs that computerization and automation will similarly reduce the number of workers needed in industry and commerce.

Computerization is so recent that, at present, we can only see the beginnings of its impact; but when the Second Industrial Revolution is complete, how will it affect society? When our children finish their education, will they face technological unemployment?

The initial stages of the First Industrial Revolution produced much suffering, because labor was regarded as a commodity to be bought and sold according to the laws of supply and demand, with almost no consideration for the needs of the workers. Will we repeat this mistake? Or will society learn from its earlier experience, and use the technology of automation to achieve widely-shared human happiness?

The Nobel-laureate economist, Wassily W. Leontief, has made the following comment on the problem of technological unemployment:

“Adam and Eve enjoyed, before they were expelled from Paradise, a high standard of living without working. After their expulsion, they and their successors were condemned to eke out a miserable existence, working from dawn to dusk. The history of technological progress over the last 200 years is essentially the story of the human species working its way slowly and steadily back into Paradise. What would happen, however, if we suddenly found ourselves in it? With all goods and services provided without work, no one would be gainfully employed. Being unemployed means receiving no wages. As a result, until appropriate new income policies were formulated to fit the changed technological conditions,

everyone would starve in Paradise.”

To say the same thing in a slightly different way: consider what will happen when a factory which now employs a thousand workers introduces microprocessor-controlled industrial robots and reduces its work force to only fifty. What will the nine hundred and fifty redundant workers do? They will not be able to find jobs elsewhere in industry, commerce or agriculture, because all over the economic landscape, the scene will be the same.

There will still be much socially useful work to be done - for example, taking care of elderly people, beautifying the cities, starting youth centers, planting forests, cleaning up pollution, building schools in developing countries, and so on. These socially beneficial goals are not commercially “profitable”. They are rather the sort of projects which governments sometimes support if they have the funds for it. However, the money needed to usefully employ the nine hundred and fifty workers will not be in the hands of the government. It will be in the hands of the factory owner who has just automated his production line.

In order to make the economic system function again, either the factory owner will have to be persuaded to support socially beneficial but commercially unprofitable projects, or else an appreciable fraction of his profits will have to be transferred to the government, which will then be able to constructively re-employ the redundant workers.

The future problems of automation and technological unemployment may force us to rethink some of our economic ideas. It is possible that helping young people to make a smooth transition from education to secure jobs will become one of the important responsibilities of governments, even in countries whose economies are based on free enterprise. If such a change does take place in the future, while at the same time socialistic countries are adopting a few of the better features of free enterprise, then one can hope that the world will become less sharply divided by contrasting economic systems.

A.11 Neural networks

Physiologists have begun to make use of insights derived from computer design in their efforts to understand the mechanism of the brain; and computer designers are beginning to construct computers modeled after neural networks. We may soon see the development of computers capable of learning complex ideas, generalization, value judgements, artistic creativity, and much else that was once thought to be uniquely characteristic of the human mind. Efforts to design such computers will undoubtedly give us a better understanding of the way in which the brain performs its astonishing functions.

Much of our understanding of the nervous systems of higher animals is due to the Spanish microscopist, Ramón y Cajal, and to the English physiologists, Alan Hodgkin and Andrew Huxley. Cajal’s work, which has been confirmed and elaborated by modern electron microscopy, showed that the central nervous system is a network of nerve cells (neurons) and threadlike fibers growing from them. Each neuron has many input fibers (dendrites), and one output fiber (the axon), which may have several branches.

It is possible the computers of the future will have pattern-recognition and learning abilities derived from architecture inspired by our understanding of the synapse, by Young's model, or by other biological models. However, pattern recognition and learning can also be achieved by programming, using computers of conventional architecture. Programs already exist which allow computers to understand both handwriting and human speech; and a recent chess-playing program was able to learn by studying a large number of championship games. Having optimized its parameters by means of this learning experience, the chess-playing program was able to win against grand masters!

Like nuclear physics and genesplicing, artificial intelligence presents a challenge: Will society use its new powers wisely and humanely? The computer technology of the future can liberate us from dull and repetitive work, and allow us to use our energies creatively; or it can produce unemployment and misery, depending on how we organize our society. Which will we choose?

Suggestions for further reading

1. H. Babbage, *Babbages Calculating Engines: A Collection of Papers by Henry Prevost Babbage*, MIT Press, (1984).
2. A.M. Turing, *The Enigma of Intelligence*, Burnett, London (1983).
3. Ft. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford University Press, (1989).
4. S. Wolfram, *A New Kind of Science*, Wolfram Media, Champaign IL, (2002).
5. A.M. Turing, *On computable numbers, with an application to the Entscheidungsproblem*, Proc. Lond. Math. Soc. Ser 2, **42**, (1937). Reprinted in M. David Ed., *The Undecidable*, Raven Press, Hewlett N.Y., (1965).
6. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
7. J. Shurkin, *Engines of the Mind: A History of Computers*, W.W. Norton, (1984).
8. J. Palfreman and D. Swade, *The Dream Machine: Exploring the Computer Age*, BBC Press (UK), (1991).
9. T.J. Watson, Jr. and P. Petre, *Father, Son, and Co.*, Bantam Books, New York, (1991).
10. A. Hodges, Alan Turing: *The Enigma*, Simon and Schuster, (1983).
11. H.H. Goldstein, *The Computer from Pascal to Von Neumann*, Princeton University Press, (1972).
12. C.J. Bashe, L.R. Johnson, J.H. Palmer, and E.W. Pugh, *IBM's Early Computers*, Vol. 3 in the History of Computing Series, MIT Press, (1986).
13. K.D. Fishman, *The Computer Establishment*, McGraw-Hill, (1982).
14. S. Levy, *Hackers*, Doubleday, (1984).
15. S. Franklin, *Artificial Minds*, MIT Press, (1997).
16. P. Freiberger and M. Swaine, *Fire in the Valley: The Making of the Personal Computer*, Osborne/McGraw-Hill, (1984).
17. R.X. Cringely, *Accidental Empires*, Addison-Wesley, (1992).

18. R. Randell editor, *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
19. H. Lukoff, *From Dits to Bits*, Robotics Press, (1979).
20. D.E. Lundstrom, *A Few Good Men from Univac*, MIT Press, (1987).
21. D. Rutland, *Why Computers Are Computers (The SWAC and the PC)*, Wren Publishers, (1995).
22. P.E. Ceruzzi, *Reckoners: The Prehistory of the Digital Computer, from Relays to the Stored Program Concept, 1935-1945*, Greenwood Press, Westport, (1983)
23. S.G. Nash, *A History of Scientific Computing*, Adison-Wesley, Reading Mass., (1990).
24. P.E. Ceruzzi, *Crossing the divide: Architectural issues and the emergence of stored programme computers, 1935-1953*, IEEE Annals of the History of Computing, **19**, 5-12, January-March (1997).
25. P.E. Ceruzzi, *A History of Modern Computing*, MIT Press, Cambridge MA, (1998).
26. K. Zuse, *Some remarks on the history of computing in Germany*, in *A History of Computing in the 20th Century*, N. Metropolis et al. editors, 611-627, Academic Press, New York, (1980).
27. A.R. Mackintosh, *The First Electronic Computer*, Physics Today, March, (1987).
28. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
29. A. Hodges, *Alan Turing: The Enigma*, Simon and Schuster, New York, (1983).
30. A. Turing, *On computable numbers with reference to the Entscheidungsproblem*, Journal of the London Mathematical Society, **II, 2. 42**, 230-265 (1937).
31. J. von Neumann, *The Computer and the Brain*, Yale University Press, (1958).
32. I.E. Sutherland, *Microelectronics and computer science*, Scientific American, 210-228, September (1977).
33. W. Aspray, *John von Neumann and the Origins of Modern Computing*, M.I.T. Press, Cambridge MA, (1990, 2nd ed. 1992).
34. W. Aspray, *The history of computing within the history of information technology*, History and Technology, **11**, 7-19 (1994).
35. G.F. Luger, *Computation and Intelligence: Collected Readings*, MIT Press, (1995).
36. Z.W. Pylyshyn, *Computation and Cognition: Towards a Foundation for Cognitive Science*, MIT Press, (1986).
37. D.E. Shasha and C. Lazere, *Out of Their Minds: The Creators of Computer Science*, Copernicus, New York, (1995).
38. W. Aspray, *An annotated bibliography of secondary sources on the history of software*, Annals of the History of Computing **9**, 291-243 (1988).
39. R. Kurzweil, *The Age of Intelligent Machines*, MIT Press, (1992).
40. S.L. Garfinkel and H. Abelson, eds., *Architects of the Information Society: Thirty-Five Years of the Laboratory for Computer Sciences at MIT*, MIT Press, (1999).
41. J. Haugeland, *Artificial Intelligence: The Very Idea*, MIT Press, (1989).
42. M.A. Boden, *Artificial Intelligence in Psychology: Interdisciplinary Essays*, MIT Press, (1989).
43. J.W. Cortada, *A Bibliographic Guide to the History of Computer Applications, 1950-1990*, Greenwood Press, Westport Conn., (1996).

44. M. Campbell-Kelly and W. Aspry, *Computer: A History of the Information Machine*, Basic Books, New York, (1996).
45. B.I. Blum and K. Duncan, editors, *A History of Medical Informatics*, ACM Press, New York, (1990).
46. J.-C. Guedon, La Planete Cyber, *Internet et Cyberspace*, Gallimard, (1996).
47. S. Augarten, *Bit by Bit: An Illustrated History of Computers*, Unwin, London, (1985).
48. N. Wiener, *Cybernetics; or Control and Communication in the Animal and the Machine*, The Technology Press, John Wiley and Sons, New York, (1948).
49. W.R. Ashby, *An Introduction to Cybernetics*, Chapman and Hall, London, (1956).
50. M.A. Arbib, *A partial survey of cybernetics in eastern Europe and the Soviet Union*, Behavioral Sci., **11**, 193-216, (1966).
51. A. Rosenblueth, N. Wiener and J. Bigelow, *Behavior, purpose and teleology*, Phil. Soc. **10** (1), 18-24 (1943).
52. N. Wiener and A. Rosenblueth, *Conduction of impulses in cardiac muscle*, Arch. Inst. Cardiol. Mex., **16**, 205-265 (1946).
53. H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems. Transactions of sixth-tenth conferences*, Josiah J. Macy Jr. Foundation, New York, (1950- 1954).
54. W.S. McCulloch and W. Pitts, *A logical calculus of ideas immanent in nervous activity*, Bull. Math. Biophys., **5**, 115-133 (1943).
55. W.S. McCulloch, *An Account of the First Three Conferences on Teleological Mechanisms*, Josiah Macy Jr. Foundation, (1947).
56. G.A. Miller, *Languages and Communication*, McGraw-Hill, New York, (1951).
57. G.A. Miller, *Statistical behavioristics and sequences of responses*, Psychol. Rev. **56**, 6 (1949).
58. G. Bateson, *Bali - the value system of a steady state*, in M. Fortes, editor, *Social Structure Studies Presented to A.R. Radcliffe-Brown*, Clarendon Press, Oxford, (1949).
59. G. Bateson, *Communication, the Social Matrix of Psychiatry*, Norton, (1951).
60. G. Bateson, *Steps to an Ecology of Mind*, Chandler, San Francisco, (1972).
61. G. Bateson, *Communication et Societe*, Seuil, Paris, (1988).
62. S. Heims, *Gregory Bateson and the mathematicians: From interdisciplinary interactions to societal functions*, J. History Behavioral Sci., **13**, 141-159 (1977).
63. S. Heims, *John von Neumann and Norbert Wiener. From Mathematics to the Technology of Life and Death*, MIT Press, Cambridge MA, (1980).
64. S. Heims, *The Cybernetics Group*, MIT Press, Cambridge MA, (1991).
65. G. van de Vijver, *New Perspectives on Cybernetics (Self-Organization, Autonomy and Connectionism)*, Kluwer, Dordrecht, (1992).
66. A. Bavelas, *A mathematical model for group structures*, Appl. Anthropol. **7** (3), 16 (1948).
67. P. de Latil, *La Pensee Artificielle - Introduction a la Cybernetique*, Gallimard, Paris, (1953).
68. L.K. Frank, G.E. Hutchinson, W.K. Livingston, W.S. McCulloch and N. Wiener, *Teleological Mechanisms*, Ann. N.Y. Acad. Sci. **50**, 187- 277 (1948).

69. H. von Foerster, *Quantum theory of memory*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the sixth conferences, Josiah J. Macy Jr. Foundation, New York, (1950).
70. H. von Foerster, *Observing Systems*, Intersystems Publications, California, (1984).
71. H. von Foerster, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer, New York, (2002).
72. M. Newborn, *Kasparov vs. Deep Blue: Computer Chess Comes of age*, Springer Verlag, (1996).
73. K.M. Colby, *Artificial Paranoia: A Computer Simulation of the Paranoid Process*, Pergamon Press, New York, (1975).
74. J.Z. Young, *Discrimination and learning in the octopus*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the ninth conference, Josiah J. Macy Jr. Foundation, New York, (1953).
75. M.J. Apter and L. Wolpert, *Cybernetics and development*. I. Information theory, *J. Theor. Biol.* **8**, 244-257 (1965).
76. H. Atlan, *L'Organization Biologique et la Theorie de l'Information*, Hermann, Paris, (1972).
77. H. Atlan, *On a formal definition of organization*, *J. Theor. Biol.* **45**, 295-304 (1974).
78. H. Atlan, *Organization du vivant, information et auto-organization*, in Volume Symposium 1986 de l'Encyclopediea Universalis, pp. 355-361, Paris, (1986).
79. E.R. Kandel, Nerve cells and behavior, *Scientific American*, **223**, 57-70, July, (1970).
80. E.R. Kandel, *Small systems of neurons*, *Scientific American*, **241 no.3**, 66-76 (1979).
81. A.K. Katchalsky et al., *Dynamic patterns of brain cell assemblies*, *Neurosciences Res. Prog. Bull.*, **12 no.1**, (1974).
82. G.E. Moore, *Cramming more components onto integrated circuits*, *Electronics*, April 19, (1965).
83. P. Gelsinger, P. Gargini, G. Parker and A. Yu, *Microprocessors circa 2000*, *IEEE Spectrum*, October, (1989).
84. P. Baron, *On distributed communications networks*, *IEEE Trans. Comm. Systems*, March (1964).
85. V.G. Cerf and R.E. Khan, *A protocol for packet network intercommunication*, *Trans. Comm. Tech.* **COM-22, V5**, 627-641, May (1974).
86. L. Kleinrock, *Communication Nets: Stochastic Message Flow and Delay*, McGraw-Hill, New York, (1964).
87. L. Kleinrock, *Queueing Systems: Vol. II, Computer Applications*, Wiley, New York, (1976).
88. R. Kahn, editor, *Special Issue on Packet Communication Networks*, *Proc. IEEE*, **66**, November, (1978).
89. L.G. Roberts, *The evolution of packet switching*, *Proc. of the IEEE* **66**, 1307-13, (1978).
90. J. Abbate, *The electrical century: Inventing the web*, *Proc. IEEE* **87**, November, (1999).

91. J. Abbate, *Inventing the Internet*, MIT Press, Cambridge MA, (1999).
92. J.C. McDonald, editor, *Fundamentals of Digital Switching, 2nd Edition*, Plenum, New York, (1990).
93. B. Metcalfe, *Packet Communication, Peer-to-Peer Communication*, San Jose Calif, (1996).
94. T. Berners-Lee, *The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*, Harper San Francisco, (1999).
95. J. Clark, *Netscape Time: The Making of the Billion-Dollar Start-Up That Took On Microsoft*, St. Martin's Press, New York, (1999).
96. J. Wallace, *Overdrive: Bill Gates and the Race to Control Cyberspace*, Wiley, New York, (1997).
97. P. Cunningham and F. Froschl, *The Electronic Business Revolution*, Springer Verlag, New York, (1999).
98. J.L. McKenny, *Waves of Change: Business Evolution Through Information Technology*, Harvard Business School Press, (1995).
99. M.A. Cosumano, *Competing on Internet Time: Lessons From Netscape and Its Battle with Microsoft*, Free Press, New York, (1998).
100. F.J. Dyson, *The Sun, the Genome and the Internet: Tools of Scientific Revolutions*, Oxford University Press, (1999).
101. L. Bruno, *Fiber Optimism: Nortel, Lucent and Cisco are battling to win the high-stakes fiber-optics game*, Red Herring, June (2000).
102. N. Cochrane, *We're insatiable: Now it's 20 million million bytes a day*, Melbourne Age, January 15, (2001).
103. K.G. Coffman and A.N. Odlyzko, *The size and growth rate of the Internet*, First Monday, October, (1998).
104. C.A. Eldering, M.L. Sylla, and J.A. Eisenach, *Is there a Moore's law for bandwidth?*, IEEE Comm. Mag., 2-7, October, (1999).
105. G. Gilder, *Fiber keeps its promise: Get ready, bandwidth will triple each year for the next 25 years*, Forbes, April 7, (1997).
106. A.M. Noll, *Does data traffic exceed voice traffic?*, Comm. ACM, 121- 124, June, (1999).
107. B. St. Arnaud, J. Coulter, J. Fitchett, and S. Mokbel, *Architectural and engineering issues for building an optical Internet*, Proc. Soc. Optical Eng. (1998).
108. M. Weisner, *The computer for the 21st century*, Scientific American, September, (1991).
109. R. Wright, *Three Scientists and Their Gods*, Time Books, (1988).
110. S. Nora and A. Mine, *The Computerization of Society*, MIT Press, (1981).
111. T. Forester, *Computers in the Human Context: Information Theory, Productivity, and People*, MIT Press, (1989).
112. J. Hoffmeyer, *Some semiotic aspects of the psycho-physical relation: the endo-exosemiotic boundary*, in Biosemiotics. The Semiotic Web, T.A. Sebeok and J. Umiker-Sebeok, editors, Mouton de Gruyter, Berlin/New York, (1991).

113. J. Hoffmeyer, *The swarming cyberspace of the body*, Cybernetics and Human Knowing, **3(1)**, 1-10 (1995).
114. J. Hoffmeyer, *Signs of Meaning in the Universe*, Indiana University Press, Bloomington IN, (1996).
115. J. Hoffmeyer, *Biosemiotics: Towards a new synthesis in biology*, European J. Semiotic Stud. **9(2)**, 355-376 (1997).
116. J. Hoffmeyer and C. Emmeche, *Code-duality and the semiotics of nature*, in On Semiotic Modeling, M. Anderson and F. Merrell, editors, Mouton de Gruyter, New York, (1991).
117. C. Emmeche and J. Hoffmeyer, *From language to nature - The semiotic metaphor in biology*, Semiotica, **84**, 1-42 (1991).
118. C. Emmeche, *The biosemiotics of emergent properties in a pluralist ontology*, in Semiosis, Evolution, Energy: Towards a Reconceptualization of the Sign, E. Taborsky, editor, Shaker Verlag, Aachen, (1999).
119. S. Brier, *Information and consciousness: A critique of the mechanistic concept of information*, in Cybernetics and Human Knowing, **1(2/3)**, 71-94 (1992).
120. S. Brier, *Ciber-Semiotics: Second-order cybernetics and the semiotics of C.S. Peirce*, Proceedings from the Second European Congress on System Science, Prague, October 5-8, 1993, AFCET, (1993).
121. S. Brier, *A cybernetic and semiotic view on a Galilean theory of psychology*, Cybernetics and Human Knowing, **2 (2)**, 31-46 (1993).
122. S. Brier, *Cybersemiotics: A suggestion for a transdisciplinary framework for description of observing, anticipatory, and meaning producing systems*, in D.M. Dubois, editor, Computing Anticipatory Systems, CASYS - First International Conference, Liege, Belgium 1997, AIP Conference Proceedings no. 437, (1997).
123. S. Oyama, *The Ontogeny of Information*, Cambridge University Press, (1985).
124. J. Hoffmeyer, *The swarming cyberspace of the body*, Cybernetics and Human Knowing, **3(1)**, 1-10 (1995).
125. J.L. Casti and A. Karlqvist, editors, *Complexity, Language, and Life: Mathematical Approaches*, Springer, Berlin, (1985).
126. H. Maturana and F. Varela, *Autopoiesis and Cognition: The Realization of the Living*, Reidel, London, (1980).
127. J. Mingers, *Self-Producing Systems: Implications and Application of Autopoiesis*, Plenum Press, New York, (1995).
128. J. Buchler, editor, *Philosophical Writings of Peirce: Selected and Edited with an Introduction by Justus Buchler*, Dover Publications, New York, (1955).
129. T.L. Short, *Peirce's semiotic theory of the self*, Semiotica, **91 (1/2)**, 109-131 (1992).
130. J. von Uexküll, *Umwelt und Innenwelt der Tiere. 2. verm. und verb. Aufl.*, Springer, Berlin, (1921).
131. J. von Uexküll, *The theory of meaning*, Semiotica, **42(1)**, 25-87 (1982 [1940]).
132. T. von Uexküll, *Introduction: Meaning and science in Jacob von Uexküll's concept of biology*, Semiotica, **42**, 1-24 (1982).
133. T. von Uexküll, *Medicine and semiotics*, Semiotica, **61** , 201-217 (1986).

134. G. Bateson, *Form, substance, and difference. Nineteenth Annual Korzybski Memorial Lecture*, (1970). Reprinted in G. Bateson, *Steps to an Ecology of Mind*, Balentine Books, New York, (1972), pp. 448-464.
135. G. Bateson, *Mind and Nature: A Necessary Unity*, Bantam Books, New York, (1980).
136. G. Bateson, *Sacred Unity: Further Steps to an Ecology of Mind*, Harper Collins, New York, (1991).
137. J. Ruesch and G. Bateson, *Communication*, Norton, New York, (1987).
138. E.F. Yates, *Semiotics as a bridge between information (biology) and dynamics (physics)*, *Recherches Semiotiques/Semiotic Inquiry* **5**, 347- 360 (1985).
139. T.A. Sebeok, *Communication in animals and men*, *Language*, **39**, 448-466 (1963).
140. T.A. Sebeok, *The Sign and its Masters*, University of Texas Press, (1979).
141. P. Bouissac, *Ecology of semiotic space: Competition, exploitation, and the evolution of arbitrary signs*, *Am. J. Semiotics*, **10**, 145-166 (1972).
142. F. Varla, *Autopoiesis: A Theory of Living Organization*, North Holland, New York, (1986).
143. R. Posner, K. Robins and T.A. Sebeok, editors, *Semiotics: A Handbook of the Sign-Theoretic Foundations of Nature and Culture*, Walter de Gruyter, Berlin, (1992).
144. R. Paton, *The ecologies of hereditary information*, *Cybernetics and Human Knowing*, **5(4)**, 31-44 (1998).
145. T. Stonier, *Information and the Internal Structure of the Universe*, Springer, Berlin, (1990).
146. T. Stonier, *Information and Meaning: An Evolutionary Perspective*, Springer, Berlin, (1997).
147. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
148. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
149. Alan Turing, *The Enigma of Intelligence*, Burnett, London (1983).
150. R. Randell (editor), *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
151. Allan R. Mackintosh, *The First Electronic Computer*, *Physics Today*, March, (1987).
152. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
153. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
154. R. Randell (editor), *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
155. Allan R. Mackintosh, *The First Electronic Computer*, *Physics Today*, March, (1987).
156. H. Babbage, *Babbages Calculating Engines: A Collection of Papers by Henry Prevost Babbage*, MIT Press, (1984).
157. A.M. Turing, *The Enigma of Intelligence*, Burnett, London (1983).
158. Ft. Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford University Press, (1989).
159. S. Wolfram, *A New Kind of Science*, Wolfram Media, Champaign IL, (2002).

160. A.M. Turing, *On computable numbers, with an application to the Entscheidungsproblem*, Proc. Lond. Math. Soc. Ser 2, **42**, (1937). Reprinted in M. David Ed., *The Undecidable*, Raven Press, Hewlett N.Y., (1965).
161. N. Metropolis, J. Howlett, and Gian-Carlo Rota (editors), *A History of Computing in the Twentieth Century*, Academic Press (1980).
162. J. Shurkin, *Engines of the Mind: A History of Computers*, W.W. Norton, (1984).
163. J. Palfreman and D. Swade, *The Dream Machine: Exploring the Computer Age*, BBC Press (UK), (1991).
164. T.J. Watson, Jr. and P. Petre, *Father, Son, and Co.*, Bantam Books, New York, (1991).
165. A. Hodges, Alan Turing: *The Enigma*, Simon and Schuster, (1983).
166. H.H. Goldstein, *The Computer from Pascal to Von Neumann*, Princeton University Press, (1972).
167. C.J. Bashe, L.R. Johnson, J.H. Palmer, and E.W. Pugh, *IBM's Early Computers*, Vol. 3 in the History of Computing Series, MIT Press, (1986).
168. K.D. Fishman, *The Computer Establishment*, McGraw-Hill, (1982).
169. S. Levy, *Hackers*, Doubleday, (1984).
170. S. Franklin, *Artificial Minds*, MIT Press, (1997).
171. P. Freiburger and M. Swaine, *Fire in the Valley: The Making of the Personal Computer*, Osborne/McGraw-Hill, (1984).
172. R.X. Cringely, *Accidental Empires*, Addison-Wesley, (1992).
173. R. Randell editor, *The Origins of Digital Computers, Selected Papers*, Springer-Verlag, New York (1973).
174. H. Lukoff, *From Dits to Bits*, Robotics Press, (1979).
175. D.E. Lundstrom, *A Few Good Men from Univac*, MIT Press, (1987).
176. D. Rutland, *Why Computers Are Computers (The SWAC and the PC)*, Wren Publishers, (1995).
177. P.E. Ceruzzi, *Reckoners: The Prehistory of the Digital Computer, from Relays to the Stored Program Concept, 1935-1945*, Greenwood Press, Westport, (1983)
178. S.G. Nash, *A History of Scientific Computing*, Addison-Wesley, Reading Mass., (1990).
179. P.E. Ceruzzi, *Crossing the divide: Architectural issues and the emergence of stored programme computers, 1935-1953*, IEEE Annals of the History of Computing, **19**, 5-12, January-March (1997).
180. P.E. Ceruzzi, *A History of Modern Computing*, MIT Press, Cambridge MA, (1998).
181. K. Zuse, *Some remarks on the history of computing in Germany*, in *A History of Computing in the 20th Century*, N. Metropolis et al. editors, 611-627, Academic Press, New York, (1980).
182. A.R. Mackintosh, *The First Electronic Computer*, Physics Today, March, (1987).
183. S.H. Hollingdale and G.C. Tootil, *Electronic Computers*, Penguin Books Ltd. (1970).
184. A. Hodges, *Alan Turing: The Enigma*, Simon and Schuster, New York, (1983).
185. A. Turing, *On computable numbers with reference to the Entscheidungsproblem*, Journal of the London Mathematical Society, **II**, **2**, **42**, 230-265 (1937).
186. J. von Neumann, *The Computer and the Brain*, Yale University Press, (1958).

187. I.E. Sutherland, *Microelectronics and computer science*, Scientific American, 210-228, September (1977).
188. W. Aspray, *John von Neumann and the Origins of Modern Computing*, M.I.T. Press, Cambridge MA, (1990, 2nd ed. 1992).
189. W. Aspray, *The history of computing within the history of information technology*, History and Technology, **11**, 7-19 (1994).
190. G.F. Luger, *Computation and Intelligence: Collected Readings*, MIT Press, (1995).
191. Z.W. Pylyshyn, *Computation and Cognition: Towards a Foundation for Cognitive Science*, MIT Press, (1986).
192. D.E. Shasha and C. Lazere, *Out of Their Minds: The Creators of Computer Science*, Copernicus, New York, (1995).
193. W. Aspray, *An annotated bibliography of secondary sources on the history of software*, Annals of the History of Computing **9**, 291-243 (1988).
194. R. Kurzweil, *The Age of Intelligent Machines*, MIT Press, (1992).
195. S.L. Garfinkel and H. Abelson, eds., *Architects of the Information Society: Thirty-Five Years of the Laboratory for Computer Sciences at MIT*, MIT Press, (1999).
196. J. Haugeland, *Artificial Intelligence: The Very Idea*, MIT Press, (1989).
197. M.A. Boden, *Artificial Intelligence in Psychology: Interdisciplinary Essays*, MIT Press, (1989).
198. J.W. Cortada, *A Bibliographic Guide to the History of Computer Applications, 1950-1990*, Greenwood Press, Westport Conn., (1996).
199. M. Campbell-Kelly and W. Aspry, *Computer: A History of the Information Machine*, Basic Books, New York, (1996).
200. B.I. Blum and K. Duncan, editors, *A History of Medical Informatics*, ACM Press, New York, (1990).
201. J.-C. Guedon, La Planete Cyber, *Internet et Cyberspace*, Gallimard, (1996).
202. S. Augarten, *Bit by Bit: An Illustrated History of Computers*, Unwin, London, (1985).
203. N. Wiener, *Cybernetics; or Control and Communication in the Animal and the Machine*, The Technology Press, John Wiley and Sons, New York, (1948).
204. W.R. Ashby, *An Introduction to Cybernetics*, Chapman and Hall, London, (1956).
205. M.A. Arbib, *A partial survey of cybernetics in eastern Europe and the Soviet Union*, Behavioral Sci., **11**, 193-216, (1966).
206. A. Rosenblueth, N. Weiner and J. Bigelow, *Behavior, purpose and teleology*, Phil. Soc. **10** (1), 18-24 (1943).
207. N. Weiner and A. Rosenblueth, *Conduction of impulses in cardiac muscle*, Arch. Inst. Cardiol. Mex., **16**, 205-265 (1946).
208. H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems. Transactions of sixth-tenth conferences*, Josiah J. Macy Jr. Foundation, New York, (1950- 1954).
209. W.S. McCulloch and W. Pitts, *A logical calculus of ideas immanent in nervous activity*, Bull. Math. Biophys., **5**, 115-133 (1943).
210. W.S. McCulloch, *An Account of the First Three Conferences on Teleological Mechanisms*, Josiah Macy Jr. Foundation, (1947).

211. G.A. Miller, *Languages and Communication*, McGraw-Hill, New York, (1951).
212. G.A. Miller, *Statistical behavioristics and sequences of responses*, Psychol. Rev. **56**, 6 (1949).
213. G. Bateson, *Bali - the value system of a steady state*, in M. Fortes, editor, *Social Structure Studies Presented to A.R. Radcliffe-Brown*, Clarendon Press, Oxford, (1949).
214. G. Bateson, *Communication, the Social Matrix of Psychiatry*, Norton, (1951).
215. G. Bateson, *Steps to an Ecology of Mind*, Chandler, San Francisco, (1972).
216. G. Bateson, *Communication et Societe*, Seuil, Paris, (1988).
217. S. Heims, *Gregory Bateson and the mathematicians: From interdisciplinary interactions to societal functions*, J. History Behavioral Sci., **13**, 141-159 (1977).
218. S. Heims, *John von Neumann and Norbert Wiener. From Mathematics to the Technology of Life and Death*, MIT Press, Cambridge MA, (1980).
219. S. Heims, *The Cybernetics Group*, MIT Press, Cambridge MA, (1991).
220. G. van de Vijver, *New Perspectives on Cybernetics (Self-Organization, Autonomy and Connectionism)*, Kluwer, Dordrecht, (1992).
221. A. Bavelas, *A mathematical model for group structures*, Appl. Anthropol. **7 (3)**, 16 (1948).
222. P. de Latil, *La Pensee Artificielle - Introduction a la Cybernetique*, Gallimard, Paris, (1953).
223. L.K. Frank, G.E. Hutchinson, W.K. Livingston, W.S. McCulloch and N. Wiener, *Teleological Mechanisms*, Ann. N.Y. Acad. Sci. **50**, 187- 277 (1948).
224. H. von Foerster, *Quantum theory of memory*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the sixth conferences, Josiah J. Macy Jr. Foundation, New York, (1950).
225. H. von Foerster, *Observing Systems*, Intersystems Publications, California, (1984).
226. H. von Foerster, *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer, New York, (2002).
227. M. Newborn, *Kasparov vs. Deep Blue: Computer Chess Comes of age*, Springer Verlag, (1996).
228. K.M. Colby, *Artificial Paranoia: A Computer Simulation of the Paranoid Process*, Pergamon Press, New York, (1975).
229. J.Z. Young, *Discrimination and learning in the octopus*, in H. von Foerster, editor, *Cybernetics - circular, causal and feed-back mechanisms in biological and social systems*. Transactions of the ninth conference, Josiah J. Macy Jr. Foundation, New York, (1953).
230. M.J. Apter and L. Wolpert, *Cybernetics and development*. I. Information theory, J. Theor. Biol. **8**, 244-257 (1965).
231. H. Atlan, *L'Organisation Biologique et la Theorie de l'Information*, Hermann, Paris, (1972).
232. H. Atlan, *On a formal definition of organization*, J. Theor. Biol. **45**, 295-304 (1974).
233. H. Atlan, *Organization du vivant, information et auto-organization*, in Volume Symposium 1986 de l'Encyclopediea Universalis, pp. 355-361, Paris, (1986).
234. E.R. Kandel, *Nerve cells and behavior*, Scientific American, **223**, 57-70, July, (1970).

235. E.R. Kandel, *Small systems of neurons*, Scientific American, **241 no.3**, 66-76 (1979).
236. A.K. Katchalsky et al., *Dynamic patterns of brain cell assemblies*, Neurosciences Res. Prog. Bull., **12 no.1**, (1974).
237. G.E. Moore, *Cramming more components onto integrated circuits*, Electronics, April 19, (1965).
238. P. Gelsinger, P. Gargini, G. Parker and A. Yu, *Microprocessors circa 2000*, IEEE Spectrum, October, (1989).
239. P. Baron, *On distributed communications networks*, IEEE Trans. Comm. Systems, March (1964).
240. V.G. Cerf and R.E. Khan, *A protocol for packet network intercommunication*, Trans. Comm. Tech. **COM-22, V5**, 627-641, May (1974).
241. L. Kleinrock, *Communication Nets: Stochastic Message Flow and Delay*, McGraw-Hill, New York, (1964).
242. L. Kleinrock, *Queueing Systems: Vol. II, Computer Applications*, Wiley, New York, (1976).
243. R. Kahn, editor, *Special Issue on Packet Communication Networks*, Proc. IEEE, **66**, November, (1978).
244. L.G. Roberts, *The evolution of packet switching*, Proc. of the IEEE **66**, 1307-13, (1978).
245. J. Abbate, *The electrical century: Inventing the web*, Proc. IEEE **87**, November, (1999).
246. J. Abbate, *Inventing the Internet*, MIT Press, Cambridge MA, (1999).
247. J.C. McDonald, editor, *Fundamentals of Digital Switching, 2nd Edition*, Plenum, New York, (1990).
248. B. Metcalfe, *Packet Communication, Peer-to-Peer Communication*, San Jose Calif, (1996).
249. T. Berners-Lee, *The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*, Harper San Francisco, (1999).
250. J. Clark, *Netscape Time: The Making of the Billion-Dollar Start-Up That Took On Microsoft*, St. Martin's Press, New York, (1999).
251. J. Wallace, *Overdrive: Bill Gates and the Race to Control Cyberspace*, Wiley, New York, (1997).
252. P. Cunningham and F. Froschl, *The Electronic Business Revolution*, Springer Verlag, New York, (1999).
253. J.L. McKenny, *Waves of Change: Business Evolution Through Information Technology*, Harvard Business School Press, (1995).
254. M.A. Cosumano, *Competing on Internet Time: Lessons From Netscape and Its Battle with Microsoft*, Free Press, New York, (1998).
255. F.J. Dyson, *The Sun, the Genome and the Internet: Tools of Scientific Revolutions*, Oxford University Press, (1999).
256. L. Bruno, *Fiber Optimism: Nortel, Lucent and Cisco are battling to win the high-stakes fiber-optics game*, Red Herring, June (2000).

257. N. Cochrane, *We're insatiable: Now it's 20 million million bytes a day*, Melbourne Age, January 15, (2001).
258. K.G. Coffman and A.N. Odlyzko, The size and growth rate of the Internet, First Monday, October, (1998).
259. C.A. Eldering, M.L. Sylla, and J.A. Eisenach, *Is there a Moore's law for bandwidth?*, IEEE Comm. Mag., 2-7, October, (1999).
260. G. Gilder, *Fiber keeps its promise: Get ready, bandwidth will triple each year for the next 25 years*, Forbes, April 7, (1997).
261. A.M. Noll, *Does data traffic exceed voice traffic?*, Comm. ACM, 121- 124, June, (1999).
262. B. St. Arnaud, J. Coulter, J. Fitchett, and S. Mokbel, *Architectural and engineering issues for building an optical Internet*, Proc. Soc. Optical Eng. (1998).
263. M. Weisner, *The computer for the 21st century*, Scientific American, September, (1991).
264. R. Wright, *Three Scientists and Their Gods*, Time Books, (1988).
265. S. Nora and A. Mine, *The Computerization of Society*, MIT Press, (1981).
266. T. Forester, *Computers in the Human Context: Information Theory, Productivity, and People*, MIT Press, (1989).

Appendix B

GROUP THEORY

B.1 Finite Groups

The definition of a finite group

A finite group is defined by the following conditions:

1. If any two elements belonging to the group are multiplied together, the product is another element belonging to the group.
2. There is an identity element.
3. Each element has an inverse.
4. Multiplication of the elements is associative¹, but necessarily commutative.
5. The group contains g elements, where g is a finite positive integer called the *order* of the group.

As a simple example, we might think of a molecule which is symmetric with respect to rotations through an angle of $2\pi/3$ about some axis but which has no other symmetry. Then the set of geometrical operations that leave the molecule invariant form a group containing 3 elements: the identity element; a rotation through an angle $2\pi/3$ about the axis of symmetry, and a rotation through an angle $4\pi/3$ about the same axis. Let us denote these operations respectively by E , C_3 , and C_3^{-1} . We can easily construct a multiplication table for the group. If we do so, *each element of the group will appear once and only once in any row or column of the multiplication table*. This follows from the fact that $AX = B$ has one and only one solution among the group elements. Since A^{-1} and B belong to the group, and since the product of any two elements belongs to the group, $X = A^{-1}B$ is also a uniquely-defined element. Now suppose that the element B appears more than once in the A th row of the multiplication table. Then $AX = B$ will have more than one

¹ $A(BC)=(AB)C$

solution which is impossible. Since no element can appear more than once, each element must appear once because there are g elements and g places in the row, all of which have to be filled.

B.2 Representations of geometrical symmetry groups

The elements of a geometrical symmetry group are linear coordinate transformations. Such transformations have the form

$$X^i = \sum_{j=1}^d \frac{\partial X^i}{\partial x^j} x^j + b^i \quad (\text{B.1})$$

where $\partial X^i/\partial x^j$ and b^i are constants.

Now consider a set of functions $\Phi_1, \Phi_2, \dots, \Phi_M$. We can use equation (B.1) to express $\Phi_1(\mathbf{x})$ as a function of \mathbf{X} . If we then expand the resulting function of \mathbf{X} in terms of the other $|\Phi_n\rangle$'s, we shall obtain a relation of the form

$$\Phi_n(\mathbf{x}) = \sum_{n'} \Phi_{n'}(\mathbf{X}) D_{n',n} \quad (\text{B.2})$$

If we denote the coordinate transformation in equation (B.1) by the symbol G , we can rewrite equations (B.1) and (B.2) in the form:

$$\begin{aligned} \mathbf{X} &= G_j \mathbf{x} \\ \Phi_n(\mathbf{x}) &\equiv \Phi_n(G_j^{-1} \mathbf{X}) \equiv G_j \Phi_n(\mathbf{X}) \\ &= \sum_{n'} \Phi_{n'}(\mathbf{X}) D_{n',n}(G) \end{aligned} \quad (\text{B.3})$$

In this sense, the coordinate transformation defines an operator G_j , and $D_{n',n}(G_j)$ is a matrix representing G_j . It can easily be shown that the matrices representing a set of operators G_1, G_2, \dots, G_g in a given basis, obey the same multiplication table as the operators themselves. For example, if we know that

$$C_3 C_3^{-1} = E \quad (\text{B.4})$$

and that

$$\begin{aligned} C_3 \Phi_n &= \sum_{n'} \Phi_{n'} D_{n',n}(C_3) \\ C_3^{-1} \Phi_n &= \sum_{n'} \Phi_{n'} D_{n',n}(C_3^{-1}) \\ E \Phi_n &= \sum_{n'} \Phi_{n'} D_{n',n}(E) \end{aligned} \quad (\text{B.5})$$

then it follows that:

$$\begin{aligned}
C_3 C_3^{-1} \Phi_n &= \sum_{n'} C_3 \Phi_{n'} D_{n',n}(C_3^{-1}) \\
&= \sum_{n''} \Phi_{n''} \left\{ \sum_{n'} D_{n'',n'}(C_3) D_{n',n}(C_3^{-1}) \right\} \\
&= E \Phi_n = \sum_{n''} \Phi_{n''} D_{n'',n}(E)
\end{aligned} \tag{B.6}$$

so that we must have

$$D_{n'',n}(E) = \sum_{n'} D_{n'',n'}(C_3) D_{n',n}(C_3^{-1}) \tag{B.7}$$

Thus given any set of basis functions $\Phi_1, \Phi_2, \dots, \Phi_M$ which mix together under the elements of a group G_1, G_2, \dots, G_g , we can obtain a set of matrices $D_{n',n}(G_j)$ defined by the relationships

$$G_j \Phi_n = \sum_{n'} \Phi_{n'} D_{n',n}(G_j) \quad j = 1, 2, \dots, g \tag{B.8}$$

These matrices will obey the same multiplication table as the operators G_1, G_2, \dots, G_g , and they are said to form a matrix representation of the group.

B.3 Similarity transformations

Now let us consider another representation, $D'_{m',m}(G_j)$, based on a set of functions $\Phi'_1, \Phi'_2, \dots, \Phi'_M$ which are related to our original set $\Phi_1, \Phi_2, \dots, \Phi_M$ by the transformation:

$$\begin{aligned}
\Phi'_m &= \sum_n \Phi_n S_{n,m} \\
\Phi_n &= \sum_m \Phi'_m S_{m,n}^{-1}
\end{aligned} \tag{B.9}$$

The primed representation is defined by the relationship

$$G_j \Phi'_m = \sum_{m'} \Phi'_{m'} D'_{m',m}(G_j) \quad j = 1, 2, \dots, g \tag{B.10}$$

Then from equations (B.8)-(B.10) we have

$$\begin{aligned}
G_j \Phi'_m &= \sum_{m'} \Phi'_{m'} D'_{m',m}(G_j) \\
&= G_j \sum_n \Phi_n S_{n,m} \\
&= \sum_{n,n'} \Phi_{n'} D_{n',n}(G_j) S_{n,m} \\
&= \sum_{m',n,n'} \Phi'_{m'} S_{m',n'}^{-1} D_{n',n}(G_j) S_{n,m}
\end{aligned} \tag{B.11}$$

so that we must have

$$D'_{m',m}(G_j) = \sum_{n,n'} S_{m',n'}^{-1} D_{n',n}(G_j) S_{n,m} \quad (\text{B.12})$$

or

$$D' = S^{-1} D S \quad (\text{B.13})$$

A transformation of this type, where the matrix S need not be unitary, is called a ‘similarity transformation’.

B.4 Characters and reducibility

The character $\chi(G_j)$ of the matrix $D_{n',n}(G_j)$ is defined as the sum of the diagonal elements:

$$\chi(G_j) \equiv \sum_n D_{n,n}(G_j) \quad (\text{B.14})$$

We would like to show that the character of each element in a representation of a finite group is invariant under a similarity transformation. From equations (B.12) and (B.14) we have:

$$\begin{aligned} \chi'(G_j) &\equiv \sum_m D'_{m,m}(G_j) \\ &= \sum_{m,n,n'} S_{m',n'}^{-1} D_{n',n}(G_j) S_{n,m} \\ &= \sum_{n,n'} \left(\sum_m S_{n,m} S_{m',n'}^{-1} \right) D_{n',n}(G_j) \\ &= \sum_{n,n'} \delta_{n',n} D_{n',n}(G_j) \\ &= \sum_n D_{n,n}(G_j) = \chi(G_j) \quad \text{q.e.d.} \end{aligned} \quad (\text{B.15})$$

If two representations are connected by a similarity transformation, then they are said to be ‘equivalent’. From (B.15) it follows that when two representations are equivalent, then $\chi'(G_j) = \chi(G_j)$ for $j = 1, 2, \dots, g$.

Sometimes it is possible by means of a similarity transformation to bring all the elements of a representation into a block-diagonal form. In other words it may be possible to bring $D'_{m',m}(G_j)$ into a form where the non-zero elements are confined blocks along the diagonal, the blocks being the same for all the group elements. To express the same idea differently, it is sometimes possible to go over by means of a similarity transformation from the original basis set, $\Phi_1, \Phi_2, \dots, \Phi_M$ to a new basis set $\Phi'_1, \Phi'_2, \dots, \Phi'_M$ which can be divided into two

Table A.1 Multiplication table for the group C_3

	E	C_3	C_3^{-1}
E	E	C_3	C_3^{-1}
C_3	C_3	C_3^{-1}	E
C_3^{-1}	C_3^{-1}	E	C_3

or more subsets, each of which mixes only with itself under the operations G_1, G_2, \dots, G_g . A representation based on two or more subsets of basis functions which mix only with themselves under the operations of the group is said to be 'reduced'. Whenever it is possible to bring a representation into a reduced form by means of a similarity transformation, it is said to be 'reducible'. Whenever this is not possible, the representation is said to be 'irreducible'.

Table A.2 Character table for the group C_3

	E	C_3	C_3^{-1}
A	1	1	1
Γ_c	1	$e^{i(2\pi/3)}$	$e^{-i(2\pi/3)}$
Γ_c^*	1	$e^{-i(2\pi/3)}$	$e^{i(2\pi/3)}$

B.5 The great orthogonality theorem

A unitary matrix is a matrix whose conjugate transpose (Hermitian adjoint) is equal to its inverse. It is always possible, by means of a similarity transformation, to bring the matrix representations of a finite group into unitary form. Now let $D_{n',n}^\alpha(G_j)$ and $D_{m',m}^\beta(G_j)$ be two unitary irreducible representations of a finite group of order g . The great orthogonality theorem, from which much of the power of group theory is derived, then states that

$$\sum_{j=1}^g D_{n',n}^{\alpha*}(G_j) D_{m',m}^\beta(G_j) = \frac{g}{d_\alpha} \delta_{\alpha,\beta} \delta_{n',m'} \delta_{n,m} \quad (\text{B.16})$$

where d_α is the dimension of the matrices $D_{n',n}^\alpha(G_j)$. The proof of the great orthogonality theorem depends on Schur's lemma, which states that if A is a matrix that commutes with every matrix $D_{n',n}^\alpha(G_j)$, $j = 1, 2, \dots, g$ in a unitary irreducible representation of a finite group, then A must be a multiple of the unit matrix, i.e., if

$$AD(G_j) - D(G_j)A = 0, \quad j = 1, 2, \dots, g \quad (\text{B.17})$$

then

$$A \sim I \quad (\text{B.18})$$

The proof of Schur's lemma is as follows: If A commutes with $D_{n',n}^\alpha(G_j)$, $j = 1, 2, \dots, g$, then so does its conjugate transpose A^\dagger . Therefore we can let A be Hermitian without loss of generality, and we can diagonalize A by means of a unitary transformation:

$$UAU^{-1} = A^{(d)} \quad (\text{B.19})$$

where $A^{(d)}$ is diagonal. Then

$$U^{-1}A^{(d)}UD(G_j) - D(G_j)U^{-1}A^{(d)}U = 0, \quad j = 1, 2, \dots, g \quad (\text{B.20})$$

Multiplying on the left by U and on the right by U^{-1} then yields

$$A^{(d)}UD(G_j)U^{-1} - UD(G_j)U^{-1}A^{(d)} = 0, \quad j = 1, 2, \dots, g \quad (\text{B.21})$$

Thus we can write

$$A^{(d)}D'(G_j) - D'(G_j)A^{(d)} = 0, \quad j = 1, 2, \dots, g \quad (\text{B.22})$$

where

$$D'(G_j) \equiv UD(G_j)U^{-1} \quad (\text{B.23})$$

Since $A^{(d)}$ is diagonal we can write $A_{n',n}^{(d)} = A_n^{(d)}\delta_{n',n}$. Thus with the indices written out, (B.22) becomes:

$$\sum_{n'} \left(A_{n'}^{(d)}\delta_{n'',n'} D_{n',n}^{\prime\alpha}(G_j) - D_{n'',n'}^{\prime\alpha}(G_j)A_n^{(d)}\delta_{n',n} \right) = 0, \quad j = 1, \dots, g \quad (\text{B.24})$$

from which it follows that

$$\left(A_{n'}^{(d)} - A_n^{(d)} \right) D_{n',n}^{\prime\alpha}(G_j) = 0, \quad j = 1, 2, \dots, g \quad (\text{B.25})$$

Without loss of generality, we can choose U in such a way that repeated eigenvalues of $A^{(d)}$ are grouped together along the diagonal. Then $A_{n''}^{(d)} \neq A_n^{(d)}$ implies that

$$D_{n',n}^{\prime\alpha}(G_j) = D_{n'',n}^{\prime\alpha}(G_j) = D_{n,n'}^{\alpha}(G_j^{-1}) = 0, \quad j = 1, 2, \dots, g \quad (\text{B.26})$$

Thus $D_{n',n}^{\prime\alpha}(G_j)$ can only have non-zero elements in the blocks that correspond to repeated eigenvalues of $A^{(d)}$ and it would therefore be reducible unless all of the eigenvalues are equal, which would contradict the original assumption of irreducibility. This proves Schur's lemma.

Having demonstrated the validity of Schur's lemma, we are now in a position to prove the great orthogonality relation. To do so we define the matrix M by the relationship

$$M \equiv \sum_{j=1}^g D^\alpha(G_j)X D^\beta(G_j^{-1}) \quad (\text{B.27})$$

where X is an arbitrary matrix of appropriate dimensions to make matrix multiplication possible and where $D^\alpha(G_j)$ and $D^\beta(G_j)$ are unitary irreducible representations of the finite group. Then

$$\begin{aligned} D^\alpha(G_i)MD^\beta(G_i^{-1}) &= \sum_{j=1}^g D^\alpha(G_i)D^\alpha(G_j)XD^\beta(G_j^{-1})D^\beta(G_i^{-1}) \\ &= \sum_{k=1}^g D^\alpha(G_k)XD^\beta(G_k^{-1}) = M \end{aligned} \quad (\text{B.28})$$

where $G_iG_j = G_k$ and where we have used the fact that each group element appears once and only once in every row of the multiplication table to replace the sum over j by a sum over k . Multiplying (B.28) from the right by $D^\beta(G_i)$ we obtain:

$$D^\alpha(G_i)M = MD^\beta(G_i) \quad i = 1, 2, \dots, g \quad (\text{B.29})$$

Then, according to Schur's lemma, M must be a multiple of the unit matrix. It may of course be a square matrix consisting entirely of zeros, since such a matrix is also a multiple of the unit matrix. Multiplying (B.29) from the left by M^{-1} we obtain:

$$M^{-1}D^\alpha(G_i)M = D^\beta(G_i) \quad i = 1, 2, \dots, g \quad (\text{B.30})$$

from which we can see that if M is not the null matrix, then the irreducible representations $D^\alpha(G_i)$ and $D^\beta(G_i)$ must be the same, i.e., if M is not the null matrix, $\alpha = \beta$.

Let us first consider the case where M is the null matrix and where $\alpha \neq \beta$. Then putting indices into (B.27) we have:

$$\sum_{j=1}^g \sum_{n=1}^{d_\alpha} \sum_{m'=1}^{d_\beta} D_{n',n}^\alpha(G_j)X_{n,m'}D_{m',m}^\beta(G_j^{-1}) = 0 \quad (\text{B.31})$$

But $X_{n,m'}$ is arbitrary, and therefore (B.31) can only hold for all cases if

$$\sum_{j=1}^g D_{n',n}^\alpha(G_j)D_{m',m}^\beta(G_j^{-1}) = 0 \quad (\text{B.32})$$

Now let us consider the second possibility: Suppose that $\alpha = \beta$. Then

$$\delta_{\alpha,\beta}M = \sum_{j=1}^g D^\beta(G_j)XD^\beta(G_j^{-1}) \quad (\text{B.33})$$

Putting indices into (B.33) we have

$$\delta_{\alpha,\beta}M_{n',m} = \sum_{j=1}^g \sum_{n=1}^{d_\beta} \sum_{m'=1}^{d_\beta} D_{n',n}^\alpha(G_j)X_{n,m'}D_{m',m}^\beta(G_j^{-1}) \quad (\text{B.34})$$

Taking the trace of both sides of (B.34) yields

$$\begin{aligned} \delta_{\alpha,\beta} \text{tr}(M) &= \sum_{j=1}^g \sum_{m=1}^{d_\beta} \sum_{n=1}^{d_\beta} \sum_{m'=1}^{d_\beta} D_{m,n}^\beta(G_j) X_{n,m'} D_{m',m}^\beta(G_j^{-1}) \\ &= \sum_{j=1}^g \sum_{n=1}^{d_\beta} \sum_{m'=1}^{d_\beta} \delta_{n,m'} X_{n,m'} = g \text{tr}(X) \end{aligned} \quad (\text{B.35})$$

so that

$$I \delta_{\alpha,\beta} \frac{g}{d_\alpha} \text{tr} X = \sum_{j=1}^g \sum_{n=1}^{d_\beta} \sum_{m'=1}^{d_\beta} D_{n',n}^\alpha(G_j) X_{n,m'} D_{m',m}^\beta(G_j^{-1}) \quad (\text{B.36})$$

where I is the identity matrix. Because X is arbitrary, this relationship can only hold in all cases if (B.16) is valid.

The great orthogonality relation is very central, and almost all of the results of group theory depend upon it. For example, combining (B.16) with the definition of characters (B.14), we obtain:

$$\begin{aligned} \sum_{j=1}^g \chi^{\alpha^*}(G_j) \chi^\beta(G_j) &\equiv \sum_{j=1}^g \left\{ \sum_n D_{n,n}^{\alpha^*}(G_j) \right\} \left\{ \sum_m D_{m,m}^\beta(G_j) \right\} \\ &= \frac{g}{d_\alpha} \delta_{\alpha,\beta} \sum_n \sum_m \delta_{n,m} \delta_{n,m} = g \delta_{\alpha,\beta} \end{aligned} \quad (\text{B.37})$$

Equation (B.37) holds only for unitary representations, but every representation is equivalent to a unitary representation since it is always possible to perform a similarity transformation that orthonormalizes the basis functions. Therefore, since characters are invariant under similarity transformations, the orthonormality of characters

$$\frac{1}{g} \sum_{j=1}^g \chi^{\alpha^*}(G_j) \chi^\beta(G_j) = \epsilon_{\alpha,\beta} \equiv \begin{cases} 0 & \text{if the representations are inequivalent} \\ 1 & \text{if the representations are equivalent} \end{cases} \quad (\text{B.38})$$

holds even for non-unitary irreducible representations.

Now consider a representation $D_{n',n}(G_j)$ which may be reducible. If we reduce it by means of a similarity transformation, then in its reduced form it will be block-diagonal, each block being irreducible. Taking the trace, we find that the character of an element in the reduced representation $D'_{n',n}(G_j)$ is the sum of the characters of the irreducible representations of which it is composed. Thus

$$\begin{aligned} \chi(G_j) &\equiv \sum_{n=1}^d D'_{n,n}(G_j) \\ &= \chi^1(G_j) + \chi^2(G_j) + \dots \\ &= \sum_{\beta} n_{\beta} \chi^{\beta}(G_j) \end{aligned} \quad (\text{B.39})$$

where n_β is the number of times that the irreducible representation D^β occurs among the diagonal blocks of D' . Then from (B.38) we have

$$\begin{aligned} \frac{1}{g} \sum_{j=1}^g \chi^{\alpha^*}(G_j) \chi(G_j) &= \sum_{\beta} n_{\beta} \sum_{j=1}^g \chi^{\alpha^*}(G_j) \chi^{\beta}(G_j) \\ &= \sum_{\beta} n_{\beta} \epsilon_{\alpha, \beta} = n_{\alpha} \end{aligned} \quad (\text{B.40})$$

This gives us a way to find out how many times a particular irreducible representation D^α occurs in a reducible representation D . According to (B.40), we just have to take the scalar product of the characters and divide by the order of the group. When we say that D^α ‘occurs’ n_α times in D , we mean that it is possible by means of a similarity transformation to bring D into block-diagonal form where D^α occurs n_α times along the diagonal blocks. The relationship is sometimes written in the form

$$D = n_1 D^1 + n_2 D^2 + \dots \quad (\text{B.41})$$

Obviously in this decomposition we do not need to distinguish between different equivalent forms of an irreducible representation D^α , since all of them have the same character, and it is possible to go from one to another by means of a similarity transformation.

B.6 Classes

Two elements of a group G_i and G_j are said to be in the same ‘class’ if there exists another element G_l in the group such that

$$G_i = G_l^{-1} G_j G_l \quad (\text{B.42})$$

Thus, if we start with a particular element G_j , we can generate the set of elements in the same class by keeping j fixed in (B.42) and letting G_l run through all the elements of the group. It also follows from (B.42) that we can construct an operator M_k which commutes with all the elements of the group by summing the elements of a particular class:

$$M_k \equiv \sum_{\text{class } k} G_j \quad (\text{B.43})$$

Then for an arbitrary group element G_l we have

$$\begin{aligned} G_l^{-1} [M_k, G_l] &= \sum_{\text{class } k} G_l^{-1} [G_j, G_l] \\ &= \sum_{\text{class } k} (G_l^{-1} G_j G_l - G_j) \\ &= \sum_{\text{class } k} (G_i - G_j) = 0 \end{aligned} \quad (\text{B.44})$$

Equation (B.44) can hold only if $[M_k, G_l] = 0$. An operator, such as M_k , which commutes with every element of the group is called an ‘invariant’. If there are r classes in a group, there will be r linearly independent invariants that can be constructed in this way.

For any representation of two elements G_i and G_j in the same class, it follows from (B.42) that

$$D(G_i) = D(G_i^{-1})D(G_j)D(G_l) = D(G_l)^{-1}D(G_j)D(G_l) \quad (\text{B.45})$$

Thus if $D(G_i)$ and $D(G_j)$ represent two elements in the same class, they are connected by a similarity transformation, and therefore they have the same character. In other words, *all elements in the same class have the same character*. This means that in applying equation (B.40) we do not need to go through quite so much work. Instead of summing over all of the elements in the group, we can take the product of characters for a representative element in each class, multiply by the number of elements in the class, and then sum over the classes. If g_k represents the number of elements in the class k , then the orthogonality relation for characters, equation (B.38), can be written in the form

$$\sum_{k=1}^r \sqrt{\frac{g_k}{g}} \chi_k^{\alpha*}(G_j) \sqrt{\frac{g_k}{g}} \chi_k^\beta(G_j) = \delta_{\alpha,\beta} \quad (\text{B.46})$$

where χ_k^α is the character of a representative element in class k .

B.7 Projection operators

The great orthogonality theorem, equation (B.16), can be used to construct group-theoretical projection operators. Suppose that the sets of functions $(\Phi_1^1, \Phi_2^1, \dots, \Phi_{d_1}^1)$, $(\Phi_1^2, \Phi_2^2, \dots, \Phi_{d_2}^2)$, etc. each form the basis for an irreducible representation of a group, and that there are r' nonequivalent irreducible representations. Then

$$G_j \Phi_n^\beta = \sum_{n'=1}^{d_\beta} \Phi_{n'}^\beta D_{n',n}^\beta(G_j) \quad (\text{B.47})$$

Then from (B.16) we have

$$\begin{aligned} \sum_{j=1}^g D_{m,m}^{\alpha*}(G_j) G_j \Phi_n^\beta &= \sum_{n'=1}^{d_\beta} \Phi_{n'}^\beta \sum_{j=1}^g D_{m,m}^{\alpha*}(G_j) D_{n',n}^\beta(G_j) \\ &= \delta_{\alpha,\beta} \frac{g}{d_\alpha} \sum_{n'=1}^{d_\beta} \Phi_{n'}^\beta \delta_{m,n'} \delta_{m,n} \\ &= \delta_{\alpha,\beta} \frac{g}{d_\alpha} \Phi_m^\beta \delta_{m,n} \end{aligned} \quad (\text{B.48})$$

From (B.48) it follows that if we let

$$P_m^\alpha \equiv \frac{d_\alpha}{g} \sum_{j=1}^g D_{m,m}^{\alpha*}(G_j) G_j \quad (\text{B.49})$$

then

$$P_m^\alpha \Phi_n^\beta = \delta_{\alpha,\beta} \delta_{m,n} \Phi_m^\beta \quad (\text{B.50})$$

In other words, when the operator P_m^α defined by equation (B.49) acts on any function in the set $(\Phi_1^1, \Phi_2^1, \dots, \Phi_{d_1}^1), (\Phi_1^2, \Phi_2^2, \dots, \Phi_{d_2}^2), \dots$, the function is given back unchanged, provided that $m = n$ and $\alpha = \beta$. Otherwise the function is annihilated. Thus, P_m^α is a projection operator corresponding to the m th basis function of the α th irreducible representation of the group in a standard unitary representation. If P_m^α acts on an arbitrary function, it will annihilate all of it except the component that transforms like the m th basis function of D^α .

A second type of group-theoretical projection operator can be defined by the relationship

$$P^\alpha \equiv \sum_{m=1}^{d_\alpha} P_m^\alpha = \frac{d_\alpha}{g} \sum_{j=1}^g \sum_{m=1}^{d_\alpha} D_{m,m}^{\alpha*}(G_j) G_j \quad (\text{B.51})$$

which can be rewritten as

$$P^\alpha \equiv \frac{d_\alpha}{g} \sum_{j=1}^g \chi^{\alpha*}(G_j) G_j \quad (\text{B.52})$$

From (B.50) it follows that

$$P^\alpha \Phi_n^\beta = \sum_{m=1}^{d_\alpha} P_m^\alpha \Phi_n^\beta = \delta_{\alpha,\beta} \sum_{m=1}^{d_\alpha} \delta_{m,n} \Phi_m^\beta = \delta_{\alpha,\beta} \Phi_n^\beta \quad (\text{B.53})$$

When P^α acts on an arbitrary function, it annihilates everything except the component which can be expressed as a linear combination of basis functions of the irreducible representation D^α . If we sum (B.53) over all of the irreducible representations of the group, we obtain

$$\sum_{\alpha=1}^{r'} P^\alpha \Phi_n^\beta = \sum_{\alpha=1}^{r'} \delta_{\alpha,\beta} \Phi_n^\beta = \Phi_n^\beta \quad (\text{B.54})$$

Therefore the sum acts like the identity operator and we can write

$$\sum_{\alpha=1}^{r'} P^\alpha = E \quad (\text{B.55})$$

Combining (B.55) with (B.52), we obtain

$$\sum_{j=1}^g \sum_{\alpha=1}^{r'} \frac{d_\alpha}{g} \chi^{\alpha*}(G_j) G_j = E \equiv G_1 \quad (\text{B.56})$$

Since the group elements G_1, \dots, G_g are linearly independent, equation (B.55) implies that

$$\sum_{\alpha=1}^{r'} \frac{d_\alpha}{g} \chi^{\alpha*}(G_j) = \delta_{j,1} \quad (\text{B.57})$$

The character of the identity element in any representation is equal to the dimension of that representation:

$$\chi^{\alpha*}(E) = \chi^\alpha(E) = d_\alpha \quad (\text{B.58})$$

Therefore, when $j = 1$, we obtain from (B.57) the relationship

$$\sum_{\alpha=1}^{r'} d_\alpha^2 = g \quad (\text{B.59})$$

i.e., *the sum of the squares of the dimensions of the irreducible representations is equal to the order of the group.*

B.8 The regular representation

The ‘regular representation’ of a finite group is a reducible representation D^{reg} in which the basis consists of the group elements themselves:

$$G_j G_n = \sum_{n'=1}^g G_{n'} D_{n',n}^{\text{reg}}(G_j) \quad (\text{B.60})$$

D^{reg} must thus be a set of $g \times g$ matrices. If we know the multiplication table for a finite group, we can construct the regular representation. For example, the multiplication table for the group C_3 is shown above. It can easily be verified that if we let

$$\begin{aligned} D^{\text{reg}}(E) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ D^{\text{reg}}(C_3) &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ D^{\text{reg}}(C_3^{-1}) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{aligned} \quad (\text{B.61})$$

then the matrices will be the regular representation of the group C_3 according to the definition shown in (B.60) and the multiplication table (A.1). Since $G_i G_j \neq G_j$ for $G_i \neq E$, it follows that the character of every group element except the identity element vanishes in the regular representation. (We can notice that this holds in the example given above.) Therefore in the case of the regular representation, equation (B.40) becomes:

$$n_\alpha = \frac{1}{g} \sum_{j=1}^g \chi^{\alpha*}(G_j) \chi^{\text{reg}}(G_j) = \frac{1}{g} \chi^{\alpha*}(E) \chi^{\text{reg}}(E) = d_\alpha \quad (\text{B.62})$$

Thus each irreducible representation of a finite group appears d_α times in the regular representation.

When each element of a group commutes with every other one, a group is said to be Abelian. Then from the definition of classes, (B.42), it follows that in an Abelian group, every element is in a class by itself, so that an Abelian group contains g classes, i.e. $r = g$. We can next ask how many non-equivalent irreducible representations an Abelian group contains. To answer this question, we remember from Schur's lemma that the only matrix that commutes with every matrix in an irreducible representation of a group must be a multiple of the unit matrix. But in an Abelian group, all of the elements commute with each other, and therefore their irreducible representations must all be multiples of the unit matrix. This can happen only if all the irreducible representations are 1-dimensional. Thus for an Abelian group, $d_\alpha = 1$, $\alpha = 1, 2, \dots, r'$ and $r' = g$. It can be seen from the multiplication table of the group C_3 that it is Abelian. In the example of C_3 , (B.59) becomes $1 + 1 + 1 = 3$.

B.9 Classification of basis functions

We can use the group-theoretical projection operators to classify basis sets into basis functions for the various irreducible representations of a group. For example, we can construct the projection operators of the group C_3 from the character table:

$$\begin{aligned} P^1 &= \frac{1}{3} (E + C_3 + C_3^{-1}) \\ P^2 &= \frac{1}{3} (E + e^{-i2\pi/3} C_3 + e^{i2\pi/3} C_3^{-1}) \\ P^3 &= \frac{1}{3} (E + e^{i2\pi/3} C_3 + e^{-i2\pi/3} C_3^{-1}) \end{aligned} \quad (\text{B.63})$$

Since the group C_3 is Abelian, all of its irreducible representations are 1-dimensional, and hence there is no difference between projection operators of the type P^α and those of the type P_n^α . Notice that $P^1 + P^2 + P^3 = E$ in accordance with (B.55), and that the projection operators are idempotent, i.e., $P^\alpha P^\beta = \delta_{\alpha,\beta} P^\alpha$. All projection operators must be idempotent, since projecting out a subspace of a Hilbert space twice has the same effect as doing it once, and acting in succession with projection operators corresponding to different subspaces must yield zero.

Now consider the set of functions $\Phi_m = e^{im\varphi}$ where m is an integer. We can use the projection operators of (B.63) to split the Hilbert space spanned by this set of functions into three subspaces. Using the relationships

$$\begin{aligned} Ee^{im\varphi} &= e^{im\varphi} \\ C_3e^{im\varphi} &= e^{im(\varphi-2\pi/3)} \\ C_3^{-1}e^{im\varphi} &= e^{im(\varphi+2\pi/3)} \end{aligned} \tag{B.64}$$

we obtain

$$\begin{aligned} P^1e^{im\varphi} &= \frac{1}{3}e^{im\varphi} (1 + e^{-im2\pi/3} + e^{im2\pi/3}) \\ &= \begin{cases} 0 & \text{if } m = \pm 1, \pm 2, \pm 4, \pm 5, \dots \\ e^{im\varphi} & \text{if } m = 0, \pm 3, \pm 6, \pm 9, \dots \end{cases} \end{aligned} \tag{B.65}$$

and similarly

$$\begin{aligned} P^2e^{im\varphi} &= \begin{cases} 0 & \text{if } m + 1 = \pm 1, \pm 2, \pm 4, \pm 5, \dots \\ e^{im\varphi} & \text{if } m + 1 = 0, \pm 3, \pm 6, \pm 9, \dots \end{cases} \\ P^3e^{im\varphi} &= \begin{cases} 0 & \text{if } m - 1 = \pm 1, \pm 2, \pm 4, \pm 5, \dots \\ e^{im\varphi} & \text{if } m - 1 = 0, \pm 3, \pm 6, \pm 9, \dots \end{cases} \end{aligned} \tag{B.66}$$

Thus the Hilbert space spanned by the functions $\Phi_m = e^{im\varphi}$ is divided into three subspaces each of which consists of basis functions for one of the irreducible representations of C_3 . For non-Abelian groups the Hilbert space spanned by a set of basis functions can be divided into still smaller subspaces through the use of projection operators of the type P_n^α defined in equation (B.49). If we wish to have names for the the two types of projection operators, we might call P_m^α ‘strong’ and P^α ‘weak’, since P_n^α has a stronger effect than P^α .

Now suppose that we have divided the Hilbert space spanned by a set of basis functions into small subspaces by means of the strong projection operators P_n^α , so that

$$P_n^\alpha \Phi_j = p_j \Phi_j \quad p_j = 0 \text{ or } 1 \tag{B.67}$$

We will now show that if an operator T commutes with every element of the group, then the matrix elements of T linking functions belonging to different subspaces must necessarily vanish. The proof is as follows: Since T commutes with every element of the group, and since the projection operators are constructed from group elements, we have

$$[P_n^\alpha, T] = 0 \tag{B.68}$$

Then

$$\langle \Phi_j | [P_n^\alpha, T] | \Phi_k \rangle = (p_j - p_k) \langle \Phi_j | T | \Phi_k \rangle = 0 \tag{B.69}$$

Thus if Φ_j and Φ_k belong to different subspaces when the basis set is classified by the action of the projection operators P_n^α , i.e., if $p_j \neq p_k$, then $\langle \Phi_j | T | \Phi_k \rangle = 0$. *It follows that*

a matrix representation of the operator T will be block-diagonal if it is based on functions that have been classified by means of the projection operators P_n^α , i.e. if it is based on a set of functions that satisfy (B.67). Such a basis set is said to be ‘symmetry-adapted’.

We can introduce a special notation to represent fully symmetry-adapted basis functions. Let $|\eta_j^{\alpha,n}\rangle$ be such a function.² By this we indicate that the function transforms under the action of the group elements like n th basis function of the α th standard irreducible representation of the group, while the index j distinguishes between the various linearly independent functions that have this property. With this notation we can write:

$$P_n^\alpha |\eta_j^{\beta,m}\rangle = \delta_{\alpha,\beta} \delta_{n,m} |\eta_j^{\beta,m}\rangle \quad (\text{B.70})$$

Using this notation, the statement that a matrix representation of the operator T based on symmetry-adapted functions will be block-diagonal can be written in the form:

$$\langle \eta_i^{\alpha,n} | T | \eta_j^{\beta,m} \rangle = \delta_{\alpha,\beta} \delta_{n,m} \langle \eta_i^{\alpha,n} | T | \eta_j^{\beta,m} \rangle \quad (\text{B.71})$$

The eigenvalues and eigenfunctions of T can also be expressed in this notation:

$$T | \Psi_\kappa^{\alpha,m} \rangle = \lambda_\kappa^{\alpha,m} | \Psi_\kappa^{\alpha,m} \rangle \quad (\text{B.72})$$

where

$$| \Psi_\kappa^{\alpha,m} \rangle = \sum_j |\eta_j^{\alpha,m}\rangle C_{j,\kappa} \quad (\text{B.73})$$

In other words, a set of functions all of which transform like the n th basis function of the α th irreducible representation of a group combine to form an eigenfunction of an operator T that commutes with all of the group elements.

We will now try to find a relationship between the degeneracy of the root $\lambda_\kappa^{\alpha,n}$ and the dimension d_α of the irreducible representation D^α . To do this, we introduce the ‘shift operator’

$$P_{m',m}^\alpha \equiv \frac{d_\alpha}{g} \sum_{j=1}^g D_{m',m}^{\alpha*}(G_j) G_j \quad m' \neq m \quad (\text{B.74})$$

Then by an argument similar to (B.48) we have

$$\begin{aligned} P_{m',m}^\alpha |\eta_j^{\alpha,m}\rangle &= \frac{d_\alpha}{g} \sum_{j=1}^g D_{m',m}^{\alpha*}(G_j) G_j |\eta_j^{\alpha,m}\rangle \\ &= \sum_{m''=1}^{d_\alpha} |\eta_j^{\alpha,m''}\rangle \frac{d_\alpha}{g} \sum_{j=1}^g D_{m',m}^{\alpha*}(G_j) D_{m'',m}^\alpha(G_j) \\ &= \sum_{m''=1}^{d_\alpha} |\eta_j^{\alpha,m''}\rangle \delta_{m'',m'} = |\eta_j^{\alpha,m'}\rangle \end{aligned} \quad (\text{B.75})$$

²We also introduce the Dirac notation here, since it is useful in the discussion of matrix elements.

where we have made use of the great orthogonality relation (B.16). Since $P_{m',m}^\alpha$ is a linear combination of group elements, it must commute with T :

$$[P_{m',m}^\alpha, T] = 0 \quad (\text{B.76})$$

Therefore

$$\begin{aligned} \langle \Psi_\kappa^{\alpha,m'} | [P_{m',m}^\alpha, T] | \Psi_\kappa^{\alpha,m} \rangle &= \left(\lambda_\kappa^{\alpha,m'} - \lambda_\kappa^{\alpha,m} \right) \langle \Psi_\kappa^{\alpha,m'} | P_{m',m}^\alpha | \Psi_\kappa^{\alpha,m} \rangle \\ &= \left(\lambda_\kappa^{\alpha,m'} - \lambda_\kappa^{\alpha,m} \right) = 0 \end{aligned} \quad (\text{B.77})$$

so that the roots corresponding to the d_α eigenfunctions $|\Psi_\kappa^{\alpha,1}\rangle, \dots, |\Psi_\kappa^{\alpha,d_\alpha}\rangle$ must be degenerate. Such a degeneracy is called a 'due degeneracy' because it is due to the symmetry properties of the system. If there are other degeneracies, they are termed 'accidental'.

Appendix C

NON-EUCLIDIAN GEOMETRY

C.1 Albert Einstein and non-Euclidian geometry

Albert Einstein's famous general theory of relativity has a central place in modern cosmology. We will see below how Einstein, with the help of his friend Marcel Grossman, came to realize that non-Euclidian geometry was exactly what he needed to turn his principle of the equivalence of gravitational and inertial mass into a quantitative theory.

C.2 Family background

Albert Einstein was born in Ulm, Germany, in 1879. He was the son of middle-class, irreligious Jewish parents, who sent him to a Catholic school. Einstein was slow in learning to speak, and at first his parents feared that he might be retarded; but by the time he was eight, his grandfather could say in a letter: "Dear Albert has been back in school for a week. I just love that boy, because you cannot imagine how good and intelligent he has become."

Remembering his boyhood, Einstein himself later wrote: "When I was 12, a little book dealing with Euclidean plane geometry came into my hands at the beginning of the school year. Here were assertions, as for example the intersection of the altitudes of a triangle in one point, which, though by no means self-evident, could nevertheless be proved with such certainty that any doubt appeared to be out of the question. The lucidity and certainty made an indescribable impression on me."

When Albert Einstein was in his teens, the factory owned by his father and uncle began to encounter hard times. The two Einstein families moved to Italy, leaving Albert alone and miserable in Munich, where he was supposed to finish his course at the gymnasium. Einstein's classmates had given him the nickname "Beidermeier", which means something like "Honest John"; and his tactlessness in criticizing authority soon got him into trouble. In Einstein's words, what happened next was the following: "When I was in the seventh grade at the Lutpold Gymnasium, I was summoned by my home-room teacher, who expressed the wish that I leave the school. To my remark that I had done nothing wrong, he

replied only, ‘Your mere presence spoils the respect of the class for me’.”

Einstein left gymnasium without graduating, and followed his parents to Italy, where he spent a joyous and carefree year. He also decided to change his citizenship. “The over-emphasized military mentality of the German State was alien to me, even as a boy”, Einstein wrote later. “When my father moved to Italy, he took steps, at my request, to have me released from German citizenship, because I wanted to be a Swiss citizen.”

The financial circumstances of the Einstein family were now precarious, and it was clear that Albert would have to think seriously about a practical career. In 1896, he entered the famous Zürich Polytechnic Institute with the intention of becoming a teacher of mathematics and physics. However, his undisciplined and nonconformist attitudes again got him into trouble. His mathematics professor, Hermann Minkowski (1864-1909), considered Einstein to be a “lazy dog”; and his physics professor, Heinrich Weber, who originally had gone out of his way to help Einstein, said to him in anger and exasperation: “You’re a clever fellow, but you have one fault: You won’t let anyone tell you a thing! You won’t let anyone tell you a thing!”

Einstein missed most of his classes, and read only the subjects which interested him. He was interested most of all in Maxwell’s theory of electro-magnetism, a subject which was too “modern” for Weber. There were two major examinations at the Zürich Polytechnic Institute, and Einstein would certainly have failed them had it not been for the help of his loyal friend, the mathematician Marcel Grossman.

Grossman was an excellent and conscientious student, who attended every class and took meticulous notes. With the help of these notes, Einstein managed to pass his examinations; but because he had alienated Weber and the other professors who could have helped him, he found himself completely unable to get a job. In a letter to Professor F. Ostwald on behalf of his son, Einstein’s father wrote: “My son is profoundly unhappy because of his present joblessness; and every day the idea becomes more firmly implanted in his mind that he is a failure, and will not be able to find the way back again.”

From this painful situation, Einstein was rescued (again!) by his friend Marcel Grossman, whose influential father obtained for Einstein a position at the Swiss Patent Office: Technical Expert (Third Class). Anchored at last in a safe, though humble, position, Einstein married one of his classmates. He learned to do his work at the Patent Office very efficiently; and he used the remainder of his time on his own calculations, hiding them guiltily in a drawer when footsteps approached.

In 1905, this Technical Expert (Third Class) astonished the world of science with five papers, written within a few weeks of each other, and published in the *Annalen der Physik*. Of these five papers, three were classics: One of these was the paper in which Einstein applied Planck’s quantum hypothesis to the photoelectric effect. The second paper discussed “Brownian motion”, the zig-zag motion of small particles suspended in a liquid and hit randomly by the molecules of the liquid. This paper supplied a direct proof of the validity of atomic ideas and of Boltzmann’s kinetic theory. The third paper was destined to establish Einstein’s reputation as one of the greatest physicists of all time. It was entitled “On the Electrodynamics of Moving Bodies”, and in this paper, Albert Einstein formulated his

special theory of relativity. Essentially, this theory maintained that all of the fundamental laws of nature exhibit a symmetry with respect to rotations in a 4-dimensional space-time continuum.

C.3 Special relativity theory

The theory of relativity grew out of problems connected with Maxwell's electromagnetic theory of light. Ever since the wavelike nature of light had first been demonstrated, it had been supposed that there must be some medium to carry the light waves, just as there must be some medium (for example air) to carry sound waves. A word was even invented for the medium which was supposed to carry electromagnetic waves: It was called the "ether".

By analogy with sound, it was believed that the velocity of light would depend on the velocity of the observer relative to the "ether". However, all attempts to measure differences in the velocity of light in different directions had failed, including an especially sensitive experiment which was performed in America in 1887 by A.A. Michelson and E.W. Morley.

Even if the earth had, by a coincidence, been stationary with respect to the "ether" when Michelson and Morley first performed their experiment, they should have found an "ether wind" when they repeated their experiment half a year later, with the earth at the other side of its orbit. Strangely, the observed velocity of light seemed to be completely independent of the motion of the observer!

In his famous 1905 paper on relativity, Einstein made the negative result of the Michelson-Morley experiment the basis of a far-reaching principle: He asserted that no experiment whatever can tell us whether we are at rest or whether we are in a state of uniform motion. With this assumption, the Michelson-Morley experiment of course had to fail, and the measured velocity of light had to be independent of the motion of the observer.

Einstein's Principle of Special Relativity had other extremely important consequences: He soon saw that if his principle were to hold, then Newtonian mechanics would have to be modified. In fact, Einstein's Principle of Special Relativity required that *all* fundamental physical laws exhibit a symmetry between space and time. The three space dimensions, and a fourth dimension, ict , had to enter every fundamental physical law in a symmetrical way. (Here i is the square root of -1 , c is the velocity of light, and t is time.)

When this symmetry requirement is fulfilled, a physical law is said to be "Lorentz-invariant" (in honor of the Dutch physicist H.A. Lorentz, who anticipated some of Einstein's ideas). Today, we would express Einstein's principle by saying that every fundamental physical law must be Lorentz-invariant (i.e. symmetrical in the space and time coordinates). The law will then be independent of the motion of the observer, provided that the observer is moving uniformly.

Einstein was able to show that, when properly expressed, Maxwell's equations are already Lorentz-invariant; but Newton's equations of motion have to be modified. When the needed modifications are made, Einstein found, then the mass of a moving particle appears to increase as it is accelerated. A particle can never be accelerated to a velocity

greater than the velocity of light; it merely becomes heavier and heavier, the added energy being converted into mass.

From his 1905 theory, Einstein deduced his famous formula equating the energy of a system to its mass multiplied by the square of the velocity of light. As we shall see, his formula was soon used to explain the source of the energy produced by decaying uranium and radium; and eventually it led to the construction of the atomic bomb. Thus Einstein, a lifelong pacifist, who renounced his German citizenship as a protest against militarism, became instrumental in the construction of the most destructive weapon ever invented - a weapon which casts an ominous shadow over the future of humankind.

Just as Einstein was one of the first to take Planck's quantum hypothesis seriously, so Planck was one of the first physicists to take Einstein's relativity seriously. Another early enthusiast for relativity was Hermann Minkowski, Einstein's former professor of mathematics. Although he once had characterized Einstein as a "lazy dog", Minkowski now contributed importantly to the mathematical formalism of Einstein's theory; and in 1907, he published the first book on relativity. In honor of Minkowski's contributions to relativity, the 4-dimensional space-time continuum in which we live is sometimes called "Minkowski space".

In 1908, Minkowski began a lecture to the Eightieth Congress of German Scientists and Physicians with the following words:

"From now on, space by itself, and time by itself, are destined to sink completely into the shadows; and only a kind of union of both will retain an independent existence."

Gradually, the importance of Einstein's work began to be realized, and he was much sought after. He was first made Assistant Professor at the University of Zürich, then full Professor in Prague, then Professor at the Zürich Polytechnic Institute; and finally, in 1913, Planck and Nernst persuaded Einstein to become Director of Scientific Research at the Kaiser Wilhelm Institute in Berlin. He was at this post when the First World War broke out

While many other German intellectuals produced manifestos justifying Germany's invasion of Belgium, Einstein dared to write and sign an anti-war manifesto. Einstein's manifesto appealed for cooperation and understanding among the scholars of Europe for the sake of the future; and it proposed the eventual establishment of a League of Europeans. During the war, Einstein remained in Berlin, doing whatever he could for the cause of peace, burying himself unhappily in his work, and trying to forget the agony of Europe, whose civilization was dying in a rain of shells, machine-gun bullets, and poison gas.

C.4 General relativity

The work into which Einstein threw himself during this period was an extension of his theory of relativity. He already had modified Newton's equations of motion so that they exhibited the space-time symmetry required by his Principle of Special Relativity. However, Newton's law of gravitation remained a problem.

Obviously it had to be modified, since it disagreed with his Special Theory of Relativity;

but how should it be changed? What principles could Einstein use in his search for a more correct law of gravitation? Certainly whatever new law he found would have to give results very close to Newton's law, since Newton's theory could predict the motions of the planets with almost perfect accuracy. This was the deep problem with which he struggled.

In 1907, Einstein had found one of the principles which was to guide him, the Principle of Equivalence of inertial and gravitational mass. After turning Newton's theory over and over in his mind, Einstein realized that Newton had used mass in two distinct ways: His laws of motion stated that the force acting on a body is equal to the mass of the body multiplied by its acceleration; but according to Newton, the gravitational force on a body is also proportional to its mass. In Newton's theory, gravitational mass, by a coincidence, is equal to inertial mass; and this holds for all bodies. Einstein decided to construct a theory in which gravitational and inertial mass necessarily have to be the same.

He then imagined an experimenter inside a box, unable to see anything outside it. If the box is on the surface of the earth, the person inside it will feel the pull of the earth's gravitational field. If the experimenter drops an object, it will fall to the floor with an acceleration of 32 feet per second per second. Now suppose that the box is taken out into empty space, far away from strong gravitational fields, and accelerated by exactly 32 feet per second per second. Will the enclosed experimenter be able to tell the difference between these two situations? Certainly no difference can be detected by dropping an object, since in the accelerated box, the object will fall to the floor in exactly the same way as before.

With this "thought experiment" in mind, Einstein formulated a general Principle of Equivalence: He asserted that no experiment whatever can tell an observer enclosed in a small box whether the box is being accelerated, or whether it is in a gravitational field. According to this principle, gravitation and acceleration are locally equivalent, or, to say the same thing in different words, gravitational mass and inertial mass are equivalent.

Einstein soon realized that his Principle of Equivalence implied that a ray of light must be bent by a gravitational field. This conclusion followed because, to an observer in an accelerated frame, a light beam which would appear straight to a stationary observer, must necessarily appear very slightly curved. If the Principle of Equivalence held, then the same slight bending of the light ray would be observed by an experimenter in a stationary frame in a gravitational field.

Another consequence of the Principle of Equivalence was that a light wave propagating upwards in a gravitational field should be very slightly shifted to the red. This followed because in an accelerated frame, the wave crests would be slightly farther apart than they normally would be, and the same must then be true for a stationary frame in a gravitational field. It seemed to Einstein that it ought to be possible to test experimentally both the gravitational bending of a light ray and the gravitational red shift.

This seemed promising; but how was Einstein to proceed from the Principle of Equivalence to a formulation of the law of gravitation? Perhaps the theory ought to be modeled after Maxwell's electromagnetic theory, which was a field theory, rather than an "action at a distance" theory. Part of the trouble with Newton's law of gravitation was that it allowed a signal to be propagated instantaneously, contrary to the Principle of Special Relativity. A field theory of gravitation might cure this defect, but how was Einstein to find such a

theory? There seemed to be no way.

From these troubles Albert Einstein was rescued (a third time!) by his staunch friend Marcel Grossman. By this time, Grossman had become a professor of mathematics in Zürich, after having written a doctoral dissertation on tensor analysis and non-Euclidean geometry, the very things that Einstein needed. The year was then 1912, and Einstein had just returned to Zürich as Professor of Physics at the Polytechnic Institute. For two years, Einstein and Grossman worked together; and by the time Einstein left for Berlin in 1914, the way was clear. With Grossman's help, Einstein saw that the gravitational field could be expressed as a curvature of the 4-dimensional space-time continuum.

In 1919, a British expedition, headed by Sir Arthur Eddington, sailed to a small island off the coast of West Africa. Their purpose was to test Einstein's prediction of the bending of light in a gravitational field by observing stars close to the sun during a total eclipse. The observed bending agreed exactly with Einstein's predictions; and as a result he became world-famous. The general public was fascinated by relativity, in spite of the abstruseness of the theory (or perhaps because of it). Einstein, the absent-minded professor, with long, uncombed hair, became a symbol of science. The world was tired of war, and wanted something else to think about.

Einstein met President Harding, Winston Churchill and Charlie Chaplin; and he was invited to lunch by the Archbishop of Canterbury. Although adulated elsewhere, he was soon attacked in Germany. Many Germans, looking for an excuse for the defeat of their nation, blamed it on the pacifists and Jews; and Einstein was both these things.

C.5 Metric tensors

Let us consider a coordinate system x^1, x^2, \dots, x^d labelling the points in a d -dimensional space. We can label the points in a different way by going to a new coordinate system X^1, X^2, \dots, X^d where the new coordinates are expressed as functions of the old ones.

$$\begin{aligned} X^1 &= X^1(x^1, x^2, \dots, x^d) \\ X^2 &= X^2(x^1, x^2, \dots, x^d) \\ &\vdots \\ X^d &= X^d(x^1, x^2, \dots, x^d) \end{aligned} \tag{C.1}$$

For example, (C.1) might represent a transformation from Cartesian coordinates to spherical polar coordinates. If we have an equation written in terms of the old coordinates, we might ask how to rewrite it in terms of the new ones. More generally, we can try to write a physical equation in such a way that it will look the same in every coordinate system. Suppose that the space is Euclidean (flat), so that in terms of the Cartesian coordinates x^1, x^2, \dots, x^d , the infinitesimal element of length separating two points is given by the Pythagorean rule:

$$ds^2 = \delta_{i,j} dx^i dx^j \equiv g_{i,j} dx^i dx^j \tag{C.2}$$

(In equation (C.2) and in the remainder of this section, we use the Einstein convention, in which a sum over repeated indices is understood, although not written explicitly.) The symbol $g_{i,j}$ which appears in the definition of the infinitesimal length ds^2 is called the *covariant metric tensor*, and for Cartesian coordinates in a Euclidean space, it is just the Kronecker delta function. Using the identity

$$dx^i = \frac{\partial x^i}{\partial X^\mu} dX^\mu \quad (\text{C.3})$$

we can rewrite (C.2) as

$$ds^2 = \delta_{i,j} \frac{\partial x^i}{\partial X^\mu} \frac{\partial x^j}{\partial X^\nu} dX^\mu dX^\nu \equiv G_{\mu,\nu} dX^\mu dX^\nu \quad (\text{C.4})$$

where

$$G_{\mu,\nu} \equiv g_{i,j} \frac{\partial x^i}{\partial X^\mu} \frac{\partial x^j}{\partial X^\nu} \quad (\text{C.5})$$

The quantity $G_{\mu,\nu}$ which appears in equations (C.4) and (C.5) is the covariant metric tensor in the new coordinate system. In any space, whether Euclidean or not, the covariant metric tensor is defined by the expression which yields ds^2 , the square of the infinitesimal distance between two points, as in equation (C.2) or (C.4). The word *tensor* refers to the way in which a quantity transforms under changes in the coordinate system. The *rank* of a tensor is the number of indices. The covariant metric tensor is the prototype of a covariant tensor of second rank. Any physical quantity which must be transformed according to the rule

$$A_{\mu,\nu} = a_{i,j} \frac{\partial x^i}{\partial X^\mu} \frac{\partial x^j}{\partial X^\nu} \quad (\text{C.6})$$

under the coordinate transformation $x^1, x^2, \dots, x^d \rightarrow X^1, X^2, \dots, X^d$ is said to be a covariant tensor of second rank. The d -component entity

$$dX^\mu = \frac{\partial X^\mu}{\partial x^i} dx^i \quad (\text{C.7})$$

is the prototype of a contravariant tensor of first rank. Any quantity that transforms according to the rule

$$A^\mu = \frac{\partial X^\mu}{\partial x^i} a^i \quad (\text{C.8})$$

is said to be a contravariant tensor of first rank (or contravariant vector). The distance element ds is the prototype of an invariant or *scalar*. Any quantity ϕ which is invariant under coordinate transformations is said to be a scalar. The gradient of a scalar

$$\frac{\partial \phi}{\partial X^\mu} = \frac{\partial x^i}{\partial X^\mu} \frac{\partial \phi}{\partial x^i} \quad (\text{C.9})$$

is the prototype of a covariant tensor of first rank, or covariant vector. Any quantity which transforms according to the rule

$$A_\mu = \frac{\partial x^i}{\partial X^\mu} a_i \quad (\text{C.10})$$

is said to be a covariant vector. We can also define tensors of higher rank. For example,

$$A^{\mu\nu\sigma} = \frac{\partial X^\mu}{\partial x^i} \frac{\partial X^\nu}{\partial x^j} \frac{\partial X^\sigma}{\partial x^k} a^{ijk} \quad (\text{C.11})$$

is said to be a contravariant tensor of third rank. A covariant vector and a contravariant vector can be *contracted* into a scalar:

$$A_\mu B^\mu = \frac{\partial x^i}{\partial X^\mu} \frac{\partial X^\mu}{\partial x^j} a_i b^j = \delta_j^i a_i b^j = a_i b^i \quad (\text{C.12})$$

Similarly, if we contract a contravariant vector with the covariant metric tensor, we obtain a covariant vector:

$$\begin{aligned} G_{\mu\nu} A^\nu &= A_\mu \\ g_{ij} a^j &= a_i \end{aligned} \quad (\text{C.13})$$

It is useful to define a quantity called the *contravariant metric tensor*, which gives the Kronecker δ -function when it is contracted with the covariant metric tensor:

$$\begin{aligned} G^{\mu\nu} G_{\nu\sigma} &= \delta_\sigma^\mu \\ g^{ij} g_{jk} &= \delta_k^i \\ G^{\mu\nu} &= \frac{\partial X^\mu}{\partial x^i} \frac{\partial X^\nu}{\partial x^j} g^{ij} \end{aligned} \quad (\text{C.14})$$

If we contract a covariant vector with the contravariant metric tensor, we obtain a contravariant vector:

$$G^{\mu\nu} A_\nu = A^\mu \quad (\text{C.15})$$

In a similar way, we can raise or lower the indices of a tensor of higher rank. For example, it is easy to show that

$$G_{\mu\nu} A^{\nu\sigma\rho} = A_\mu^{\sigma\rho} \quad (\text{C.16})$$

In a Cartesian coordinate system with unit metric we are accustomed to writing the volume element as

$$dv = dx^1 dx^2 \cdots dx^d \quad (\text{C.17})$$

This is obviously unsatisfactory from the standpoint of tensor analysis, since the right-hand side of equation (C.17) appears to be a contravariant tensor of rank d (or rather a

particular component of such a tensor), while the left-hand side has no indices at all. In order to write the volume element in an invariant way, the Italian mathematician Tullio Levi-Civita (1873-1941) introduced a totally antisymmetric covariant tensor of rank d . In a Cartesian coordinate system, for a flat space, the Levi-Civita tensor is given by

$$e_{ijkl\dots} = \begin{cases} (-1)^\sigma & \text{if } ijkl\dots = \sigma(1234\dots) \\ 0 & \text{otherwise} \end{cases} \quad (\text{C.18})$$

In other words, the Levi-Civita tensor is ± 1 if $ijkl\dots$ is a permutation of $1234\dots$, with the sign depending on whether the permutation is even or odd, and it is zero otherwise. In terms of this tensor, the volume element of equation (C.17) becomes

$$dv = \frac{1}{d!} e_{ijkl\dots} dx^i dx^j dx^k dx^l \dots \quad (\text{C.19})$$

while in a transformed coordinate system it is

$$dV = \frac{1}{d!} E_{\mu\nu\sigma\rho\dots} dX^\mu dX^\nu dX^\sigma dX^\rho \dots \quad (\text{C.20})$$

where

$$E_{\mu\nu\sigma\dots} = e_{ijk\dots} \frac{\partial x^i}{\partial X^\mu} \frac{\partial x^j}{\partial X^\nu} \frac{\partial x^k}{\partial X^\sigma} \dots \quad (\text{C.21})$$

In this way, Levi-Civita used the formalism of tensor calculus to re-derive the previous result of the German mathematician Carl Gustav Jacobi (1804-1851), who had shown that in a curvilinear coordinate system, the volume element is given by

$$dV = \left| \frac{\partial x^i}{\partial X^\mu} \right| dX^1 dX^2 \dots dX^d \quad (\text{C.22})$$

where $|\partial x^i / \partial X^\mu|$ is the determinant of the $d \times d$ square matrix of transformation coefficients from Cartesian coordinates to curvilinear coordinates. This determinant is called the *Jacobian* of the transformation. From the relationship

$$G_{\mu,\nu} = \frac{\partial x^i}{\partial X^\mu} \delta_{ij} \frac{\partial x^j}{\partial X^\nu} \quad (\text{C.23})$$

one can show that the Jacobian

$$\sqrt{|G_{\mu\nu}|} \equiv \left| \frac{\partial x^i}{\partial X^\mu} \right| \equiv \sqrt{|G|} \quad (\text{C.24})$$

is the square root of the determinant of the covariant metric tensor. The Jacobian is usually represented by the symbol $\sqrt{|G|}$. Levi-Civita's book *Absolute Differential Calculus* has been translated into many languages. It is still in print, and it remains one of the

best textbooks in the field, along with Schrödinger's *Space-Time Structure*, Brillouin's *Les Tenseurs* and Landau and Lifshitz's *The Classical Theory of Fields*.

The Jacobian, $\sqrt{|G|}$, is the prototype of a *scalar density*. We can construct *tensor densities* by multiplying tensors by the Jacobian appropriate for the coordinate system. When a tensor density is transformed to another coordinate system, the Jacobian has to be recalculated from the transformed covariant metric tensor. *Tensor capacities* can be constructed by dividing tensors by the Jacobian. Now consider a scalar function ψ . Its gradient is a covariant vector, and therefore

$$G^{\mu\nu} \frac{\partial\psi}{\partial X^\mu} \frac{\partial\psi}{\partial X^\nu} = \text{scalar} \quad (\text{C.25})$$

It follows that if we let

$$\mathcal{L} = \sqrt{|G|} \left[G^{\mu\nu} \frac{\partial\psi}{\partial X^\mu} \frac{\partial\psi}{\partial X^\nu} + \kappa\psi^2 \right] \quad (\text{C.26})$$

where κ is a constant, then the variational principle

$$\delta \int \int \cdots \int \mathcal{L} dX^1 dX^2 \cdots dX^d = 0 \quad (\text{C.27})$$

will be invariant under a curvilinear coordinate transformation. As we saw above, the Euler-Lagrange equations that follow from this variational principle are

$$\frac{\partial}{\partial X^\mu} \frac{\partial\mathcal{L}}{\partial(\partial\psi/\partial X^\mu)} - \frac{\partial\mathcal{L}}{\partial\psi} = 0 \quad (\text{C.28})$$

With the Lagrangian density of equation (C.26), this becomes

$$\frac{1}{\sqrt{|G|}} \frac{\partial}{\partial X^\mu} \sqrt{|G|} G^{\mu\nu} \frac{\partial\psi}{\partial X^\nu} = \kappa \psi \quad (\text{C.29})$$

C.6 The Laplace-Beltrami operator

The operator

$$\Delta = \sum_{\mu=1}^d \sum_{\nu=1}^d \frac{1}{\sqrt{|G|}} \frac{\partial}{\partial X^\mu} \sqrt{|G|} G^{\mu\nu} \frac{\partial}{\partial X^\nu} \quad (\text{C.30})$$

is the generalized Laplacian operator, which plays such an important role in the theory of hyperspherical harmonics, but here it is written in a form due to Eugenio Beltrami (1835-1899), which is invariant under coordinate transformations. (In equation (C.29), we have

abandoned the Einstein convention, and have re-introduced explicit sums.) To illustrate this equation, let us consider some examples. In a d -dimensional space, we can let

$$\begin{aligned}
 x^1 &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{d-2} \cos \theta_{d-1} \\
 x^2 &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{d-2} \sin \theta_{d-1} \\
 x^3 &= r \sin \theta_1 \sin \theta_2 \dots \cos \theta_{d-2} \\
 &\vdots \\
 x^{d-1} &= r \sin \theta_1 \cos \theta_2 \\
 x^d &= r \cos \theta_1
 \end{aligned} \tag{C.31}$$

while

$$\begin{aligned}
 X^1 &= r \\
 X^2 &= \theta_1 \\
 X^3 &= \theta_2 \\
 &\vdots \\
 X^{d-1} &= \theta_{d-2} \\
 X^d &= \theta_{d-1}
 \end{aligned} \tag{C.32}$$

Then the Jacobians for various values of d are

$$\begin{aligned}
 d = 3 \quad \sqrt{|G|} &= r^2 \sin \theta_1 \\
 d = 4 \quad \sqrt{|G|} &= r^3 \sin^2 \theta_1 \sin \theta_2 \\
 d = 5 \quad \sqrt{|G|} &= r^4 \sin^3 \theta_1 \sin^2 \theta_2 \sin \theta_3 \\
 d = 6 \quad \sqrt{|G|} &= r^5 \sin^4 \theta_1 \sin^3 \theta_2 \sin^2 \theta_3 \sin \theta_4 \\
 &\vdots \\
 d = d \quad \sqrt{|G|} &= r^{d-1} \sin^{d-2} \theta_1 \sin^{d-3} \theta_2 \dots \sin^2 \theta_{d-3} \sin \theta_{d-2}
 \end{aligned} \tag{C.33}$$

The covariant metric tensor for $d = 3$ is

$$G_{\mu,\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta_1 \end{pmatrix} \tag{C.34}$$

while for $d = 4$

$$G_{\mu,\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & r^2 & 0 & 0 \\ 0 & 0 & r^2 \sin^2 \theta_1 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta_1 \sin^2 \theta_2 \end{pmatrix} \tag{C.35}$$

for $d = 5$

$$G_{\mu,\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & r^2 & 0 & 0 & 0 \\ 0 & 0 & r^2 \sin^2 \theta_1 & 0 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta_1 \sin^2 \theta_2 & 0 \\ 0 & 0 & 0 & 0 & r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3 \end{pmatrix} \quad (\text{C.36})$$

The contravariant metric tensors are just their reciprocals.

$$G^{\mu,\nu} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta_1} \end{pmatrix} \quad (\text{C.37})$$

$$G^{\mu,\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta_1} & 0 \\ 0 & 0 & 0 & \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} \end{pmatrix} \quad (\text{C.38})$$

$$G^{\mu,\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta_1} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3} \end{pmatrix} \quad (\text{C.39})$$

and so on. Combining these results, we obtain the Laplace-Beltrami operators:

$$\begin{aligned} & \sum_{\nu} G^{\mu,\nu} \frac{\partial}{\partial X^{\nu}} \\ &= \left(\frac{\partial}{\partial r}, \frac{1}{r^2} \frac{\partial}{\partial \theta_1}, \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial}{\partial \theta_2}, \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial}{\partial \theta_3}, \dots \right) \end{aligned} \quad (\text{C.40})$$

For $d = 3$,

$$\begin{aligned} & \sum_{\nu=1}^3 \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^{\nu}} \\ &= r^2 \sin \theta_1 \left(\frac{\partial}{\partial r}, \frac{1}{r^2} \frac{\partial}{\partial \theta_1}, \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial}{\partial \theta_2} \right) \end{aligned} \quad (\text{C.41})$$

$$\begin{aligned}
& \sum_{\mu=1}^3 \sum_{\nu=1}^3 \frac{1}{\sqrt{|G|}} \frac{\partial}{\partial X^\mu} \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
&= \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \theta_1} \frac{\partial}{\partial \theta_1} \sin \theta_1 \frac{\partial}{\partial \theta_1} + \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial^2}{\partial \theta_2^2}
\end{aligned} \tag{C.42}$$

For $d = 4$,

$$\begin{aligned}
& \sum_{\nu=1}^4 \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
&= r^3 \sin^2 \theta_1 \sin \theta_2 \left(\frac{\partial}{\partial r}, \frac{1}{r^2} \frac{\partial}{\partial \theta_1}, \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial}{\partial \theta_2}, \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial}{\partial \theta_3} \right)
\end{aligned} \tag{C.43}$$

$$\begin{aligned}
& \sum_{\mu=1}^4 \sum_{\nu=1}^4 \frac{1}{\sqrt{|G|}} \frac{\partial}{\partial X^\mu} \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
&= \frac{1}{r^3} \frac{\partial}{\partial r} r^3 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial}{\partial \theta_1} \sin^2 \theta_1 \frac{\partial}{\partial \theta_1} \\
&+ \frac{1}{r^2 \sin^2 \theta_1 \sin \theta_2} \frac{\partial}{\partial \theta_2} \sin \theta_2 \frac{\partial}{\partial \theta_2} + \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial^2}{\partial \theta_3^2}
\end{aligned} \tag{C.44}$$

For $d = 5$,

$$\begin{aligned}
& \sum_{\nu=1}^5 \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
&= r^4 \sin^3 \theta_1 \sin^2 \theta_2 \sin \theta_3 \\
&\times \left(\frac{\partial}{\partial r}, \frac{1}{r^2} \frac{\partial}{\partial \theta_1}, \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial}{\partial \theta_2}, \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial}{\partial \theta_3}, \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3} \frac{\partial}{\partial \theta_4} \right)
\end{aligned} \tag{C.45}$$

$$\begin{aligned}
& \sum_{\mu=1}^5 \sum_{\nu=1}^5 \frac{1}{\sqrt{|G|}} \frac{\partial}{\partial X^\mu} \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
&= \frac{1}{r^4} \frac{\partial}{\partial r} r^4 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin^3 \theta_1} \frac{\partial}{\partial \theta_1} \sin^3 \theta_1 \frac{\partial}{\partial \theta_1} + \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial}{\partial \theta_2} \sin^2 \theta_2 \frac{\partial}{\partial \theta_2} \\
&+ \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin \theta_3} \frac{\partial}{\partial \theta_3} \sin \theta_3 \frac{\partial}{\partial \theta_3} + \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3} \frac{\partial^2}{\partial \theta_4^2}
\end{aligned} \tag{C.46}$$

For general values of d ,

$$\begin{aligned}
& \sum_{\nu=1}^d \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
= & r^{d-1} \sin^{d-2} \theta_1 \sin^{d-3} \theta_2 \cdots \sin \theta_{d-2} \\
\times & \left(\frac{\partial}{\partial r}, \frac{1}{r^2} \frac{\partial}{\partial \theta_1}, \frac{1}{r^2 \sin^2 \theta_1} \frac{\partial}{\partial \theta_2}, \cdots, \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \cdots \sin^2 \theta_{d-2}} \frac{\partial}{\partial \theta_{d-1}} \right)
\end{aligned} \tag{C.47}$$

$$\begin{aligned}
& \sum_{\mu=1}^d \sum_{\nu=1}^d \frac{1}{\sqrt{|G|}} \frac{\partial}{\partial X^\mu} \sqrt{|G|} G^{\mu,\nu} \frac{\partial}{\partial X^\nu} \\
= & \frac{1}{r^{d-1}} \frac{\partial}{\partial r} r^{d-1} \frac{\partial}{\partial r} + \frac{1}{r^2 \sin^{d-2} \theta_1} \frac{\partial}{\partial \theta_1} \sin^{d-2} \theta_1 \frac{\partial}{\partial \theta_1} \\
& + \frac{1}{r^2 \sin^2 \theta_1 \sin^{d-3} \theta_2} \frac{\partial}{\partial \theta_2} \sin^{d-3} \theta_2 \frac{\partial}{\partial \theta_2} + \cdots \\
& + \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3 \cdots \sin \theta_{d-2}} \frac{\partial}{\partial \theta_{d-2}} \sin \theta_{d-2} \frac{\partial}{\partial \theta_{d-2}} \\
& + \frac{1}{r^2 \sin^2 \theta_1 \sin^2 \theta_2 \cdots \sin^2 \theta_{d-2}} \frac{\partial^2}{\partial \theta_{d-1}^2}
\end{aligned} \tag{C.48}$$

As we saw in equation (??), the Laplace-Beltrami operator in hyperspherical coordinates can be written as

$$\Delta = \frac{1}{r^{d-1}} \frac{\partial}{\partial r} r^{d-1} \frac{\partial}{\partial r} - \frac{\Lambda^2}{r^2} \tag{C.49}$$

where r is the hyperradius and where Λ^2 is the generalized angular momentum operator. Comparing this with the results that we have just been discussing, we can see that for $d = 3$,

$$-\Lambda^2 = \frac{1}{\sin \theta_1} \frac{\partial}{\partial \theta_1} \sin \theta_1 \frac{\partial}{\partial \theta_1} + \frac{1}{\sin^2 \theta_1} \frac{\partial^2}{\partial \theta_2^2} \tag{C.50}$$

while for $d = 4$,

$$\begin{aligned}
-\Lambda^2 = & \frac{1}{\sin^2 \theta_1} \frac{\partial}{\partial \theta_1} \sin^2 \theta_1 \frac{\partial}{\partial \theta_1} \\
& + \frac{1}{\sin^2 \theta_1 \sin \theta_2} \frac{\partial}{\partial \theta_2} \sin \theta_2 \frac{\partial}{\partial \theta_2} + \frac{1}{\sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial^2}{\partial \theta_3^2}
\end{aligned} \tag{C.51}$$

and for $d = 5$,

$$\begin{aligned}
-\Lambda^2 = & \frac{1}{\sin^3 \theta_1} \frac{\partial}{\partial \theta_1} \sin^3 \theta_1 \frac{\partial}{\partial \theta_1} + \frac{1}{\sin^2 \theta_1 \sin^2 \theta_2} \frac{\partial}{\partial \theta_2} \sin^2 \theta_2 \frac{\partial}{\partial \theta_2} \\
& + \frac{1}{\sin^2 \theta_1 \sin^2 \theta_2 \sin \theta_3} \frac{\partial}{\partial \theta_3} \sin \theta_3 \frac{\partial}{\partial \theta_3} + \frac{1}{\sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3} \frac{\partial^2}{\partial \theta_4^2}
\end{aligned} \tag{C.52}$$

For general values of d , we have

$$\begin{aligned}
-\Lambda^2 &= \frac{1}{\sin^{d-2} \theta_1} \frac{\partial}{\partial \theta_1} \sin^{d-2} \theta_1 \frac{\partial}{\partial \theta_1} \\
&+ \frac{1}{\sin^2 \theta_1 \sin^{d-3} \theta_2} \frac{\partial}{\partial \theta_2} \sin^{d-3} \theta_2 \frac{\partial}{\partial \theta_2} + \dots \\
&+ \frac{1}{\sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3 \dots \sin \theta_{d-2}} \frac{\partial}{\partial \theta_{d-2}} \sin \theta_{d-2} \frac{\partial}{\partial \theta_{d-2}} \\
&+ \frac{1}{\sin^2 \theta_1 \sin^2 \theta_2 \dots \sin^2 \theta_{d-2}} \frac{\partial^2}{\partial \theta_{d-1}^2}
\end{aligned} \tag{C.53}$$

We have until now been considering spaces that are intrinsically flat, but a d -dimensional hyperspherical surface embedded in a $d + 1$ -dimensional space has intrinsic curvature. If the hyperradius r is regarded as a constant, then the Laplace-Beltrami operator for such a surface is given by

$$\Delta = -\frac{\Lambda^2}{r^2} \tag{C.54}$$

while the covariant metric tensor on the surface is

$$G_{\mu,\nu} = \begin{pmatrix} r^2 & 0 & 0 & 0 & \dots \\ 0 & r^2 \sin^2 \theta_1 & 0 & 0 & \dots \\ 0 & 0 & r^2 \sin^2 \theta_1 \sin^2 \theta_2 & 0 & \dots \\ 0 & 0 & 0 & r^2 \sin^2 \theta_1 \sin^2 \theta_2 \sin^2 \theta_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix} \tag{C.55}$$

The infinitesimal element of length on the surface, ds is given by

$$\begin{aligned}
ds^2 &= \sum_{\mu=1}^d \sum_{\nu=1}^d G_{\mu,\nu} dX^\mu dX^\nu \\
&= r^2 (d\theta_1^2 + \sin^2 \theta_1 d\theta_2^2 + \sin^2 \theta_1 \sin^2 \theta_2 d\theta_3^2 + \dots)
\end{aligned} \tag{C.56}$$

C.7 Geodesics

In the geometry of curved spaces, *geodesics* play the role that straight lines play in Euclidean geometry. The geodesic curves are local minima of path length. The *minimal geodesics* between two points are the shortest paths through the curved space, and play an important role when analyzing physical systems in curved space. They can be determined

by the variational principle

$$\begin{aligned}
 s &= \int ds = \int \sqrt{\sum_{\mu=1}^d \sum_{\nu=1}^d G_{\mu,\nu} \frac{dX^\mu}{ds} \frac{dX^\nu}{ds}} ds \\
 &= \int \sum_{\mu=1}^d \sum_{\nu=1}^d G_{\mu,\nu} \frac{dX^\mu}{ds} \frac{dX^\nu}{ds} ds = \text{minimum}
 \end{aligned} \tag{C.57}$$

The Euler-Lagrange equations which follow from this variational principle are

$$\frac{d}{ds} \frac{\partial L}{\partial (dX^\mu/ds)} - \frac{\partial L}{\partial X^\mu} = 0 \quad \mu = 1, 2, \dots, d \tag{C.58}$$

with

$$L = \sum_{\mu=1}^d \sum_{\nu=1}^d G_{\mu,\nu}(\mathbf{X}) \frac{dX^\mu}{ds} \frac{dX^\nu}{ds} \tag{C.59}$$

The Euler-Lagrange equations for geodesics can be written in the form

$$\frac{d^2 X^\sigma}{ds^2} = \Gamma_{\mu\nu}^\sigma \frac{dX^\mu}{ds} \frac{dX^\nu}{ds} \tag{C.60}$$

Here $\Gamma_{\mu\nu}^\sigma$ is a Christoffel symbol, which is related to the metric tensors by

$$\Gamma_{\mu\nu}^\sigma = \frac{1}{2} G^{\sigma\rho} \left(\frac{\partial G_{\rho\mu}}{\partial X^\nu} + \frac{\partial G_{\rho\nu}}{\partial X^\mu} - \frac{\partial G_{\mu\nu}}{\partial X^\rho} \right) \tag{C.61}$$

In general relativity theory, the trajectories of particles are geodesics in a space-time continuum, whose metric is affected by the presence of other masses.

Appendix D

Sturmian basis sets

D.1 One-electron Coulomb Sturmians

Because of their completeness properties, one-electron Sturmian basis sets have long been used in theoretical atomic physics [Shull and Löwdin, 1959], [Rotenberg, 1962], [Rotenberg, 1970], [Avery, 2003], [Gasaneo et al., 2009]. Their form is identical with that of the familiar hydrogenlike atomic orbitals, except that the factor Z/n is replaced by a constant k . The one-electron Coulomb Sturmians can be written as

$$\chi_{nlm}(\mathbf{x}) = R_{nl}(r)Y_{lm}(\theta, \phi) \quad (\text{D.1})$$

where Y_{lm} is a spherical harmonic, and where the radial function has the form

$$R_{nl}(r) = \mathcal{N}_{nl}(2kr)^l e^{-kr} F(l+1-n|2l+2|2kr) \quad (\text{D.2})$$

Here

$$\mathcal{N}_{nl} = \frac{2k^{3/2}}{(2l+1)!} \sqrt{\frac{(l+n)!}{n(n-l-1)!}} \quad (\text{D.3})$$

is a normalizing constant, while

$$F(a|b|x) \equiv \sum_{t=0}^{\infty} \frac{a^{\bar{t}}}{t!b^{\bar{t}}} x^t = 1 + \frac{a}{b}x + \frac{a(a+1)}{2b(b+1)}x^2 + \dots \quad (\text{D.4})$$

is a confluent hypergeometric function. Coulomb Sturmian basis functions obey the following one-electron Schrödinger equation (in atomic units):

$$\left[-\frac{1}{2}\nabla^2 - \frac{nk}{r} + \frac{1}{2}k^2 \right] \chi_{nlm}(\mathbf{x}) = 0 \quad (\text{D.5})$$

which is just the Schrödinger equation for an electron in a hydrogenlike atom with the replacement $Z/n \rightarrow k$. All of the functions in a such a basis set correspond to the same energy,

$$\epsilon = -\frac{1}{2}k^2 \quad (\text{D.6})$$

Table B.1: One-electron Coulomb Sturmian radial functions. If k is replaced by Z/n they are identical to the familiar hydrogenlike radial wave functions.

n	l	$R_{n,l}(r)$
1	0	$2k^{3/2}e^{-kr}$
2	0	$2k^{3/2}(1 - kr)e^{-kr}$
2	1	$\frac{2k^{3/2}}{\sqrt{3}} kr e^{-kr}$
3	0	$2k^{3/2} \left(1 - 2kr + \frac{2(kr)^2}{3} \right) e^{-kr}$
3	1	$2k^{3/2} \frac{2\sqrt{2}}{3} kr \left(1 - \frac{kr}{2} \right) e^{-kr}$
3	2	$2k^{3/2} \frac{\sqrt{2}}{3\sqrt{5}} (kr)^2 e^{-kr}$

In other words the basis set is isoenergetic. In the wave equation obeyed by the Sturmians, (D.5), the potential is weighted differently for members of the basis set corresponding to different values of n . Equation (D.5) can be written in the form:

$$\left[-\frac{1}{2}\nabla^2 - \beta_n \frac{Z}{r} + \frac{1}{2}k^2 \right] \chi_{nlm}(\mathbf{x}) = 0 \quad \beta_n = \frac{kn}{Z} \quad (\text{D.7})$$

The weighting factors β_n are chosen in such a way as to make all of the solutions isoenergetic. All solutions correspond to the energy $\epsilon = -k^2/2$. In the Hamiltonian formulation of physics, the eigenvalues of the wave equation are a spectrum of allowed energies, but here all of the solutions of the wave equation correspond to the same energy, and the weighting factors play the role of eigenvalues. The functions in a Coulomb Sturmian basis set can be shown to obey and obey a potential-weighted orthonormality relation: To see this, we consider two solutions, $\chi_{nlm}(\mathbf{x})$ and $\chi_{n'l'm'}(\mathbf{x})$, obeying the equations:

$$\begin{aligned} \left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 \right] \chi_{nlm}(\mathbf{x}) &= \frac{nk}{r} \chi_{nlm}(\mathbf{x}) \\ \left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 \right] \chi_{n'l'm'}^*(\mathbf{x}) &= \frac{n'k}{r} \chi_{n'l'm'}^*(\mathbf{x}) \end{aligned} \quad (\text{D.8})$$

Multiplying the two equations from the left respectively by $\chi_{n'l'm'}^*(\mathbf{x})$ and $\chi_{nlm}(\mathbf{x})$, integrating over the coordinates, and subtracting the two equations, we obtain:

$$(n - n') \int d^3x \chi_{n'l'm'}^*(\mathbf{x}) \frac{1}{r} \chi_{nlm}(\mathbf{x}) = 0 \quad (\text{D.9})$$

where we have also made use of the fact that (from Hermiticity)

$$\begin{aligned} &\int d^3x \chi_{n'l'm'}^*(\mathbf{x}) \left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 \right] \chi_{nlm}(\mathbf{x}) \\ &- \int d^3x \chi_{nlm}(\mathbf{x}) \left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 \right] \chi_{n'l'm'}^*(\mathbf{x}) = 0 \end{aligned} \quad (\text{D.10})$$

Thus for $n \neq n'$, the potential-weighted scalar product vanishes, and it vanishes also when $l' \neq l$ or $m' \neq m$ because of the orthogonality of the spherical harmonics. The Coulomb Sturmians are normalized in such a way that the orthonormality relation is:

$$\int d^3x \chi_{n'l'm'}^*(\mathbf{x}) \frac{1}{r} \chi_{nlm}(\mathbf{x}) = \frac{k}{n} \delta_{n'n} \delta_{l'l} \delta_{m'm} \quad (\text{D.11})$$

Because of their completeness and their close relationship with Coulomb potentials, Coulomb Sturmians are widely used in atomic physics.

D.2 Löwdin-orthogonalized Coulomb Sturmians

The Coulomb Sturmians form a complete set in the sense that any square-integrable function of \mathbf{x} can be expanded in terms of them. For this reason, they are useful as basis functions in many applications. Sometimes it may be convenient to use Coulomb Sturmian basis functions in a form that is orthonormalized in the conventional way. Let us denote the orthogonalized Coulomb Sturmians by $\tilde{\chi}_\mu(\mathbf{x})$, where $\mu \equiv (n, l, m)$. This new basis set is related to the original set of Coulomb Sturmians discussed above by

$$\tilde{\chi}_\mu(\mathbf{x}) = \sum_{\mu'} \chi_\mu(\mathbf{x}) W_{\mu',\mu} \quad (\text{D.12})$$

where $W_{\mu',\mu}$ is a transformation matrix. We wish the transformation to be such that

$$\int d^3x \tilde{\chi}_{\mu'}^*(\mathbf{x}) \tilde{\chi}_\mu(\mathbf{x}) \equiv \tilde{S}_{\mu',\mu} = \delta_{\mu',\mu} \quad (\text{D.13})$$

Suppose that

$$\int d^3x \chi_{\mu'}^*(\mathbf{x}) \chi_\mu(\mathbf{x}) = S_{\mu',\mu} \quad (\text{D.14})$$

Then, in matrix notation, the condition that the transformation matrix W must satisfy is

$$W^\dagger S W = I \quad (\text{D.15})$$

where the dagger denotes the Hermitian adjoint, i.e., the conjugate transpose. Following Löwdin and Wannier, we can choose from all the possible solutions to the matrix equation (D.15) the one for which

$$W^\dagger = W \quad (\text{D.16})$$

(This is sometimes called *symmetrical orthogonalization*.) Then (D.15) will be satisfied if

$$W = S^{-1/2} \quad (\text{D.17})$$

In order to find the square root of the overlap matrix S , we diagonalize it, take the inverse square root in the diagonal representation, and then transform back to the original representation. This gives us $W = S^{-1/2}$, which we then use to perform the transformation shown in equation (D.12).

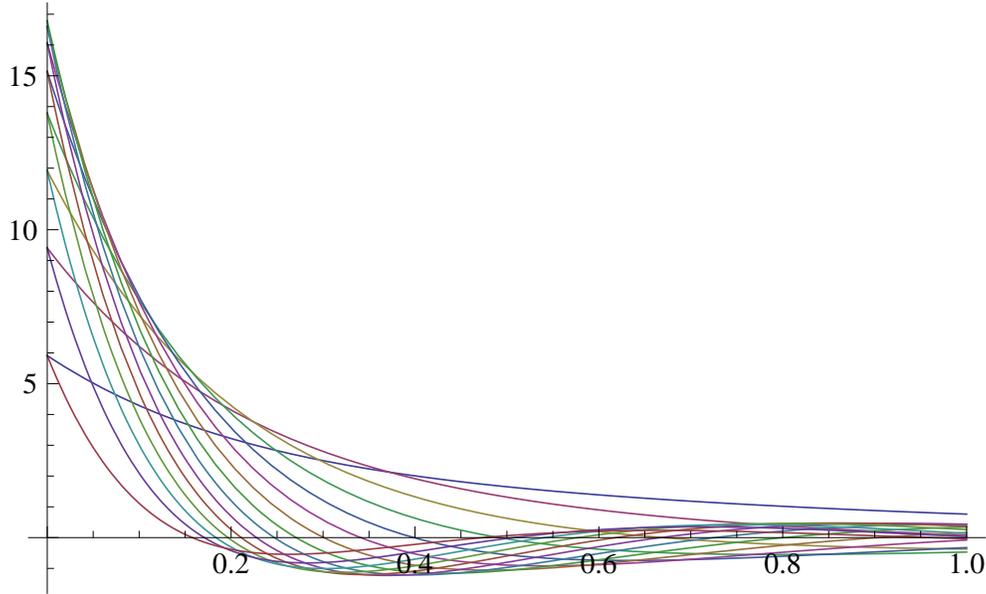


Figure D.1: A set of 15 Löwdin-orthogonalized Coulomb Sturmians corresponding to $l = 0$ and $k = 1$. The radial parts are shown as functions of r . If an arbitrary radial function is to be expanded in terms of this set, the value of k for the set can be adjusted in such a way as to give maximum accuracy. Löwdin-orthogonalized Coulomb Sturmians are used in the Hartree-Fock calculations of Chapter 2.

D.3 The Fock projection

Coulomb Sturmian basis functions and their Fourier transforms are related by

$$\chi_{nlm}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^3}} \int d^3x e^{i\mathbf{p}\cdot\mathbf{x}} \chi_{nlm}^t(\mathbf{p}) \quad (\text{D.18})$$

and by the inverse transform

$$\chi_{nlm}^t(\mathbf{p}) = \frac{1}{\sqrt{(2\pi)^3}} \int d^3x e^{-i\mathbf{p}\cdot\mathbf{x}} \chi_{nlm}(\mathbf{x}) \quad (\text{D.19})$$

By projecting momentum-space onto the surface of a 4-dimensional hypersphere, V. Fock [Fock, 1935], [Fock, 1958] was able to show that the Fourier-transformed Coulomb Sturmians can be very simply expressed in terms of 4-dimensional hyperspherical harmonics through the relationship

$$\chi_{n,l,m}^t(\mathbf{p}) = M(p) Y_{n-1,l,m}(\hat{\mathbf{u}}) \quad (\text{D.20})$$

where

$$M(p) \equiv \frac{4k^{5/2}}{(k^2 + p^2)^2} \quad (\text{D.21})$$

and

$$\begin{aligned}
 u_1 &= \frac{2kp_1}{k^2 + p^2} \\
 u_2 &= \frac{2kp_2}{k^2 + p^2} \\
 u_3 &= \frac{2kp_3}{k^2 + p^2} \\
 u_4 &= \frac{k^2 - p^2}{k^2 + p^2}
 \end{aligned} \tag{D.22}$$

The 4-dimensional hyperspherical harmonics are given by [Judd, 1975], [Avery, 1989], [Avery, 2000], [Avery and Avery, 2006]

$$Y_{\lambda,l,m}(\hat{\mathbf{u}}) = \mathcal{N}_{\lambda,l} C_{\lambda-l}^{1+l}(u_4) Y_{l,m}(u_1, u_2, u_3) \tag{D.23}$$

where $Y_{l,m}$ is a spherical harmonic of the familiar type, while

$$\mathcal{N}_{\lambda,l} = (-1)^{\lambda-l} (2l)!! \sqrt{\frac{2(\lambda+1)(\lambda-l)!}{\pi(\lambda+l+1)!}} \tag{D.24}$$

is a normalizing constant, and

$$C_j^\alpha(u_4) = \sum_{t=0}^{\lfloor j/2 \rfloor} \frac{(-1)^t \Gamma(j+\alpha-t)}{t!(j-2t)!\Gamma(\alpha)} (2u_4)^{j-2t} \tag{D.25}$$

is a Gegenbauer polynomial. The first few The relationships between hyperspherical harmonics, harmonic polynomials, and harmonic projection will be discussed in Appendix C. Table 5.1 in Chapter 5 shows the first few hyperspherical harmonics.

D.4 Generalized Sturmians and many-particle problems

In 1968, Osvaldo Goscinski [[Goscinski, 1968, 2003]] generalized the concept of Sturmian basis sets by considering isoenergetic sets of solutions to a many-particle Schrödinger equation with a weighted potential:

$$\left[-\frac{1}{2}\Delta + \beta_\nu V_0(\mathbf{x}) - E_\kappa \right] |\Phi_\nu\rangle = 0 \tag{D.26}$$

The weighting factors β_ν are chosen in such a way as to make all of the functions in the set correspond to the same energy, E_κ , and this energy is usually chosen to be that of the quantum mechanical state which is to be represented by a superposition of generalized

D.5. USE OF GENERALIZED STURMIAN BASIS SETS TO SOLVE THE MANY-PARTICLE SCHRÖD

Sturmian basis functions. If the basis set is used to treat N -particle systems where the particles have different masses, the operator Δ in the kinetic energy term is given by

$$\Delta \equiv \sum_{j=1}^N \frac{1}{m_j} \nabla_j^2 \quad (\text{D.27})$$

while if the masses are all equal, it is given by the generalized Laplacian operator:

$$\Delta \equiv \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2} \quad (\text{D.28})$$

with $d = 3N$ and

$$\mathbf{x} = (x_1, x_2, \dots, x_d) \quad (\text{D.29})$$

Like the one-electron Coulomb Sturmians, the functions in generalized Sturmian basis sets can be shown to obey a potential-weighted orthonormality relation [Avery and Avery, 2006]:

$$\langle \Phi_{\nu'} | V_0(\mathbf{x}) | \Phi_{\nu} \rangle = \delta_{\nu', \nu} \frac{2E_{\kappa}}{\beta_{\nu}} \quad (\text{D.30})$$

D.5 Use of generalized Sturmian basis sets to solve the many-particle Schrödinger equation

If we wish to solve a many-particle Schrödinger equation of the form

$$\left[-\frac{1}{2} \Delta + V(\mathbf{x}) - E_{\kappa} \right] |\Psi_{\kappa}\rangle = 0 \quad (\text{D.31})$$

we can approximate a solution as a superposition of generalized Sturmian basis functions

$$|\Psi_{\kappa}\rangle \approx \sum_{\nu} |\Phi_{\nu}\rangle B_{\nu, \kappa} \quad (\text{D.32})$$

Substituting this superposition into the Schrödinger equation and remembering that each of the basis functions satisfies eq.(D.26), we obtain:

$$\begin{aligned} & \sum_{\nu} \left[-\frac{1}{2} \Delta + V(\mathbf{x}) - E_{\kappa} \right] |\Phi_{\nu}\rangle B_{\nu, \kappa} \\ &= \sum_{\nu} [V(\mathbf{x}) - \beta_{\nu} V_0(\mathbf{x})] |\Phi_{\nu}\rangle B_{\nu, \kappa} \approx 0 \end{aligned} \quad (\text{D.33})$$

If we multiply from the left by a conjugate function from our generalized Sturmian basis set and integrate over all coordinates, we obtain a set of secular equations from which the kinetic energy term has disappeared:

$$\sum_{\nu} \langle \Phi_{\nu'}^* | [V(\mathbf{x}) - \beta_{\nu} V_0(\mathbf{x})] |\Phi_{\nu}\rangle B_{\nu, \kappa} = 0 \quad (\text{D.34})$$

If we introduce the definition

$$T_{\nu',\nu} \equiv -\frac{1}{p_\kappa} \langle \Phi_{\nu'}^* | V(\mathbf{x}) | \Phi_\nu \rangle \quad (\text{D.35})$$

where

$$p_\kappa \equiv \sqrt{-2E_\kappa} \quad (\text{D.36})$$

and make use of the potential-weighted orthonormality relations (D.30), we can rewrite the secular equations in the form:

$$\sum_\nu [T_{\nu',\nu} - p_\kappa \delta_{\nu',\nu}] B_{\nu,\kappa} = 0 \quad (\text{D.37})$$

The generalized Sturmian secular equations are strikingly different from conventional Hamiltonian secular equations in several ways:

- The kinetic energy term has disappeared.
- The matrix representing the approximate potential $V_0(\mathbf{x})$ is diagonal.
- The roots of the secular equations are not energies, but values of the scaling parameter p_κ , from which the energy can be obtained through the relationship $E_\kappa = -p_\kappa^2/2$.
- For Coulomb potentials, the matrix $T_{\nu',\nu}$ is energy-independent.

D.6 Momentum-space orthonormality relations for Sturmian basis sets

By arguments similar to those used in equations (D.8)-(D.11), a set of generalized Sturmian basis functions can be shown to obey a potential-weighted orthonormality relation in direct space

$$\int dx \Phi_{\nu'}^*(\mathbf{x}) V_0(\mathbf{x}) \Phi_\nu(\mathbf{x}) = \delta_{\nu',\nu} \frac{2E_\kappa}{\beta_\nu} = -\delta_{\nu',\nu} \frac{p_\kappa^2}{\beta_\nu} \quad (\text{D.38})$$

where

$$p_\kappa^2 \equiv -2E_\kappa \quad (\text{D.39})$$

(In equation (D.38) and in the remainder of this appendix, we abandon the Dirac bra and ket notation in order to distinguish between functions of $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ and functions of $\mathbf{p} \equiv (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$). We would now like to find the momentum-space orthonormality relations obeyed by Fourier transforms of the generalized Sturmian basis set. Because the Fourier transform is unitary, the inner product of any two functions in L_2 is preserved under the operation of taking their Fourier transforms, i.e.

$$\int dx f^*(\mathbf{x}) g(\mathbf{x}) = \int dp f^{t*}(\mathbf{p}) g^t(\mathbf{p}) \quad (\text{D.40})$$

Using this well-known relationship with $f^*(\mathbf{x}) = \Phi_{\nu'}^*(\mathbf{x})$ and $g(\mathbf{x}) = V_0(\mathbf{x})\Phi_\nu(\mathbf{x})$, we have

$$\int dx \Phi_{\nu'}^*(\mathbf{x})V_0(\mathbf{x})\Phi_\nu(\mathbf{x}) = \int dp \Phi_{\nu'}^{t*}(\mathbf{p}) [V_0\Phi_\nu]^t(\mathbf{p}) \quad (\text{D.41})$$

In order to evaluate $[V_0\Phi_\nu]^t(\mathbf{p})$, we remember the Fourier convolution theorem, which states that the Fourier transform of the product of two functions is the convolution of their Fourier transforms. Thus if a and b are any two functions in L_2 ,

$$[ab]^t(\mathbf{p}') \equiv \frac{1}{(2\pi)^{d/2}} \int dx e^{-i\mathbf{p}'\cdot\mathbf{x}} a(\mathbf{x})b(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}} \int dp a^t(\mathbf{p}' - \mathbf{p})b^t(\mathbf{p}) \quad (\text{D.42})$$

Letting $a(\mathbf{x}) = V_0(\mathbf{x})$ and $b(\mathbf{x}) = \Phi_\nu(\mathbf{x})$ we have

$$[V_0\Phi_\nu]^t(\mathbf{p}') = \frac{1}{(2\pi)^{d/2}} \int dp V_0^t(\mathbf{p}' - \mathbf{p})\Phi_\nu^t(\mathbf{p}) \quad (\text{D.43})$$

Since the momentum-space integral equation corresponding to (D.26) has the form

$$(p'^2 + p_\kappa^2)\Phi_\nu^t(\mathbf{p}') = -\frac{2\beta_\nu}{(2\pi)^{d/2}} \int dp V_0^t(\mathbf{p}' - \mathbf{p})\Phi_\nu^t(\mathbf{p}) \quad (\text{D.44})$$

it follows that

$$[V_0\Phi_\nu]^t(\mathbf{p}) = -\frac{(p^2 + p_\kappa^2)}{2\beta_\nu} \Phi_\nu^t(\mathbf{p}) \quad (\text{D.45})$$

Finally, substituting (D.45) into (D.41), we obtain the momentum-space orthonormality relations for a set of generalized Sturmian basis functions:

$$\int dp \Phi_{\nu'}^{t*}(\mathbf{p}) \left(\frac{p^2 + p_\kappa^2}{2p_\kappa^2} \right) \Phi_\nu^t(\mathbf{p}) = \delta_{\nu',\nu} \quad (\text{D.46})$$

Because all of the functions $\Phi_\nu(\mathbf{x})$ in the generalized Sturmian basis set obey equation (D.26), the potential-weighted direct space orthonormality relations shown in equation (D.38) can be rewritten in the form

$$\int dx \Phi_{\nu'}^*(\mathbf{x}) \left(\frac{-\Delta + p_\kappa^2}{2p_\kappa^2} \right) \Phi_\nu(\mathbf{x}) = \delta_{\nu',\nu} \quad (\text{D.47})$$

so that the momentum-space and direct-space orthonormality relations can be seen to be related to each other in a symmetrical way. These weighted orthonormality relations in $L_2(\mathbb{R}^d)$ are the usual orthonormality relations in the Sobolev space $W_2^{(1)}(\mathbb{R}^d)$ (see [Weniger, 1985]). For the case of unequal masses, where

$$\Delta \equiv \sum_{j=1}^d \frac{1}{m_j} \frac{\partial^2}{\partial x_j^2} \quad (\text{D.48})$$

the momentum-space orthonormality relations for generalized Sturmians (D.46) takes on the slightly modified form

$$\int dp \Phi_{\nu'}^{t*}(\mathbf{p}) \left(\frac{\sum_j p_j^2/m_j + p_\kappa^2}{2p_\kappa^2} \right) \Phi_\nu^t(\mathbf{p}) = \delta_{\nu',\nu} \quad (\text{D.49})$$

D.7 Sturmian expansions of d -dimensional plane waves

If the set of generalized Sturmian basis functions is complete in the sense of spanning the Sobolev space $W_2^{(1)}(\mathbb{R}^d)$, we can use it to construct a weakly convergent expansion of a d -dimensional plane wave (valid only in the sense of distributions). Suppose that we let

$$e^{i\mathbf{p}\cdot\mathbf{x}} = \left(\frac{p_\kappa^2 + p^2}{2p_\kappa^2} \right) \sum_\nu \Phi_\nu^{t*}(\mathbf{p}) a_\nu(\mathbf{x}) \quad (\text{D.50})$$

We can then determine the unknown functions $a_\nu(\mathbf{x})$ by means of the orthonormality relations (D.46). Multiplying (D.50) on the left by $\Phi_{\nu'}^{t*}(\mathbf{p})$ and integrating over $d\mathbf{p}$ making use of (D.46), we obtain

$$\int d\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} \Phi_{\nu'}^{t*}(\mathbf{p}) = \sum_\nu \delta_{\nu',\nu} a_\nu(\mathbf{x}) = a_{\nu'}(\mathbf{x}) \quad (\text{D.51})$$

so that

$$a_\nu(\mathbf{x}) = \int d\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} \Phi_\nu^t(\mathbf{p}) = (2\pi)^{d/2} \Phi_\nu(\mathbf{x}) \quad (\text{D.52})$$

Thus finally we obtain an expansion of the form

$$e^{i\mathbf{p}\cdot\mathbf{x}} = (2\pi)^{d/2} \left(\frac{p_\kappa^2 + p^2}{2p_\kappa^2} \right) \sum_\nu \Phi_\nu^{t*}(\mathbf{p}) \Phi_\nu(\mathbf{x}) \quad (\text{D.53})$$

If the set of generalized Sturmians $\Phi_\nu(\mathbf{x})$ does not span $W_2^{(1)}(\mathbb{R}^d)$, equation (D.53) becomes

$$P [e^{i\mathbf{p}\cdot\mathbf{x}}] = (2\pi)^{d/2} \left(\frac{p_\kappa^2 + p^2}{2p_\kappa^2} \right) \sum_\nu \Phi_\nu^{t*}(\mathbf{p}) \Phi_\nu(\mathbf{x}) \quad (\text{D.54})$$

where $P [e^{i\mathbf{p}\cdot\mathbf{x}}]$ is the projection of the d -dimensional plane wave onto the subspace spanned by the set $\{\Phi_\nu(\mathbf{x})\}$. For example, if we are considering a system of N electrons, with $d = 3N$, the generalized Sturmian basis set might be antisymmetric with respect to exchange of the N electron coordinates but otherwise complete. In that case, $P [e^{i\mathbf{p}\cdot\mathbf{x}}]$ would represent the projection of the plane wave onto that part of Hilbert space corresponding to functions of \mathbf{x} that are antisymmetric with respect to exchange of the N electron coordinates. Neither the expansion shown in equation (D.53) nor that shown in equation (D.54) is point-wise convergent. In other words, we cannot perform the sums shown on the right-hand sides of these equations and expect them to give point-wise convergent representations of the plane wave or its projection. However, the expansions are valid in the sense of distributions. For the case of unequal masses, the generalized Sturmian plane wave expansion takes on the slightly modified form

$$e^{i\mathbf{p}\cdot\mathbf{x}} = (2\pi)^{d/2} \left(\frac{p_\kappa^2 + \sum_j p_j^2/m_j}{2p_\kappa^2} \right) \sum_\nu \Phi_\nu^{t*}(\mathbf{p}) \Phi_\nu(\mathbf{x}) \quad (\text{D.55})$$

D.8 An alternative expansion of a d-dimensional plane wave

In the Hamiltonian formulation of physics, one typically obtains sets of functions whose orthonormality relation has the form

$$\int dx \Phi_{\nu'}^*(\mathbf{x})\Phi_{\nu}(\mathbf{x}) = \delta_{\nu',\nu} \quad (\text{D.56})$$

Such a set of basis functions might, for example be the configurations resulting from the solution of the N -electron approximate Schrödinger equation

$$\left[-\frac{1}{2}\Delta + V_0(\mathbf{x}) - E_{\nu}\right]\Phi_{\nu}(\mathbf{x}) = 0 \quad (\text{D.57})$$

with $\mathbf{x} \equiv (x_1, x_2, \dots, x_d)$ and $d = 3N$. It is interesting to notice that a d -dimensional plane wave can also be expanded in terms of a basis set with orthonormality relations of the form shown in equation (D.56). To see this we write

$$e^{-i\mathbf{p}\cdot\mathbf{x}} = \sum_{\nu} a_{\nu}(\mathbf{p})\Phi_{\nu}^*(\mathbf{x}) \quad (\text{D.58})$$

Multiplying from the left by $\Phi_{\nu'}(\mathbf{x})$ and integrating over the coordinates, we obtain the relation

$$\begin{aligned} \int dx e^{-i\mathbf{p}\cdot\mathbf{x}}\Phi_{\nu'}(\mathbf{x}) &= \sum_{\nu} a_{\nu}(\mathbf{p}) \int dx \Phi_{\nu}^*(\mathbf{x})\Phi_{\nu'}(\mathbf{x}) \\ &= \sum_{\nu} a_{\nu}(\mathbf{p})\delta_{\nu',\nu} = a_{\nu'}(\mathbf{p}) = (2\pi)^{d/2}\Phi_{\nu'}^t(\mathbf{p}) \end{aligned} \quad (\text{D.59})$$

Thus we obtain the alternative expansion

$$e^{-i\mathbf{p}\cdot\mathbf{x}} = (2\pi)^{d/2} \sum_{\nu} \Phi_{\nu'}^t(\mathbf{p})\Phi_{\nu}^*(\mathbf{x}) \quad (\text{D.60})$$

or

$$e^{i\mathbf{p}\cdot\mathbf{x}} = (2\pi)^{d/2} \sum_{\nu} \Phi_{\nu'}^{t*}(\mathbf{p})\Phi_{\nu}(\mathbf{x}) \quad (\text{D.61})$$

The expansion (D.53) was a consequence of the weighted orthonormality relations obeyed by generalized Sturmian basis sets, while the expansion (D.61) resulted from the more conventional orthonormality relations (D.56). Both forms of the expansion are used in Chapter 8.

Appendix E

GENERALIZED STURMIANS APPLIED TO ATOMS

E.1 Goscinskian configurations

The Generalized Sturmian Method (Appendix D) is a newly-developed direct method for performing Configuration Interaction calculations on bound states [Avery, 1989]-[Avery and Avery, 2006]. It avoids the initial Hartree-Fock-Roothaan SCF calculation, and it is especially suitable for calculating large numbers of excited states of few-electron atoms or ions.

When the Generalized Sturmian Method is applied to atoms or atomic ions, it is convenient to use basis functions that are Slater determinants:

$$|\Phi_\nu\rangle = |\chi_\mu\chi_{\mu'}\chi_{\mu''}\cdots\rangle \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_\mu(\mathbf{x}_1) & \chi_{\mu'}(\mathbf{x}_1) & \chi_{\mu''}(\mathbf{x}_1) & \cdots \\ \chi_\mu(\mathbf{x}_2) & \chi_{\mu'}(\mathbf{x}_2) & \chi_{\mu''}(\mathbf{x}_2) & \cdots \\ \chi_\mu(\mathbf{x}_3) & \chi_{\mu'}(\mathbf{x}_3) & \chi_{\mu''}(\mathbf{x}_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \quad (\text{E.1})$$

built from hydrogenlike atomic spin-orbitals of the form

$$\chi_\mu(\mathbf{x}_i) \equiv \chi_{n,l,m,m_s}(\mathbf{x}_i) \equiv R_{n,l}(r_i)Y_{l,m}(\theta_i, \phi_i) \begin{cases} \alpha_i & m_s = 1/2 \\ \beta_i & m_s = -1/2 \end{cases} \quad (\text{E.2})$$

with weighted nuclear charges Q_ν . In other words, the atomic spin-orbitals have the form

Since the Slater determinant $|\Phi_\nu\rangle$ is an antisymmetrized product of atomic orbitals, all of which obey (E.8), it follows that

$$\begin{aligned} \left[-\frac{1}{2} \sum_{j=1}^N \nabla_j^2 \right] |\Phi_\nu\rangle &= \left[-\left(\frac{Q_\nu^2}{2n^2} + \frac{Q_\nu^2}{2n'^2} + \dots \right) + \left(\frac{Q_\nu}{r_1} + \frac{Q_\nu}{r_2} + \dots \right) \right] |\Phi_\nu\rangle \\ &= [E_\kappa - \beta_\nu V_0(\mathbf{x})] |\Phi_\nu\rangle \end{aligned} \quad (\text{E.9})$$

and thus equation (E.6) is satisfied. Each configuration $|\Phi_\nu\rangle$ has its own effective nuclear charge Q_ν . Within a particular configuration, the hydrogenlike atomic orbitals are orthonormal

$$\int d\tau_j \chi_{\mu'}^*(\mathbf{x}_j) \chi_\mu(\mathbf{x}_j) = \delta_{\mu',\mu} \quad (\text{E.10})$$

and they also obey the virial relationship

$$- \int d\tau_j |\chi_\mu(\mathbf{x}_j)|^2 \frac{Q_\nu}{r_j} = -\frac{Q_\nu^2}{n^2} \quad (\text{E.11})$$

From equations (E.6), (E.10) and (E.11), it can be shown [Avery, 2000], [Avery and Avery, 2006] that the generalized Sturmian configurations $|\Phi_\nu\rangle$ obey the potential-weighted orthonormality relation

$$\langle \Phi_{\nu'}^* | V_0 | \Phi_\nu \rangle = \delta_{\nu',\nu} \frac{2E_\kappa}{\beta_\nu} \quad (\text{E.12})$$

We next introduce the definitions

$$p_\kappa \equiv \sqrt{-2E_\kappa} \quad (\text{E.13})$$

and

$$\mathcal{R}_\nu \equiv \sqrt{\frac{1}{n^2} + \frac{1}{n'^2} + \dots} \quad (\text{E.14})$$

With the help of these definitions, equation (E.4) can be written in the form

$$Q_\nu = \beta_\nu Z = \frac{p_\kappa}{\mathcal{R}_\nu} \quad (\text{E.15})$$

The set of Sturmian configurations forms a set of isoenergetic solutions of the approximate Schrödinger equation (E.6), where the potential is weighted, and the weighting factors β_ν are chosen in such a way as to insure that all the solutions correspond to a common energy. From (E.13) we can see that their common energy E_κ is related to p_κ by

$$E_\kappa = -\frac{p_\kappa^2}{2} \quad (\text{E.16})$$

In previous publications we have called such atomic configurations *Goscinskian configurations* to recognize Prof. Osvaldo Goscinski's pioneering work in generalizing the concept

of Sturmian basis sets [Goscinski, 1968, 2003]. The non-relativistic Schrödinger equation of an N -electron atom has the form:

$$\left[-\frac{1}{2} \sum_{j=1}^N \nabla_j^2 + V(\mathbf{x}) - E_\kappa \right] |\Psi_\kappa\rangle = 0 \quad (\text{E.17})$$

where

$$V(\mathbf{x}) = V_0(\mathbf{x}) + V'(\mathbf{x}) \quad (\text{E.18})$$

Here $V_0(\mathbf{x})$ is the nuclear attraction potential shown in equation (E.7) while $V'(\mathbf{x})$ is the interelectron repulsion potential

$$V'(\mathbf{x}) = \sum_{i>j}^N \frac{1}{r_{ij}} \quad (\text{E.19})$$

We can try to build up the wave function from a superposition of Goscinskian configurations, i.e. from a superposition of isoenergetic solutions of the approximate wave equation (E.6), where V_0 is the nuclear attraction potential of the atom. Thus we write:

$$|\Psi_\kappa\rangle \approx \sum_{\nu} |\Phi_\nu\rangle C_{\nu,\kappa} \quad (\text{E.20})$$

Inserting this superposition into (E.17) we have

$$\sum_{\nu} \left[-\frac{1}{2} \Delta + V(\mathbf{x}) - E_\kappa \right] |\Phi_\nu\rangle C_{\nu,\kappa} \approx 0 \quad (\text{E.21})$$

However, each of the basis functions obeys (E.6), and therefore we can rewrite (E.21) in the form

$$\sum_{\nu} [V(\mathbf{x}) - \beta_\nu V_0(\mathbf{x})] |\Phi_\nu\rangle C_{\nu,\kappa} \approx 0 \quad (\text{E.22})$$

The energy term E_κ is now nowhere to be seen, and a remark is perhaps needed here to explain what has happened to it: The configurations in our Generalized Sturmian basis set are isoenergetic. They all correspond to the same energy, E_κ , since the weighting factors β_ν are chosen especially to make them do so. What we have done in going from (E.21) to (E.22) is to choose this energy to be the same as that which appears in (E.21). In other words, the energy to which all the members of our basis set correspond is chosen to be equal to the energy of the state that we are trying to approximate.

If we take the scalar product of (E.22) with a conjugate function from our basis set, we obtain the set of secular equations:

$$\sum_{\nu} \langle \Phi_{\nu'} | [V(\mathbf{x}) - \beta_\nu V_0(\mathbf{x})] |\Phi_\nu\rangle C_{\nu,\kappa} = 0 \quad (\text{E.23})$$

We now introduce the definitions:

$$T_{\nu',\nu}^0 \equiv -\frac{1}{p_\kappa} \langle \Phi_{\nu'}^* | V_0 | \Phi_\nu \rangle \quad (\text{E.24})$$

and

$$T'_{\nu',\nu} \equiv -\frac{1}{p_\kappa} \langle \Phi_{\nu'}^* | V' | \Phi_\nu \rangle \quad (\text{E.25})$$

From the potential-weighted orthonormality relations (E.12) we can see that

$$T_{\nu',\nu}^0 = \delta_{\nu'\nu} Z \mathcal{R}_\nu \quad (\text{E.26})$$

Notice that the nuclear attraction matrix $T_{\nu',\nu}^0$ is both diagonal and energy-independent. The interelectron repulsion matrix $T'_{\nu',\nu}$ can be evaluated using methods discussed in Appendix D, and it is also energy-independent. In order to see that $T'_{\nu',\nu}$ really is energy-independent, we notice that it is built up from terms of the form

$$\frac{1}{p_\kappa} J_{\mu_1,\mu_2,\mu_3,\mu_4} = \frac{1}{p_\kappa} \int d^3x \int d^3x' \rho_{\mu_1,\mu_2}(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{x}'|} \rho_{\mu_3,\mu_4}(\mathbf{x}') \quad (\text{E.27})$$

where densities are defined by

$$\begin{aligned} \rho_{\mu_1,\mu_2}(\mathbf{x}) &\equiv \chi_{\mu_1}^*(\mathbf{x}) \chi_{\mu_2}(\mathbf{x}) \\ \rho_{\mu_3,\mu_4}(\mathbf{x}') &\equiv \chi_{\mu_3}^*(\mathbf{x}') \chi_{\mu_4}(\mathbf{x}') \end{aligned} \quad (\text{E.28})$$

and where the orbitals are the hydrogenlike orbitals with weighted nuclear charge shown in equations (E.2) and (E.3). We now let

$$\begin{aligned} \mathbf{s} &\equiv p_\kappa \mathbf{x} \\ \mathbf{s}' &\equiv p_\kappa \mathbf{x}' \end{aligned} \quad (\text{E.29})$$

Then, making the substitution $Q_\nu \rightarrow p_\kappa/R_\nu$ in (E.3) we have

$$\begin{aligned} \rho_{\mu_1,\mu_2}(\mathbf{x}) &= p_\kappa^3 \tilde{\rho}_{\mu_1,\mu_2}(\mathbf{s}) \\ \rho_{\mu_3,\mu_4}(\mathbf{x}') &= p_\kappa^3 \tilde{\rho}_{\mu_3,\mu_4}(\mathbf{s}') \end{aligned} \quad (\text{E.30})$$

where $\tilde{\rho}_{\mu_1,\mu_2}(\mathbf{s})$ and $\tilde{\rho}_{\mu_3,\mu_4}(\mathbf{s}')$ are pure functions of \mathbf{s} and \mathbf{s}' respectively. Finally, noticing that

$$\frac{1}{p_\kappa |\mathbf{x} - \mathbf{x}'|} = \frac{1}{|\mathbf{s} - \mathbf{s}'|} \quad (\text{E.31})$$

we can write

$$\frac{1}{p_\kappa} J_{\mu_1,\mu_2,\mu_3,\mu_4} = \int d^3s \int d^3s' \tilde{\rho}_{\mu_1,\mu_2}(\mathbf{s}) \frac{1}{|\mathbf{s} - \mathbf{s}'|} \tilde{\rho}_{\mu_3,\mu_4}(\mathbf{s}') \quad (\text{E.32})$$

Since the building-blocks from which it is composed are independent of p_κ , the interelectron repulsion matrix $T'_{\nu',\nu}$ is also independent of p_κ and hence independent of energy. The energy-independent interelectron repulsion matrix $T'_{\nu',\nu}$ consists of pure numbers (in atomic units) which can be evaluated once and for all and stored.

With the help of equations (E.24)-(E.26), the secular equation (E.23) can be rewritten in the form:

$$\sum_{\nu} [-p_\kappa \delta_{\nu',\nu} Z \mathcal{R}_\nu - p_\kappa T'_{\nu',\nu} + \beta_\nu p_\kappa \delta_{\nu',\nu} Z \mathcal{R}_\nu] C_{\nu,\kappa} = 0 \quad (\text{E.33})$$

Finally, using the relationship

$$\beta_\nu Z \mathcal{R}_\nu = p_\kappa \quad (\text{E.34})$$

and dividing by p_κ , and reversing the signs, we obtain

$$\sum_{\nu} [\delta_{\nu',\nu} Z \mathcal{R}_\nu + T'_{\nu',\nu} - p_\kappa \delta_{\nu',\nu}] C_{\nu,\kappa} = 0 \quad (\text{E.35})$$

The Generalized Sturmian secular equation for atoms and atomic ions (E.35) differs in several remarkable ways from the secular equations that would be obtained using a Hamiltonian method:

1. The kinetic energy term has disappeared.
2. The nuclear attraction term, $\delta_{\nu',\nu} Z \mathcal{R}_\nu$, is diagonal.
3. The interelectron repulsion matrix $T'_{\nu',\nu}$ is energy-independent. It consists of dimensionless pure numbers.
4. Finally, the roots of the secular equations are not energies but values of the parameter p_κ , which is related to the energy spectrum through equation (E.16). The parameter $p_\kappa = \beta_\nu Z \mathcal{R}_\nu = Q_\nu \mathcal{R}_\nu$ can be thought of as a scaling parameter, since the effective nuclear charges associated with the Goscinskian configurations are proportional to it.
5. The configurations $|\Phi_\nu\rangle$ in the basis set are not fully determined until the secular equations have been solved. Only the form of the basis functions is known in advance, but not the scale. When the secular equation is solved, the resulting spectrum of p_κ values yields not only a spectrum of energies but a nearly optimum set of basis functions for the representation of each state. The basis set for the representation of highly excited states is diffuse, while the set for representation of tightly-bound states is contracted. The step of optimizing Slater exponents for each problem is thus not needed.
6. Once the energy-independent interelectron repulsion matrix $T'_{\nu',\nu}$ has been constructed, the properties of an entire isoelectronic series can be calculated with almost no additional effort.

E.2 Relativistic corrections

If the number of electrons N is kept constant while Z is allowed to increase, the energies calculated from the Generalized Sturmian secular equation approach those found by solution of the non-relativistic Schrödinger equation, but a relativistic correction must be added in order for the energies to approach experimental values. A crude relativistic correction can be found for a multiconfigurational state $\Psi_\kappa(\mathbf{x}) = \sum_\nu \Phi_\nu(\mathbf{x})C_{\nu\kappa}$ by calculating the ratio of the relativistic energy of the with interelectron repulsion entirely neglected to the non-relativistic energy, again with interelectron repulsion entirely neglected. The ratio can be written in the form

$$f_\kappa(Z) = \frac{E_{\kappa,\text{rel}}}{E_{\kappa,\text{nonrel}}} = \frac{\sum_\nu C_{\nu\kappa}^2 \langle \Phi_\nu | H_0 | \Phi_\nu \rangle_{\text{rel}}}{-\frac{1}{2}Z^2 \sum_\nu C_{\nu\kappa}^2 \mathcal{R}_\nu^2} \quad (\text{E.36})$$

Here

$$\langle \Phi_\nu | H_0 | \Phi_\nu \rangle_{\text{rel}} = \sum_{\mu \in \nu} \epsilon_{\mu,\text{rel}} \quad \mu = (n, l, m, m_s) \quad (\text{E.37})$$

is the relativistic energy of the configuration $\Phi_\nu(\mathbf{x})$ with interelectron repulsion entirely neglected, while

$$-\sum_{\mu \in \nu} \frac{1}{2} \frac{Z^2}{n^2} = -\frac{1}{2} Z^2 R_\nu^2 \quad \mu = (n, l, m, m_s) \quad (\text{E.38})$$

is the nonrelativistic energy of $\Phi_\nu(\mathbf{x})$. The quantity $\epsilon_{\mu,\text{rel}}$ represents the relativistic energy of a single electron moving in the attractive Coulomb potential of a nucleus with charge Z . This energy is easy to calculate exactly [Akhiezer and Berestetskii, 1965], if effects such as vacuum polarization and the Lamb shift are neglected. It is given by:

$$\epsilon_{\mu,\text{rel}} = \frac{c^2}{\left[1 + \left(\frac{Z}{c(\gamma+n-|j+1/2|)}\right)^2\right]^{1/2}} - c^2 \quad (\text{E.39})$$

$$\gamma \equiv \sqrt{\left(j + \frac{1}{2}\right)^2 - \left(\frac{Z}{c}\right)^2} \quad c = 137.036 \quad (\text{E.40})$$

where j is the total angular momentum (orbital plus spin) of a single electron, i.e. $l \pm \frac{1}{2}$. The corrected energy, $f_\kappa(Z)E_{\kappa,\text{nonrel}}$, agrees closely with the experimental values of energies, especially when Z is large compared with N .

The approximate relativistic correction discussed here is by no means confined to the Generalized Sturmian Method. It can be used in quantum calculations of every kind, performed on atoms and molecules. The assumption behind the correction is that relativistic effects are due mainly to the nuclear attraction part of the Hamiltonian, and only to a lesser extent to interelectron repulsion terms.

Table E.1: This table shows the relativistic correction for a single electron moving in the field of a nucleus with charge Z , i.e. the relativistic energy without the rest energy, divided by the non-relativistic energy. It is interesting to notice that the correction affects the 4th significant figure of the energy for values of Z as low as 10. In all cases the effect of the relativistic correction is to increase the binding energy.

n	j	$Z=1$	$Z=10$	$Z=20$	$Z=30$
1	$\frac{1}{2}$	1.00001	1.00133	1.00538	1.01228
2	$\frac{1}{2}$	1.00002	1.00167	1.00673	1.01537
2	$\frac{3}{2}$	1.00000	1.00033	1.00133	1.00301
3	$\frac{1}{2}$	1.00001	1.00133	1.00538	1.01226
3	$\frac{3}{2}$	1.00000	1.00044	1.00178	1.00402
3	$\frac{5}{2}$	1.00000	1.00015	1.00059	1.00133

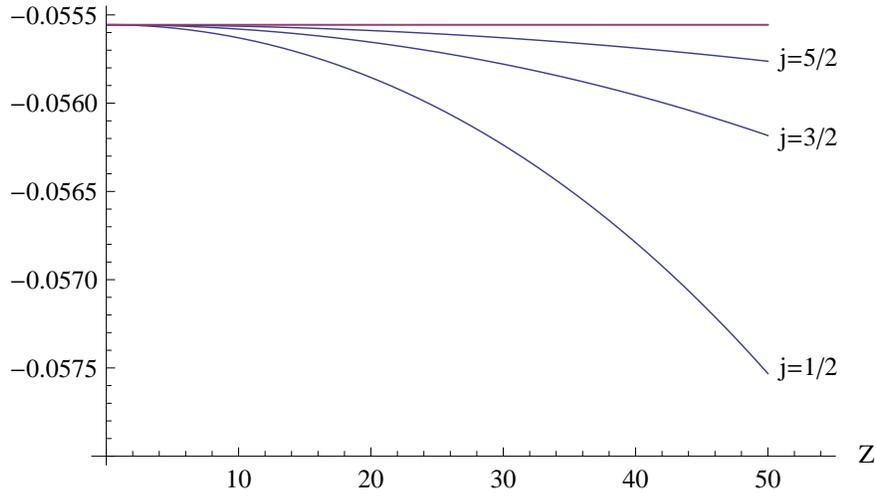


Figure E.1: This figure shows ϵ_μ/Z^2 for $n=3$. The non-relativistic energy is the horizontal line $\epsilon_{nonrel} = -1/18$, while the relativistic energies are shown by curves.

E.3 The large- Z approximation: Restriction of the basis set to an \mathcal{R} -block

If interelectron repulsion is entirely neglected, i.e. when disregarding the second term in Eq. (E.35), the calculated energies E_κ become those of a set of N completely independent electrons moving in the field of the bare nucleus:

$$E_\kappa = -\frac{p_\kappa^2}{2} \longrightarrow -\frac{1}{2}Z^2\mathcal{R}_\nu^2 = -\frac{Z^2}{2n_1^2} - \frac{Z^2}{2n_2^2} - \dots - \frac{Z^2}{2n_N^2} \quad (\text{E.41})$$

In the large- Z approximation, we do not neglect interelectron repulsion, but we restrict the basis set to those Goscinskian configurations that would be degenerate if interelectron repulsion were entirely neglected, i.e., we restrict the basis to a set of configurations all of which correspond to the same value of \mathcal{R}_ν . In that case, the first term in (E.35) is a multiple of the identity matrix, and the eigenvectors $C_{\nu\kappa}$ are the same as those that would be obtained by diagonalizing the energy-independent interelectron repulsion matrix $T'_{\nu'\nu}$, since the eigenfunctions of any matrix are unchanged by adding a multiple of the unit matrix. The simplified secular equation then becomes:

$$\sum_{\nu'} [T'_{\nu'\nu} - \lambda_\kappa \delta_{\nu'\nu}] C_{\nu\kappa} = 0 \quad (\text{E.42})$$

The roots are shifted by an amount equal to the constant by which the identity matrix is multiplied:

$$p_\kappa = Z\mathcal{R}_\nu + \lambda_\kappa = Z\mathcal{R}_\nu - |\lambda_\kappa| \quad (\text{E.43})$$

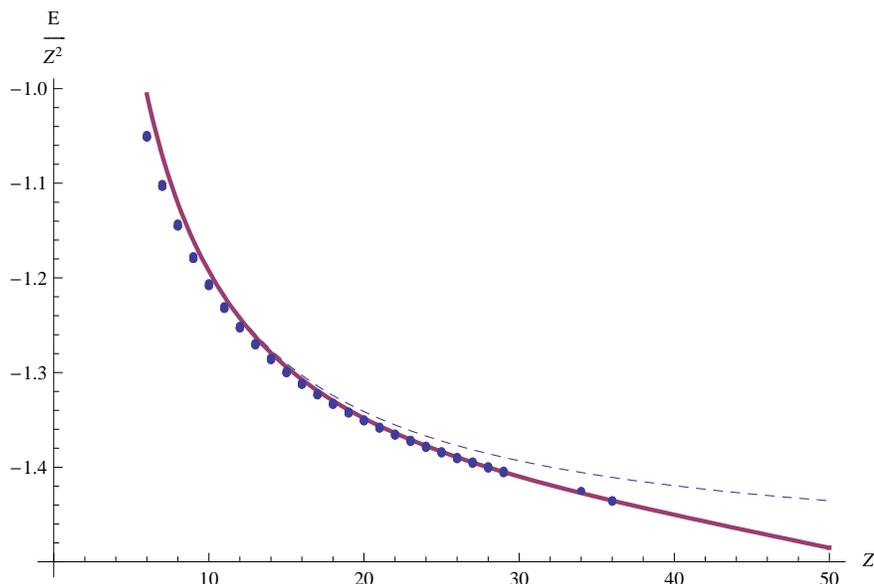


Figure E.2: The ground state of the carbon-like isoelectronic series, calculated in the large- Z approximation. The energies divided by Z^2 are shown as functions of Z . Experimental values are indicated by dots, while the energies calculated from equation (E.44) are shown as curves. The lower (solid) curve, which approaches the experimental values with increasing Z , has been corrected for relativistic effects. The upper (dashed) curve is uncorrected.

and the energies become

$$E_{\kappa} = -\frac{1}{2}(Z\mathcal{R}_{\nu} - |\lambda_{\kappa}|)^2 \quad (\text{E.44})$$

With the relativistic correction of equation (E.36), this becomes

$$E_{\kappa} = -f(Z)\frac{1}{2}(Z\mathcal{R}_{\nu} - |\lambda_{\kappa}|)^2 \quad (\text{E.45})$$

Since the roots λ_{κ} are always negative, we may use the form $-|\lambda_{\kappa}|$ in place of λ_{κ} to make explicit the fact that interelectron repulsion reduces the binding energies, as of course it must. The roots λ_{κ} are pure numbers that can be calculated once and for all and stored. From these roots, a great deal of information about atomic states can be found with very little effort.

E.4 Electronic potential at the nucleus in the large- Z approximation

The electronic potential $\varphi(\mathbf{x}_1)$ is related to the electronic density distribution by

$$\varphi(\mathbf{x}_1) = \int d^3x'_1 \frac{\rho(\mathbf{x}'_1)}{|\mathbf{x}_1 - \mathbf{x}'_1|} \quad (\text{E.46})$$

If the coordinate system is centered on the nucleus, the electronic potential at the nucleus is then given by

$$\varphi(0) = \int d^3x'_1 \frac{\rho(\mathbf{x}'_1)}{|\mathbf{x}'_1|} \quad (\text{E.47})$$

But the electron density corresponding to the state Ψ_κ is defined as

$$\rho(\mathbf{x}_1) = N \int ds_1 \int d^3x_2 \int ds_2 \cdots \int d^3x_N \int ds_N \Psi_\kappa^*(\mathbf{x}) \Psi_\kappa(\mathbf{x}) \quad (\text{E.48})$$

where the integral is taken over the spin coordinate of the first electron and over the space and spin coordinates of all the other electrons. The wave function $\Psi_\kappa(\mathbf{x}) = \sum_\nu \Phi_\nu(\mathbf{x}) B_{\nu\kappa}$ is a linear combination of Goscinskian configurations. Thus the density is given by

$$\rho(\mathbf{x}_1) = \sum_{\nu', \nu} \rho_{\nu'\nu}(\mathbf{x}_1) B_{\nu\kappa}^* B_{\nu\kappa} \quad (\text{E.49})$$

where

$$\begin{aligned} \rho_{\nu'\nu}(\mathbf{x}_1) &= N \int ds_1 \int d^3x_2 \int ds_2 \cdots \int d^3x_N \int ds_N \Phi_{\nu'}^*(\mathbf{x}) \Phi_\nu(\mathbf{x}) \\ &= \begin{cases} 0 & \text{if } \nu' \text{ and } \nu \text{ differ by 2 or more orbitals} \\ \chi_{\mu'}^*(\mathbf{x}_1) \chi_\mu(\mathbf{x}_1) & \text{if } \nu' \text{ and } \nu \text{ differ only by } \mu \rightarrow \mu' \\ \sum_{i=1}^N |\chi_{\mu_i}(\mathbf{x}_1)|^2 & \text{if } \nu' = \nu \end{cases} \end{aligned} \quad (\text{E.50})$$

In equation (E.50) we have made use of the fact that within an \mathcal{R} -block, the atomic spin-orbitals are orthonormal.

Within the framework of the large- Z approximation we have

$$\begin{aligned} \int dx \Psi_\kappa^*(\mathbf{x}) V_0(\mathbf{x}) \Psi_\kappa(\mathbf{x}) &= \sum_{\nu'} \sum_{\nu} B_{\nu'\kappa}^* B_{\nu\kappa} \int dx \Phi_{\nu'}^*(\mathbf{x}) V_0(\mathbf{x}) \Phi_\nu(\mathbf{x}) \\ &= -\frac{p_\kappa^2}{\beta_\nu} \sum_{\nu} |B_{\nu\kappa}|^2 \end{aligned} \quad (\text{E.51})$$

In the second step above, we make use of the potential weighted orthonormality relation (E.12). Further, since $\sum_{\nu} |B_{\nu\kappa}|^2 = 1$, equation (E.51) reduces to

$$\int dx \Psi_\kappa^*(\mathbf{x}) V_0(\mathbf{x}) \Psi_\kappa(\mathbf{x}) = -\frac{p_\kappa^2}{\beta_\nu} = -p_\kappa Z \mathcal{R}_\nu \quad (\text{E.52})$$

This result can be used to express the electronic potential at the nucleus in a very simple form. Combining (E.47) and (E.48), we obtain

$$\varphi(0) = N \int dx \frac{1}{|\mathbf{x}_1|} \Psi_\kappa^*(\mathbf{x}) \Psi_\kappa(\mathbf{x}) \quad (\text{E.53})$$

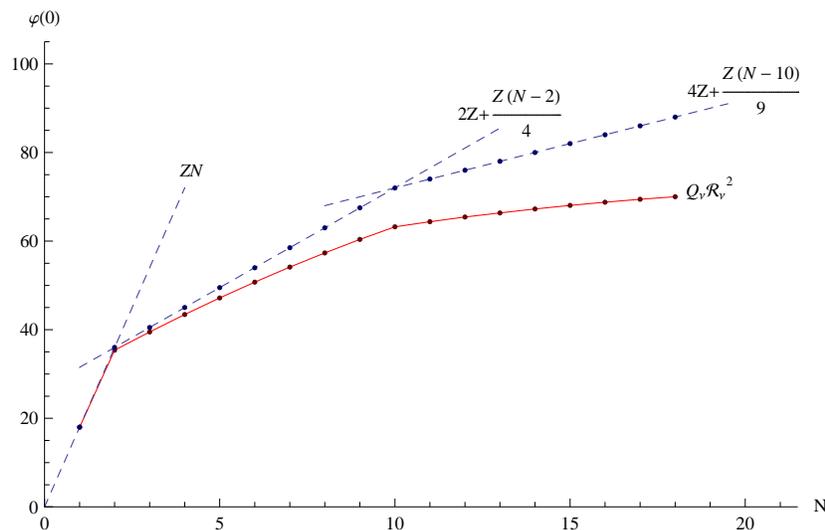


Figure E.3: When interelectron repulsion is entirely neglected, the electronic potential at the nucleus is given by $Z\mathcal{R}_\nu^2$, which is exactly piecewise linear in N . The effect of interelectron repulsion is to decrease $\varphi(0)$ and to make the dependence only approximately piecewise linear. The figure shows $\varphi(0)$ neglecting interelectron repulsion (upper values) and including it (lower values). The dots are calculated from the electronic densities of the ground state wave functions, whereas the lines are the closed form expressions found in equations (E.58) and (E.56).

From the definition of V_0 , equation (E.7), and from the fact that each term in the sum in (E.7) gives the same contribution, we have

$$\varphi(0) = -\frac{1}{Z} \int dx \Psi_\kappa^*(\mathbf{x}) V_0(\mathbf{x}) \Psi_\kappa(\mathbf{x}) \quad (\text{E.54})$$

Combining equations (E.54) and (E.52) we obtain the extremely simple result:

$$\varphi(0) = p_\kappa \mathcal{R}_\nu \quad (\text{E.55})$$

which can alternatively be written in the form:

$$\varphi(0) = Z\mathcal{R}_\nu^2 - |\lambda_\kappa| \mathcal{R}_\nu \quad (\text{E.56})$$

or in a third form:

$$\varphi(0) = Q_\nu \mathcal{R}_\nu^2 \quad (\text{E.57})$$

since $Q_\nu = Z - |\lambda_\kappa|/\mathcal{R}_\nu$. From equations (E.55)-(E.57) it follows that for an isonuclear series, the electronic potential at the nucleus depends on N in an approximately piecewise linear way. For example, let us consider the isonuclear series where $Z = 18$. Keeping the

nuclear charge Z constant at this value, we begin to add electrons. For the ground state we have:

$$\mathcal{R}_\nu^2 \equiv \frac{1}{n_1^2} + \frac{1}{n_2^2} + \cdots + \frac{1}{n_N^2} = \begin{cases} \frac{N}{1} & N \leq 2 \\ \frac{2}{1} + \frac{N-2}{4} & 2 \leq N \leq 10 \\ \frac{2}{1} + \frac{8}{4} + \frac{N-10}{9} & 10 \leq N \leq 18 \end{cases} \quad (\text{E.58})$$

E.5 Core ionization energies

The large- Z approximation can be used to calculate core-ionization energies, i.e. the energies required to remove an electron from the inner shell of an atom. From (E.44) we can see that this energy will be given by

$$\Delta E = \frac{1}{2} [(Z\mathcal{R}_\nu - |\lambda_\kappa|)^2 - (Z\mathcal{R}'_\nu - |\lambda'_\kappa|)^2] \quad (\text{E.59})$$

where the unprimed quantities refer to the original ground state, while the primed quantities refer to the core-ionized states. Since

$$\mathcal{R}_\nu^2 - \mathcal{R}'_\nu{}^2 = 1 \quad (\text{E.60})$$

Equation (E.59) can be written in the form

$$\Delta E - \frac{Z^2}{2} = Z [\mathcal{R}'_\nu |\lambda'_\kappa| - \mathcal{R}_\nu |\lambda_\kappa|] + \frac{|\lambda_\kappa|^2 - |\lambda'_\kappa|^2}{2} \quad (\text{E.61})$$

Thus we can see that within the framework of the large- Z approximation, the quantity $\Delta E - Z^2/2$ is linear in Z for an isoelectronic series. This quantity represents the contribution of interelectron repulsion to the core ionization energy, since if interelectron repulsion is completely neglected, the core ionization energy is given by $\Delta E = Z^2/2$. Core ionization energies calculated from equations (E.59)-(E.61) are shown in Figures E.4 through E.6.

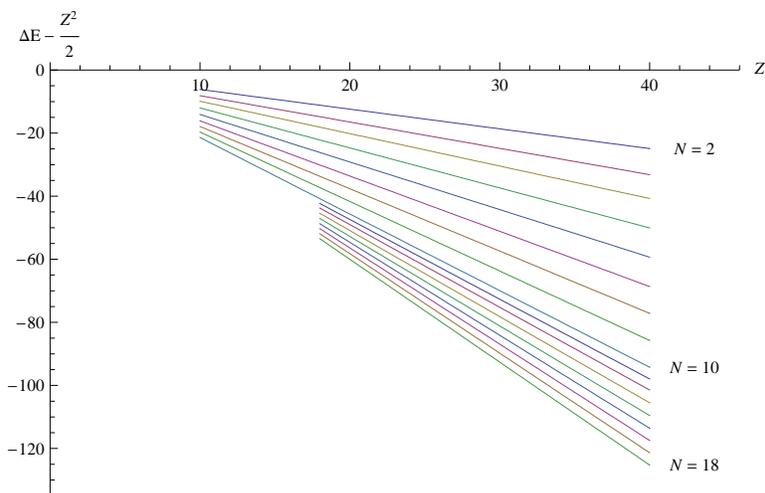


Figure E.4: For isoelectronic series, equation (E.61) indicates that within the large- Z approximation, the quantity $\Delta E - Z^2/2$ is exactly linear in Z , as is illustrated above. ΔE is the core ionization energy.

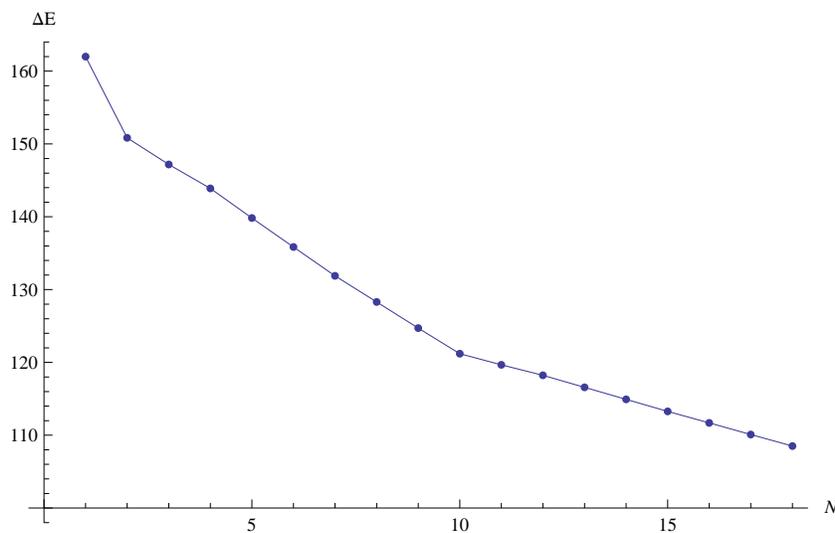


Figure E.5: For isonuclear series, the dependence of the core ionization energy on N is approximately piecewise linear. Whenever a new shell starts to fill, the slope of the line changes. The dots in the figure were calculated using equation (E.61), where it is not obvious that the dependence ought to be approximately piecewise linear. However, equations (E.58) and (E.56) can give us some insight into the approximately piecewise linear relationship.

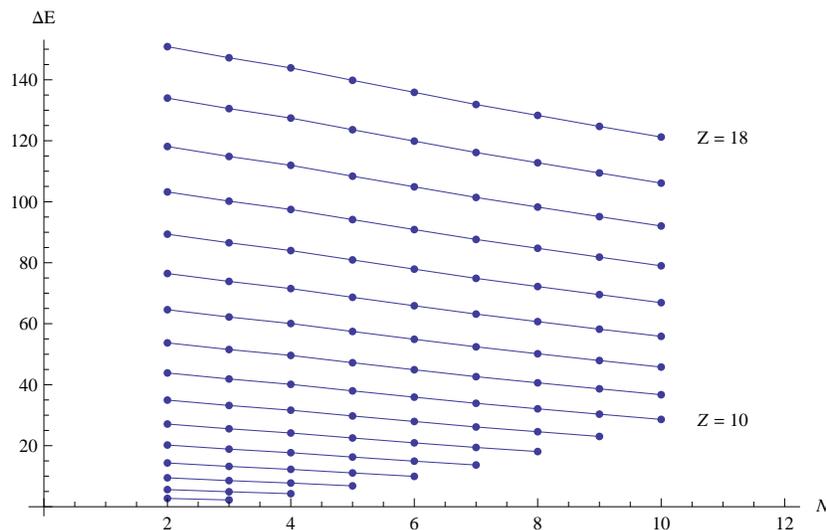


Figure E.6: This figure shows the dependence of the core-ionization energy on both N and Z for the filling of the $n = 2$ shell. Points with $N > Z$ are omitted because the large- Z approximation cannot be used for these points. The energies are in Hartrees.

E.6 Advantages and disadvantages of Goscinskian configurations

We seen that when $V_0(\mathbf{x})$ is chosen to be the the Coulomb attraction of the bare nucleus, the approximate Schrödinger equation

$$\left[-\frac{1}{2} \sum_{j=1}^N \nabla_j^2 + \beta_\nu V_0(\mathbf{x}) - E_\kappa \right] |\Phi_\nu\rangle = 0 \quad \beta_\nu V_0(\mathbf{x}) = - \sum_{j=1}^N \frac{Q_\nu}{r_j} \quad (\text{E.62})$$

can be solved exactly using configurations composed of hydrogenlike spin-orbitals with the especially chosen weighted charges Q_ν shown in equation (E.4). There is no need to calculate the weighting factors β_ν . These are obtained automatically when the secular equation is solved. Nor is there a need to normalize the configurations. This is also achieved automatically. Thus the choice of $V_0(\mathbf{x})$ as the potential of the bare nucleus has many advantages; but it also has disadvantages. Just as is the case in perturbation theory, convergence is most rapid if $V_0(\mathbf{x})$ is chosen to be as close as possible to the actual potential. By choosing $V_0(\mathbf{x})$ to be the Coulomb attraction of the bare nucleus, we have neglected interelectron repulsion. This is why the Generalized Sturmian Method with Goscinskian configurations works best when the number of electrons in an atom or ion is small, and why it works especially well when $Z \gg N$, i.e. when the Coulomb attraction of the nucleus dominates over the effects of interelectron repulsion.

To extend the range of applicability of the method to atoms and ions with large values of N , we would need to choose a $V_0(\mathbf{x})$ which included some of the effects of interelec-

tron repulsion. For example, we could let it be the Hartree potential. The approximate Schrödinger equation (E.6) can always be solved provided that it is separable, and it is separable whenever the approximate potential has the form

$$V_0(\mathbf{x}) = \sum_{j=1}^N v(\mathbf{x}_j) \quad (\text{E.63})$$

The separated form of (E.6) becomes:

$$\left[-\frac{1}{2} \nabla_j^2 + \beta_\nu v(\mathbf{x}_j) - \epsilon_\zeta \right] \varphi_\zeta(\mathbf{x}_j) = 0 \quad (\text{E.64})$$

where the weighting factors β_ν must be chosen in such a way that

$$\sum_{\zeta \in \nu} \epsilon_\zeta = E_\nu \quad (\text{E.65})$$

If the spin-orbitals $\varphi_\zeta(\mathbf{x}_j)$ satisfy (E.64), then configurations of the form

$$|\Phi_\nu\rangle = |\varphi_\zeta \varphi_{\zeta'} \varphi_{\zeta''} \cdots\rangle \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_\zeta(\mathbf{x}_1) & \varphi_{\zeta'}(\mathbf{x}_1) & \varphi_{\zeta''}(\mathbf{x}_1) & \cdots \\ \varphi_\zeta(\mathbf{x}_2) & \varphi_{\zeta'}(\mathbf{x}_2) & \varphi_{\zeta''}(\mathbf{x}_2) & \cdots \\ \varphi_\zeta(\mathbf{x}_3) & \varphi_{\zeta'}(\mathbf{x}_3) & \varphi_{\zeta''}(\mathbf{x}_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \quad (\text{E.66})$$

will satisfy the approximate Schrödinger equation (E.6). Some of the neatness of the Generalized Sturmian Method with Goscinskian configurations is certainly lost by choosing a $V_0(\mathbf{x})$ that includes effects of interelectron repulsion, but it could be worth paying this price in order to extend the method to atoms and atomic ions with larger values of N . We are at present exploring these possibilities, and some work in this direction is also being done by Prof. Gustavo Gasaneo and his group in Argentina.

E.7 \mathcal{R} -blocks, invariant subsets and invariant blocks

To tie the discussion of this chapter in with the general principles discussed in Chapter 1, we identify T with the operator whose roots and eigenfunctions we wish to study. The group of symmetry operations \mathcal{G} that leave the nuclear attraction and interelectron repulsion matrix of an atom invariant consists of rotations of the entire system about the nucleus, together with reflections and inversions that do not affect the interelectron distances. These operations do not affect the radial parts of the atomic orbitals from which the Goscinskian configurations are constructed, nor do they affect the spin. Thus the set of configurations, all of which are characterized by the same value of

$$\mathcal{R}_\nu \equiv \sqrt{\frac{1}{n^2} + \frac{1}{n'^2} + \frac{1}{n''^2} + \cdots} \quad (\text{E.67})$$

i.e. configurations all of which are built from hydrogenlike atomic spin-orbitals with a particular set of principal quantum numbers (n, n', n'', \dots) , is closed under \mathcal{G} , and it corresponds to an invariant subset as discussed in Chapter 1. The block of T' based on it corresponds to an invariant block. As expected, the eigenfunctions of interelectron repulsion matrix for the \mathcal{R} -blocks are the symmetry-adapted basis functions that we desire. In Chapter 1, we mentioned that when the roots of an invariant block are degenerate, then in order to take full advantage of the symmetry of the problem, we need to add an extremely small perturbation which will slightly remove the degeneracy. In the present case, this slight perturbation is given by

$$T_p = aL_z + bS_zF \quad (\text{E.68})$$

where a and b are two very small irrational numbers. (They are chosen to be irrational in order to avoid accidental degeneracies.) When this small perturbation is added to T' , the degeneracy is slightly removed. The eigenfunctions of $T' + T_p$ for an \mathcal{R} -block are then Russell-Saunders states, i.e. they are simultaneous eigenfunctions of the total angular momentum operator L^2 , its z -component L_z , the total spin operator S^2 , and its z -component S_z . We can ask how many linearly independent configurations there are in a ground-state \mathcal{R} -block. The answer is that when the Pauli principle is taken into account, the number of configurations m_k in an \mathcal{R} -block is given by the binomial coefficient

$$\binom{n_s}{N_v} = \frac{n_s!}{N_v!(n_s - N_v)!} \equiv m_k \quad (\text{E.69})$$

where n_s is the number of atomic spin-orbitals in the highest filled shell, while N_v is the number of valence electrons. For the ground-state configurations illustrated in Tables 4.3 and 4.4, N_v is the number of valence electrons. For example, for the ground state of lithium, we are starting to fill the $n = 2$ shell, for which $n_s = 8$. Since there is only one valence electron, $N_v = 1$ and

$$m_k \text{ for lithium ground state: } \binom{n_s}{N_v} = \binom{8}{1} = \frac{8!}{1!(8-1)!} = 8 \quad (\text{E.70})$$

Tables 4.3 and 4.4 show the roots of T' and their corresponding spectral terms for the first row of the periodic table. For the lithiumlike isoelectronic series, one of the roots corresponds to a 2S state and the other corresponds to a 2P state. We now remember that the total degeneracy of a Russell-Saunders state is given by $(2L + 1) \times (2S + 1)$, where L and S are respectively the quantum numbers of total orbital angular momentum and total spin. The 2S state is thus 2-fold degenerate, while the 2P state is 6-fold degenerate. By diagonalizing the ground-state \mathcal{R} -block for lithium, we thus obtain 8 Russell-Saunders states, the same as the number of Pauli-allowed states in the block. With a little effort, the reader can verify that the number of Russell-Saunders states shown in Tables 4.3 and 4.4 corresponds to the number of Pauli-allowed configurations given by the binomial coefficient in equation (E.69). Of course for the lithium ground state, the construction of

Table E.2: Eigenvalues of the 2-electron interelectron repulsion matrix $T'_{\nu',\nu}$ for $S=1$, $M_S=1$, $n = 2$ and $n'=3, 4, 5$.

$n'=3$ $ \lambda_\kappa $	term	$n' = 4$ $ \lambda_\kappa $	term	$n'=5$ $ \lambda_\kappa $	term
.108252	³ S	.077484	³ S	.056075	³ S
.134734	³ P	.087582	³ P	.065019	³ P
.135408	³ D	.090845	³ D	.061128	³ P
.138421	³ P	.093401	³ P	.063370	³ D
.155155	³ F	.099235	³ F	.067758	³ F
.160439	³ P	.099991	³ P	.067934	³ P
.165613	³ D	.104253	³ D	.070494	³ D
.168814	³ S	.106271	³ D	.071269	³ D
.173917	³ D	.107976	³ S	.072413	³ F
.186893	³ P	.108188	³ F	.072857	³ S
		.111210	³ G	.073295	³ G
		.111264	³ F	.073588	³ G
		.113313	³ P	.073920	³ F
		.114381	³ D	.074306	³ G
				.074578	³ H
				.074963	³ F
				.075173	³ P
				.075545	³ D

Russell-Saunders states is trivial, while for beryllium it can be accomplished with the help of Clebsch-Gordan coefficients. However for carbon, where

$$m_k \text{ for carbon ground state: } \binom{n_s}{N_v} = \binom{8}{4} = \frac{8!}{4!(8-4)!} = 70 \quad (\text{E.71})$$

the construction of the Russell-Saunders states is non-trivial. An indication of the nature of the states obtained is given in Table 4.5. As can easily be verified, the sum of the degeneracies shown in the table is equal to 70.

Table E.3: Roots of the ground state \mathcal{R} -block of the interelectron repulsion matrix for the Li-like, Be-like, B-like and C-like isoelectronic series.

Li-like $ \lambda_\kappa $	term	Be-like $ \lambda_\kappa $	term	B-like $ \lambda_\kappa $	term	C-like $ \lambda_\kappa $	term
0.681870	2S	0.986172	1S	1.40355	2P	1.88151	3P
0.729017	2P	1.02720	3P	1.44095	4P	1.89369	1D
		1.06426	1P	1.47134	2D	1.90681	1S
		1.09169	3P	1.49042	2S	1.91623	5S
		1.10503	1D	1.49395	2P	1.995141	3D
		1.13246	1S	1.52129	4S	1.96359	3P
				1.54037	2D	1.98389	3S
				1.55726	2P	1.98524	1D
						1.99742	1P
						2.04342	3P
						2.05560	1D
						2.07900	1S

Table E.4: Roots of the ground state \mathcal{R} -block of the interelectron repulsion matrix $T'_{\nu'\nu}$ for the N-like, O-like, F-like and Ne-like isoelectronic series.

N-like $ \lambda_\kappa $	term	O-like $ \lambda_\kappa $	term	F-like $ \lambda_\kappa $	term	Ne-like $ \lambda_\kappa $	term
2.41491	⁴ S	3.02641	³ P	3.68415	² P	4.38541	¹ S
2.43246	² D	3.03769	¹ D	3.78926	² S		
2.44111	² P	3.05065	¹ S				
2.49314	⁴ P	3.11850	³ P				
2.52109	² D	3.14982	¹ P				
2.53864	² S	3.24065	¹ S				
2.54189	² P						
2.61775	² P						

Table E.5: Eigenvalues of $T'_{\nu',\nu}$ for the carbon-like $\mathcal{R}_\nu = \sqrt{3}$ block.

$ \lambda_\kappa $	term	degen.	configuration
1.88151	^3P	9	$.994467(1s)^2(2s)^2(2p)^2 + .105047(1s)^2(2p)^4$
1.89369	^1D	5	$.994467(1s)^2(2s)^2(2p)^2 - .105047(1s)^2(2p)^4$
1.90681	^1S	1	$.979686(1s)^2(2s)^2(2p)^2 + .200537(1s)^2(2p)^4$
1.91623	^5S	5	$(1s)^2(2s)(2p)^3$
1.95141	^3D	15	$(1s)^2(2s)(2p)^3$
1.96359	^3P	9	$(1s)^2(2s)(2p)^3$
1.98389	^3S	3	$(1s)^2(2s)(2p)^3$
1.98524	^1D	5	$(1s)^2(2s)(2p)^3$
1.99742	^1P	3	$(1s)^2(2s)(2p)^3$
2.04342	^3P	9	$.105047(1s)^2(2s)^2(2p)^2 - .994467(1s)^2(2p)^4$
2.05560	^1D	5	$.105047(1s)^2(2s)^2(2p)^2 + .994467(1s)^2(2p)^4$
2.07900	^1S	1	$.200537(1s)^2(2s)^2(2p)^2 - .979686(1s)^2(2p)^4$

E.8 Invariant subsets based on subshells; Classification according to M_L and M_s

When we are dealing with configurations corresponding to excited states of an atom, the number of Pauli-allowed states in an \mathcal{R} -block may be very large, and we can ask whether our large basis set of Goscinskian configurations W can be divided into smaller invariant subsets W_k . A little thought reveals that this is indeed possible. The symmetry operations that leave T' invariant do not affect the radial parts $R_{n,l}(r)$ of the atomic orbitals, so sets of configurations built from the subshell sets $((n, l), (n'l'), (n'', l''), \dots)$ will be invariant subsets of our large basis set W under the operations of \mathcal{G} . The results will then differ slightly from those obtained in the large- Z approximation, which allows mixing between subshells. As an example of an invariant subset based on subshells we can consider the set of configurations corresponding to $(1s)(2d)^2$ for the lithium-like isoelectronic series:

$$(1s)(3d)^2 \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 10 \\ 2 \end{pmatrix} = 90 \quad (\text{E.72})$$

The invariant subset contains 90 configurations. Diagonalization of the 90×90 block yields the Russell Saunders states shown in Table 4.5.

We can also pick subsets W_k characterized by particular eigenvalues of S_z and L_z . These reductions in the size of the invariant subsets and the invariant blocks make it feasible to generate symmetry-adapted basis sets automatically also in the case of highly excited configurations. The use of symmetry-adapted basis sets leads to accurate calculations as is illustrated in Tables 4.7-4.9.

E.8. INVARIANT SUBSETS BASED ON SUBSHELLS; CLASSIFICATION ACCORDING TO M_L AND

Table E.6: This table shows the multiplets generated by diagonalizing the energy-invariant interelectron repulsion matrix T' for the 90×90 block the Hamiltonian based on neutral lithium configurations corresponding to $(1s)(3d)^2$, with $R_\nu = \sqrt{11}/3$. The reader can verify that the sum of the degeneracies of the multiplets is 90. The energies shown are for neutral lithium.

$ \lambda_\kappa $	term	degen.	energy
.270978	4F	28	-4.63798
.271649	2F	14	-4.63594
.278128	2D	10	-4.63526
.278998	4P	12	-4.61623
.279669	2P	6	-4.61359
.281871	2G	18	-4.61155
.297850	2S	2	-4.55650

Table E.7: 3S excited state energies calculated with 78 Goscinskians, using the crude relativistic correction described in the text. The calculation of similar tables for 1P , 3P , 1D , 3D , doubly excited autoionizing states, etc., is equally easy, rapid, and of comparable accuracy. Tables are given in Chapter 4 in [Avery and Avery, 2006], but may easily be reproduced using our programs, as shown in Tutorial 1 on [Avery and Avery, 2006a].

	He	Li ⁺	Be ²⁺	B ³⁺	C ⁴⁺	N ⁵⁺
1s2s 3S	-2.1737	-5.1085	-9.2957	-14.735	-21.427	-29.373
expt.	-2.1750	-5.1109	-9.2983	-14.738	-21.429	-29.375
1s3s 3S	-2.0683	-4.7509	-8.5459	-13.454	-19.476	-26.612
expt.	-2.0685	-4.7522	-8.5480	-13.457	-19.478	-26.614
1s4s 3S	-2.0364	-4.6365	-8.2999	-13.027	-18.820	-25.678
expt.	-2.0363	-4.6373	-8.3015	-13.030	-18.822	-25.680
1s5s 3S	-2.0226	-4.5859	-8.1896	-12.835	-18.522	-25.253
expt.	-2.0224	-4.5862	-8.1905		-18.524	-25.254
1s6s 3S	-2.0154	-4.5591	-8.1309	-12.732	-18.363	-25.024
expt.	-2.0152	-4.5592			-18.364	
1s7s 3S	-2.0112	-4.5432	-8.096	-12.67	-18.267	-24.888
expt.	-2.0109	-4.5431			-18.268	
1s8s 3S	-2.0085	-4.5330	-8.0736	-12.631	-18.206	-24.799
expt.	-2.0082	-4.5328			-18.206	
1s9s 3S	-2.0067	-4.5261	-8.0583	-12.604	-18.164	-24.739
expt.	-2.0064					
1s10s 3S	-2.0051	-4.5212	-8.0475	-12.585	-18.134	-24.696
expt.	-2.0051					

E.8. INVARIANT SUBSETS BASED ON SUBSHELLS; CLASSIFICATION ACCORDING TO M_L AND

Table E.8: 1S ground and excited state energies (in Hartrees) for the 2-electron isoelectronic series. The basis set used consisted of 592 generalized Sturmians of the Goscinski type, using the crude relativistic correction described in the text. The whole table was computed approximately a second. Experimental values are taken from the NIST tables [National Institute for Standards and Technology (NIST)] (<http://physics.nist.gov/asd>).

	He	Li ⁺	Be ²⁺	B ³⁺	C ⁴⁺
1s ² ¹ S	-2.8956	-7.2716	-13.649	-22.028	-32.412
expt.	-2.9034	-7.2798	-13.657	-22.035	-32.416
1s2s ¹ S	-2.1441	-5.0350	-9.1768	-14.571	-21.218
expt.	-2.1458	-5.0410	-9.1860	-14.582	-21.230
1s3s ¹ S	-2.0607	-4.7303	-8.5112	-13.405	-19.414
expt.	-2.0611	-4.7339	-8.5183	-13.415	-19.425
1s4s ¹ S	-2.0333	-4.6280	-8.2844	-13.005	-18.791
expt.	-2.0334	-4.6299	-8.2891		-18.800
1s5s ¹ S	-2.0221	-4.5858	-8.1889	-12.833	-18.520
expt.	-2.0210	-4.5825			-18.513
1s6s ¹ S	-2.0147	-4.5579	-8.129	-12.729	-18.359
expt.	-2.0144	-4.5571			
1s7s ³ S	-2.0109	-4.5426	-8.0951	-12.67	-18.281
expt.	-2.0104	-4.5418			
1s8s ³ S	-2.0083	-4.5326	-8.0732	-12.641	-18.26
expt.	-2.0079				
1s9s ³ S	-2.0065	-4.5258	-8.0583	-12.626	-18.203
expt.	-2.0062				
1s10s ³ S	-2.0049	-4.521	-8.0476	-12.602	-18.162
expt.	-2.0050				

Table E.9: 1D excited state energies for the 2-electron isoelectronic series, compared with experimental values taken from the NIST tables.

	He	Li ⁺	Be ²⁺	B ³⁺	C ⁴⁺	N ⁵⁺
1s3d 1D	-2.0555	-4.7218	-8.4990	-13.388	-19.387	-26.498
expt.	-2.0554	-4.7225	-8.5012	-13.392	-19.396	-26.514
1s4d 1D	-2.0312	-4.6246	-8.2801	-12.998	-18.779	-25.622
expt.	-2.0311	-4.6252	-8.2824	-13.003	-18.788	-25.639
1s5d 1D	-2.0200	-4.5797	-8.1790	-12.818	-18.497	-25.217
expt.	-2.0198	-4.5801	-8.1807		-18.507	-25.234
1s6d 1D	-2.0139	-4.5554	-8.1242	-12.721	-18.345	-24.997
expt.	-2.0137	-4.5557			-18.354	
1s7d 1D	-2.0102	-4.5407	-8.0912	-12.662	-18.253	-24.865
expt.	-2.0100	-4.5409			-18.262	
1s8d 1D	-2.0078	-4.5311	-8.0699	-12.624	-18.194	-24.779
expt.	-2.0076	-4.5314			-18.202	
1s9d 1D	-2.0062	-4.5246	-8.0552	-12.598	-18.153	-24.720
expt.	-2.0060					
1s10d 1D	-2.0050	-4.5199	-8.0447	-12.579	-18.124	-24.678
expt.	-2.0048					
1s11d 1D	-2.0041	-4.5165	-8.0370	-12.566	-18.102	-24.647
expt.	-2.0035					
1s12d 1D	-2.0032	-4.5139	-8.0311	-12.555	-18.086	-24.624
expt.	-2.0033					

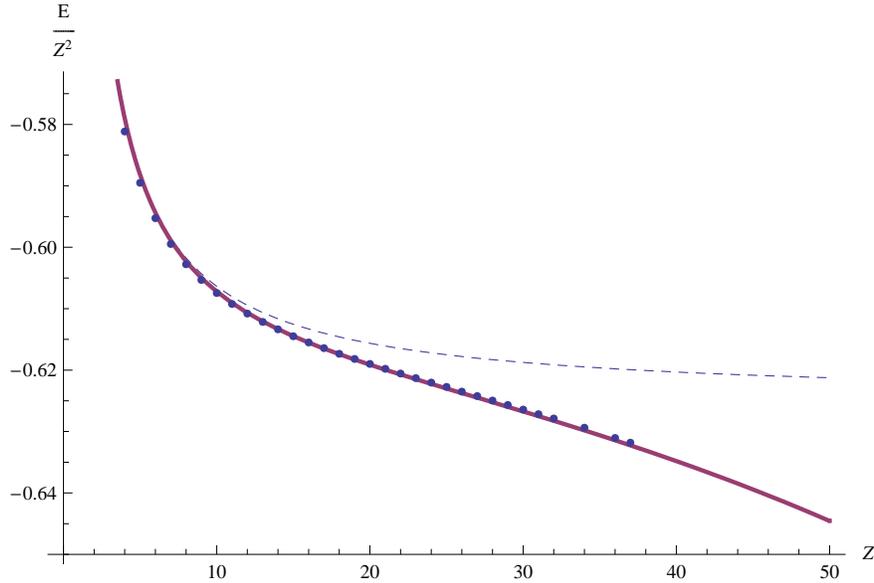


Figure E.7: Here the lowest 3S energies of the heliumlike isoelectronic series are divided by Z^2 . The lower line is corrected for relativistic effects. The dots are experimental values.

E.9 An atom surrounded by point charges

For a heavy atom surrounded by lighter atoms (ligands), the effect of the ligands is sometimes approximated by considering the atom to be surrounded by an array of point charges with an appropriate symmetry. The potential experienced by such an atom is

$$V(\mathbf{x}) = V_0(\mathbf{x}) + V'(\mathbf{x}) + V''(\mathbf{x}) \quad (\text{E.73})$$

where V_0 is the nuclear attraction potential

$$V_0(\mathbf{x}) = - \sum_{i=1}^N \frac{Z_0}{r_i} \quad (\text{E.74})$$

and V' is the interelectron repulsion potential

$$V'(\mathbf{x}) = \sum_{j>i}^N \sum_{i=1}^N \frac{1}{r_{ij}} \quad (\text{E.75})$$

while V'' (the “crystal field”) expresses the effect of a set of charges q_a located at the positions \mathbf{X}_a :

$$V''(\mathbf{x}) = - \sum_{i=1}^{\mathcal{N}} \sum_a \frac{q_a}{|\mathbf{x}_i - \mathbf{X}_a|} \quad (\text{E.76})$$

The generalized Sturmian secular equation analogous to (E.35) then becomes

$$\sum_{\nu} [\delta_{\nu',\nu} Z \mathcal{R}_{\nu} + T'_{\nu',\nu} + T''_{\nu',\nu} - p_{\kappa} \delta_{\nu',\nu}] C_{\nu,\kappa} = 0 \quad (\text{E.77})$$

where

$$T''_{\nu',\nu} \equiv -\frac{1}{p_\kappa} \langle \Phi_{\nu'}^* | V'' | \Phi_\nu \rangle \quad (\text{E.78})$$

In order to evaluate $T''_{\nu',\nu}$ we must first calculate one-electron matrix elements of the form

$$\begin{aligned} v''_{\mu_1,\mu_2} &\equiv \int d^3x \chi_{\mu_1}^*(\mathbf{x}) \chi_{\mu_2}(\mathbf{x}) \sum_a \frac{q_a}{|\mathbf{x} - \mathbf{X}_a|} \\ &= \sum_a \sum_l q_a \int d^3x_i \chi_{\mu_1}^*(\mathbf{x}) \chi_{\mu_2}(\mathbf{x}) \left(\frac{r_{<}^l}{r_{>}^{l+1}} \right)_a P_l(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) \\ &= \sum_a \sum_l q_a \int_0^\infty dr r^2 \left(\frac{r_{<}^l}{r_{>}^{l+1}} \right)_a R_{n_1,l_1}(r) R_{n_2,l_2}(r) \\ &\quad \times \int d\Omega Y_{l_1,m_1}^*(\hat{\mathbf{x}}) Y_{l_2,m_2}(\hat{\mathbf{x}}) P_l(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) \end{aligned} \quad (\text{E.79})$$

where

$$\left(\frac{r_{<}^l}{r_{>}^{l+1}} \right)_a \equiv \begin{cases} r^l / R_a^{l+1} & r < R_a \\ R_a^l / r^{l+1} & R_a < r \end{cases} \quad (\text{E.80})$$

If the points \mathbf{X}_a are all equidistant from the central atom and if the charges are all equal, then equation (E.79) can be rewritten in the form:

$$\begin{aligned} v''_{\mu_1,\mu_2} &= \sum_a \sum_l \frac{4\pi q_a}{2l+1} \int_0^\infty dr r^2 \left(\frac{r_{<}^l}{r_{>}^{l+1}} \right)_a R_{n_1,l_1}(r) R_{n_2,l_2}(r) \\ &\quad \times Y_{l,m}^*(\hat{\mathbf{X}}_a) \int d\Omega Y_{l_1,m_1}^*(\hat{\mathbf{x}}) Y_{l_2,m_2}(\hat{\mathbf{x}}) Y_{l,m}(\hat{\mathbf{x}}) \\ &\equiv \sum_l F_l \sum_a Y_{l,m}^*(\hat{\mathbf{X}}_a) \int d\Omega Y_{l_1,m_1}^*(\hat{\mathbf{x}}) Y_{l_2,m_2}(\hat{\mathbf{x}}) Y_{l,m}(\hat{\mathbf{x}}) \end{aligned} \quad (\text{E.81})$$

where

$$F_l \equiv \frac{4\pi q_a}{2l+1} \int_0^\infty dr r^2 \left(\frac{r_{<}^l}{r_{>}^{l+1}} \right)_a R_{n_1,l_1}(r) R_{n_2,l_2}(r) \quad (\text{E.82})$$

Once we are in possession of v''_{μ_1,μ_2} , the matrix element T''_{ν_1,ν_2} can be evaluated by means of the generalized Slater-Condon rules discussed in Appendix D, Section D.1. For fixed angular geometry, the matrix T''_{ν_1,ν_2} turns out to be a function of the parameters

$$s_a \equiv p_\kappa R_a \equiv p_\kappa |\mathbf{X}_a| \quad (\text{E.83})$$

where $|\mathbf{X}_a|$ is the distance of the a th charge q_a from the central atom. The generalized Sturmian secular equations (E.77) can be solved as follows: We begin by picking values of s_a . The secular equation (E.77) is then solved, yielding roots p_κ for the ground state and the excited states. The corresponding values of $|\mathbf{X}_a|$ are then known. This can be repeated for a number of s_a values, yielding solutions as functions of the distances $|\mathbf{X}_a|$. As an example, we can consider an atom surrounded by 8 equal point charges q at the positions

$$\begin{aligned}\hat{\mathbf{X}}_1 &= (+1, +1, +1)/\sqrt{3} & \hat{\mathbf{X}}_5 &= (+1, -1, -1)/\sqrt{3} \\ \hat{\mathbf{X}}_2 &= (-1, +1, +1)/\sqrt{3} & \hat{\mathbf{X}}_6 &= (-1, +1, -1)/\sqrt{3} \\ \hat{\mathbf{X}}_3 &= (+1, -1, +1)/\sqrt{3} & \hat{\mathbf{X}}_7 &= (-1, -1, +1)/\sqrt{3} \\ \hat{\mathbf{X}}_4 &= (+1, +1, -1)/\sqrt{3} & \hat{\mathbf{X}}_8 &= (-1, -1, -1)/\sqrt{3}\end{aligned}\quad (\text{E.84})$$

which has cubic symmetry. We can also discuss the effect of a square-planar array

$$\begin{aligned}\hat{\mathbf{X}}_1 &= (+1, 0, 0) & \hat{\mathbf{X}}_3 &= (0, +1, 0) \\ \hat{\mathbf{X}}_2 &= (-1, 0, 0) & \hat{\mathbf{X}}_4 &= (0, -1, 0)\end{aligned}\quad (\text{E.85})$$

a linear array:

$$\hat{\mathbf{X}}_1 = (0, 0, +1) \quad \hat{\mathbf{X}}_2 = (0, 0, -1) \quad (\text{E.86})$$

a trigonal array:

$$\begin{aligned}\hat{\mathbf{X}}_1 &= (2\sqrt{2}/3, 0, -1/3) \\ \hat{\mathbf{X}}_2 &= (-\sqrt{2}/3, \sqrt{6}/3, -1/3) \\ \hat{\mathbf{X}}_3 &= (-\sqrt{2}/3, -\sqrt{6}/3, -1/3)\end{aligned}\quad (\text{E.87})$$

and a tetrahedral array

$$\begin{aligned}\hat{\mathbf{X}}_1 &= (0, 0, 1) \\ \hat{\mathbf{X}}_2 &= (2\sqrt{2}/3, 0, -1/3) \\ \hat{\mathbf{X}}_3 &= (-\sqrt{2}/3, \sqrt{2}/3, -1/3) \\ \hat{\mathbf{X}}_4 &= (-\sqrt{2}/3, -\sqrt{2}/3, -1/3)\end{aligned}\quad (\text{E.88})$$

The presence of a symmetrical set of charges surrounding a central atom reduces the symmetry of the system: Instead of being invariant under the the full group of rotations about the central atom (as well as rotations in spin space) the system is now invariant only under a restricted group of rotations. However, subsets of configurations that are closed under the operations of the full rotation group, are also closed under the more restricted group of rotations. Therefore, if we are willing to have invariant subsets of configurations that are larger than minimal, we can choose them almost as though the

perturbing surrounding charges were not there. Thus the invariant subsets in the perturbed case can be established by following the prescription:

$$\begin{aligned}
 \nu &= (a; b) \\
 a &= (M_S, (n, l), (n', l'), (n'', l''), \dots) \\
 b &= ((m, m_s), (m', m'_s), (m'', m''_s), \dots) \\
 W_a &= (|\Phi_{a,b_1}\rangle, |\Phi_{a,b_2}\rangle, |\Phi_{a,b_3}\rangle, \dots)
 \end{aligned} \tag{E.89}$$

The symmetry-adapted basis functions found by diagonalizing the invariant blocks can be used as basis functions for a more ambitious configuration interaction calculation. For example, let us think of a neutral lithium atom, surrounded by a square planar array of point charges. To make the example more specific, let us think of the set of configurations characterized by

$$(1s)^2(3d) \quad \begin{pmatrix} 10 \\ 1 \end{pmatrix} \tag{E.90}$$

This set has 10 members, and they can be used as a basis for solving the generalized Sturmian secular equations in the presence of the square-planar array of point charges. In the absence of the charges, the 10 configurations are members of a ²D multiplet, and in the large-Z approximation its energy is

$$E_\kappa = -\frac{1}{2}(ZR_\nu - |\lambda|)^2 = -\frac{1}{2}(\sqrt{19} - .58281)^2 = -7.1294 \tag{E.91}$$

This corresponds to the effective charge

$$Q = \frac{p_\kappa}{R_\nu} = \frac{ZR_\nu - |\lambda|}{R_\nu} = 2.5989 \tag{E.92}$$

where

$$R_\nu = \sqrt{\frac{1}{1^2} + \frac{1}{1^2} + \frac{1}{3^2}} = \frac{\sqrt{19}}{3} \quad |\lambda| = .58281 \tag{E.93}$$

In order to calculate T'' and find the effect of the point charges (which we treat here as a small perturbation), we need to evaluate the radial integrals:

$$\int_0^\infty dr r^2 \left(\frac{r^l}{r^{l+1}} \right)_a |R_{3,2}(r)|^2 = \begin{cases} .19193 & l = 0 \\ .10363 & l = 2 \\ .06826 & l = 4 \end{cases} \tag{E.94}$$

with $|\mathbf{X}_a| = 5$ and

$$R_{3,2}(r) = \frac{2Q^{7/2}}{81} \sqrt{\frac{2}{15}} r^2 e^{-Qr/3} \tag{E.95}$$

The angular integrals needed for constructing T''' are

$$\begin{aligned}
\sum_a \int d\Omega Y_{2,m_1}^*(\hat{\mathbf{x}}) Y_{2,m_2}(\hat{\mathbf{x}}) P_0(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) &= 4\delta_{m_1,m_2} \\
\sum_a \int d\Omega Y_{2,m_1}^*(\hat{\mathbf{x}}) Y_{2,m_2}(\hat{\mathbf{x}}) P_1(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) &= 0 \\
\sum_a \int d\Omega Y_{2,m_1}^*(\hat{\mathbf{x}}) Y_{2,m_2}(\hat{\mathbf{x}}) P_3(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) &= 0
\end{aligned} \tag{E.96}$$

and

$$\begin{aligned}
&\sum_a \int d\Omega Y_{2,m_1}^*(\hat{\mathbf{x}}) Y_{2,m_2}(\hat{\mathbf{x}}) P_2(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) \\
&= \begin{pmatrix} \frac{4}{7} & 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{7} & 0 & 0 & 0 \\ 0 & 0 & -\frac{4}{7} & 0 & 0 \\ 0 & 0 & 0 & -\frac{2}{7} & 0 \\ 0 & 0 & 0 & 0 & \frac{4}{7} \end{pmatrix}
\end{aligned} \tag{E.97}$$

while

$$\begin{aligned}
&\sum_a \int d\Omega Y_{2,m_1}^*(\hat{\mathbf{x}}) Y_{2,m_2}(\hat{\mathbf{x}}) P_4(\hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_a) \\
&= \begin{pmatrix} \frac{1}{14} & 0 & 0 & 0 & \frac{5}{6} \\ 0 & -\frac{2}{7} & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{7} & 0 & 0 \\ 0 & 0 & 0 & -\frac{2}{7} & 0 \\ \frac{5}{6} & 0 & 0 & 0 & \frac{1}{14} \end{pmatrix}
\end{aligned} \tag{E.98}$$

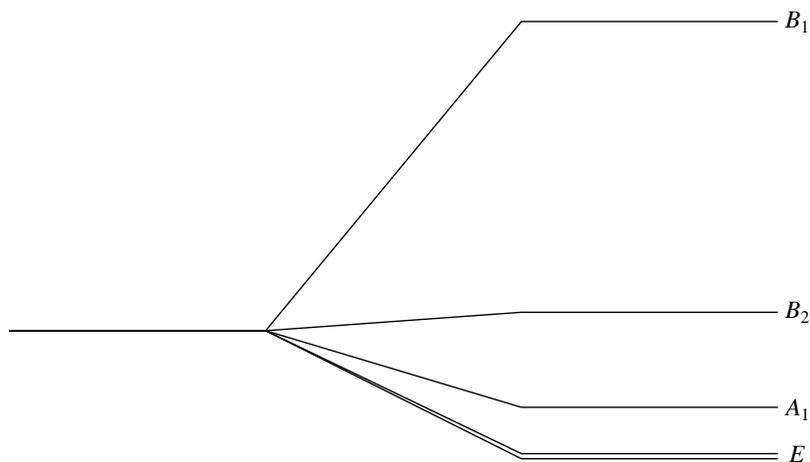


Figure E.8: This diagram shows schematically the splitting of the ${}^2D_{1/2}$ multiplet of neutral lithium discussed above in the field of a square-planar array of point charges. The ${}^2D_{-1/2}$ multiplet which is similarly split is not shown. Besides being split by the presence of the charges, the multiplet is also displaced in energy, as is the lithium ground state.

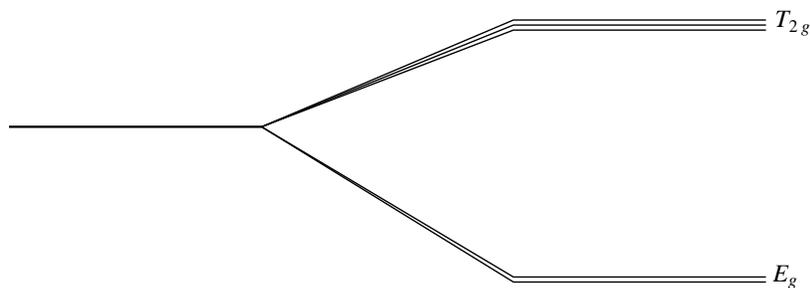


Figure E.9: This figure shows the same multiplet as Figure 4.8, but the splitting is due to an octohedral array of point charges.

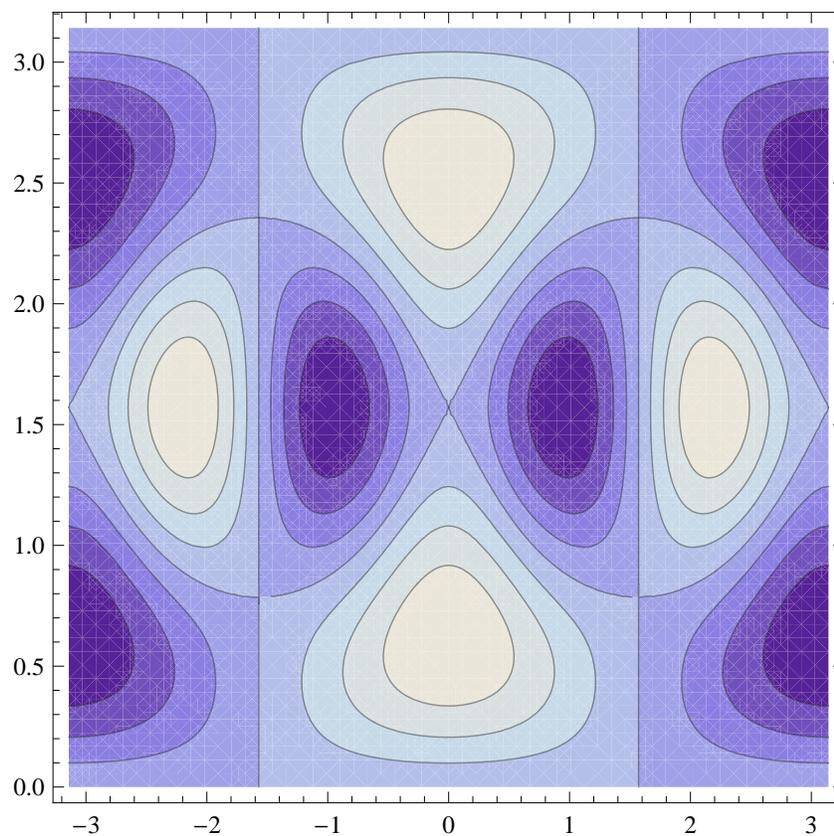


Figure E.10: This figure shows the angular dependence of one of the T_{1u} orbitals produced when degenerate f-orbitals are split by the presence of an octohedral arrangement of point charges. The T_{1u} irreducible representation of the group O_h is a 3-dimensional representation. The other two components are shown in Figure 4.11.

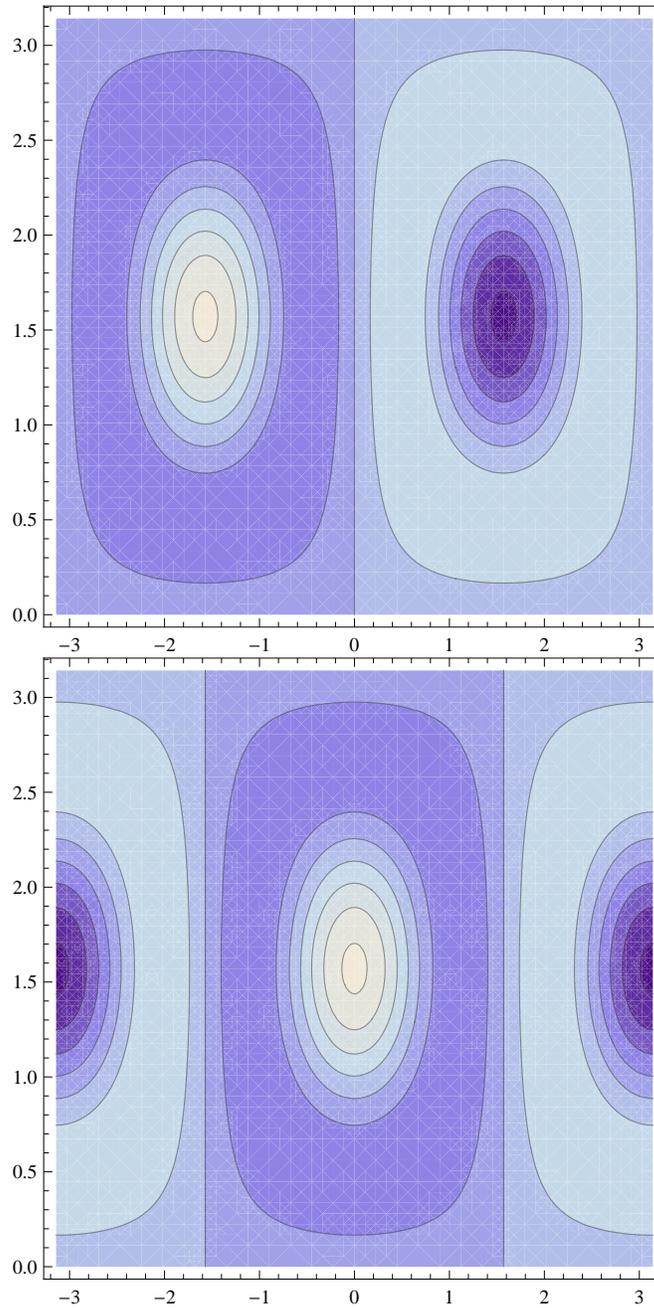


Figure E.11: This figure shows the angular dependence of the remaining two T_{1u} orbitals. Both in this figure and in Figure 4.10, the vertical axis represents θ and the horizontal axis represents ϕ .

Appendix F

MOLECULAR ORBITALS BASED ON STURMIANS

F.1 The one-electron secular equation

Molecular orbitals may be represented as superpositions of Coulomb Sturmian basis functions (Appendix D) centered on the nuclei of a molecule. These basis functions are an example of Exponential-Type Orbitals (ETO's) [Harris and Michels, 1967], [Pinchon and Hoggan 2007], [Pinchon and Hoggan, 2009], [Weatherford, 1982], and calculations using them can potentially be much more accurate than calculations based on Gaussians.

Gaussian basis functions have serious drawbacks, since very many of them are needed to approximate the molecular orbitals, and since the cusp at the nucleus is never adequately represented. Furthermore, Gaussian basis functions cannot accurately represent the exponential decay of the orbitals at large distances from the nuclei. Thus while the mainstream effort of quantum chemistry today follows the path of Gaussian technology, there exists a small group of researchers who struggle with the difficult mathematical problems involved in using exponential-type orbitals (ETO's) as basis functions, and we hope that the present chapter will make a contribution to this effort. We will see that the automatic scaling properties associated with the Generalized Sturmian Method have advantages also in the case of molecules, and that molecular orbitals based on many-center Coulomb Sturmians have advantages over other ETO's with respect to the ease of evaluation of interelectron repulsion integrals.

We will first consider the use of Coulomb Sturmian basis functions located on the different atoms of a molecule to solve the 1-electron molecular Schrödinger equation, an endeavor which was pioneered by C.E. Wulfman, B. Judd, T. Koga, V. Aquilanti and others [Shibuya and Wulfman, 1965], [Wulfman, 2011], [Judd, 1975], [Koga et al., 1984]-[Koga et al., 1991], [Aquilanti et al., 1996],[Aquilanti et al., 1997]. These authors solved the Schrödinger equation in momentum space, but here we will use a direct-space treatment to reach the same results. In this approach to molecular orbital theory, we search for

solutions to the one-electron Schrödinger equation

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{x}) - \epsilon_\zeta \right] \varphi_\zeta(\mathbf{x}) = 0 \quad (\text{F.1})$$

where $v(\mathbf{x})$ is the Coulomb attraction potential of the nuclei:

$$v(\mathbf{x}) = - \sum_a \frac{Z_a}{|\mathbf{x} - \mathbf{X}_a|} \quad (\text{F.2})$$

We will approximate the molecular orbitals $\varphi_\zeta(\mathbf{x})$ by superpositions of Coulomb Sturmian atomic orbitals centered on the various atoms of the molecule. To do this it is convenient to introduce a notation where τ stands for a set of four indices, the first three being the quantum numbers of a one-electron Coulomb Sturmian basis function of the type discussed in Appendix B, while the final index, a , is the index of the nucleus on which the atomic orbital is centered:

$$\tau \equiv (n, l, m, a) \quad (\text{F.3})$$

In this notation we can write

$$\chi_\tau(\mathbf{x}) \equiv \chi_{nlm}(\mathbf{x} - \mathbf{X}_a) \quad (\text{F.4})$$

A molecular orbital is then represented by a superposition of the form

$$\varphi_\zeta(\mathbf{x}) = \sum_{nlma} \chi_{nlm}(\mathbf{x} - \mathbf{X}_a) C_{\tau,\zeta} \equiv \sum_{\tau} \chi_\tau(\mathbf{x}) C_{\tau,\zeta} \quad (\text{F.5})$$

The normalization condition for the molecular orbitals is

$$1 = \int d^3x_j \varphi_\zeta^*(\mathbf{x}_j) \varphi_\zeta(\mathbf{x}_j) = \sum_{\tau'} \sum_{\tau} C_{\tau',\zeta}^* m_{\tau',\tau} C_{\tau,\zeta} \quad (\text{F.6})$$

where

$$m_{\tau',\tau} \equiv \int d^3x_j \chi_{\tau'}^*(\mathbf{x}_j) \chi_\tau(\mathbf{x}_j) \quad (\text{F.7})$$

is the matrix of many-center Sturmian overlap integrals. The matrix $m_{\tau',\tau}$ may be evaluated using the properties of hyperspherical harmonics, and we will discuss below the details of how this may be done.

Coulomb Sturmian basis functions are discussed in detail in Appendix B. They have exactly the same form as the familiar hydrogenlike atomic orbitals,

$$\chi_{nlm}(\mathbf{x}) = R_{n,l}(r) Y_{lm}(\theta, \phi) \quad (\text{F.8})$$

except in the radial part, $R_{n,l}(r)$, the factor Z/n is replaced by a constant, k . The first few Coulomb Sturmian radial functions are

$$\begin{aligned} R_{1,0}(r) &= 2k^{3/2} e^{-kr} \\ R_{2,0}(r) &= 2k^{3/2} (1 - kr) e^{-kr} \\ R_{2,1}(r) &= \frac{1}{\sqrt{3}} 2k^{3/2} kr e^{-kr} \\ R_{3,1}(r) &= 2k^{3/3} \left(1 - 2kr + \frac{2k^2 r^2}{3} \right) e^{-kr} \end{aligned} \quad (\text{F.9})$$

The reader can verify that these are precisely the same as hydrogenlike atomic orbitals with the replacement $Z/n \rightarrow k$. We now substitute the superposition (F.5) into the one-electron Schrödinger equation (F.1). This gives us:

$$\begin{aligned} & \sum_{nlma} \left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 + v(\mathbf{x}) \right] \chi_{nlm}(\mathbf{x} - \mathbf{X}_a) C_{\tau,\zeta} \\ \equiv & \sum_{\tau} \left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 + v(\mathbf{x}) \right] \chi_{\tau}(\mathbf{x}) C_{\tau,\zeta} = 0 \end{aligned} \quad (\text{F.10})$$

with

$$\epsilon_{\zeta} \equiv -\frac{1}{2}k^2 \quad (\text{F.11})$$

where each of the Coulomb Sturmian atomic orbitals $\chi_{\tau}(\mathbf{x}) \equiv \chi_{nlm}(\mathbf{x} - \mathbf{X}_a)$ obeys a one-electron Schrödinger equation of the form

$$\left[-\frac{1}{2}\nabla^2 + \frac{1}{2}k^2 - \frac{nk}{|\mathbf{x} - \mathbf{X}_a|} \right] \chi_{nlm}(\mathbf{x} - \mathbf{X}_a) = 0 \quad (\text{F.12})$$

Taking the scalar product of (F.10) with a conjugate function in our basis set, we obtain

$$\sum_{\tau} \int d^3x \chi_{\tau'}^*(\mathbf{x}) \left[-\frac{1}{2}\nabla^2 + \frac{k^2}{2} + v(\mathbf{x}) \right] \chi_{\tau}(\mathbf{x}) C_{\tau,\zeta} = 0 \quad (\text{F.13})$$

With the notation

$$W_{\tau',\tau} \equiv -\frac{1}{k} \int d^3x \chi_{\tau'}^*(\mathbf{x}) v(\mathbf{x}) \chi_{\tau}(\mathbf{x}) \quad (\text{F.14})$$

and

$$\begin{aligned} S_{\tau',\tau} & \equiv \frac{1}{k^2} \int d^3x \chi_{\tau'}^*(\mathbf{x}) \left(-\frac{1}{2}\nabla^2 + \frac{k^2}{2} \right) \chi_{\tau}(\mathbf{x}) \\ & = \frac{n}{k} \int d^3x \chi_{\tau'}^*(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{X}_a|} \chi_{\tau}(\mathbf{x}) \end{aligned} \quad (\text{F.15})$$

we obtain a secular equation of the form

$$\sum_{\tau} [W_{\tau',\tau} - kS_{\tau',\tau}] C_{\tau,\zeta} = 0 \quad (\text{F.16})$$

We can call $\mathfrak{W}_{\tau',\tau}$ the *Wulfman integrals* to honor the pioneering contributions of Prof. C.E. Wulfman. The integrals $S_{\tau',\tau}$ are called *Shibuya-Wulfman integrals* [Avery, 1989], [Avery, 2000], [Avery and Avery, 2006], [Judd, 1975], [Shibuya and Wulfman, 1965], and methods for their evaluation are discussed below. It can be shown [Koga and Matsushashi, 1987] that the matrix elements of the many-center potential $\mathfrak{W}_{\tau',\tau}$ can be expressed in terms of the Shibuya-Wulfman integrals by means of the sum rule

$$\mathfrak{W}_{\tau',\tau} = \sqrt{\frac{n'n}{Z_{a'}Z_a}} \sum_{\tau''} K_{\tau',\tau''} K_{\tau'',\tau} \quad (\text{F.17})$$

where

$$K_{\tau',\tau} \equiv \sqrt{\frac{Z_{a'}Z_a}{n'n}} \mathfrak{G}_{\tau',\tau} \quad (\text{F.18})$$

With the help of this sum rule, the secular equations (F.16) can be rewritten in the form

$$\sum_{\tau} \left[\sum_{\tau''} K_{\tau',\tau''} K_{\tau'',\tau} - k K_{\tau',\tau} \right] C'_{\tau,\zeta} = 0 \quad (\text{F.19})$$

with

$$C'_{\tau,\zeta} = \sqrt{\frac{Z_a}{n}} C_{\tau,\zeta} \quad (\text{F.20})$$

Now suppose that we have solved the secular equation

$$\sum_{\tau} [K_{\tau',\tau} - k \delta_{\tau',\tau}] C'_{\tau,\zeta} \quad (\text{F.21})$$

The values of k and $C'_{\tau,\zeta}$ thus obtained will also be solutions to (F.19). To see this, we perform the sum over τ'' in (F.19), making use of (F.21):

$$\begin{aligned} & \sum_{\tau} \left[\sum_{\tau''} K_{\tau',\tau''} K_{\tau'',\tau} - k K_{\tau',\tau} \right] C'_{\tau,\zeta} \\ &= k \sum_{\tau} \left[\sum_{\tau''} K_{\tau',\tau''} \delta_{\tau'',\tau} C'_{\tau,\zeta} - k \delta_{\tau',\tau} \right] C'_{\tau,\zeta} \\ &= k \sum_{\tau} [K_{\tau',\tau} - k \delta_{\tau',\tau}] C'_{\tau,\zeta} = 0 \end{aligned} \quad (\text{F.22})$$

Since we have two forms of the molecular Sturmian secular equation, (F.16) and (F.21), one might ask which form is the best. The answer is that if the number of basis functions used is small, (F.16) give the most accurate results. However, particularly for small values of the parameter S , equation (F.21) suffers from problems of overcompleteness when the number of basis functions is increased. By contrast, as Monkhorst and Jeziorski have pointed out [Monkhorst and Jeziorski, 1979], equation (F.16) has no such problems, and therefore it is the method of choice when the basis set used is very large or when S is small. We can call the matrix $K_{\tau',\tau}$ the *Koga matrix* to honor the contributions of Prof. T. Koga and his group.

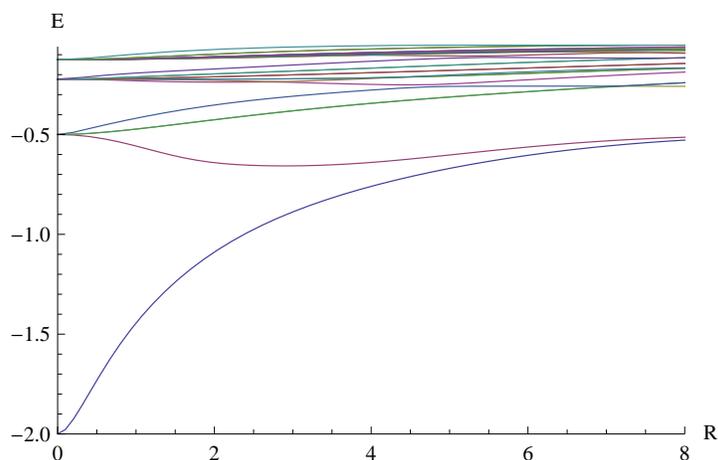


Figure F.1: Energies of the ground state and excited states of H_2^+ , calculated by solving equation (F.21). The energies are expressed in Hartrees and are given as a function of the internuclear separation R , expressed in Bohrs.

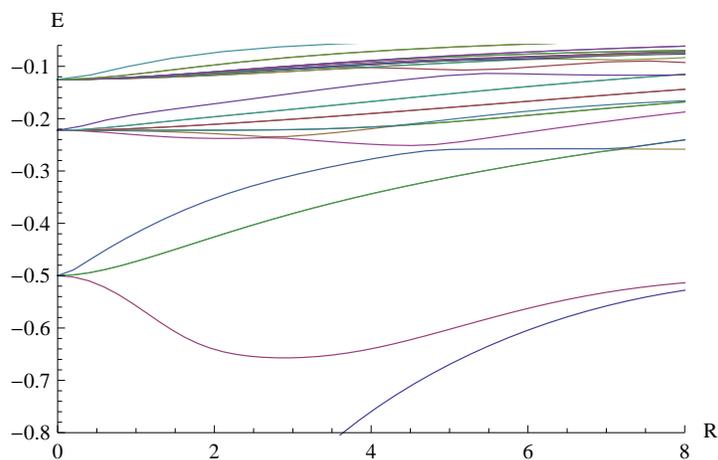


Figure F.2: A closer view of the excited state energies of H_2^+ . In the united-atom limit, these energies approach those of the excited states of He^+ , i.e., $4/(2n^2)$.

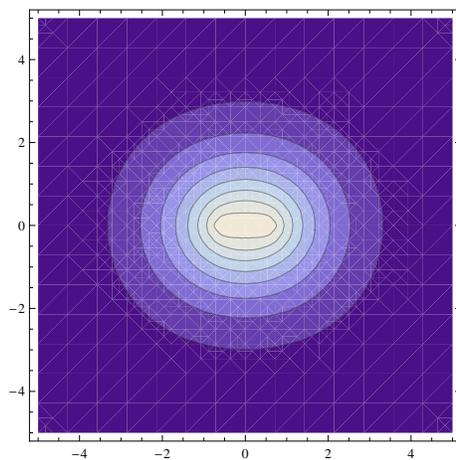


Figure F.3: The ground state molecular orbital of the H_2^+ ion at nuclear separation $R=1.21702$ Bohr ($S=2$, $k=1.64335$). In the united-atom limit, $k=2$.

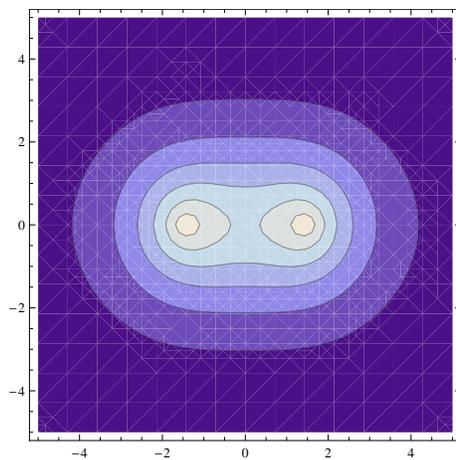


Figure F.4: The same state at nuclear separation 2.98216 Bohr ($S=4$, $k=1.34131$).

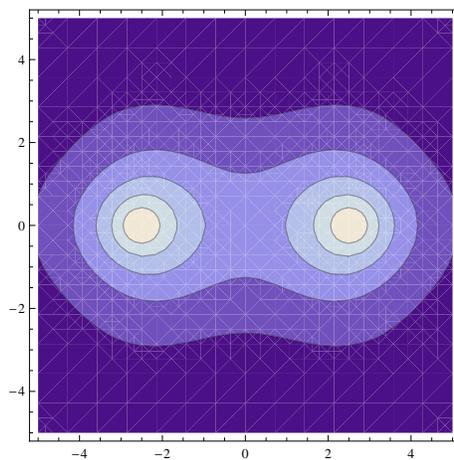


Figure F.5: Here the internuclear distance has been increased to 5.13325 Bohrs ($S=6$, $k=1.16885$).

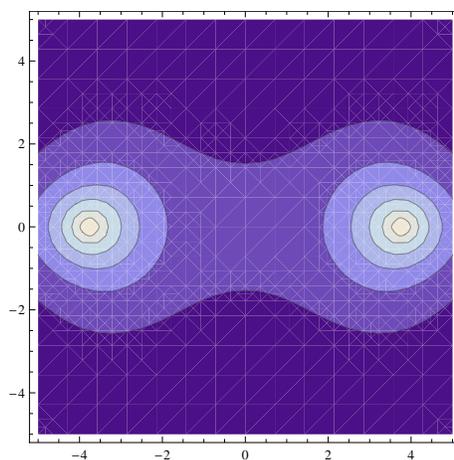


Figure F.6: The same state with nuclear separation 7.50577 Bohrs ($S=8$, $k=1.06585$). As the nuclear separation increases, k approaches 1.

F.2 Shibuya-Wulfman integrals and Sturmian overlap integrals evaluated in terms of hyperpherical harmonics

The Shibuya-Wulfman integrals $S_{\tau',\tau}$ defined by equation (F.15) as well as the molecular Sturmian overlap integrals

$$m_{\tau',\tau} \equiv \int d^3x \chi_{\tau'}^*(\mathbf{x})\chi_{\tau}(\mathbf{x}) \quad (\text{F.23})$$

can conveniently be evaluated in reciprocal space. Let us first consider the Sturmian overlap integrals. Using the fact that

$$m_{\tau',\tau} \equiv \int d^3x \chi_{\tau'}^*(\mathbf{x})\chi_{\tau}(\mathbf{x}) = \int d^3p \chi_{\tau'}^{t*}(\mathbf{p})\chi_{\tau}^t(\mathbf{p}) \quad (\text{F.24})$$

where, if we let

$$\mu \equiv (n, l, m) \quad (\text{F.25})$$

while $\tau \equiv (n, l, m, a)$, then

$$\begin{aligned} \chi_{\tau}^t(\mathbf{p}) &\equiv \frac{1}{(2\pi)^{3/2}} \int d^3x e^{-i\mathbf{p}\cdot\mathbf{x}} \chi_{\tau}(\mathbf{x}) = e^{-i\mathbf{p}\cdot\mathbf{X}_a} \chi_{\mu}^t(\mathbf{p}) \\ \chi_{\tau'}^{t*}(\mathbf{p}) &\equiv \frac{1}{(2\pi)^{3/2}} \int d^3x e^{i\mathbf{p}\cdot\mathbf{x}} \chi_{\tau'}^*(\mathbf{x}) = e^{i\mathbf{p}\cdot\mathbf{X}_{a'}} \chi_{\mu}^{t*}(\mathbf{p}) \end{aligned} \quad (\text{F.26})$$

We thus obtain

$$m_{\tau',\tau} = \int d^3p e^{i\mathbf{p}\cdot\mathbf{R}} \chi_{\mu'}^{t*}(\mathbf{p})\chi_{\mu}^t(\mathbf{p}) \quad (\text{F.27})$$

where

$$\mathbf{R} \equiv \mathbf{X}_{a'} - \mathbf{X}_a \quad (\text{F.28})$$

We now make use of V. Fock's relationship [Fock, 1935], [Fock, 1958]

$$\chi_{\mu}^t(\mathbf{p}) = M(p)Y_{n-1,l,m}(\hat{\mathbf{u}}) \quad (\text{F.29})$$

where

$$M(p) \equiv \frac{4k^{5/2}}{(k^2 + p^2)^2} \quad (\text{F.30})$$

In equation (F.29), $\hat{\mathbf{u}}$ is a 4-dimensional unit vector that defines Fock's projection of momentum space onto the surface of a 4-dimensional hypersphere.

$$\hat{\mathbf{u}} = (u_1, u_2, u_3, u_4) = \left(\frac{2kp_1}{k^2 + p^2}, \frac{2kp_2}{k^2 + p^2}, \frac{2kp_3}{k^2 + p^2}, \frac{k^2 - p^2}{k^2 + p^2} \right) \quad (\text{F.31})$$

F.2. SHIBUYA-WULFMAN INTEGRALS AND STURMIAN OVERLAP INTEGRALS EVALUATED IN

while $Y_\mu(\hat{\mathbf{u}})$ is a 4-dimensional hyperspherical harmonic defined by:

$$Y_{\lambda,l,m}(\hat{\mathbf{u}}) = \mathcal{N}_{\lambda,l} C_{\lambda-l}^{1+l}(u_4) Y_{l,m}(u_1, u_2, u_3) \quad (\text{F.32})$$

where

$$\mathcal{N}_{\lambda,l} = (-1)^{\lambda-l} (2l)!! \sqrt{\frac{2(\lambda+1)(\lambda-l)!}{\pi(\lambda+l+1)!}} \quad (\text{F.33})$$

is a normalizing constant while

$$C_j^\alpha(u_4) = \sum_{t=0}^{[j/2]} \frac{(-1)^t \Gamma(n+\alpha-t)}{t!(j-2t)!\Gamma(\alpha)} (2u_4)^{j-2t} \quad (\text{F.34})$$

is a Gegenbauer polynomial, and where $Y_{l,m}$ is a familiar 3-dimensional spherical harmonic. The first few hyperspherical harmonics are shown in Table 5.1. The index λ corresponds to $n-1$ so that the Fourier transform of $\chi_{1,0,0}(\mathbf{x})$ is given by $\chi_{1,0,0}^t(\mathbf{p}) = M(p) Y_{0,0,0}(\hat{\mathbf{u}}) = M(p)/(\sqrt{2\pi})$, and so on. Substituting (F.29) into (F.27), we obtain

$$\begin{aligned} m_{\tau',\tau} &= \int d^3p e^{i\mathbf{p}\cdot\mathbf{R}} M(p)^2 Y_{n'-1,l',m'}^*(\hat{\mathbf{u}}) Y_{n-1,l,m}(\hat{\mathbf{u}}) \\ &\equiv \int d^3p e^{i\mathbf{p}\cdot\mathbf{R}} M(p)^2 Y_{\mu'}^*(\hat{\mathbf{u}}) Y_\mu(\hat{\mathbf{u}}) \end{aligned} \quad (\text{F.35})$$

(Here, and throughout the book, a unit vector is indicated by a “hat”.) One can show [Avery, 1989] that the Shibuya-Wulfman integrals can be written in a similar form:

$$\mathfrak{S}_{\tau',\tau} = \int d^3p e^{i\mathbf{p}\cdot\mathbf{R}} \left(\frac{2k}{k^2+p^2} \right)^3 Y_{\mu'}^*(\hat{\mathbf{u}}) Y_\mu(\hat{\mathbf{u}}) \quad (\text{F.36})$$

One can also show [Caligiana, 2003] that

$$\int d^3p e^{i\mathbf{p}\cdot\mathbf{R}} \left(\frac{2k}{k^2+p^2} \right)^3 Y_\mu^*(\hat{\mathbf{u}}) = (2\pi)^{3/2} f_{n,l}(S) Y_{l,m}(\hat{\mathbf{S}}) \quad (\text{F.37})$$

where $Y_{l,m}$ is an ordinary 3-dimensional spherical harmonic and where

$$\mathbf{S} = \{S_x, S_y, S_z\} \equiv k\mathbf{R} \quad S = k|\mathbf{R}| \quad (\text{F.38})$$

The function $f_{n,l}(S)$ is defined by

$$\begin{aligned} k^{3/2} f_{n,l} &\equiv R_{n,l} - \frac{1}{2} \sqrt{\frac{(n-l)(n+l+1)}{n(n+1)}} R_{n+1,l} \\ &\quad - \frac{1}{2} \sqrt{\frac{(n+l)(n-l-1)}{n(n-1)}} R_{n-1,l} \end{aligned} \quad (\text{F.39})$$

where $R_{n,l}$ is the radial function of the Coulomb Sturmians given in equation (F.9), and where

$$R_{n-1,l} \equiv 0 \quad \text{if } l > n - 1 \quad (\text{F.40})$$

Similarly, one can show [Caligiana, 2003] that

$$\int d^3p e^{i\mathbf{p}\cdot\mathbf{R}} M(p)^2 Y_\mu(\hat{\mathbf{u}}) = (2\pi)^{3/2} g_{n,l}(S) Y_{l,m}(\hat{\mathbf{S}}) \quad (\text{F.41})$$

where

$$\begin{aligned} g_{n,l} \equiv & f_{n,l} - \frac{1}{2} \sqrt{\frac{(n-l)(n+l+1)}{n(n+1)}} f_{n+1,l} \\ & - \frac{1}{2} \sqrt{\frac{(n+l)(n-l-1)}{n(n-1)}} f_{n-1,l} \end{aligned} \quad (\text{F.42})$$

where we define

$$f_{n-1,l} \equiv 0 \quad \text{if } l > n - 1 \quad (\text{F.43})$$

The first few values of $f_{n,l}(S)$ and $g_{n,l}(S)$ are shown in Table 5.2.

Equations (F.37) and (F.41) are respectively identical with the Shibuya Wulfman integrals and the molecular Sturmian overlap integrals except that they contain only one 4-dimensional hyperspherical harmonic instead of a product of two. Thus the problem of evaluating both $S_{\tau',\tau}$ and $m_{\tau',\tau}$ reduces to the problem of evaluating the coefficients

$$c_{\mu'';\mu',\mu} = \int d\Omega_4 Y_{\mu''}^*(\hat{\mathbf{u}}) Y_{\mu'}^*(\hat{\mathbf{u}}) Y_\mu(\hat{\mathbf{u}}) \quad (\text{F.44})$$

These coefficients can readily be pre-evaluated once and for all using the hyperangular integration theorems discussed in Appendix C, and they can be stored as a large but very sparse matrix. We then obtain the relationships:

$$Y_{\mu'}^*(\hat{\mathbf{u}}) Y_\mu(\hat{\mathbf{u}}) = \sum_{\mu''} Y_{\mu''}(\hat{\mathbf{u}}) c_{\mu'';\mu',\mu} \quad (\text{F.45})$$

$$S_{\tau',\tau} = (2\pi)^{3/2} \sum_{\mu''} Y_{l',m''}(\hat{\mathbf{S}}) f_{n'',l''}(S) c_{\mu'';\mu',\mu} \quad (\text{F.46})$$

and

$$m_{\tau',\tau} = (2\pi)^{3/2} \sum_{\mu''} Y_{l'',m''}(\hat{\mathbf{S}}) g_{n'',l''}(S) c_{\mu'';\mu',\mu} \quad (\text{F.47})$$

Similar methods can be used to calculate the Wulfman integrals $\mathfrak{W}_{\tau',\tau}$ [Avery, 2000]. The first few Shibuya-Wulfman integrals are shown in Table 5.3. We can notice that when $S = 0$ the diagonal elements are 1, while the off-diagonal elements vanish. The first few displaced Coulomb Sturmian overlap integrals $m_{\tau',\tau}$ are shown in Table 5.4.

F.2. SHIBUYA-WULFMAN INTEGRALS AND STURMIAN OVERLAP INTEGRALS EVALUATED IN

Table F.1: $g_{n,l}(S)$ and $f_{n,l}(S)$, where $S \equiv k|\mathbf{X}_{a'} - \mathbf{X}_a|$. The functions $g_{n,l}(S)$ and $f_{n,l}(S)$ appear respectively in the two-center overlap integrals and the Shibuya-Wulfman integrals.

n	l	$g_{n,l}(S)$	$f_{n,l}(S)$
1	0	$\frac{e^{-S} (3 + 3S + S^2)}{3}$	$e^{-S} (1 + S)$
2	0	$-\frac{1}{6}e^{-S} (3 + 3S + 2S^2 + S^3)$	$-\frac{2}{3}e^{-S}S^2$
2	1	$\frac{e^{-S}S (3 + 3S + S^2)}{6\sqrt{3}}$	$\frac{2e^{-S}S(1 + S)}{3\sqrt{3}}$
3	0	$\frac{1}{15}e^{-S}S^4$	$\frac{1}{3}e^{-S}S^2(-2 + S)$
3	1	$-\frac{e^{-S}S^3(1 + S)}{15\sqrt{2}}$	$\frac{e^{-S}S(1 + S - S^2)}{3\sqrt{2}}$
3	2	$\frac{e^{-S}S^2(3 + 3S + S^2)}{15\sqrt{10}}$	$\frac{e^{-S}S^2(1 + S)}{3\sqrt{10}}$

Table F.2: This table shows the first few Shibuya-Wulfman integrals $\mathfrak{S}_{\tau',\tau}$, as functions of $\mathbf{S} \equiv k(\mathbf{X}_{a'} - \mathbf{X}_a)$, with $S \equiv |\mathbf{S}|$ and $\mathbf{S} \equiv (S \sin \theta \cos \phi, S \sin \theta \sin \phi, S \cos \theta)$. The integrals were generated by means of equation (F.46).

τ'	$\tau = (1, 0, 0, a)$	$\tau = (2, 0, 0, a)$
$(1, 0, 0, a')$	$e^{-S}(1 + S)$	$-\frac{2}{3}e^{-S}S^2$
$(2, 0, 0, a')$	$-\frac{2}{3}e^{-S}S^2$	$\frac{1}{3}e^{-S}(3 + 3S - 2S^2 + S^3)$
$(2, 1, -1, a')$	$-\frac{\sqrt{2}}{3}e^{-S}S(1 + S) \sin \theta e^{i\phi}$	$\frac{1}{3\sqrt{2}}e^{-S}S(-1 - S + S^2) \sin \theta e^{i\phi}$
$(2, 1, 0, a')$	$-\frac{2}{3}e^{-S}S(1 + S) \cos \theta$	$\frac{1}{3}e^{-S}S(-1 - S + S^2) \cos \theta$
$(2, 1, 1, a')$	$\frac{\sqrt{2}}{3}e^{-S}S(1 + S) \sin \theta e^{-i\phi}$	$-\frac{1}{3\sqrt{2}}e^{-S}S(-1 - S + S^2) \sin \theta e^{-i\phi}$

F.2. SHIBUYA-WULFMAN INTEGRALS AND STURMIAN OVERLAP INTEGRALS EVALUATED IN

Table F.3: The first few overlap integrals $m_{\tau',\tau} \equiv \int d^3x \chi_{\tau'}^*(\mathbf{x}) \chi_{\tau}(\mathbf{x})$ between displaced Coulomb Sturmians. The definitions of S , θ and ϕ are the same as in Table 5.3. The integrals were evaluated by means of equation (F.47).

τ'	$\tau = (1, 0, 0, a)$	$\tau = (2, 0, 0, a)$
$(1, 0, 0, a')$	$\frac{1}{3}e^{-S}(3 + 3S + S^2)$	$-\frac{1}{6}e^{-S}(3 + 3S + 2S^2 + S^3)$
$(2, 0, 0, a')$	$-\frac{1}{6}e^{-S}(3 + 3S + 2S^2 + S^3)$	$\frac{1}{15}e^{-S}(15 + 15S + 5S^2 + S^4)$
$(2, 1, -1, a')$	$-\frac{1}{6\sqrt{2}}e^{-S}S(3 + 3S + S^2) \sin \theta e^{i\phi}$	$\frac{1}{15\sqrt{2}}e^{-S}S^3(1 + S) \sin \theta e^{i\phi}$
$(2, 1, 0, a')$	$-\frac{1}{6}e^{-S}S(3 + 3S + S^2) \cos \theta$	$\frac{1}{15}e^{-S}S^3(1 + S) \cos \theta$
$(2, 1, 1, a')$	$\frac{1}{6\sqrt{2}}e^{-S}S(3 + 3S + S^2) \sin \theta e^{-i\phi}$	$-\frac{1}{15\sqrt{2}}e^{-S}S^3(1 + S) \sin \theta e^{-i\phi}$

F.3 Molecular calculations using the isoenergetic configurations

We now introduce N -electron configurations which are Slater determinants of the form:

$$|\Phi_\nu\rangle = |\varphi_\zeta \varphi_{\zeta'} \varphi_{\zeta''} \cdots\rangle \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_\zeta(\mathbf{x}_1) & \varphi_{\zeta'}(\mathbf{x}_1) & \varphi_{\zeta''}(\mathbf{x}_1) & \cdots \\ \varphi_\zeta(\mathbf{x}_2) & \varphi_{\zeta'}(\mathbf{x}_2) & \varphi_{\zeta''}(\mathbf{x}_2) & \cdots \\ \varphi_\zeta(\mathbf{x}_3) & \varphi_{\zeta'}(\mathbf{x}_3) & \varphi_{\zeta''}(\mathbf{x}_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \quad (\text{F.48})$$

where the molecular spin-orbitals $\varphi_\zeta(\mathbf{x})$ satisfy

$$\left[-\frac{1}{2} \nabla_j^2 + \frac{k^2}{2} + \beta_\nu v(\mathbf{x}_j) \right] \varphi_\zeta(\mathbf{x}_j) = 0 \quad v(\mathbf{x}_j) = \sum_a \frac{Z_a}{|\mathbf{x}_j - \mathbf{X}_a|} \quad (\text{F.49})$$

Since the individual molecular orbitals satisfy (F.49), the configurations $|\Phi_\nu\rangle$ are solutions to the separable N -electron equation:

$$\sum_{j=1}^N \left[-\frac{1}{2} \nabla_j^2 + \frac{k^2}{2} + \beta_\nu v(\mathbf{x}_j) \right] |\Phi_\nu\rangle = 0 \quad (\text{F.50})$$

which can also be written in the form:

$$\begin{aligned} & \left[\sum_{j=1}^N \left(-\frac{1}{2} \nabla_j^2 + \frac{k^2}{2} \right) + \beta_\nu V_0(\mathbf{x}) \right] |\Phi_\nu\rangle \\ &= \left[-\frac{1}{2} \sum_{j=1}^N \nabla_j^2 + \beta_\nu V_0(\mathbf{x}) - E_\kappa \right] |\Phi_\nu\rangle = 0 \end{aligned} \quad (\text{F.51})$$

where $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ and

$$E_\kappa = -\sum_{j=1}^N \frac{k^2}{2} = -\frac{Nk^2}{2} \quad (\text{F.52})$$

and where

$$V_0(\mathbf{x}) \equiv \sum_{j=1}^N v(\mathbf{x}_j) = \sum_{j=1}^N \sum_a \frac{Z_a}{|\mathbf{x}_j - \mathbf{X}_a|} \quad (\text{F.53})$$

Comparing (F.51) with (E.6), we can see that they are the same. Thus the isoenergetic solutions to the approximate N -electron Schrödinger equation (F.51) form a generalized Sturmian basis set. We would like to use these configurations to build up solutions to the N -electron Schrödinger equation

$$\left[\sum_{j=1}^N \left(-\frac{1}{2} \nabla_j^2 + \frac{k^2}{2} \right) + V(\mathbf{x}) \right] |\Psi_\kappa\rangle = 0 \quad (\text{F.54})$$

with

$$V(\mathbf{x}) = V_0(\mathbf{x}) + \sum_{i>j}^N \frac{1}{r_{ij}} \quad (\text{F.55})$$

Thus we write

$$|\Psi_\kappa\rangle \approx \sum_{\nu} |\Phi_\nu\rangle B_{\nu\kappa} \quad (\text{F.56})$$

Substituting this into the N -electron Schrödinger equation, and taking the scalar product with a conjugate configuration, we obtain the secular equations:

$$\sum_{\nu} \langle \Phi_{\nu'} | \left[\sum_{j=1}^N \left(-\frac{1}{2} \nabla_j^2 + \frac{k^2}{2} \right) + V(\mathbf{x}) \right] | \Phi_{\nu} \rangle B_{\nu\kappa} = 0 \quad (\text{F.57})$$

We now introduce a k -independent matrix representing the total potential based on the configurations $|\Phi_\nu\rangle$:

$$T_{\nu'\nu}^{(N)} \equiv -\frac{1}{k} \langle \Phi_{\nu'} | V(\mathbf{x}) | \Phi_{\nu} \rangle \quad (\text{F.58})$$

and another k -independent matrix

$$S_{\nu'\nu}^{(N)} \equiv \frac{1}{k^2} \langle \Phi_{\nu'}^* | \sum_{j=1}^N \left(-\frac{1}{2} \nabla_j^2 + \frac{k^2}{2} \right) | \Phi_{\nu} \rangle \quad (\text{F.59})$$

In terms of these matrices, the secular equations become:

$$\sum_{\nu} \left[T_{\nu'\nu}^{(N)} - k S_{\nu'\nu}^{(N)} \right] B_{\nu\kappa} = 0 \quad (\text{F.60})$$

Solving equation (F.60), we obtain k for each state κ and thus the energy $E_\kappa = -\frac{Nk^2}{2}$. For a given state κ , the value of k then determines the weighting factors $\beta_{\nu_1}, \beta_{\nu_2}, \dots$ needed to make each configuration $|\Phi_{\nu_1}\rangle, |\Phi_{\nu_2}\rangle, \dots$ correspond to the same energy E_κ .

In order to build the N -electron matrices $T_{\nu'\nu}^{(N)}$ and $S_{\nu'\nu}^{(N)}$ and solve equation (F.60), we must first obtain the coefficients $C_{\tau\zeta}$ by solving (F.16) or (F.21). In the case of diatomic molecules, we begin by picking a value of the parameter $S = kR$, where R is the interatomic distance and k is the exponent of the Coulomb Sturmian basis set. Neither R nor k is known at this point, but only their product S . As we shall see below, for the diatomic case, all of the integrals involved in equations (F.16) and (F.21) are pure functions of S . Having chosen S , we can thus solve the one-electron secular equations and obtain the coefficients $C_{\tau\zeta}$ and the spectrum of ratios k/β_ν . We are then able to solve equation F.60, which gives us a spectrum of k -values, and thus energies $-Nk^2/2$, and the eigenvectors $B_{\nu\kappa}$. From a k -value, we also get the unscaled distance $R = S/k$. We repeat the procedure for a range of S -values and interpolate to find the solutions as functions of R .

In the case of polyatomic molecules, one can choose a set of angles between the nuclei; these are left fixed under scaling of the coordinate system. The procedure is then similar to that described for the diatomic case.

F.4 Building $T_{\nu'\nu}^{(N)}$ and $\mathfrak{S}_{\nu'\nu}^{(N)}$ from 1-electron components

We have already discussed how the matrix of many-center Sturmian overlap integrals

$$m_{\tau'\tau} \equiv \int d^3x_j \chi_{\tau'}^*(\mathbf{x}_j) \chi_{\tau}(\mathbf{x}_j) \quad (\text{F.61})$$

may be evaluated using the properties of hyperspherical harmonics (F.47). The matrix $m_{\tau'\tau}$ is needed in order to normalize the molecular orbitals $\varphi_{\zeta}(\mathbf{x}_j)$, the normalization condition on the coefficients $C_{\tau,\zeta}$ being

$$1 = \int d^3x_j \varphi_{\zeta}^*(\mathbf{x}_j) \varphi_{\zeta}(\mathbf{x}_j) = \sum_{\tau'} \sum_{\tau} C_{\tau',\zeta}^* m_{\tau',\tau} C_{\tau,\zeta} \quad (\text{F.62})$$

Having performed the normalization, we then need to transform the nuclear attraction matrix elements $\mathfrak{W}_{\tau'\tau}$ to a representation based on the molecular orbitals:

$$\tilde{v}_{\zeta'\zeta} \equiv \int d^3x_j \varphi_{\zeta'}^*(\mathbf{x}_j) v(\mathbf{x}_j) \varphi_{\zeta}(\mathbf{x}_j) = -k \sum_{\tau'} \sum_{\tau} C_{\tau',\zeta'}^* \mathfrak{W}_{\tau'\tau} C_{\tau,\zeta} \quad (\text{F.63})$$

Once we are in possession of the 1-electron matrix elements $\tilde{v}_{\zeta'\zeta}$, we can evaluate

$$T_{\nu'\nu}^{0,(N)} \equiv -\frac{1}{k} \langle \Phi_{\nu'}^* | V_0(\mathbf{x}) | \Phi_{\nu} \rangle \quad (\text{F.64})$$

by means of the Slater-Condon rules. Because of the potential-weighted orthonormality relations obeyed by generalized Sturmian basis sets (Appendix B), we expect the matrix $T_{\nu'\nu}^{0,(N)}$ to be diagonal. We next transform the matrix of Shibuya-Wulfman integrals to a representation based on the molecular orbitals:

$$\tilde{S}_{\zeta'\zeta} \equiv \sum_{\tau'} \sum_{\tau} C_{\tau',\zeta'}^* \mathfrak{S}_{\tau'\tau} C_{\tau,\zeta} \quad (\text{F.65})$$

From these 1-electron matrix elements, the N -electron matrix $S_{\nu'\nu}^{(N)}$ can be constructed with the help of the Slater-Condon rules. Finally we must deal with the difficult term

$$T_{\nu'\nu}'^{(N)} \equiv -\frac{1}{k} \langle \Phi_{\nu'}^* | \sum_{i>j}^N \frac{1}{r_{ij}} | \Phi_{\nu} \rangle \quad (\text{F.66})$$

which is a k -independent matrix representing the effects of interelectron repulsion, and this will be discussed in the next section.

F.5 Interelectron repulsion integrals for molecular Sturmians from hyperspherical harmonics

We will now show that just as the theory of hyperspherical harmonics can be used to facilitate the calculation of Shibuya-Wulfman and molecular Sturmian overlap integrals, it also provides a method for very rapid calculation of the most important interelectron repulsion integrals involving molecular Sturmians. We again make use of momentum space:

Let $\rho_{\mu_1, \mu_2}(\mathbf{x} - \mathbf{X}_a)$ and $\rho_{\mu_3, \mu_4}(\mathbf{x}' - \mathbf{X}_{a'})$ be two electron density distributions, centered respectively on nuclei at the positions \mathbf{X}_a and $\mathbf{X}_{a'}$. Then the interelectron repulsion between them is given by the integral:

$$J_{\mu_1, \mu_2, \mu_3, \mu_4} = \int d^3x \int d^3x' \rho_{\mu_1, \mu_2}(\mathbf{x} - \mathbf{X}_a) \frac{1}{|\mathbf{x} - \mathbf{x}'|} \rho_{\mu_3, \mu_4}(\mathbf{x}' - \mathbf{X}_{a'}) \quad (\text{F.67})$$

If we introduce the Fourier transform representation

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{2\pi^2} \int d^3p \frac{1}{p^2} e^{-i\mathbf{p} \cdot (\mathbf{x} - \mathbf{x}')} \quad (\text{F.68})$$

we can rewrite $J_{\mu_1, \mu_2, \mu_3, \mu_4}$ in the form

$$J_{\mu_1, \mu_2, \mu_3, \mu_4} = 4\pi \int d^3p \frac{1}{p^2} e^{i\mathbf{p} \cdot \mathbf{R}} \rho_{\mu_1, \mu_2}^t(\mathbf{p}) \rho_{\mu_3, \mu_4}^t(-\mathbf{p}) \quad (\text{F.69})$$

where $\mathbf{R} = \mathbf{X}_{a'} - \mathbf{X}_a$ and

$$\rho_{\mu_i, \mu_j}^t(\mathbf{p}) = \frac{1}{(2\pi)^{3/2}} \int d^3x \rho_{\mu_i, \mu_j}(\mathbf{x}) e^{-i\mathbf{p} \cdot \mathbf{x}} \quad (\text{F.70})$$

Now let $R_{n,l}(2r)$ be a Coulomb Sturmian radial function with r replaced by $2r$, and we let $g(r)$ be any function of r . From the completeness property of Sturmian basis sets we know that it is possible to make an expansion of the form

$$g(r) = \sum_n a_n R_{n,l}(2r) \quad (\text{F.71})$$

and from the potential-weighted orthonormality relations (D.11) it follows that the expansion coefficients will be given by

$$a_n = \frac{n}{2} \int_0^\infty dr r R_{n,l}(2r) g(r) \quad (\text{F.72})$$

We now let the density be composed of a product of two Coulomb Sturmian basis functions:

$$\rho_{\mu_1, \mu_2}(\mathbf{x}) = \chi_{\mu_1}^*(\mathbf{x}) \chi_{\mu_2}(\mathbf{x}) = R_{n_1, l_1}(r) R_{n_2, l_2}(r) Y_{l_1, m_1}^*(\hat{\mathbf{x}}) Y_{l_2, m_2}(\hat{\mathbf{x}}) \quad (\text{F.73})$$

If we make the expansion

$$\rho_{\mu_1, \mu_2}(\mathbf{x}) = \sum_{\mu''} R_{n'', l''}(2r) Y_{l'', m''}(\hat{\mathbf{x}}) \mathcal{C}_{\mu'', \mu_1, \mu_2} \equiv \sum_{\mu''} \chi_{\mu''}(2\mathbf{x}) \mathcal{C}_{\mu'', \mu_1, \mu_2} \quad (\text{F.74})$$

then the coefficients in the expansion will be given by

$$\begin{aligned} \mathcal{C}_{\mu'', \mu_1, \mu_2} &= \frac{n''}{2} \int_0^\infty dr r R_{n'', l''}(2r) R_{n_1, l_1}(r) R_{n_2, l_2}(r) \\ &\quad \times \int d\Omega_3 Y_{l'', m''}^*(\hat{\mathbf{x}}) Y_{l_1, m_1}^*(\hat{\mathbf{x}}) Y_{l_2, m_2}(\hat{\mathbf{x}}) \end{aligned} \quad (\text{F.75})$$

Like the coefficients $c_{\mu''; \mu', \mu}$, the coefficients $\mathcal{C}_{\mu''; \mu', \mu}$ form a large but very sparse matrix which can be pre-calculated once and for all and stored. The series in (F.74) terminates and the expansion is exact. Making use of the relationships (F.73) and (F.74), we obtain the result

$$\begin{aligned} J_{\mu_1, \mu_2, \mu_3, \mu_4} &= \int d^3x \int d^3x' \rho_{\mu_1, \mu_2}(\mathbf{x} - \mathbf{X}_a) \frac{1}{|\mathbf{x} - \mathbf{x}'|} \rho_{\mu_3, \mu_4}(\mathbf{x}' - \mathbf{X}_{a'}) \\ &= \sum_{\mu', \mu} J_{\mu', \mu} C_{\mu', \mu_1, \mu_2} C_{\mu, \mu_3, \mu_4} \end{aligned} \quad (\text{F.76})$$

where

$$J_{\mu', \mu} = 4\pi \int d^3p \frac{1}{p^2} e^{i\mathbf{p} \cdot \mathbf{R}} \rho_{\mu'}^t(\mathbf{p}) \rho_{\mu}^t(-\mathbf{p}) \quad (\text{F.77})$$

and where

$$\begin{aligned} \rho_{\mu'}(\mathbf{x}) &= R_{n', l'}(2r) Y_{l', m'}(\hat{\mathbf{x}}) \\ \rho_{\mu}(\mathbf{x}) &= R_{n, l}(2r) Y_{l, m}(\hat{\mathbf{x}}) \end{aligned} \quad (\text{F.78})$$

Then, making use of Fock's relationship we have:

$$\begin{aligned} \rho_{\mu'}^t(\mathbf{p}) &= \tilde{M}(p) Y_{n'-1, l', m'}(\hat{\mathbf{w}}) \equiv \tilde{M}(p) Y_{\mu'}(\hat{\mathbf{w}}) \\ \rho_{\mu}^t(-\mathbf{p}) &= (-1)^l \tilde{M}(p) Y_{n-1, l, m}(\hat{\mathbf{w}}) \equiv (-1)^l \tilde{M}(p) Y_{\mu}(\hat{\mathbf{w}}) \end{aligned} \quad (\text{F.79})$$

where $\tilde{M}(p)$ and $\hat{\mathbf{w}}$ are similar to $M(p)$ and $\hat{\mathbf{u}}$, but with double the k value.

$$\hat{\mathbf{w}} = (w_1, w_2, w_3, w_4) = \left(\frac{4kp_1}{4k^2 + p^2}, \frac{4kp_2}{4k^2 + p^2}, \frac{4kp_3}{4k^2 + p^2}, \frac{4k^2 - p^2}{4k^2 + p^2} \right) \quad (\text{F.80})$$

and

$$\tilde{M}(p) \equiv \frac{4(2k)^{5/2}}{(4k^2 + p^2)^2} \quad (\text{F.81})$$

F.5. INTERELECTRON REPULSION INTEGRALS FOR MOLECULAR STURMIANS FROM HYPER

Then

$$J_{\mu',\mu} = 4\pi \int d^3p \frac{1}{p^2} e^{i\mathbf{p}\cdot\mathbf{R}} \tilde{M}^2(p) (-1)^l Y_{\mu'}(\hat{\mathbf{w}}) Y_{\mu}(\hat{\mathbf{w}}) \quad (\text{F.82})$$

Since we know how to represent the product $(-1)^l Y_{\mu'}(\hat{\mathbf{w}}) Y_{\mu}(\hat{\mathbf{w}})$ in terms of $Y_{\mu''}(\hat{\mathbf{w}})$, we can express the matrix $J_{\mu',\mu}$ in terms of a single vector, $J_{\mu''}$: Let

$$\tilde{c}_{\mu'';\mu',\mu} \equiv (-1)^l \int d\Omega_4 Y_{\mu''}^*(\hat{\mathbf{w}}) Y_{\mu'}(\hat{\mathbf{w}}) Y_{\mu}(\hat{\mathbf{w}}) \quad (\text{F.83})$$

Then

$$J_{\mu',\mu} = \sum_{\mu''} J_{\mu''} \tilde{c}_{\mu'';\mu',\mu} \quad (\text{F.84})$$

where

$$J_{\mu} \equiv 4\pi \int d^3p \frac{1}{p^2} e^{i\mathbf{p}\cdot\mathbf{R}} \tilde{M}^2(p) Y_{\mu}(\hat{\mathbf{w}}) \quad (\text{F.85})$$

The coefficients $\tilde{c}_{\mu'';\mu',\mu}$ differ slightly from the coefficients $c_{\mu'';\mu',\mu}$, but they too form a large but very sparse matrix that can be pre-calculated and stored. We must now evaluate J_{μ} . To do so, it is convenient to introduce the notation

$$\rho_{\mu}^t(\mathbf{p}) = \tilde{M}(p) Y_{\mu}(\hat{\mathbf{w}}) \equiv i^{-l} R_{n,l}^t(p) Y_{l,m}(\hat{\mathbf{p}}) \quad (\text{F.86})$$

where

$$R_{n,l}^t(p) \equiv \sqrt{\frac{2}{\pi}} \int_0^{\infty} dr r^2 j_l(pr) R_{n,l}(2r) \quad (\text{F.87})$$

Then, expanding the plane wave in terms of spherical harmonics and spherical Bessel functions and integrating over the solid angle in momentum space, we have:

$$\begin{aligned} J_{\mu} &= 4\pi \int d^3p \frac{1}{p^2} e^{i\mathbf{p}\cdot\mathbf{R}} \tilde{M}^2(p) Y_{\mu}(\hat{\mathbf{w}}) \\ &= (4\pi)^2 Y_{l,m}(\hat{\mathbf{S}}) \int_0^{\infty} dp \tilde{M}(p) R_{n,l}^t(p) j_l(pR) \end{aligned} \quad (\text{F.88})$$

The radial p -integrals in equation (F.88) are simple enough to be evaluated exactly by Mathematica, and they depend only on n and l . They can conveniently be stored as interpolation functions. It is also convenient to initialize by performing the sum shown in equation (F.84). This sum, and the sums required for the evaluation of $J_{\mu_1,\mu_2,\mu_3,\mu_4}$ from $J_{\mu',\mu}$ can be performed very rapidly because of the sparseness of the coefficients.

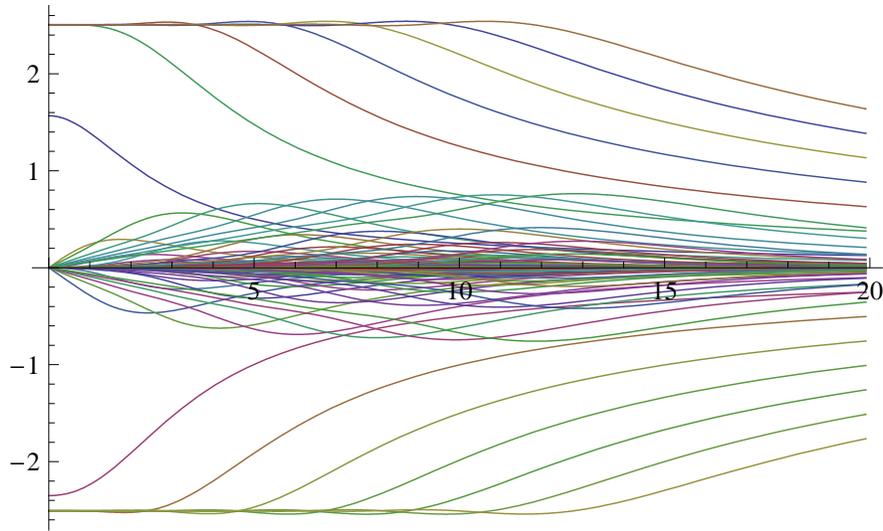


Figure F.7: The integrals $\int_0^\infty dp \tilde{M}(p) R_{n,l}^t(p) j_l(pR)/k$ of equation (F.88) are shown here plotted as functions of $S \equiv kR$. There are 105 functions, corresponding $n = 1, 2, \dots, 14$ and $l = 0, 1, \dots, n - 1$.

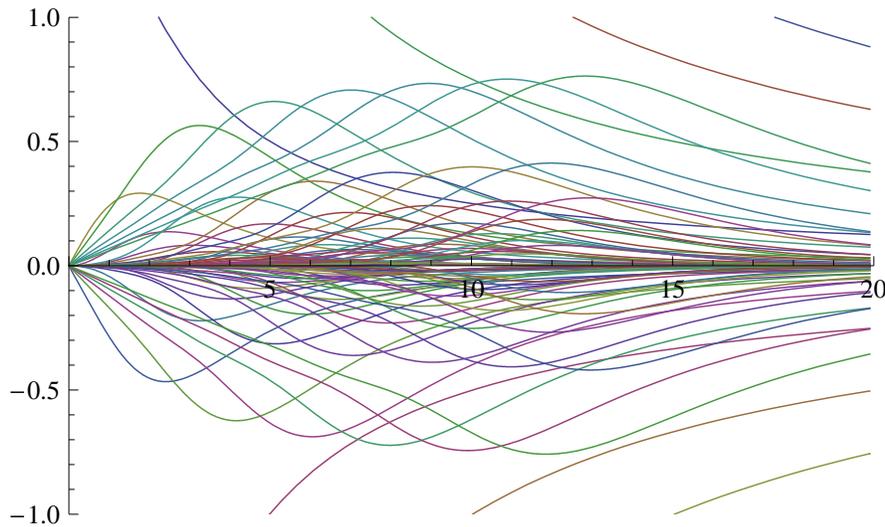


Figure F.8: The integrals $\int_0^\infty dp \tilde{M}(p) R_{n,l}^t(p) j_l(pR)/k$ shown in more detail. For small values of S the integrals are proportional to S^l , while for large values they are proportional to $1/S^{l+1}$.

F.6 Many-center integrals treated by Gaussian expansions (Appendix E)

One also needs to calculate 3-center and 4-center integrals of the form

$$J_{\tau_1, \tau_2, \tau_3, \tau_4} = \int d^3x \int d^3x' \chi_{\tau_1}^*(\mathbf{x}) \chi_{\tau_2}(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{x}'|} \chi_{\tau_3}^*(\mathbf{x}') \chi_{\tau_4}(\mathbf{x}') \quad (\text{F.89})$$

where the centers $\mathbf{X}_{a_1}, \dots, \mathbf{X}_{a_4}$ may in general be 4 different points. Even in this difficult case, molecular Sturmian basis functions have very marked advantages. One can show (Appendix E) that $J_{\tau_1, \tau_2, \tau_3, \tau_4}/k$ is independent of k . They may therefore be calculated once and for all and stored. One can also show (Appendix E) that the Coulomb Sturmian atomic orbitals can be expressed in terms of Gaussian expansions, where the Gaussian exponents α_i are universals that need never be changed despite changes in scaling due changes in the value of k . The coefficients $\gamma_{0,i}, \gamma_{1,i}, \dots$ in the following expression

$$s^j e^{-s} \approx \sum_i \gamma_{j,i} e^{-\alpha_i s^2} \quad s \equiv kr \quad (\text{F.90})$$

are also universals, and they too need never be changed, despite changes in scaling. We make the expansion

$$\chi_\tau(\mathbf{x}) = \chi_{n,l,m}(\mathbf{x} - \mathbf{X}_a) \approx k^{3/2} \sum_i \Gamma_{n,l,i} e^{-\alpha_i |k\mathbf{x} - k\mathbf{X}_a|^2} R_l^m(k\mathbf{x} - k\mathbf{X}_a) \quad (\text{F.91})$$

where the coefficients $\Gamma_{n,l,i}$ are defined by the relationship

$$\sqrt{\frac{2l+1}{4\pi}} \tilde{R}_{n,l}(s) s^{-l} \approx \sum_i \Gamma_{n,l,i} e^{-\alpha_i s^2} \quad (\text{F.92})$$

with $\tilde{R}_{n,l}(s) \equiv R_{n,l}(r)/k^{3/2}$ and where R_l^m is a regular solid harmonic (Appendix E). Figure 5.9 shows the Gaussian expansion of $e^{-s} \equiv e^{-kr}$, while the table shows coefficients in the expansion.

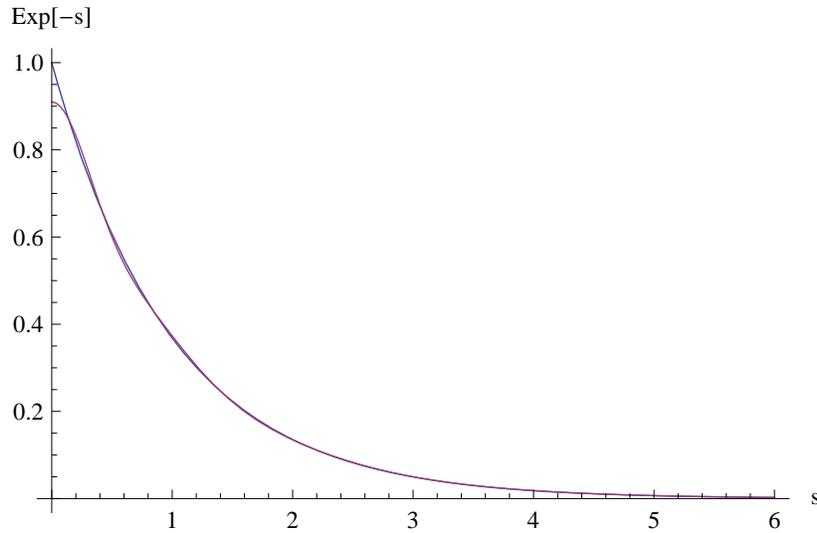


Figure F.9: This figure shows the Gaussian expansion $e^{-s} \approx \sum_i \gamma_{0,i} e^{-\alpha_i s^2}$, using the 10 coefficients and exponents shown in Table 5.5. The expansion is reasonably accurate throughout most of the range, but it fails to produce the sharp cusp near $s = 0$.

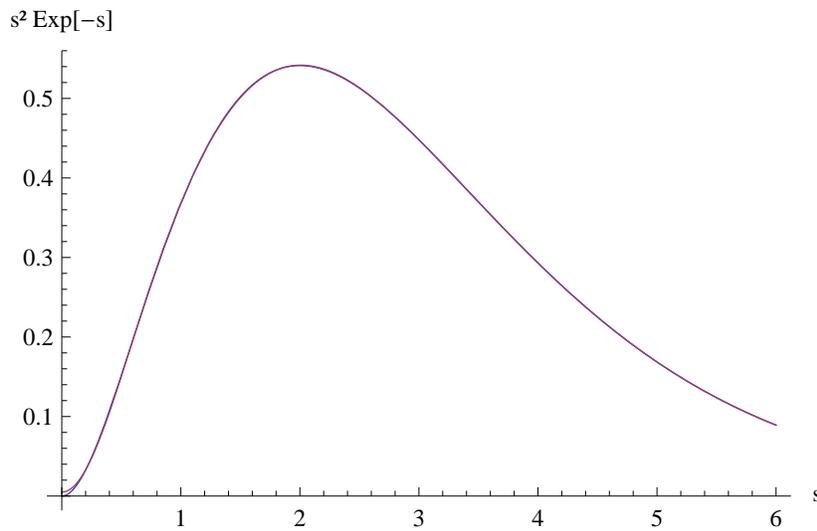


Figure F.10: Here we see the Gaussian expansion $s^2 e^{-s} \approx \sum_i \gamma_{2,i} e^{-\alpha_i s^2}$. As in Figure 5.9, the expansion is compared with the exact function.

Table F.4: Universal coefficients for Gaussian expansions of Coulomb Sturmians: They are used in the relationship $s^j e^{-s} = \sum_i \gamma_{j,i} e^{-\alpha_i s^2}$, where $s \equiv kr$. When k changes with scaling, the Gaussian expansion changes scale automatically.

i	α_i	$\gamma_{0,i}$	$\gamma_{1,i}$	$\gamma_{2,i}$
1	5.12	0.474589	-0.456553	-0.011253
2	2.56	-0.409842	0.420846	-0.135640
3	1.28	0.522704	-0.461490	-0.030952
4	0.64	-0.028869	0.157189	-0.390496
5	0.32	0.237377	0.008340	-0.284720
6	0.16	0.074194	0.248277	0.001174
7	0.08	0.039810	0.147977	0.631545
8	0.04	-0.001091	0.025882	0.224411
9	0.02	0.000808	-0.001018	0.000462
10	0.01	-0.000129	0.000170	0.000468

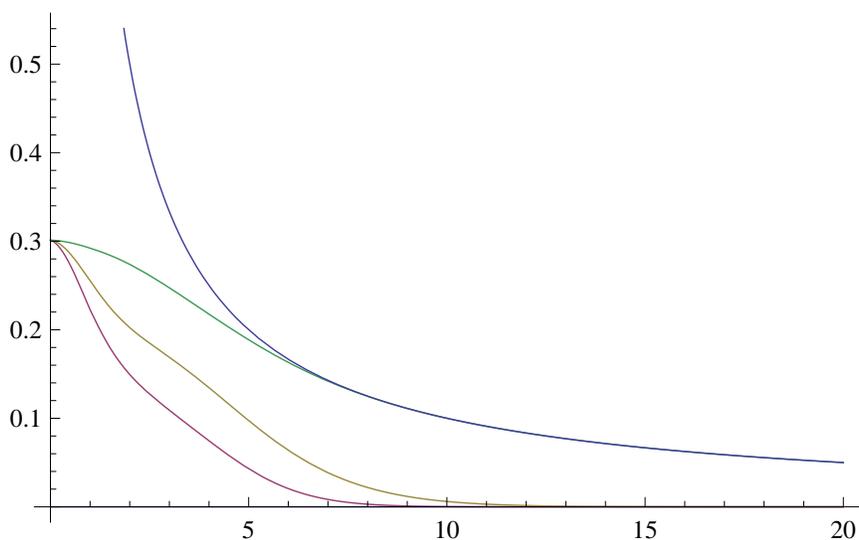


Figure F.11: Interelectron repulsion integrals $J_{\tau_1, \tau_2, \tau_3, \tau_4} / k$ for diatomic molecules for $\mathbf{n} = (2, 2, 2, 2)$ and $\mathbf{l} = (0, 0, 0, 0)$ as functions of $S = kR$. The lowest curve shows the results when $(a_1, a_2, a_3, a_4) = (S, 0, S, 0)$. The next higher curve shows the case where $(a_1, a_2, a_3, a_4) = (S, S, S, 0)$. The highest curve, which is the $(a_1, a_2, a_3, a_4) = (S, S, 0, 0)$ case, is compared with $1/S$, which it approaches asymptotically as S becomes large.

F.7 A pilot calculation

We have made a small pilot calculation on the dissociation of the hydrogen molecule, using a very restricted basis set. In the calculation shown in Figure 5.12, the ground state wave function changes character as a function of the internuclear separation, R . As $R \rightarrow 0$, the wave function becomes more and more dominated by a configuration which is built from two *gerade* molecular orbitals. But as the molecule dissociates, the wave function becomes the linear combination of configurations representing two isolated neutral hydrogen atoms, each with its own electron, and the total energy corresponds to that of two isolated hydrogen atoms.

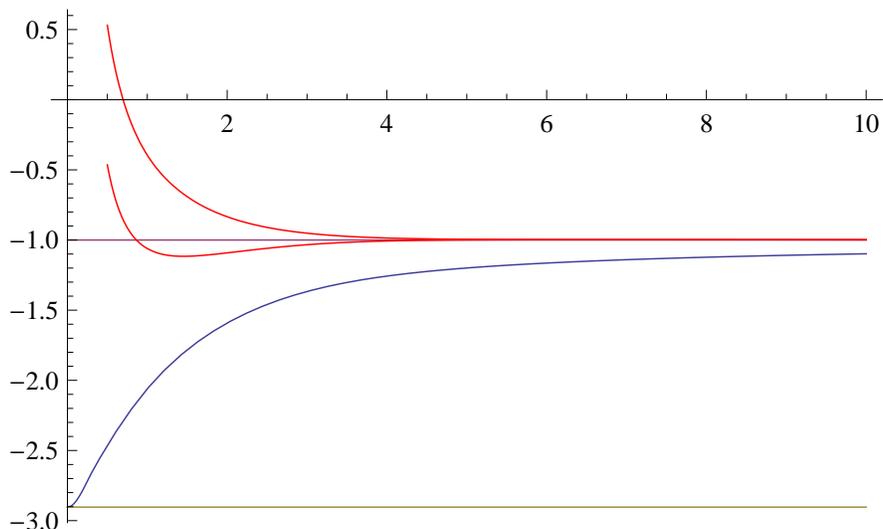


Figure F.12: This figure shows the results of a preliminary calculation on the dissociation of the hydrogen molecule using a very restricted basis set. Energies are shown in Hartrees as functions of the internuclear separation, measured in Bohrs. The lowest curve shows the ground-state electronic energy by itself, without internuclear repulsion. The two upper curves show the ground state and first excited singlet state electronic energies with nuclear repulsion added, i.e. the total energies of the two states. The calculated equilibrium bond length is 1.41 Bohrs, which can be compared with the experimental value, 1.40 Bohrs. It can be seen from the figure that at a separation of 5 Bohrs or more, the molecule is completely dissociated, and in fact the calculated wave function at that internuclear separation corresponds to two neutral hydrogen atoms, each with its own electron, while the total energy corresponds to that of two isolated hydrogen atoms.

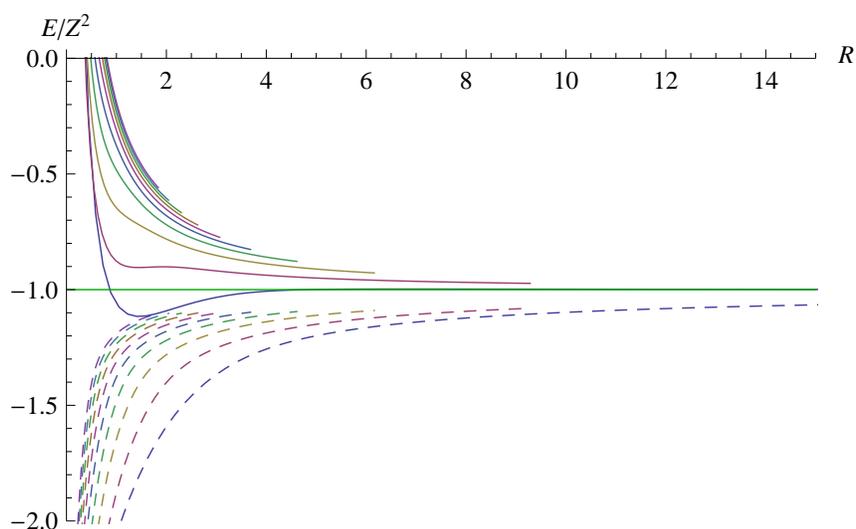


Figure F.13: This figure shows ground-state energies divided by Z^2 for the 2-electron isoelectronic series for homonuclear diatomic molecules, Z being the nuclear charges. The energies in Hartrees are shown as functions of the interatomic distance R , measured in Bohrs. The dotted curves are electronic energies alone, while the solid curves also include internuclear repulsion. For both the solid and dotted curves the lowest curve corresponds to $Z=1$, the next lowest to $Z = 2$, and so on. As in Figure 5.12, a very restricted basis set was used for the calculation.

F.8 Automatic generation of symmetry-adapted basis functions

In Chapter 4, Section 4.2, we discussed a large- Z approximation that could be used when the Generalized Sturmian method is applied to atoms. In the large- Z approximation, the basis set used to treat a particular state is restricted to the set of configurations that become degenerate if interelectron repulsion is completely neglected. We also saw in Chapter 4 that such a set of states could be used as an invariant subset, and thus be used to generate symmetry-adapted basis functions for more accurate calculations where the large- Z approximation is abandoned.

If we compare equation (F.51) with equation (E.6) of Chapter 4, we can see that the two equations are closely analogous. In each case, the configuration $|\Phi_\nu\rangle$ satisfies an approximate Schrödinger equation with a weighted potential V_0 , the weighting factors being chosen in such a way as to make all of the configurations in the basis set isoenergetic. Thus both the Goscinskian configurations of Chapter 4 and the isoenergetic configurations of Chapter 5 are examples of generalized Sturmians, as discussed in Appendix B. Therefore it is interesting to ask whether something analogous to the large- Z approximation exists in the case of molecules. What happens if we decide to use as a basis only those molecular configurations that become degenerate if we completely neglect interelectron repulsion? Let us suppose that this degeneracy is not accidental, but is a due degeneracy. It then follows that any set of configurations that become degenerate if interelectron repulsion is completely neglected is closed under the operations of the symmetry group of the molecule, and it can be used as an invariant subset for the automatic generation of symmetry-adapted basis functions needed in a large and accurate calculation. The method can thus be summarized as follows:

1. Construct the matrices $T_{\nu'\nu}^{(0,N)} \equiv -\frac{1}{k} \langle \Phi_{\nu'} | V_0(\mathbf{x}) | \Phi_\nu \rangle$ and $S_{\nu'\nu}^{(N)}$, based on configurations that are solutions to (F.51). These will already be diagonal, as was discussed above.
2. Search for configurations corresponding to the same value of $T_{\nu,\nu}^{(0,N)} / \mathfrak{S}_{\nu,\nu}^{(N)}$. Such a set of degenerate configurations is an invariant subset provided that the search has been sufficiently complete.
3. For each invariant subset, solve equation (F.60). The eigenfunctions will then be the symmetry-adapted configurations needed for a larger and more accurate calculation.

Use of the *S*-Matrix in the Relativistic Treatment of Resonance Energy Transfer

JOHN AVERY

Department of Physical Chemistry, H. C. Ørsted Institute, University of Copenhagen, Copenhagen, Denmark

Abstract

The *S*-matrix formalism is used to treat the phenomenon of resonance energy transfer (sensitized fluorescence). It is shown that for dipole-allowed transitions and short sensitizer-acceptor separations, the relativistic treatment yields the same result as the nonrelativistic Perrin-Förster theory. For large sensitizer-acceptor separations, long-range coupling terms appear in the relativistic treatment. Resonance energy transfer through these long-range coupling terms is compared with spontaneous photon emission, and direct-interaction theories of electromagnetism are discussed. In the Appendix it is shown that the relativistic theory predicts resonance transfer of triplet excitation energy through the spin-spin coupling term in the Breit interaction.

Experimental Discovery of Resonance Energy Transfer and Early Theories

The phenomenon of resonance energy transfer was discovered experimentally in 1923 by Cario and Franck [1], who exposed a mixture of mercury and thallium vapors to a frequency of light which only the mercury could absorb. In the emission spectrum they observed a frequency of light which only the thallium could emit, which demonstrated that energy had been transferred from one species to the other.

A theory for the mechanism of energy transfer in the Cario-Franck experiment was first proposed by Perrin [2, 3], and later improved by Perrin, Förster, and others [4-9]. In the Perrin-Förster theory of resonance energy transfer, the system is divided into two subunits, and its initial state is represented by a wave function of the form:

$$|\Psi_i\rangle = |\psi_\xi(1) \psi_g(2)\rangle, \quad (1)$$

where $\psi_\xi(1)$ represents subunit 1 in the excited state ξ , and $\psi_g(2)$ represents subunit 2 in the ground state. (For example, subunit 1 might be a mercury atom, while subunit 2 might represent a thallium atom.)

The two subunits are assumed to be separated by a distance R which is large enough so that exchange can be neglected. In the final state of the total system, which is represented by a wave function of the form

$$|\Psi_f\rangle = |\psi_{g'}(1) \psi_{\xi'}(2)\rangle, \quad (2)$$

a transfer of excitation energy has taken place, so that subunit 1 is in its ground state g' , while subunit 2 is in an excited state ξ' .

The Hamiltonian of the total system can be written in the form:

$$H = H_0 + H', \quad (3)$$

where the unperturbed Hamiltonian

$$H_0 = H_1 + H_2 \quad (4)$$

is the sum of the Hamiltonians of the isolated subunits 1 and 2, and where

$$H' = \sum_i \sum_j \frac{q_i q_j}{|\mathbf{x}_i - \mathbf{x}_j|} \quad (5)$$

represents the Coulomb interaction between the charged particles i of subunit 1 and the charged particles j of subunit 2, q_i and q_j being the respective charges.

In the Perrin-Förster theory of resonance energy transfer, $|\Psi_i\rangle$ and $|\Psi_j\rangle$ are eigenfunctions of H_0 , and a transition between them is induced by H' . In order to calculate the probability per unit time of the transition $i \rightarrow f$ as a function of the separation R , one makes a bipolar expansion of $1/|\mathbf{x}_i - \mathbf{x}_j|$ around the center-of-mass positions \mathbf{R}_1 and \mathbf{R}_2 of the two subsystems:

$$\frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} = \left[1 + (\mathbf{x}_i - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} + (\mathbf{x}_j - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} + \left((\mathbf{x}_i - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} \right) \left((\mathbf{x}_j - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} \right) + \dots \right] \frac{1}{|\mathbf{R}_1 - \mathbf{R}_2|}. \quad (6)$$

Because of the orthogonality relationships

$$\int d\tau_1 \psi_\xi^*(1) \psi_\eta(1) = 0, \quad (7)$$

$$\int d\tau_2 \psi_{\xi'}^*(2) \psi_{\eta'}(2) = 0, \quad (8)$$

the matrix elements of the first three terms in Eq. (6) vanish, and the leading term in the matrix element for the transition is the dipole-dipole term:

$$\begin{aligned} & \int d\tau \Psi_i^* \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \Psi_f \\ &= \sum_{\mu, \nu=1}^3 \int d\tau_1 \psi_\xi^*(1) (\mathbf{x}_i - \mathbf{R}_1)_\mu \psi_\eta(1) \int d\tau_2 \psi_{\xi'}^*(2) (\mathbf{x}_j - \mathbf{R}_2)_\nu \psi_{\eta'}(2) \\ & \quad \times \frac{\partial}{\partial R_{1\mu}} \frac{1}{\partial R_{2\nu}} \frac{1}{|\mathbf{R}_1 - \mathbf{R}_2|}. \end{aligned} \quad (9)$$

If we let

$$\mathbf{D}_1 \equiv \int d\tau_1 \psi_\xi^*(1) \sum_i q_i (\mathbf{x}_i - \mathbf{R}_1) \psi_\eta(1) \quad (10)$$

and

$$\mathbf{D}_2 \equiv \int d\tau_2 \psi_g^*(2) \sum_j q_j(\mathbf{x}_j - \mathbf{R}_2) \psi_{g'}(2), \quad (11)$$

and if we note that

$$\frac{\partial^2}{\partial R_{1\mu} \partial R_{2\nu}} \frac{1}{R} = \frac{\delta_{\mu\nu}}{R^3} - \frac{3R_\mu R_\nu}{R^5}, \quad (12)$$

where

$$\mathbf{R} \equiv \mathbf{R}_1 - \mathbf{R}_2, \quad (13)$$

then Eqs. (5) and (9) yield:

$$\int d\tau \Psi_i^* H' \Psi_f = \frac{\mathbf{D}_1 \cdot \mathbf{D}_2}{R^3} - \frac{3(\mathbf{D}_1 \cdot \mathbf{R})(\mathbf{D}_2 \cdot \mathbf{R})}{R^5} + \dots \quad (14)$$

Since the matrix element for the transition falls off as $1/R^3$, the transition probability per unit time falls off as $1/R^6$. Because of this, one sometimes expresses the Perrin-Förster rate constant for resonance energy transfer in the form:

$$\frac{1}{\tau_{i \rightarrow f}} = \frac{1}{\tau_0} \left(\frac{R_0}{R} \right)^6, \quad (15)$$

where τ_0 is the time needed for spontaneous emission of a photon, and R_0 is the separation of 1 and 2 for resonance energy transfer which is equal to the probability for spontaneous photon emission. Typically, for dipole-allowed transitions, R_0 is of the order of 50 Å. The validity of the Perrin-Förster theory for dipole-allowed resonance energy transfer over distances of this order has been amply confirmed by experiments [11-14].

Relativistic Treatment of Resonance Energy Transfer

Let us now try to improve the Perrin-Förster theory by taking relativistic effects into account [15, 16]. To do this, we begin with the second-order electron-electron S -matrix [17]:

$$S = \sum_{r,s,t,u} S_{rs|tu} b_r^\dagger b_s^\dagger b_t b_u, \quad (16)$$

where

$$S_{rs|tu} = e^2 \int d^4x_1 \int d^4x_2 \bar{\Phi}_r(1) \gamma_\mu^1 \Phi_u(1) G(1, 2) \bar{\Phi}_s(2) \gamma_\mu^2 \Phi_t(2), \quad (17)$$

$$d^4x_j = d^3x_j dt_j, \quad (18)$$

and where $G(1, 2)$ is the Green's function for the d'Alembertian operator:

$$G(1, 2) = \frac{1}{4\pi^3} \int \frac{d^4k \exp i\mathbf{k} \cdot (\mathbf{x}_1 - \mathbf{x}_2) - i\omega(t_1 - t_2)}{\mathbf{k} \cdot \mathbf{k} - (\omega/c)^2 - i\epsilon}. \quad (19)$$

Here ε is a real, positive infinitesimal number which is allowed to go to zero after all the integrations have been performed, and $d^4k = d^3k d\omega$. In Eq. (16), $b_r^\dagger, b_s^\dagger, b_r$, and b_u are electron creation and annihilation operators obeying the anticommutation relations:

$$\begin{aligned} b_r^\dagger b_s^\dagger + b_s^\dagger b_r^\dagger &= 0, \\ b_r b_s + b_s b_r &= 0, \\ b_r^\dagger b_s + b_s b_r^\dagger &= \delta_{rs}. \end{aligned} \quad (20)$$

The electron creation and annihilation operators in Eq. (20) refer to spin orbitals which are eigenstates of the Dirac operator in a local potential (for example, the effective potential of an atom or molecule), and we are thus working in an independent particle approximation. In Eqs. (17) and (18) and throughout the remainder of this paper, a sum over μ running from 1 to 4 is implied whenever the index μ is repeated. The matrices γ_μ^1 and γ_μ^2 are defined as in Ref. 18, with the superscripts indicating that they act, respectively, on the spin coordinates of electrons 1 and 2. We also use the definition $\bar{\Phi}_r \equiv i\Phi_r^\dagger \gamma_4$, where the dagger means "conjugate transpose," and in general we shall follow the notation of Ref. 18.

The matrix element in Eqs. (17) and (18) can be brought into another form by performing integrations over dt_2 , d^4k , and dt_1 : If we assume that the orbitals Φ_s and Φ_t depend harmonically on time, then we can write:

$$ec \bar{\Phi}_s(2) \gamma_\mu^2 \Phi_t(2) = j_\mu^{st}(\mathbf{x}_2) \exp(i\omega_{st}t_2), \quad (21)$$

where $\omega_{st} \equiv \omega_s - \omega_t$ and

$$j_\mu^{st}(\mathbf{x}_2) \equiv ec \bar{\Phi}_s(\mathbf{x}_2) \gamma_\mu^2 \Phi_t(\mathbf{x}_2), \quad (22)$$

and similarly we can write:

$$ec \bar{\Phi}_r(1) \gamma_\mu^1 \Phi_u(1) = j_\mu^{ru}(\mathbf{x}_1) \exp(i\omega_{ru}t_1), \quad (23)$$

where

$$j_\mu^{ru}(\mathbf{x}_1) \equiv ec \bar{\Phi}_r(\mathbf{x}_1) \gamma_\mu^1 \Phi_u(\mathbf{x}_1). \quad (24)$$

Then, integrating over dt_2 , d^4k , and dt_1 , we obtain [17]:

$$\begin{aligned} S_{rs|tu} &= e^2 \int d^3x_1 dt_1 \int d^3x_2 dt_2 j_\mu^{ru}(\mathbf{x}_1) \exp(i\omega_{ru}t_1) G(1, 2) j_\mu^{st}(\mathbf{x}_2) \exp(i\omega_{st}t_2) \\ &= -2\pi\delta(\omega_r - \omega_u + \omega_s - \omega_t) U_{rs|tu}, \end{aligned} \quad (25)$$

where

$$U_{rs|tu} = -\frac{1}{c^2} \int d^3x_1 \int d^3x_2 j_\mu^{ru}(\mathbf{x}_1) \frac{\exp(i|\omega_{st}|r_{12})}{r_{12}} j_\mu^{st}(\mathbf{x}_2) \quad (26)$$

and where the transition currents $j_\mu^{ru}(\mathbf{x}_1)$ and $j_\mu^{st}(\mathbf{x}_2)$ are defined by Eqs. (22) and (24). The effective interaction energy matrix element $U_{rs|tu}$ is similar to the matrix element of a Coulomb interaction, except that the transition densities

are replaced by a scalar product of transition four-currents, and the interaction contains a retardation factor, $\exp(i|\omega_{st}|r_{12}/c)$.

In order to apply Eqs. (16)–(26) to the phenomenon of resonance energy transfer, let us represent the initial state of a system composed of two subunits by a wave function of the form:

$$|\Psi_i\rangle = |\psi_\varepsilon(1) \psi_g(2)\rangle = b_m^\dagger b_n |\psi_{g'}(1) \psi_g(2)\rangle = b_m^\dagger b_n |\Psi_0\rangle \quad (27)$$

and let us represent the final state by the wave function

$$|\Psi_f\rangle = |\psi_{g'}(1) \psi_\varepsilon(2)\rangle = b_1^\dagger b_k |\psi_{g'}(1) \psi_g(2)\rangle = b_1^\dagger b_k |\Psi_0\rangle. \quad (28)$$

Equations (27) and (28) have the same meaning as Eqs. (1) and (2). In the initial state of the system, $|\Psi_i\rangle$, subunit 1 is in a singly excited state $\psi_\varepsilon(1)$ which differs from the ground state by the excitation of an electron from the orbital n to the orbital m , while subunit 2 is in its ground state $\psi_g(2)$. In the final state, the situation is reversed, and subunit 2 is excited while subunit 1 is in its ground state. The transition probability is proportional to $|\langle\Psi_i|S|\Psi_f\rangle|^2$. Combining Eqs. (16), (27), and (28), and making use of the anticommutation relations (20), we obtain:

$$\langle\Psi_i|S|\Psi_f\rangle = S_{mk|ln} - S_{mk|nl} = S_{mk|ln}. \quad (29)$$

In Eq. (29) we have used the fact that the two subunits are assumed to be separated by a distance $R = |\mathbf{R}_1 - \mathbf{R}_2|$ which is large enough so that the exchange term can be neglected.

Let us try to find the dependence of $U_{mk|ln}$ on the separation R . We can make a bipolar expansion of $(1/r_{12}) \exp(i|\omega_{kl}|r_{12}/c)$ about the points \mathbf{R}_1 and \mathbf{R}_2 . If we let $\kappa \equiv |\omega_{kl}|/c$, then the expansion analogous to Eq. (6) is

$$\frac{\exp(i\kappa|\mathbf{x}_1 - \mathbf{x}_2|)}{|\mathbf{x}_1 - \mathbf{x}_2|} = \left(1 + (\mathbf{x}_1 - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} + (\mathbf{x}_2 - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} + \dots\right) \frac{\exp(i\kappa|\mathbf{R}_1 - \mathbf{R}_2|)}{|\mathbf{R}_1 - \mathbf{R}_2|}. \quad (30)$$

In the nonrelativistic theory of resonance energy transfer, the first three terms in the bipolar expansion [Eq. (6)] vanish because of the orthonormality relations [Eqs. (7) and (8)]. By contrast, the first three terms in Eq. (30) all give rise to nonvanishing contributions to the matrix element $U_{mk|ln}$.

Substituting Eq. (30) into Eq. (26) we obtain:

$$\begin{aligned} U_{mk|ln} = & -\frac{1}{c^2} \left(\int d^3x_1 j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2 j_\mu^{kl}(\mathbf{x}_2) + \int d^3x_1 (\mathbf{x}_1 - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} j_\mu^{mn}(\mathbf{x}_1) \right. \\ & \times \int d^3x_2 j_\mu^{kl}(\mathbf{x}_2) + \int d^3x_1 j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2 (\mathbf{x}_2 - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} j_\mu^{kl}(\mathbf{x}_2) \\ & + \int d^3x_1 (\mathbf{x}_1 - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2 (\mathbf{x}_2 - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} j_\mu^{kl}(\mathbf{x}_2) \\ & \left. + \dots \right) \frac{e^{i\kappa R}}{R}. \quad (31) \end{aligned}$$

To see that the first term in Eq. (31) does not vanish, we note that

$$\begin{aligned} \int d^3x_1 j_\mu^{mn}(\mathbf{x}_1) &= ec \int d^3x_1 \bar{\Phi}_m(\mathbf{x}_1) \gamma_\mu^1 \Phi_n(\mathbf{x}_1) \\ &= iec \int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1) \gamma_4^1 \gamma_\mu^1 \Phi_n(\mathbf{x}_1). \end{aligned} \quad (32)$$

The fourth component vanishes, since $\gamma_4^1 \gamma_4^1 = 1$ and, if $m \neq n$,

$$\int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1) \Phi_n(\mathbf{x}_1) = 0. \quad (33)$$

However, the space part of the four-vector $\int d^3x_1 j_\mu^{mn}(\mathbf{x}_1)$ does not vanish, since

$$\begin{aligned} \int d^3x_1 \bar{\Phi}_m(\mathbf{x}_1) \gamma^1 \Phi_n(\mathbf{x}_1) &= i \int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1) \gamma_4^1 \gamma^1 \Phi_n(\mathbf{x}_1) \\ &= \int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1) \alpha^1 \Phi_n(\mathbf{x}_1). \end{aligned} \quad (34)$$

Thus we find that in the relativistic theory of resonance energy transfer, there is a nonzero term of the form:

$$e^2 \int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1) \alpha^1 \Phi_n(\mathbf{x}_1) \cdot \int d^3x_2 \Phi_k^\dagger(\mathbf{x}_2) \alpha^2 \Phi_l(\mathbf{x}_2) \frac{e^{i\kappa R}}{R} \neq 0. \quad (35)$$

This term represents magnetic coupling between the two subsystems, and it is remarkable because it falls off only as $1/R$.

Other terms proportional to $1/R$ arise when the differential operators $\partial/\partial\mathbf{R}_1$ and $\partial/\partial\mathbf{R}_2$ act exclusively on the retardation factor $e^{i\kappa R}$ rather than on the denominator of $e^{i\kappa R}/R$. Thus, for example,

$$\frac{\partial^2}{\partial R_{1\sigma} \partial R_{2\tau}} \frac{e^{i\kappa R}}{R} = \frac{e^{i\kappa R}}{R} \left(\delta_{\sigma\tau} (1 - i\kappa R) - \frac{R_\sigma R_\tau (3 - i\kappa R - \kappa^2 R^2)}{R^2} \right). \quad (36)$$

The term proportional to κ^2 in Eq. (41) falls off only as $1/R$. If we substitute Eq. (36) into the fourth term of Eq. (31), we obtain:

$$\begin{aligned} &\int d^3x_1 (\mathbf{x}_1 - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2 (\mathbf{x}_2 - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} j_\mu^{kl}(\mathbf{x}_2) \frac{e^{i\kappa R}}{R} \\ &\xrightarrow{\kappa R \gg 1} e^2 c^2 \kappa^2 \frac{e^{i\kappa R}}{R} \int d^3x_1 (\mathbf{x}_1 - \mathbf{R}_1) \cdot \mathbf{R} j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2 (\mathbf{x}_2 - \mathbf{R}_2) \cdot \mathbf{R} j_\mu^{kl}(\mathbf{x}_2). \end{aligned} \quad (37)$$

If the transitions $m \rightarrow n$ and $k \rightarrow l$ are dipole allowed, then the $\mu = 4$ term predominates over the $\mu = 1, 2$, and 3 terms, and Eq. (37) becomes:

$$\int d^3x_1(\mathbf{x}_1 - \mathbf{R}_1) \cdot \frac{\partial}{\partial \mathbf{R}_1} j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2(\mathbf{x}_2 - \mathbf{R}_2) \cdot \frac{\partial}{\partial \mathbf{R}_2} j_\mu^{kl}(\mathbf{x}_2) \frac{e^{i\kappa R}}{R} \\ \xrightarrow{\kappa R \gg 1} \frac{c^2 \kappa^2 (\mathbf{D}_1 \cdot \mathbf{R})(\mathbf{D}_2 \cdot \mathbf{R})}{R^2} \frac{e^{i\kappa R}}{R} + \dots \quad (38)$$

where \mathbf{D}_1 and \mathbf{D}_2 are the transition dipole moments:

$$\mathbf{D}_1 \equiv e \int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1)(\mathbf{x}_1 - \mathbf{R}_1)\Phi_n(\mathbf{x}_1), \\ \mathbf{D}_2 = e \int d^3x_2 \Phi_k^\dagger(\mathbf{x}_2)(\mathbf{x}_2 - \mathbf{R}_2)\Phi_l(\mathbf{x}_2). \quad (39)$$

With the help of the commutation relation:

$$[H_e, \mathbf{x}] = (\hbar c/i) \boldsymbol{\alpha}, \\ H_e \equiv (\hbar c/i) \boldsymbol{\alpha} \cdot \partial/\partial \mathbf{x}, \quad (40)$$

we can express Eq. (35) in the form:

$$\int d^3x_1 j_\mu^{mn}(\mathbf{x}_1) \int d^3x_2 j_\mu^{kl}(\mathbf{x}_2) = e^2 c^2 \int d^3x_1 \Phi_m^\dagger(\mathbf{x}_1) \boldsymbol{\alpha}^1 \Phi_n(\mathbf{x}_1) \cdot \int d^3x_2 \Phi_k^\dagger(\mathbf{x}_2) \boldsymbol{\alpha}^2 \Phi_l(\mathbf{x}_2) \\ = c^2 \kappa^2 (\mathbf{D}_1 \cdot \mathbf{D}_2), \quad (41)$$

with \mathbf{D}_1 and \mathbf{D}_2 defined by Eq. (39). Thus when R is large compared with $1/\kappa = \lambda/2\pi = c/|\omega_{mn}|$, then the leading term in the multipole expansion is

$$U_{mkln} \xrightarrow{\kappa R \gg 1} \kappa^2 \left(\frac{(\mathbf{D}_1 \cdot \mathbf{R})(\mathbf{D}_2 \cdot \mathbf{R})}{R^2} - (\mathbf{D}_1 \cdot \mathbf{D}_2) \right) \frac{e^{i\kappa R}}{R}. \quad (42)$$

In the opposite case, when R is small compared with $1/\kappa$, substitution of Eq. (36) into Eq. (31) yields:

$$U_{mkln} \xrightarrow{\kappa R \ll 1} \frac{(\mathbf{D}_1 \cdot \mathbf{D}_2)}{R^3} - \frac{3(\mathbf{D}_1 \cdot \mathbf{R})(\mathbf{D}_2 \cdot \mathbf{R})}{R^5}. \quad (43)$$

Comparing this with Eq. (14), we can see that the relativistic treatment of resonance energy transfer gives the same result as the Perrin-Förster theory for dipole-allowed transitions and small values of R . For large values of R , the relativistic treatment of the dipole-allowed case yields a coupling term [Eq. (42)] that falls off as $1/R$.

Probability Per Unit Time for Resonance Energy Transfer

Let us now turn to the problem of calculating the probability per unit time for resonance energy transfer. In the nonrelativistic case, the matrix element of

the perturbation inducing the transition $|\Psi_i\rangle \rightarrow |\Psi_f\rangle$ has a leading term given by Eq. (14). In first-order time-dependent perturbation theory, the probability per unit time that the transition will take place is given by $|a_f|^2/t$ where

$$a_f = \frac{\langle \Psi_i | H' | \Psi_f \rangle}{i\hbar} \int_0^t \exp(i\omega_{if}t') dt' = \frac{H'_{if}}{i\hbar\omega_{if}} [\exp(i\omega_{if}t) - 1], \quad (44)$$

so that

$$|a_f|^2 = \frac{|H'_{if}|^2}{\hbar^2} \frac{[\exp(i\omega_{if}t) - 1][\exp(-i\omega_{if}t) - 1]}{\omega_{if}^2}. \quad (45)$$

We must now remember that the system is not completely isolated, but is coupled with its surroundings. This coupling can be taken into account by saying that the states $|\Psi_i\rangle$ and $|\Psi_f\rangle$ are not exactly stationary, but that there are certain unspecified interactions with the environment other than H' . These interactions collectively produce, on the average, Γ transitions per unit time which will take the system out of the state $|\Psi_i\rangle$ or the state $|\Psi_f\rangle$ and into some other state. This coupling of the system with its environment can be taken into account in our simplified perturbation model by saying that the states $|\Psi_i\rangle$ and $|\Psi_f\rangle$ are split by the interaction into a group of states. The density of these states (i.e., the number of states per energy interval) is given by a Lorentzian distribution:

$$\begin{aligned} \rho_i(\omega) &= \frac{\Gamma/\pi}{(\omega_i - \omega)^2 + \Gamma^2}, \\ \rho_f(\omega) &= \frac{\Gamma/\pi}{(\omega_f - \omega)^2 + \Gamma^2}. \end{aligned} \quad (46)$$

This type of distribution corresponds, for example, to the measured line shape for atomic spectra. Γ is the radiation half-life for spectra of isolated atoms in a rarified gas. In a gas at atmospheric pressure, where collision broadening predominates, $1/\Gamma$ corresponds to the average time between collisions. The factor Γ/π in the numerator in Eq. (45) normalizes the distribution in such a way that

$$\int_{-\infty}^{\infty} d\omega \rho(\omega) = 1. \quad (47)$$

We must now calculate the probability of transition from one of the group of states into which $|\Psi_i\rangle$ is split by the interaction with its environment to one of the group of states into which $|\Psi_f\rangle$ is split. Then we must sum these probabilities over the whole set of final states, and finally we must sum over the whole set of initial states. In other words, we must evaluate:

$$\begin{aligned} P &= \frac{|H'_{if}|^2}{\hbar^2} \int_{-\infty}^{\infty} \frac{(\Gamma/\pi) d\omega'}{(\omega' - \omega_i)^2 + \Gamma^2} \int_{-\infty}^{\infty} \frac{(\Gamma/\pi) d\omega}{(\omega - \omega_f)^2 + \Gamma^2} \\ &\quad \times \left(\frac{\{\exp[i(\omega - \omega')t] - 1\} \{\exp[-i(\omega - \omega')t] - 1\}}{(\omega - \omega')^2} \right). \end{aligned} \quad (48)$$

Evaluating the double integral in Eq. (48) by means of contour integration, we obtain:

$$P = \frac{|H'_{if}|^2 4\Gamma t}{\hbar^2 [(\omega_i - \omega_f)^2 + \Gamma^2]}. \quad (49)$$

This represents the probability that after an interval t the system will have undergone a transition from the initial state $|\Psi_i\rangle$ to the final state $|\Psi_f\rangle$. Thus, in the special case where $\omega_i = \omega_f$, the rate at which the transition takes place becomes:

$$\frac{1}{\tau_{i \rightarrow f}} = \left(\frac{P}{t} \right)_{\omega_i = \omega_f} = \frac{4|H'_{if}|^2}{\hbar^2 \Gamma}. \quad (50)$$

Using Eq. (14) and taking the z axis in the direction of \mathbf{R} , we can write for the nonrelativistic case:

$$\begin{aligned} |H'_{if}|^2 &= \left| \frac{1}{R^3} \{D_{1x}D_{2x} + D_{1y}D_{2y} - 2D_{1z}D_{2z}\} \right|^2 \\ &= \frac{1}{R^6} (D_{1x}^2 D_{2x}^2 + D_{1y}^2 D_{2y}^2 + 4D_{1z}^2 D_{2z}^2 - 2D_{1x}D_{1z}D_{2x}D_{2z} \\ &\quad - 2D_{1y}D_{1z}D_{2y}D_{2z} + D_{1x}D_{1y}D_{2x}D_{2y}). \end{aligned} \quad (51)$$

If our manifold of final states includes states which represent resonance energy transfer to various randomly oriented absorber molecules, we must average over the orientations of \mathbf{D}_2 with respect to \mathbf{R} . Doing this, and averaging also over the orientations of \mathbf{D}_1 with respect to \mathbf{R} , we obtain:

$$\frac{1}{\langle \tau_{i \rightarrow f} \rangle} = \frac{1}{(4\pi)^2} \int d\Omega_1 \int d\Omega_2 \frac{1}{\tau_{i \rightarrow f}} = \frac{8|\mathbf{D}_1|^2 |\mathbf{D}_2|^2}{3R^6 \Gamma \hbar^2}. \quad (52)$$

For dipole-allowed transitions in organic dyes, such as acriflavine and rhodamine B (two dyes commonly used in studying resonance energy transfer), if the transition dipole moment is expressed in the form $D = eL$, where e is the charge of an electron, then L is a length whose order of magnitude is a few Ångströms. If we let $\Gamma \approx 10^{13} \sim /s$ and $\tau_0 \approx 10^{-8}$ s and if we rewrite Eq. (52) in the form:

$$\frac{1}{\langle \tau_{i \rightarrow f} \rangle} = \frac{1}{\tau_0} \left(\frac{R_0}{R} \right)^6, \quad (53)$$

then

$$R_0 = \left(\frac{8e^4 L^4 \tau_0}{3\hbar^2 \Gamma} \right)^{1/6} \approx 100 \text{ Å}. \quad (54)$$

Now let us consider what effect the relativistic corrections have on the rate of resonance energy transfer for dipole-allowed transitions. If we replace H'_{if} by the matrix element of the effective interaction energy, $U_{mk|ln}$ [Eqs. (31),

(42), and (43)], then we find that for $\kappa R \ll 1$ the rate of resonance energy transfer is given by Eqs. (53) and (54), in agreement with the Perrin-Förster theory. However, when $\kappa R \approx 1$, the relativistic effects begin to become important, and when $\kappa R \gg 1$, they predominate. Combining Eqs. (42) and (50) and averaging over orientations of the dipole moments, we obtain an average transition rate for the region where $\kappa R \gg 1$:

$$\frac{1}{\langle \tau_{i \rightarrow f} \rangle} = \frac{8\kappa^4 |\mathbf{D}_1|^2 |\mathbf{D}_2|^2}{9\hbar^2 R^2 \Gamma}. \quad (55)$$

The R dependence of the average rate of resonance energy transfer for dipole-allowed transitions is illustrated in Figure 1.

Long-Range Resonance Energy Transfer

From Eq. (55) we can see that when $\kappa R \gg 1$, the probability for resonance transfer of excitation energy from an excited atom or molecule to an absorber

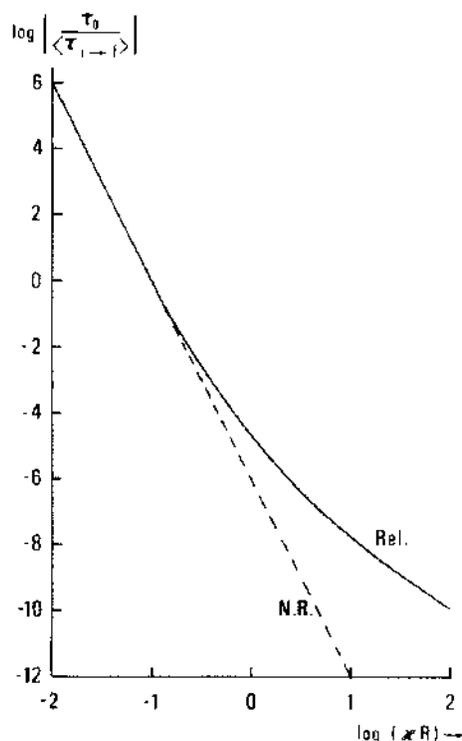


Figure 1. Dependence of the averaged rate of resonance energy transfer on the sensitizer-acceptor separation R for dipole-allowed transitions. For small separations, relativistic corrections to the Perrin-Förster theory are unimportant, and the transition rate falls off as $1/R^6$. When $R = 1/\kappa = \lambda/2\pi$, relativistic corrections become important, and for larger separations, they predominate. When $\kappa R \gg 1$, the averaged transition rate falls off as $1/R^2$. (—) Relativistic transfer rate; (---) transfer rate in the nonrelativistic Perrin-Förster theory.

is proportional to $1/R^2$. Let us now imagine an excited atom placed at the center of a sphere of radius R , where $R > 1/\kappa$. Let us also imagine that on the surface of the sphere there are a number of similar atoms in their ground states. According to the argument outlined above, the probability for resonance energy transfer to one or another of the atoms on the sphere will fall off as $1/R^2$. On the other hand, the number of absorber atoms which we can place on the sphere without changing the number per unit area will increase as R^2 . Thus, if we keep the number per unit area constant, the probability of resonance energy transfer to one or another of the absorbers will be independent of R . For example, suppose that the transition dipole moment of the atoms which we are considering is given by $D = eL$, and suppose that the number of absorber atoms per unit area on the surface of the sphere is $1/4\pi L^2$. Then the total number on the surface of the sphere will be $N = 4\pi R^2/4\pi L^2 = R^2/L^2$. The ratio of the rate of resonance energy transfer to one or another of the absorbers to the rate of spontaneous photon emission will have the order of magnitude [from Eq. (55)]:

$$\frac{\tau_0}{\langle \tau_{i \rightarrow f} \rangle} = \left(\frac{e^2}{\hbar c} \right)^2 \frac{L^2 c^2 \tau_0}{\Gamma \lambda^4}, \quad (56)$$

where $\lambda = 1/\kappa$. If we let $\tau_0 \approx 10^{-8}$ s, $\Gamma = 10^{13}$ ~ /s, and $\lambda \approx 1000$ Å, Eq. (56) gives the order of magnitude of the ratio as approximately unity. We can see from this argument that a theory of resonance energy transfer which takes relativistic effects into account leads to the surprising prediction that resonance energy transfer can take place over macroscopic distances! (It should be remembered, however, that the simple first-order perturbation theory which we have used here is inadequate in an important respect since it is unable to take into account the screening of one absorber by another. In order to describe this screening, it would be necessary to treat resonance energy transfer in a manner analogous to the treatment of delocalization of an initially localized Frenkel exciton state [19–21].)

Possible Experiments to Observe Long-Range Energy Transfer

The Perrin-Förster-type of resonance energy transfer (sometimes called "sensitized fluorescence") is usually studied in aqueous solutions of organic dyes, such as acridine orange, acriflavine, rhodamine B, etc. These dyes have the advantage of large transition dipole moments. They also have a high fluorescence efficiency, i.e., it seldom happens that an excited dye molecule loses all of its energy to vibrational modes. A little energy is always lost [22], and therefore the absorption maximum does not coincide with the fluorescence maximum. Resonance energy transfer is most likely when the fluorescence maximum of the excited molecule corresponds to the absorption maximum of the absorbers. In the usual experiments, two dyes, a "sensitizer" S and an "acceptor" A are chosen such that this resonance condition is fulfilled, and they are mixed together in solution. The solution is irradiated with light whose frequency corresponds to the absorption maximum of S , and the fluorescence spectrum is observed.

It is found that the quantum efficiency of the fluorescence of S is increased by the presence of A, and that the fluorescence efficiency of A is increased ("sensitized") by the presence of S. In moderately concentrated solutions, where the average S-A distance is of the order of 50 Å, the deduced rate of energy transfer is far too great to be explained by emission and reabsorption of photons.

The difficulty with this type of experiment as a means of detecting long-range resonance energy transfer is that the total observed fluorescence is due to the combined effect of a great many S and A molecules. These molecules are randomly distributed throughout the solution. The environment of some sensitizer molecules includes an acceptor at very close range, whereas other sensitizers have no acceptor nearby. One of Förster's important achievements is the deduction of a $1/R^6$ dependence for the transfer rate from a statistical analysis [23] of the concentration dependence of sensitized fluorescence in dye solutions. However, the indirectness and difficulty of such an analysis have led other authors to hope that experiments could be performed in which the S-A would be constant and measurable. Zwick and Kuhn [24] achieved such an experiment by absorbing the S and A molecules on opposite sides of thin layers of barium stearate, using a technique developed by Langmuir and Schaefer [25]. The dyes used by Zwick and Kuhn were acriflavine and rhodamine B. The absorption and fluorescence maxima of both these dyes are in the visible range, whereas barium stearate is transparent in the visible, having its absorption maximum in the ultraviolet. The layer of stearate was therefore optically inert and served only to separate the S monolayer from the A monolayer. Zwick and Kuhn measured the quantum fluorescence efficiency of the acceptor as a function of film thickness, and they found a deduced transfer rate in agreement with a $1/R^6$ dependence up to a film thickness of 100 Å.

The Zwick-Kuhn technique seems better suited to an attempt to measure long-range resonance energy transfer than an experiment carried out in solution. In a solution, the extremely strong short-range transfer between molecules which are accidentally close together forms a background which would mask the weaker long-range effects. On the other hand, for two layers separated by an inert film, the S-A distance is always greater than the thickness of the film. Therefore one might attempt an experiment of the following kind: Consider a sandwich-like arrangement with monolayers of S and A separated by a layer of a transparent medium as shown in Figure 2. The whole transparent film is covered on one side with the sensitizer dye. On the other side of the film, only half the area is covered with the acceptor. The film is illuminated from the S side with monochromatic light at the absorption peak of S, and the fluorescence efficiency of S is observed from the S side as a function of film thickness. If resonance energy transfer from S to A is competing with emission of a photon by S, then the fluorescence of S, observed from the S side, will be quenched in the region of the film where A is present on the opposite side. If this quenching is observed, then resonance energy transfer can be inferred, since the presence of A on the opposite side could not decrease the fluorescence of S by any other mechanism. Such an experiment has not yet been performed (see Figs. 2 and 3).

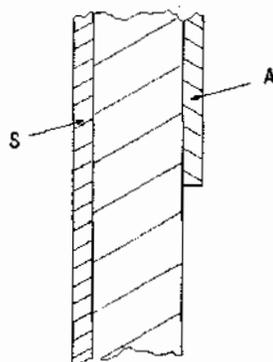


Figure 2. Arrangement of sensitizer and acceptor layers on opposite sides of a thin film of barium stearate in the experiments performed by Zwick and Kuhn [24]. These authors measured the rate of energy transfer from S to A up to film thicknesses of 100 Å. Experiments at larger film thicknesses have not yet been performed.

Direct-Interaction Theories of Electrodynamics

We saw above that a theory of resonance energy transfer which includes relativistic effects leads to the very surprising prediction that the rate of resonance energy transfer over macroscopic distances can be comparable to the rate of spontaneous photon emission. We must now ask whether this predicted long-range resonance energy transfer is a phenomenon which competes with spontaneous photon emission, or whether it is just another way of describing the spontaneous emission of a photon and its subsequent absorption by the acceptor.

Suppose that we choose the second answer and maintain that long-range resonance energy transfer is just another way of describing spontaneous photon

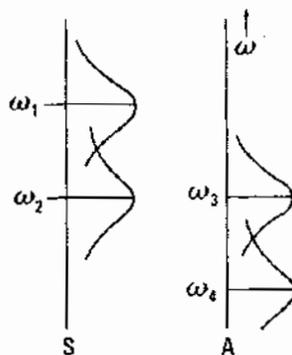


Figure 3. Absorption maxima (ω_1 and ω_3) and fluorescence maxima (ω_2 and ω_4) for experiments on resonance energy transfer performed with organic dye molecules. Because of the loss of energy to vibrational modes, the fluorescence maxima of both the sensitizer and acceptor are shifted to the red with respect to the absorption maxima. The rate of transfer is greatest when the fluorescence maximum of the sensitizers coincides with the absorption maximum of the acceptors, i.e., when $\omega_2 = \omega_3$. In our simplified theory, Γ represents the halfwidth of the bands.

emission. Then we should note an important difference between the resonance transfer description and the conventional one. In the conventional description, no absorber need be present in order for the photon to be emitted. For example, in the standard cosmological description of the early stages of the universe [26], all the matter is imagined to have been initially concentrated in a small region of space. Subsequently, the matter exploded out from this position, but with a velocity less than the velocity of light. Thus, unless space is closed, it is divided into two regions—a central region containing matter and an outer region which is entirely empty. According to the conventional description of spontaneous photon emission, photons must be streaming out from the central region into the totally empty space beyond. On the other hand, if absorbers are necessary for emission of radiation, then at all times during the evolution of the universe the radiation energy would be retained within the region occupied by the matter. Thus the two pictures lead to very different cosmological models, which could, in principle, be distinguished experimentally.

We should also note another difference between the resonance energy transfer description of radiation and the conventional one. In the conventional description, the radiation field has an infinite number of degrees of freedom, while in the direct-interaction picture, the field has no more degrees of freedom than the particles which generate it. The infinite number of degrees of freedom which the electromagnetic field has in the conventional picture lead to an embarrassing infinity in the zero-point energy of the field, and this infinity is avoided in the direct-interaction picture. For this reason, a number of authors have explored the possibility of formulating quantum electrodynamics in terms of a direct interaction [27–32, 35–42]. Interest in direct-interaction formulations of electrodynamics has also been motivated by the desire of quantum physicists and chemists to have a fully relativistic method for calculating the bound states of atoms and molecules. For example, Roothaan has recently approached the problem from this direction [34].

I hope that the present paper will contribute to the current interest in direct-interaction theories of quantum electrodynamics.

Appendix: Resonance Transfer of Triplet Excitation Energy through Spin-Spin Coupling Terms in the Breit Interaction

In the main part of this paper, we have discussed resonance energy transfer for cases where both the sensitizer and acceptor undergo dipole-allowed transitions. In this appendix, we shall treat triplet energy transfer. This is a case in which the sensitizer and acceptor undergo dipole-forbidden transitions involving a spin flip.

A theory of triplet energy transfer has been developed by Dexter for cases where the sensitizer and acceptor are close enough to each other so that there is some overlap of the wave functions [9]. However, in this appendix we shall confine our attention to cases where the sensitizer-acceptor separation R is large enough so that overlap between the wave functions can be neglected. We shall

see that in this case, the spin-spin coupling terms in the Breit interaction produce a small probability for triplet energy transfer. Although the rate of triplet energy transfer induced by this coupling is small, we shall see that it is large enough to be observed during the very long lifetimes which the triplet states of aromatic hydrocarbons exhibit when the molecules are frozen in rigid matrices.

In order to discuss the spin-spin coupling, it is convenient to expand the matrix element of the effective interaction energy $U_{rs|tu}$ [Eq. (26)] in powers of $1/c$, retaining terms up to $1/c^2$, and solving for the small components of the four-component Dirac spinors in terms of the large components. If we do this, we obtain [17]

$$U_{rs|tu} = \int d^3x_1 \int d^3x_2 \phi_r^*(\mathbf{x}_1) \phi_s^*(\mathbf{x}_2) U^{(e)} \phi_t(\mathbf{x}_2) \phi_u(\mathbf{x}_1), \quad (\text{A1})$$

where $U^{(e)}$ is the Breit interaction operator in the Pauli approximation:

$$\begin{aligned} U^{(e)} = & \frac{e^2}{r} - \pi \left(\frac{e\hbar}{mc} \right)^2 \delta(\mathbf{r}) - \left(\frac{e\hbar}{2mc} \right)^2 \frac{1}{\hbar} \\ & \times [(\mathbf{r} \times \mathbf{p}_1) \cdot \boldsymbol{\sigma}_1 - (\mathbf{r} \times \mathbf{p}_2) \cdot \boldsymbol{\sigma}_2 + 2(\mathbf{r} \times \mathbf{p}_1) \cdot \boldsymbol{\sigma}_2 - 2(\mathbf{r} \times \mathbf{p}_2) \cdot \boldsymbol{\sigma}_1] \\ & - \frac{e^2}{2m^2c^2} \left(\frac{1}{r} \mathbf{p}_1 \cdot \mathbf{p}_2 + \frac{1}{r^3} \mathbf{r}(\mathbf{r} \cdot \mathbf{p}_1) \cdot \mathbf{p}_2 \right) + \left(\frac{e\hbar}{2mc} \right)^2 \\ & \times \left(\frac{\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2}{r^3} - \frac{3(\boldsymbol{\sigma}_1 \cdot \mathbf{r})(\boldsymbol{\sigma}_2 \cdot \mathbf{r})}{r^5} - \frac{8\pi}{3} (\mathbf{a}_1 \cdot \boldsymbol{\sigma}_2) \delta(\mathbf{r}) \right). \end{aligned} \quad (\text{A2})$$

In Eq. (A2), $\mathbf{r} = \mathbf{x}_1 - \mathbf{x}_2$, while $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are three vectors whose components are the Pauli spin matrices σ_x , σ_y , and σ_z . In Eq. (A1), ϕ_r , ϕ_s , ϕ_t , and ϕ_u are two-component electron spin orbitals. The term

$$H'' = \left(\frac{e\hbar}{2mc} \right)^2 \left(\frac{\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2}{r^3} - \frac{3(\boldsymbol{\sigma}_1 \cdot \mathbf{r})(\boldsymbol{\sigma}_2 \cdot \mathbf{r})}{r^5} \right), \quad (\text{A3})$$

in Eq. (A2) is responsible for the mechanism of triplet resonance energy transfer which will be discussed in this appendix.

Let us introduce a notation where $|\uparrow, \downarrow\rangle$ denotes a spin state where electron 1 has its spin pointing parallel to the z axis, while the spin of electron 2 is antiparallel. If $|\downarrow, \uparrow\rangle$ represents the opposite situation, while $(\downarrow, \uparrow|$ denotes the adjoint of $|\downarrow, \uparrow\rangle$, then

$$(\downarrow, \uparrow| H'' |\uparrow, \downarrow\rangle = \left(\frac{e\hbar}{2mc} \right)^2 \left(\frac{1}{r^3} (\downarrow|\boldsymbol{\sigma}_1|\uparrow) \cdot (\uparrow|\boldsymbol{\sigma}_2|\downarrow) - \frac{3}{r^5} (\downarrow|\boldsymbol{\sigma}_1 \cdot \mathbf{r}|\uparrow)(\uparrow|\boldsymbol{\sigma}_2 \cdot \mathbf{r}|\downarrow) \right). \quad (\text{A4})$$

The matrix element responsible for resonance transfer of triplet excitation energy then becomes:

$$\begin{aligned} & \int d^3x_1 \int d^3x_2 \phi_r^*(\mathbf{x}_1) \phi_s^*(\mathbf{x}_2) H'' \phi_t(\mathbf{x}_2) \phi_u(\mathbf{x}_1) \\ &= \left(\frac{e\hbar}{2mc} \right)^2 \int d^3x_1 \phi_r^*(\mathbf{x}_1) \phi_u(\mathbf{x}_1) \int d^3x_2 \phi_s^*(\mathbf{x}_2) \phi_t(\mathbf{x}_2) \\ & \quad \times \left(\frac{1}{R^3} (\downarrow|\sigma_1|\uparrow) \cdot (\uparrow|\sigma_2|\downarrow) - \frac{3(\downarrow|\sigma_1 \cdot \mathbf{R}|\uparrow)(\uparrow|\sigma_2 \cdot \mathbf{R}|\downarrow)}{R^5} + \dots \right) \quad (\text{A5}) \end{aligned}$$

where we have made a bipolar expansion of $H''(r)$ around the sensitizer and acceptor centers as in Eq. (6). For triplet-singlet transitions, the integrals involving the space parts of the electron orbitals are not equal to zero because the triplet and singlet orbitals are eigenfunctions of different Fock operators, and hence not orthogonal.

Letting $\mathbf{R} = (X, Y, Z)$ we have

$$(\downarrow|\sigma_1|\uparrow) = (0, 1) \begin{pmatrix} Z & X+iY \\ X-iY & -Z \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = X - iY, \quad (\text{A6})$$

and similarly,

$$(\uparrow|\sigma_2|\downarrow) = X + iY, \quad (\text{A7})$$

while

$$(\downarrow|\sigma_1|\uparrow) \cdot (\uparrow|\sigma_2|\downarrow) = 2, \quad (\text{A8})$$

so that

$$\frac{1}{R^3} (\downarrow|\sigma_1|\uparrow) \cdot (\uparrow|\sigma_2|\downarrow) - \frac{3(\downarrow|\sigma_1 \cdot \mathbf{R}|\uparrow)(\uparrow|\sigma_2 \cdot \mathbf{R}|\downarrow)}{R^5} = \frac{1}{R^3} \left(\frac{3Z^2}{R^2} - 1 \right) \quad (\text{A9})$$

and

$$|\langle \Psi_i | H'' | \Psi_f \rangle|^2 = \left(\frac{e\hbar}{2mc} \right)^2 \left| \int d^3x_1 \phi_r^* \phi_u \right|^2 \left| \int d^3x_2 \phi_s^* \phi_t \right|^2 \frac{1}{R^6} \left(\frac{9Z^4}{R^4} + 1 - \frac{3Z^2}{R^2} \right). \quad (\text{A10})$$

Averaging over orientations and setting

$$\left| \int d^3x_1 \phi_r^* \phi_u \right|^2 \left| \int d^3x_2 \phi_s^* \phi_t \right|^2 = 10^{-4}, \quad (\text{A11})$$

we obtain the order of magnitude for the transition rate [Eq. (49)] as

$$\frac{1}{\langle \tau_{i \rightarrow f} \rangle} \approx \frac{9}{5} \left(\frac{e\hbar}{2mc} \right)^2 \frac{10^{-4}}{\hbar^2 \Gamma R^6}. \quad (\text{A12})$$

Equation (A12) can be rewritten in the form:

$$\frac{1}{\langle \tau_{i \rightarrow f} \rangle} \approx \frac{1}{\tau_0} \left(\frac{R_0''}{R} \right)^6, \quad (\text{A13})$$

where, with $\Gamma = 10^{13} \sim /s$ and $\tau_0 = 10$ s,

$$R_0'' = 10 \text{ \AA}. \quad (\text{A14})$$

Resonance transfer of triplet excitation energy over distances of this magnitude has been observed experimentally by Gay [33].

Acknowledgments

The author is extremely grateful to Lektor Lars Erik Lundberg and Professor Jens Peder Dahl for stimulating discussions. He is also very grateful to Professors D. P. Craig, E. A. Power, P. O. Löwdin, O. Goscinski, and J. Jortner for their advice and encouragement.

Bibliography

- [1] G. Cario and J. Franck, *Z. Phys.* **17**, 202 (1923).
- [2] J. Perrin, *Proces de la Conseil de Chimie Solvay* (Gauthier and Villar, Paris, 1925), p. 322.
- [3] J. Perrin, *C.R. Acad. Sci.* **184**, 1097 (1927).
- [4] H. Kallmann and F. London, *Z. Phys. Chem.* **B2**, 207 (1928).
- [5] F. Perrin, *Ann. Phys.* **17**, 283 (1932).
- [6] Th. Förster, *Naturwissenschaften* **33**, 166 (1946).
- [7] Th. Förster, *Z. Naturwiss. A* **4**, 321 (1949).
- [8] S. I. Vavilov, *J. Phys. (USSR)* **7**, 141 (1943).
- [9] D. L. Dexter, *J. Chem. Phys.* **21**, 836 (1953).
- [10] J. Franck and R. Livingston, *Rev. Mod. Phys.* **21**, 505 (1949).
- [11] S. I. Vavilov, *Microstructure of Light* (Verlag, Berlin, 1954).
- [12] E. J. Bowen and B. Brocklehurst, *Trans. Faraday Soc.* **49**, 1131 (1953).
- [13] A. N. Terenin and V. L. Ermolaev, *Usp. Fiz. Nauk.* **58**, 37 (1953).
- [14] Th. Förster, *Fluoreszinz Organischer Verbindungen* (Vandenhoeck and Ruprecht, Göttingen, 1951).
- [15] J. Avery, *Proc. Phys. Soc. (London)* **88**, 1 (1966).
- [16] R. R. McLone and E. A. Power, *Mathematica* **11**, 91 (1964).
- [17] A. I. Akhiezer and V. B. Berestetskii, *Quantum Electrodynamics* (Interscience, New York, 1965).
- [18] J. Avery, *Creation and Annihilation Operators* (McGraw-Hill, New York, 1976).
- [19] J. Avery, R. Mason, and L. G. Augenstein, *Biopolym. Symp.* **1**, 259 (1964).
- [20] J. Avery and R. Mason, *J. Phys. Chem.* **69**, 784 (1965).
- [21] J. Avery, in *Electronic Aspects of Biochemistry*, B. Pullman, Ed. (Academic, New York, 1964).
- [22] J. Avery, *The Quantum Theory of Atoms, Molecules and Photons* (McGraw-Hill, New York, 1972), pp. 251-253.
- [23] J. Avery and J. C. Packer, in *The Triplet State*, A. B. Zahlan, Ed. (Cambridge U.P., Cambridge, 1967).
- [24] M. M. Zwick and H. Kuhn, *Z. Naturforsch. A* **17**, 411 (1962).
- [25] I. Langmuir and V. J. Schaefer, *J. Am. Chem. Soc.* **59**, 1406 (1937).
- [26] S. Weinberg, *The First Three Minutes* (Bantam, New York, 1979).
- [27] J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.* **17**, 157 (1945).
- [28] J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.* **21**, 425 (1949).

- [29] G. Ludwig, *Z. Naturforsch. A* **5**, 637 (1950).
- [30] G. Süßmann, *Z. Phys.* **131**, 629 (1952).
- [31] G. N. Lewis, *Proc. Natl. Acad. Sci. USA* **12**, 22 (1926).
- [32] L. E. Lundberg, *Commun. Math. Phys.* **50**, 103 (1976).
- [33] H. Gay, "Studies of phosphorescence and energy transfer between triplet states in aromatic hydrocarbons, Ph.D. Thesis, Imperial College of Science and Technology, London, 1964.
- [34] C. C. J. Roothaan, in *The Uncertainty Principle and Foundations of Quantum Mechanics, A Fifty Year Survey—A Tribute to Werner Heisenberg*, W. C. Price and S. S. Chissick, Eds. (Wiley, London, 1977).
- [35] F. Hoyle and J. V. Narlikar, *Ann. Phys.* **54**, 207 (1969); **62**, 44 (1971).
- [36] P. C. W. Davies, *Proc. Cambr. Philos. Soc.* **68**, 751 (1970).
- [37] P. C. W. Davies, *J. Phys. A* **4**, 836 (1971); **5**, 1025 (1972).
- [38] H. J. Groenwold, *K. Ned. Akad. Wet. Versl. Gewone Vergad. Afd. Natuur. Kd.* **52**, 226 (1949); **53**, 414, 610 (1950).
- [39] J. A. Wheeler, "Pregeometry: Motivations and prospects," in *Quantum Theory and Gravitation*, A. R. Marlow, Ed. (Academic, New York, 1980).
- [40] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972), p. 619.
- [41] L.-E. Lundberg, *Commun. Math. Phys.* **50**, 103 (1976).
- [42] L.-E. Lundberg, preprint (1982).
- [43] E. A. Power, *Phys. Rev. A* **10**, 756 (1974).
- [44] E. A. Power, *J. Chem. Phys.* **46**, 4297 (1967).
- [45] E. A. Power, W. J. Meath, and J. O. Hirshfelder, *Phys. Rev. Lett.* **17**, 799 (1966).
- [46] L. Gomberoff and E. A. Power, *Proc. Phys. Soc.* **88**, 281 (1966).
- [47] L. Gomberoff, *Proc. R. Soc. London Ser. A* **295**, 476 (1966).
- [48] R. R. Mclone and E. A. Power, *Proc. R. Soc. London Ser. A* **286**, 573 (1965).

Einstein A-coefficients for acridine orange in various solvents

March 3, 2019

Definition of the Einstein A-coefficients

The Wikipedia article on the Einstein A and B coefficients defines the A coefficients as follows: “Spontaneous emission is the process by which an electron ‘spontaneously’ (i.e. without any outside influence) decays from a higher energy level to a lower one. The process is described by the Einstein coefficient $A_{21} \text{ sec}^{-1}$, which gives the probability per unit time that an electron in state 2 with energy E_2 will decay spontaneously to state 1 with energy E_1 , emitting a photon with an energy $E_2 - E_1 = h\nu$. Due to the energy-time uncertainty principle, the transition actually produces photons within a narrow range of frequencies called the spectral linewidth. If n_i is the number density of atoms in state i , then the change in the number density of atoms in state 2 per unit time due to spontaneous emission will be

$$\left(\frac{dn_2}{dt}\right)_{\text{spontaneous}} = -A_{21}n_2 \quad (1)$$

The same process results in increasing of the population of the state 1:

$$\left(\frac{dn_1}{dt}\right)_{\text{spontaneous}} = A_{21}n_2 \quad (2)$$

Tables of Einstein A-coefficients

An article by P.M. Meti et. al. gives the following values for the Einstein A-coefficients of acridine orange hemi zinc salt in various solvents as functions of concentration. These tables refer to the transition from the first electronically excited state of acridine orange to the ground state.

Table 1: Acridine orange in ethanol solvent

concentration (M)	A-coefficient, sec^{-1}
1×10^{-5}	2.5×10^6
2×10^{-6}	6.9×10^6
3×10^{-6}	4.6×10^6
4×10^{-6}	4.14×10^6

Table 2: Acridine orange in propanol solvent

concentration (M)	A-coefficient, sec^{-1}
1×10^{-5}	6.9×10^6
2×10^{-6}	1.61×10^6
3×10^{-6}	1.65×10^6
4×10^{-6}	8.6×10^6

Table 3: Acridine orange in butanol solvent

concentration (M)	A-coefficient, sec^{-1}
1×10^{-5}	4.6×10^5
2×10^{-6}	1.15×10^5
3×10^{-6}	1.03×10^5
4×10^{-6}	7.47×10^5

Table 4: Acridine orange in decanol solvent

concentration (M)	A-coefficient, sec^{-1}
1×10^{-5}	4.6×10^5
2×10^{-6}	3.2×10^6
3×10^{-6}	2.07×10^6
4×10^{-6}	9.6×10^5

Derivation of the transition dipole moment from the Einstein A-coefficient

Here is an excerpt from Notes.pdf:

In the dipole approximation, the average rate of spontaneous photon emission is given by

$$\frac{1}{\tau_0} = \frac{1}{3} \left(\frac{e^2}{\hbar c} \right) k^3 D^2 c = \frac{1}{3} \left(\frac{e^2}{\hbar c} \right) (kD)^2 kc \quad (3)$$

where

$$k = \frac{2\pi}{\lambda} = \frac{E_i}{\hbar c} \quad (4)$$

is the wave number of the emitted light, λ being the wavelength and E_i the energy difference between the Stokes-shifted initial excited state and the ground state. For acridine orange, the maximum emission takes place at a wavelength of

$$\lambda \approx 550 \text{ nanometers} = 5500 \text{ Angstroms} \quad (5)$$

Then

$$\begin{aligned} \omega_1 = kc &\approx \left(\frac{2\pi}{5.5 \times 10^{-7} \text{ m}} \right) 3 \times 10^8 \text{ m/sec} \\ &= 3.43 \times 10^{15} \text{ sec}^{-1} \end{aligned} \quad (6)$$

For an organic molecule such as acridine orange, the transition dipole moment, D , can be approximated by the size of the molecule, let us say 1 nanometer. With this very rough estimate, we can write

$$k^2 D^2 \equiv \left(\frac{2\pi D}{\lambda} \right)^2 \approx \left(\frac{2\pi}{550} \right)^2 \approx 1.3 \times 10^{-4} \quad (7)$$

a pure number. The fine structure constant is also a pure number. Putting all these elements together, we obtain a very rough estimate for the rate of spontaneous emission of an organic dye molecule such as acridine orange:

$$\begin{aligned} \frac{1}{\tau_0} &= \frac{1}{3} \left(\frac{e^2}{\hbar c} \right) k^3 D^2 c \\ &\approx \frac{1}{3} \left(\frac{1}{137} \right) (1.3 \times 10^{-4}) 3.43 \times 10^{15} \text{ sec}^{-1} \\ &\approx 10^9 \text{ sec}^{-1} \end{aligned} \quad (8)$$

For 2 molar acridine orange in ethanol, the rate of spontaneous photon emission derived from the Einstein A-coefficient is $6.9 \times 10^6 \text{ sec}^{-1}$. Thus with this solvent and concentration, equation (1) becomes

$$\left(\frac{dn_2}{dt} \right)_{\text{spontaneous}} = -6.9 \times 10^6 n_2 \text{ sec}^{-1} \quad (9)$$

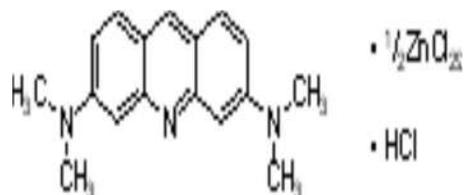


Figure 1: **Acridine orange hemi zinc salt.**

where n_2 is the number of acridine orange molecules per unit volume. In the special case where there is only one molecule present, equations (8) and (9) can be compared. We find from the ratio that

$$(kD)^2 = \left(\frac{6.9 \times 10^6}{10^9} \right) \times (1.3 \times 10^{-4}) \approx 10^{-6} \quad (10)$$

Bibliography

- [1] Abragam, A. and Bleaney B., **Electron Paramagnetic Resonance of Transition Ions**, Clarendon, 1970.
- [Akhiezer and Berestetskii, 1965] Akhiezer, A.I. and Berestetskii, V.B., **Quantum Electrodynamics**, Interscience, New York, 1965.
- [2] Anderson, R.W.; Aquilanti, V.; Cavalli, S. and Grossi, G., *Steroidirected states in molecular dynamics: a discrete basis for the quantum mechanical scattering matrix*, J. Phys. Chem., **95** 8184, 1991.
- [Aquilanti et al., 1996] Aquilanti, V., Cavalli, S., De Fazio, D, and Grossi, G. *Hyperangular momentum: Applications to atomic and molecular science*, in **New Methods in Quantum Theory**, Tsipis, C.A., Popov, V.S., Herschbach, D.R., and Avery, J.S., Eds., Kluwer, Dordrecht, 1996.
- [3] Aquilanti, V.; Cavalli, S. and Grossi, G., *Hyperspherical coordinates for molecular dynamics by the method of trees and the mapping of potential-energy surfaces for triatomic systems*, J. Chem. Phys., **85** 1362, 1986.
- [4] Aquilanti, V.; Grossi, G.; Laganá, A.; Pelikan, E., and Klar, H., *A decoupling scheme for a 3-body problem treated by expansions into hyperspherical harmonics. The hydrogen molecular ion*, Lett. Nuovo Cimento, **41** 541, 1984.
- [5] Aquilanti, V.; Grossi, G., and Laganá, A., *On hyperspherical mapping and harmonic expansions for potential energy surfaces* J. Chem. Phys., **76** 1587-8, 1982.
- [6] Aquilanti, V. and Cavalli, S., *Hyperspherical analysis of kinetic paths for elementary chemical reactions and their angular momentum dependence*, Chem. Phys. Lett., **141** 309, 1987.
- [7] Aquilanti, V.; Cavalli, S.; Grossi, G., and Anderson, R.W., *Representation in Hyperspherical and Related Coordinates of the potential-energy surface for triatomic reactions*, J. Chem. Soc. Faraday Trans., **86** 1681, 1990.
- [8] Aquilanti, V.; Benevente, L.; Grossi, G. and Vecchiocattivi, F., *Coupling schemes for atom-diatom interactions, and an adiabatic decoupling treatment of rotational temperature effects on glory scattering*, J. Chem. Phys., **89** 751-761, 1988.

- [9] Aquilanti, V. and Grossi, G., *Angular momentum coupling schemes in the quantum mechanical treatment of P-state atom collisions* J. Chem. Phys., **73** 1165-1172, 1980.
- [10] Aquilanti, V. and Grossi, G., *Discrete representations by artificial quantization in the quantum mechanics of anisotropic interactions*, Lett. Nuovo Cimento, **42** 157, 1985.
- [11] Aquilanti, V.; Cavalli, S., and Grossi, G., *Discrete analogs of spherical harmonics and their use in quantum mechanics: The hyperquantization algorithm*, Theor. Chem. Acta, **79** 283, 1991.
- [12] Aquilanti, V. and Cavalli, S., *Discrete Analogs of Hyperspherical Harmonics and their use for the Quantum Mechanical Three Body Problem*, Few Body Systems, Suppl., **6** 573, 1992.
- [13] Aquilanti, V. Cavalli, S. and De Fazio, D., *Angular and hyperangular momentum coupling coefficients as Hahn polynomials*, J. Phys. Chem., **99** 15694, 1995.
- [14] Aquilanti, V., Cavalli, S., Coletti, C. and Grossi, G., *Alternative Sturmian bases and momentum space orbitals; an application to the hydrogen molecular ion*, Chem. Phys., **209** 405, 1996.
- [Aquilanti et al., 1997] Aquilanti, V., Cavalli, S. and Coletti, C., *The d-dimensional hydrogen atom; hyperspherical harmonics as momentum space orbitals and alternative Sturmian basis sets*, Chem. Phys., **214** 1, 1997.
- [15] Aquilanti, V., and Avery, J., *Generalized potential harmonics and contracted sturmians*, Chem. Phys. Letters, **267** 1, 1997.
- [16] Aquilanti, V., Cavalli, S., Coletti, C., Di Domenico, D., and Grossi, G., *Hyperspherical harmonics as Sturmian orbitals in momentum space: a systematic approach to the few-body Coulomb problem*, Int. Rev. Phys. Chem., **20** 673, 2001.
- [17] Aquilanti, V. and Caligiana, A., *Sturmian approach to one-electron many-center systems: integrals and iteration schemes*, Chem. Phys. Letters, **366** 157 2002.
- [18] Aquilanti, V., Caligiana A., and Cavalli, S., *Hydrogenic elliptic orbitals, coulomb sturmian sets. Recoupling coefficients among alternative bases*, Int. J. Quantum Chem., **92** 99, 2003.
- [19] Aquilanti, V., Caligiana, A., Cavalli, S., and Coletti, C., *Hydrogenic orbitals in momentum space and hyperspherical harmonics. Elliptic sturmian basis sets*, Int. J. Quantum Chem., **92** 212, 2003.
- [20] Aquilanti, V., and Caligiana, A., *Sturmian orbitals in Quantum Chemistry: an introduction*, in **Fundamental World of Quantum Chemistry: A Tribute to the Memory of P.O. Löwdin**, I, Brändas, E.J. and Kryachko, E.S., Eds., Kluwer, Dordrecht, 297, 2003.

- [21] Aquilanti, V. and Avery, J., *Sturmian expansions for quantum mechanical many-body problems and hyperspherical harmonics*, Adv. Quant. Chem., **39** 72-101, 2001.
- [22] Atkins, P.W., Child, M.S. and Phillips, C.S.G., **Tables for Group Theory**, Oxford, 1970.
- [Avery, 1976] Avery, J., **Creation and Annihilation Operators**, McGraw-Hill, 1976.
- [Avery and Ørmen, 1980] Avery, J. and Ørmen, P.J., *Transferable integrals in a deformation density approach to molecular orbital calculations; IV. Evaluation of angular integrals by a vector-pairing method*, Int. J. Quantum Chem., **18** 953-971, 1980.
- [Avery, 1989] Avery, J., **Hyperspherical Harmonics; Applications in Quantum Theory**, Kluwer Academic Publishers, Dordrecht, 1989.
- [23] Avery, J., *Hyperspherical Sturmian basis functions in reciprocal space*, in **New Methods in Quantum Theory**, Tsipis, C.A., Popov, V.S., Herschbach, D.R., and Avery, J.S., Eds., Kluwer, Dordrecht, 1996.
- [24] Avery, J. and Antonsen, F., *A new approach to the quantum mechanics of atoms and small molecules*, Int. J. Quantum Chem., Symp., **23** 159, 1989.
- [25] Avery, J. and Antonsen, F., *Iteration of the Schrödinger equation, starting with Hartree-Fock wave functions*, Int. J. Quantum Chem., **42** 87, 1992.
- [26] Avery, J. and Antonsen, F., *Theor. Chim. Acta*, **85** 33, 1993.
- [27] Avery, J. and Herschbach, D. R., *Hyperspherical Sturmian basis functions*, Int. J. Quantum Chem., **41** 673, 1992.
- [28] Avery, J. and Wen, Z.-Y., *A Formulation of the quantum mechanical many-body in terms of hyperspherical coordinates*, Int. J. Quantum Chem., **25** 1069, 1984.
- [29] Avery, J., *Correlation in iterated solutions of the momentum-space Schrödinger equation*, Chem. Phys. Lett., **138** (6) 520-4, 1987.
- [30] Avery, J., *Hyperspherical harmonics; Some properties and applications*, in **Conceptual Trends in Quantum Chemistry**, Kryachko, E.S., and Calais, J.L., Eds, Kluwer, Dordrecht, 1994.
- [31] Avery, J., Hansen, T.B., Wang, M. and Antonsen, F., *Sturmian basis sets in momentum space*, Int. J. Quantum Chem., **57** 401, 1996.
- [32] Avery, J., and Hansen, T.B., *A momentum-space picture of the chemical bond* Int. J. Quantum Chem., **60** 201, 1996.
- [33] Avery, J., *Many-particle Sturmians*, J. Math. Chem., **21** 285, 1997.

- [34] Avery, J. and Antonsen, F., *Relativistic sturmian basis functions*, J. Math. Chem., **24** 175, 1998.
- [35] Avery, J., *A formula for angular and hyperangular integration*, J. Math. Chem., **24** 169, 1998.
- [Avery, 1999] Avery, J., *Many-electron Sturmians applied to atoms and ions*, J. Mol. Struct., **458** 1, 1999.
- [Avery, 1999a] Avery, J., *Many-electron Sturmians as an alternative to the SCF-CI Method*, Adv. Quantum Chem., **31** 201, 1999.
- [Avery, 2000] Avery, J., **Hyperspherical Harmonics and Generalized Sturmians**, Kluwer Academic Publishers, Dordrecht, Netherlands, 2000.
- [36] Avery, J. *Selected applications of hyperspherical harmonics in quantum theory*, J. Phys. Chem.. *97* 2406, 1993.
- [37] Avery, J. and Antonsen, F., *Evaluation of angular integrals by harmonic projection*, Theor. Chim. Acta., **85** 33, 1993.
- [38] Avery, J., *Fock transforms in reciprocal-space quantum theory*, J. Math. Chem., **15** 233, 1994.
- [39] Aquilanti V. and Avery, J., *Generalized potential harmonics and contracted Sturmians*, Chem. Phys. Letters, **267** 1-8, 1997.
- [40] Avery, J., *Many-electron Sturmians applied to atoms and ions*, J. Mol. Struct. (Theochem), **458** 1-9, 1999.
- [41] Avery, J., *A formula for angular and hyperangular integration*, J. Math. Chem., **24** 169-174, 1998.
- [42] Avery, J. and Sauer, S., *Many-electron Sturmians applied to molecules*, in **Quantum Systems in Chemistry and Physics, Volume 1**, Hernández-Laguna, A., Maruani, J., McWeeney, R. and Wilson, S, Eds., Kluwer Academic Publishers, 2000.
- [43] Avery, J. and Coletti, C., *Generalized Sturmians applied to atoms in strong external fields*, J. Math. Chem., **27** 43-51 2000.
- [44] Avery, J. and Coletti, C., *Many-electron Sturmians applied to atoms and ions in strong external fields*, in **New Trends in Quantum Systems in Chemistry and Physics**, Marauani, J. et al., Eds., 77-93, Kluwer Academic Publishers, 2001.
- [45] Avery, J. and Shim, R., *Core ionization energies of atoms and molecules calculated using the Generalized Sturmian Method*, Int. J. Quantum Chem., **79** 1-7, 2000.

- [46] Avery, J., *The Generalized Sturmian Method and inelastic scattering of fast electrons*, J. Math. Chem., **27**, 279-292, 2000.
- [47] Avery, J., *Sturmian methods in quantum theory*, in **Proc. Workshop on Concepts in Chemical Physics**, Billing, G.D. and Henriksen, N., Eds., Danish Technical University, 2001.
- [Avery, 2003] Avery, J., *Sturmians*, in **Handbook of Molecular Physics and Quantum Chemistry**, Wilson, S., Ed., Wiley, Chichester, 2003.
- [48] Avery, J., *Harmonic polynomials, hyperspherical harmonics, and Sturmians*, in **Fundamental World of Quantum Chemistry, A Tribute Volume to the Memory of Per-Olov Löwdin**, Brändas, E.J. and Kryachko, E.S., Eds., Kluwer Academic Publishers, Dordrecht, 261-296, 2003.
- [49] Avery, J. and Avery, J., *The Generalized Sturmian Method for calculating spectra of atoms and ions*, J. of Math. Chem., **33** 145-162, 2003.
- [50] Avery, J. and Avery, J., *Kramers pairs in configuration interaction*, Adv. Quant. Chem., **43** 185-206, 2003.
- [51] Avery, J., *Many-center Coulomb Sturmians and Shibuya-Wulfman integrals*, Int. J. Quantum Chem., **100** 121-130, 2003.
- [52] Avery, J., Avery J. and Goscinski, O., *Natural orbitals from generalized Sturmian calculations*, Adv. Quant. Chem., **43** 207-16, 2003.
- [Avery, Avery, Aquilanti and Caligiana] Avery, J., Avery, J., Aquilanti, V., and Caligiana, A., *Atomic densities, polarizabilities and natural orbitals derived from generalized Sturmian calculations*, Adv. Quant. Chem., **47** 156-173, 2004.
- [Avery and Avery, 2004] Avery, J. and Avery, J., *Generalized Sturmian solutions for many-particle Schrödinger equations*, J. Phys. Chem. A, **41** 8848, 2004.
- [Avery and Avery, 2006] Avery, J. and Avery, J., **Generalized Sturmians and Atomic Spectra**, World Scientific, 2006.
- [Avery and Avery, 2006a] Avery, J. and Avery, J., *The Generalized Sturmian Library*, <http://sturmian.kvante.org>, 2006a.
- [53] Avery, J. and Avery, J., *Atomic core-ionization energies; approximately piecewise-linear relationships*, J. Math. Chem., **46** 164-181, 2009.
- [Avery, 2009] Avery, J., *Harmonic polynomials, hyperspherical harmonics and atomic spectra*, J. Comp. Appl. Math., 2009.
- [Avery and Avery, 2009] Avery, J., and Avery, J., *Can Coulomb Sturmians be used as a basis for N-electron molecular calculations?*, J. Phys. Chem A., **113** 14565, 2009.

- [Avery, 2011] Avery, J.E., **New Computational Methods in the Quantum Theory of Nanostructures**, Ph.D. Thesis, University of Copenhagen, 2011.
- [54] Balhausen, C.J., **Introduction to Ligand Field Theory**, McGraw-Hill, 1962.
- [55] Ballot, L. and Farbre de la Ripelle, M., *Application of the hyperspherical formalism to trinucleon bound-state problems*, Ann. Phys., **127** 62, 1980.
- [56] Bandar, M. and Itzyksen, C., *Group theory and the H atom*, Rev. Mod. Phys., **38** 330, 1966.
- [57] Bang, J.M. and Vaagen, J.S., *The Sturmian expansion: a well-depth-method for orbitals in a deformed potential*, Z. Phys. A, **297** 223-36, 1980.
- [58] Bang, J.M., Gareev, F.G., Pinkston, W.T. and Vaagen, J.S., *One- and two-nucleon overlap functions in nuclear physics* Phys. Rep., **125** 253-399, 1985.
- [59] Baryudin, L. E. and Telnov, D. A., *Sturmian expansion of the electron density deformation for 3d-metal ions in electric field*, Vestn. Leningr. Univ., Ser 4: Fiz., Khim., **1** 83-6, 1991.
- [60] Baretty, Reinaldo; Ishikawa, Yasuyuki; and Nieves, Jose F., *Momentum space approach to relativistic atomic structure calculations*, Int. J. Quantum Chem., Quantum Chem. Symp., **20** 109-17, 1986.
- [61] Biedenharn, L.C. and Louck, J.D., **Angular Momentum in Quantum Physics**, Addison Wesley, Reading, Mass, 1981.
- [62] Biedenharn, L.C. and Louck, J.D., **The Racah-Wigner Algebra in Quantum Theory**, Addison Wesley, Reading, Mass, 1981.
- [63] Bishop, D.B., **Group Theory and Chemistry**, Courier Dover Publications, 1993.
- [64] Blinder, S.M., *On Green's functions, propagators, and Sturmians for the nonrelativistic Coulomb problem*, Int. J. Quantum Chem., Quantum Chem. Symp., **18** 293-307, 1984.
- [65] Bransden, B.H.; Noble, C.J.; and Hewitt, R.N., *On the reduction of momentum space scattering equations to Fredholm form*, J. Phys. B: At., Mol., Opt. Phys., **26** 2487-99, 1993.
- [66] Bransden, B.H., and Joachim, C.J., **Physics of Atoms and Molecules; 2nd Ed.**, Prentice-Hall, 2003.
- [67] Brink, D.M. and Satchler, D.R., **Angular Momentum**, Clarendon, Oxford, 1968.
- [Caligiana, 2003] Caligiana. A., **Sturmian Orbitals in Quantum Chemistry**, Ph.D. thesis, Department of Chemistry, University of Perugia, Italy, October, 2003.

- [68] Coleman, A.J., *The Symmetric Group Made Easy*, Adv. Quantum Chem., **4** 83-108, 1968.
- [Coletti, 1998] Coletti, C. **Struttura Atomica e Molecolare Come Tottura della Simmetria Ipersferica**, Ph.D. thesis, Department of Chemistry, University of Perugia, Italy, 1998.
- [69] Condon, E.U. and Shortley, G.H., **Theory of Atomic Spectra**, Cambridge, 1953.
- [70] Cotton, F.A., **Chemical Applications of Group Theory**, Interscience, 1963.
- [71] Coulson, C.A., *Momentum distribution in molecular systems. I. Single bond. III. Bonds of higher order*, Proc. Camb. Phys. Soc., **37** 74, 1941.
- [72] Coulson, C.A. and Duncanson, W.E., *Momentum distribution in molecular systems. II. C and C-H bond*, Proc. Camb. Phys. Soc., **37** 67, 1941.
- [73] Coxeter, H.S.M. et al. Editors, **Mathematical Expositions, Volume 21: The Collected Papers of Alfred Young, 1873-1940**, University of Toronto Press, Toronto, 1977.
- [74] Dahl, J.P., *The Wigner function*, Physica A, **114** 439, 1982.
- [75] Dahl, J.P., *On the group of translations and inversions of phase space and the Wigner function*, Phys. Scripta, **25** 499-503, 1982.
- [76] Dahl, J.P., *Dynamical equations for the Wigner functions*, in **Energy Storage and Redistribution in Molecules**, Hinze, J., Ed., Plenum, New York, 1983.
- [77] Dahl, J.P., *The phase-space representation of quantum mechanics and the Bohr-Heisenberg correspondence principle*, in **Semiclassical Description of Atomic and Nuclear Collisions**, Bang, J. and De Boer, J., Eds., North Holland, Amsterdam, 1985.
- [78] Dahl, J.P., *The dual nature of phase-space representations*, in **Classical and Quantum Systems**, Doebner, H.D. and Schroeck, F., Jr., Eds., World Scientific, Singapore, 1993.
- [79] Dahl, J.P., *A phase space essay*, in **Conceptual Trends in Quantum Chemistry**, Kryachko, E.S. and Calais, J.L., Eds., Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.
- [80] Dahl, J.P. and Springborg, M., *The Morse oscillator in position space, momentum space, and phase space* J. Chem. Phys., **88** 4535-47, 1988.
- [81] Dahl, J.P. and Springborg, M., *Wigner's phase-space function and atomic structure. I. The hydrogen atom*, J. Mol. Phys., **47** 1001, 1982.

- [82] Das, G.P.; Ghosh, S.K.; and Sahni, V.C., *On the correlation energy density functional in momentum space*, Solid State Commun., **65** (7) 719-21, 1988.
- [83] Dinesha, K.V., Rettrup, S., and Sarma, C.R., *An Indexing Scheme for Spin-Free Configurations of Electrons*, Int. J. Quantum Chem., **34** 445-455, 1988.
- [84] Dinesha, K.V., Sarma, C.R., and Rettrup, S., *Group Theoretical Techniques and the Many-Electron Problem*, Adv. Quantum Chem., **14** 125-168, P. 1981.
- [85] Englefield, M.J., **Group theory and the Coulomb problem**, Wiley-Interscience, New York, 1972.
- [Fock, 1935] Fock, V.A., *Zur Theorie des Wasserstoffatoms*, Z. Phys., **98** 145, 1935.
- [Fock, 1958] Fock, V.A., *Hydrogen atoms and non-Euclidian geometry*, Kgl. Norske Videnskab. Forh., **31** (1958) 138.
- [86] Fonseca, A.C. and Pena, M.T., *Rotational-invariant Sturmian-Faddeev ansatz for the solution of hydrogen molecular ion (H_2^+): a general approach to molecular three-body problems*, Phys. Rev. A: Gen. Phys., **38** 4967-84, 1988.
- [87] Friis-Jensen, B., Cooper, D.L., and Rettrup, S., *Normalization of Projected Spin Eigenfunctions*, J. Math. Chem., **22** 249-254, 1997.
- [88] Friis-Jensen, B., Cooper, D.L., and Rettrup, S., *Spin-Coupled Calculations based on Projected Spin Eigenfunctions*, Theor. Chem. Acc., **99** 64-67, 1998.
- [89] Friis-Jensen, B., and Rettrup, S., *A Spin-Free Approach for Evaluation of Electronic Matrix Elements using Character Operators of S_N* , Int. J. Quantum Chem., **60**, 983-991, 1996.
- [90] Friis-Jensen, B., Rettrup, S., and Sarma, C.R. *An Indexing Scheme for Classes of S_N : Partitions of N* , Int. J. Quantum Chem., **64** 421-426, 1997.
- [Gasaneo et al., 2009] Gasaneo, G. et al., *Theory of hyperspherical Sturmians for three-body reactions*, J. Phys. Chem. A, **113**, 14573-14582, 2009.
- [91] Gasaneo, G. et al., *Multivariable hypergeometric solutions for three charged particles*, J. Phys. B, **30** L265-L271, 1997.
- [92] Gasaneo, G., and Mmacek, J.H., *Hyperspherical adiabatic eigenfunctions for zero-range potentials*, J. Phys. B., **35** 2239-2250, 2002.
- [93] Gasaneo, G., and Colavecchia, F.D., *Two-body Coulomb wave functions as a kernel for three alternative integral transformations*, J. Phys. A, **36** 8443-8462, 2003.
- [94] Gerratt, J., *General theory of spin-coupled wave functions for atoms and molecules*, Adv. At. Mol. Phys., **7** 141-221, 1971.

- [Goscinski, 1968, 2003] Goscinski, O., *Upper and lower bounds and the conjugate eigenvalue problem*, Adv. Quantum Chem., **41** 51-85, 2003.
- [Gradshteyn and Ryshik, 1965] Gradshteyn, I.S. and Ryshik, I.M., **Tables of Integrals, Series and Products**, Academic Press, New York, 1965.
- [95] Griffeth, J.S., **The Theory of Transition Metal Ions**, Cambridge, 1964.
- [96] Gruber, B., Editor, **Symmetries in Science, VI: From the Rotation Group to Quantum Algebras**, Plenum, New York, 1993.
- [97] Guldberg, A., Rettrup, S., Bendazzoli, G.L., and Palmieri, P., *A New Symmetric Group Programme for Direct Configuration Interaction Studies in Molecules*, Int. J. Quantum Chem. (Symp.), **21** 513-521, 1987.
- [98] Hall, G.G., **Applied Group Theory**, Longmans, 1967.
- [99] Hammermesh, M., **Group Theory and Its Applications to Physical Problems**, Addison-Wesley, 1964.
- [100] Hansen, T.B., **The many-center one-electron problem in momentum space**, Thesis, Chemical Institute, University of Copenhagen, 1998.
- [Harris and Michels, 1967] Harris, F.E. and Michels, H.H., *Evaluation of molecular integrals for Slater-type orbitals*, Adv. Chem. Phys., **13** 205, 1967.
- [101] Hartt, K. and Yidana, P.V.A., *Analytic Sturmian functions and convergence of separable expansions*, Phys. Rev. C: Nucl. Phys., **36** 475-84, 1987.
- [102] Heine, V., **Group Theory in Quantum Mechanics**, Pergamon, 1960.
- [103] Herschbach, D. R., *Dimensional interpolation for two-electron atoms*, J. Chem. Phys., **84** 838, 1986.
- [104] Herschbach, D. R., Avery, J. and Goscinski, O., Eds., **Dimensional Scaling in Chemical Physics**, Kluwer, Dordrecht, 1993.
- [105] Hochstrasser, R.M., **Molecular Aspects of Symmetry**, Benjamin, 1966.
- [106] Hollas, J.M., **Symmetry in Molecules**, Chapman, 1972.
- [107] Jaffé, H.H. and Ochrin M., **Symmetry in Chemistry**, Wily, 1965.
- [Judd, 1975] Judd, B.R., **Angular Momentum Theory for Diatomic Molecules**, Academic Press, New York, 1975.
- [108] Judd, B.R., **Operator Techniques in Atomic Spectroscopy**, McGraw-Hill, 1963.

- [109] Judd, B.R., **Second Quantization and Atomic Spectroscopy**, Johns Hopkins Press, 1967.
- [110] Kadolkar, C., Sarma, C.R., and Rettrup, S., *Configuration Interaction Studies Using Biorthogonal Approach to VB Basis*, Int. J. Quantum Chem., **53** 183-187, 1995.
- [111] Kaplan, I.G., **Symmetry of Many-Electron Systems**, Academic Press, 1975.
- [112] Karule, E. and Pratt, R.H., *Transformed Coulomb Green function Sturmian expansion*, J. Phys. B: At., Mol. Opt. Phys., **24** 1585-91, 1991.
- [113] Knox, R.S. and Gold, A., **Symmetry in the Solid State**, Benjamin, 1964.
- [Koga et al., 1984] Koga, T. and Murai, Takeshi, *Energy-density relations in momentum space. III. Variational aspect*, Theor. Chim. Acta, **65** 311-16, 1984.
- [114] Koga, T., *Direct solution of the $H(1s) - H^+$ long-range interaction problem in momentum space*, J. Chem. Phys., **82** 2022, 1985.
- [115] Koga, T. and Matsumoto, S., *An exact solution of the interaction problem between two ground-state hydrogen atoms*, J. Chem. Phys., **82** 5127, 1985.
- [116] Koga, T. and Kawaai, R., *One-electron diatomics in momentum space. II. Second and third iterated LCAO solutions* J. Chem. Phys., **84** 5651-4, 1986.
- [117] Koga, T. and Matsushashi, T., *One-electron diatomics in momentum space. III. Nonvariational method for single-center expansion*, J. Chem. Phys., **87** 1677-80, 1987.
- [Koga and Matsushashi, 1987] Koga, T. and Matsushashi, T., *Sum rules for nuclear attraction integrals over hydrogenic orbitals*, J. Chem. Phys., **87** 4696-9, 1987.
- [118] Koga, T. and Matsushashi, T., *One-electron diatomics in momentum space. V. Non-variational LCAO approach*, J. Chem. Phys., **89** 983, 1988.
- [119] Koga, T.; Yamamoto, Y., and Matsushashi, T., *One-electron diatomics in momentum space. IV. Floating single-center expansion*, J. Chem. Phys., **88** 6675-6, 1988.
- [120] Koga, T. and Ougihara, T., *One-electron diatomics in momentum space. VI. Non-variational approach to excited states*, J. Chem. Phys., **91** 1092-5, 1989.
- [Koga et al., 1991] Koga, T.; Horiguchi, T. and Ishikawa, Y., *One-electron diatomics in momentum space. VII. Nonvariational approach to ground and excited states of heteronuclear systems*, J. Chem. Phys., **95** 1086-9, 1991.
- [121] Kupperman, A. and Hypes, P.G., *3-dimensional quantum mechanical reactive scattering using symmetrized hyperspherical coordinates*, J. Chem. Phys., **84** 5962, 1986.

- [122] la Cour Jansen, T., Rettrup, S., Sarma, C.R., and Snijders, J.G. *On the Evaluation of Spin-Orbit Coupling Matrix Elements in a Spin Adapted Basis*, Int. J. Quantum Chem., **73** 23-27, 1999.
- [123] Leech, J.W. and Newman, D.J., **How to Use Groups**, Methuen, 1969.
- [124] Lin, C.D., *Analytical channel functions for 2-electron atoms in hyperspherical coordinates*, Phys. Rev. A, **23** 1585, 1981.
- [125] Linderberg, J. and Öhrn, Y., *Kinetic energy functional in hyperspherical coordinates*, Int. J. Quantum Chem., **27** 273, 1985.
- [126] Louck, J.D., *Generalized orbital angular momentum and the n-fold degenerate quantum mechanical oscillator*, J. Mol. Spectr., **4** 298, 1960.
- [127] Loebel, E.M., Editor, **Group Theory and Its Applications**, Academic Press, 1968.
- [128] Lomont, J.S., **Applications of Finite Group Theory**, Academic Press, 1959.
- [129] Manakov, N.L.; Rapoport, L.P. and Zapryagaev, S.A., *Sturmian expansions of the relativistic Coulomb Green function*, Phys. Lett. A, **43** 139-40, 1973.
- [130] Maquet, Alfred; M., Philippe and Veniard, Valerie, *On the Coulomb Sturmian basis*, NATO ASI Ser, Ser C, **271** (Numer. Determ. Electron. Struct. At., Diat. Polyat. Mol.), 295-9, 1989.
- [131] McWeeny, R., and Coulson, C.A., *The computation of wave functions in momentum space. I. The helium atom*, Proc. Phys. Soc. (London) A, **62** 509, 1949.
- [132] McWeeny, R., *The computation of wave functions in momentum space. II. The hydrogen molecule ion*, Proc. Phys. Soc. (London) A, **62** 509, 1949.
- [133] McWeeny, R., **Methods of Molecular Quantum Mechanics**, (second edition), Academic Press, 1992.
- [134] McWeeny, R., **Symmetry**, Pergamon, 1963.
- [135] Meijer, P.H.E. and Bauer, E., **Group Theory**, North Holland, 1962.
- [Michels, 1981] Michels, M.A.J., *Evaluation of angular integrals in rotational invariants*, Int. J. Quantum Chem., **20** 951-952, 1981.
- [136] Monkhorst, H.J. and Harris, F.E., *Accurate calculation of Fourier transform of two-center Slater orbital products*, Int. J. Quantum Chem., **6** 601, 1972.
- [Monkhorst and Jeziorski, 1979] Monkhorst, H.J. and Jeziorski, B., *No linear dependence or many-center integral problems in momentum space quantum chemistry*, J. Chem. Phys., **71** 5268-9, 1979.

- [137] Moore, C.E., **Atomic Energy Levels; Circular of the National Bureau of Standards 467**, Superintendent of Documents, U.S. Government Printing Office, Washington 25 D.C., 1949.
- [138] Morse P.M. and Feshbach, H., **Methods of Theoretical Physics**, McGraw Hill, 1953.
- [139] Motenberg, M., Bivens, R., Metropolis, N., and Wooten, J.K. Jr. **The 3-j and 6-j Symbols**, Technology Press, MIT, Cambridge Mass., 1959.
- [140] Murnaghan, F.D. **The Theory of Group Representations**, Johns Hopkins Press, Baltimore, 1938.
- [141] National Institute for Fusion Science (NIFS), *Atomic and Molecular Databases: Data for Autoionizing States*, <http://dprose.nifs.ac.jp/DB/Auto/>
- [National Institute for Standards and Technology (NIST)] National Institute of Standards and Technology (NIST), *NIST Atomic Spectra Database*, <http://physics.nist.gov/asd> .
- [National Institute of Standards and Technology, 2002] National Institute of Standards and Technology (NIST), *CODATA Recommended Values of the Fundamental Physical Constants: 2002*, <http://physics.nist.gov/cuu/Constants/>
- [142] Navasa, J. and Tsoucaris, G., *Molecular wave functions in momentum space*, Phys. Rev. A, **24** 683, 1981.
- [143] Ojha, P.C., *The Jacobi-matrix method in parabolic coordinates: expansion of Coulomb functions in parabolic Sturmians*, J. Math. Phys. (N Y), **28** 392-6, 1987.
- [144] Orchin, M. and Jaffé, H.H., **Symmetry, Orbitals and Spectra**, Wiley, 1971.
- [145] Orgel, L.E., **An Introduction to Transition Metal Chemistry Ligand Field Theory**, Methuen, 1960.
- [146] Paldus, J., *Coupled Cluster Methods*, in Wilson, S., Editor, **Handbook of Molecular Physics and Quantum Chemistry**, Wiley, 2003.
- [147] Paldus, J., *Dynamical Groups*, in Drake, G.W.F. Editor, **Atomic, Molecular and Optical Physics Handbook**, American Institute of Physics, New York, 1996.
- [148] Paldus, J., and Gould, M.D., *Unitary Group Approach to Reduced Density Matrices, II*, Theor. Chim. Acta, **86** 83-96, 1993.
- [149] Paldus, J., Rettrup, S., and Sarma, C.R., *Clifford algebra realization of Rumer-Weyl basis*, J. Mol. Structure (Theochem), **199** 85-101, 1989.

- [150] Palmieri, P., and Rettrup, S., *Teoria Del Gruppo Simmetrico Per Il Calcolo Degli Elementi Di Matrice E Per La Soluzione Delle Equazioni CI*, in **Scuola di Chimica Teorica : Basi Teoriche ed Applicazioni Numeriche**, Moccia, R. and Pisani, C., Eds., Polo Editoriale Chimico, Milano, 197-276, 1990.
- [151] Palmieri, P., Tarroni, R., Mitrushenkov, A.O., and Rettrup, S., *Efficient Truncation Strategies for Multi Reference Configuration Interaction Molecular Energies and Properties*, J. Chem. Phys., **109** 7085-7091, 1998.
- [152] Pauncz, R., **The Symmetric Group in Quantum Chemistry**, CRC Press, Boca Raton, 1995.
- [153] Petrochen, M.I. and Trifanov, E.D., **Group Theory in Quantum Mechanics**, Iliffe, 1969.
- [Pinchon and Hoggan 2007] Pinchon, D. and Hoggan, P.E., *Rotation matrices for real spherical harmonics: general rotations of atomic orbitals in space-fixed axes*, J. Phys. A., **40** 1597-16100, 2007.
- [Pinchon and Hoggan, 2009] Pinchon, D. and Hoggan, P.E., *Gaussian approximation of exponential type orbitals based on B functions*, Int. J. Quantum Chem., **109** 135-148, 2009.
- [154] Rettrup, S., *A recursive formula for Young's orthogonal representation*, Chem. Phys. Lett., **47** 59-60, 1977.
- [155] Rettrup, S., *Many-body correlations using unitary groups*, in **The Unitary Group for the Evaluation of Electronic Energy Matrix Elements**, Lecture Notes in Chemistry Vol. **22**, 108-118, Hinze, J., Ed., Springer-Verlag, Berlin, 1981.
- [156] Rettrup, S., *An Iterative Method for Calculating Several of the Extreme Eigensolutions of Large Real Non-Symmetric Matrices*, J. Comp. Phys., **45** 100-107, 1982.
- [157] Rettrup, S., *Direct evaluation of spin representation matrices and ordering of permutation group elements*, Int. J. Quantum Chem., **29** 119-128, 1986.
- [158] Rettrup, S., *The permutation group in many-electron theory*, in **Conceptual Trends in Quantum Chemistry III**, Kryachko, E.S., and Calais, J.L., Eds, Kluwer, Dordrecht, 225-238, 1997.
- [Rettrup, 2003] Rettrup, S., **Lecture Notes on Quantum Chemistry**, Department of Chemistry, University of Copenhagen, 2003.
- [159] Rettrup, S., *Alternative graphical representation of determinantal many-electron states*, Int. J. Quantum Chem., **106** 2511-2517, 2006.

- [160] Rettrup, S., Bendazzoli, G.L., Evangelisti, S., and Palmieri, P., *A symmetric group approach to the calculation of electronic correlation effects in molecules*, in **Understanding Molecular Properties**, 533-546, Avery, J.S., Dahl, J.P., and Hansen, Aa.E., Eds., Reidel Publ. Co., Dordrecht, Holland, 1987.
- [161] Rettrup, S., and Chisholm, C.D.H., *On $R(4)$ Symmetries in atomic structure*, Theoret. Chim. Acta (Berl.), **57** 209-218, 1980.
- [162] Rettrup, S., and Pauncz, R., *Representations of the symmetric group generated by projected spin functions: a graphical approach*, Int. J. Quantum Chem., **60** 91-98, 1996.
- [163] Rettrup, S., and Sarma, C.R., *A new program for CI calculations in molecules*, Theoret. Chim. Acta (Berl.), **46** 73-76, 1977.
- [164] Rettrup, S., and Sarma, C.R., *A Programmable procedure for generating many-particle states*, Phys. Letters, **75A** 181-182, 1980.
- [165] Rettrup, S., Sarma, C.R., and Dahl, J.P., *Molecular point group adaptation of spin-free configurations*, Int. J. Quantum Chem., **22** 127-148, 1982.
- [166] Rettrup, S., Thorsteinsson, T., and Sarma, C.R., *A graphical approach to configuration interaction studies in molecules using determinants of non-orthogonal orbitals*, Int. J. Quantum Chem., **40** 709-717, 1991.
- [167] Robinson, G. de B., **Representation Theory of Symmetric Groups**, University of Toronto Press, Toronto, 1961.
- [168] Rohwedder, B. and Englert, B.G., *Semiclassical quantization in momentum space*, Phys. Rev. A: At., Mol., Opt. Phys., **49** 2340-6, 1994.
- [169] Rose, M.E., **Elementary Theory of Angular Momentum**, Wiley, 1957.
- [Rotenberg, 1962] Rotenberg, M., Ann. Phys. (New York), **19** 262, 1962.
- [Rotenberg, 1970] Rotenberg, M., *Theory and application of Sturmian functions*, Adv. At. Mol. Phys., **6** 233-68, 1970.
- [170] Royer, A., *Wigner function as the expectation value of a parity operator*, Phys. Rev. A, **15** 449-450, 1977.
- [171] Rudin, W., **Fourier Analysis on Groups**, Interscience, New York, 1962.
- [172] Rutherford, D.E., **Substitutional Analysis**, Edinburgh University Press, Edinburgh, 1948.

- [173] Sarma, C.R., Ahsan, M.A.H., and Rettrup, S., *A graphical approach to permutation group representations for many-electron systems*, Int. J. Quantum Chem., **58** 637-643, 1996.
- [174] Sarma, C.R., Palmieri, P., and Rettrup, S., *A graphical scheme for addressing determinants in large scale configuration interaction treatments*, Mol. Phys., **98** 1851-1856, 2000.
- [175] Sarma, C.R., Ravi Shankar, N., and Rettrup, S., *A graphical scheme for representing many-electron spin-free configurations*, Int. J. Quantum Chem., **86** 417-421, 2002.
- [176] Sarma, C.R., and Rettrup, S., *A programmable spin-free method for configuration interaction*, Theoret. Chim. Acta (Berl.), **46** 63-72, 1977.
- [177] Sarma, C.R., and Rettrup, S., *A programme for the study of many-body correlations: matrix elements of the generators of $U(n)$* , J. Phys. A: Math. Gen., **13** 2267-2274, 1980.
- [178] Schönland D.S., **Molecular Symmetry**, Van Norstrand, 1965.
- [Shibuya and Wulfman, 1965] Shibuya, T. and Wulfman, C.E., *Molecular orbitals in momentum space*, Proc. Roy. Soc. A, **286** (1965) 376.
- [Shull and Löwdin, 1959] Shull, H. and Löwdin, P.-O., *Superposition of configurations and natural spin-orbitals. Applications to the He problem*, J. Chem. Phys., **30** 617, 1959
- [179] Sinanoglu, O., and Brueckner, K., **Three Approaches to Electron Correlation in Atoms**, Yale University Press, 1970.
- [180] Slater, J.C., **Quantum Theory of Atomic Structure**, McGraw-Hill, 1960.
- [181] Smith, F.T., *Generalized angular momentum in many-body collisions*, Phys. Rev., **120** 1058, 1960.
- [182] Smith, F.T., *A symmetric representation for three-body problems. I. Motion in a plane*, J. Math. Phys., **3** 735, 1962.
- [183] Smith, F.T., *Participation of vibration in exchange reactions*, J. Chem. Phys., **31** 1352-9, 1959.
- [184] Smith, V.H. Jr., *Density functional theory and local potential approximations from momentum space considerations*, in **Local Density Approximations in Quantum Chemistry and Solid State Physics**, Plenum, New York, N. Y, Dahl, J.P. and Avery, J., Eds., 1-19, 82, 1984.
- [185] Springborg, M. and Dahl, J.P., *Wigner's phase-space function and atomic structure*, Phys. Rev. A, **36** 1050-62, 1987.

- [Sommer-Larsen, 1990] Sommer-Larsen, P., Ph.D. Thesis, University of Copenhagen, 1990.
- [186] Sugano, S., Tanabe, Y. and Kamimura, H., **Multiplets of Transition Metal Ions in Crystals**, Academic Press, 1960.
- [187] Suzuki, H., **Electronic Absorption Spectra and Geometry of Organic Molecules**, Academic Press, 1967.
- [188] Thorsteinsson, T., and Rettrup, S., *Configuration interaction studies of electronic structure*, in **Supercomputing Annual Report 1997**, J. Wasniewski, Ed., UNI•C, The Danish Computing Centre for Research and Education, Lyngby, 1998.
- [189] Tinkham, M., **Group Theory and Quantum Mechanics**, McGraw-Hill, 1964.
- [190] Urch, D.S., **Orbitals and Symmetry**, Penguin, 1970.
- [Vilenkin, 1968] Vilenkin, N.K., **Special Functions and the Theory of Group Representations**, American Mathematical Society, Providence, R.I., 1968.
- [191] Watanabe, H., **Operator Methods in Ligand Field Theory**, Prentice Hall, 1966.
- [Weatherford, 1982] Weatherford, C.A., and Jones, H.W., Editors, **International Conference on ETO Multicenter Integrals**, Reidel, Dordrecht, 1982.
- [Weyl, 1950] Weyl, H., **The Theory of Groups in Quantum Mechanics**, Dover, 1950.
- [Weyl, 1939] Weyl, H., **The Classical Groups, Their Invariants and Representations**, Princeton University Press, 1939.
- [192] Wen, Z.-Y. and Avery, J., *Some properties of hyperspherical harmonics*, J. Math. Phys., **26** 396, 1985.
- [Weniger, 1985] Weniger, E.J., *Weakly convergent expansions of a plane wave and their use in Fourier integrals*, J. Math. Phys., **26** 276, 1985.
- [193] Weniger, E.J. and Steinborn, E.O., *The Fourier transforms of some exponential-type basis functions and their relevance for multicenter problems*, J. Chem. Phys., **78** 6121, 1983.
- [194] Weniger, E.J.; Grotendorst, J., and Steinborn, E.O., *Unified analytical treatment of overlap, two-center nuclear attraction, and Coulomb integrals of B functions via the Fourier transform method*, Phys. Rev. A, **33** 3688, 1986.
- [195] Wigner, E.P., **Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra**, Academic Press, 1959.

- [196] Wilson, S., Editor, **Handbook of Molecular Physics and Quantum Chemistry**, Wiley, 2003.
- [197] Wulfman, C.E., *Semiquantitative united-atom treatment and the shape of triatomic molecules*, J. Chem. Phys., **31** 381, 1959.
- [198] Wulfman, C.E., *Dynamical groups in atomic and molecular physics*, in **Group Theory and its Applications**, Loebel, E.M. Ed., Academic Press, 1971.
- [199] Wulfman, C.E., *Approximate dynamical symmetry of two-electron atoms*, Chem. Phys. Lett., **23** 370-372, 1973.
- [200] Wulfman, C.E., *On the space of eigenvectors in molecular quantum mechanics*, Int. J. Quantum Chem., **49** 185, 1994.
- [Wulfman, 2011] Wulfman, C.E., **Dynamical Symmetry**, World Scientific, 2011.
- [201] Yurtsever, E.; Yilmaz, O., and Shillady, D.D., *Sturmian basis matrix solution of vibrational potentials*. Chem. Phys. Lett., **85** 111-116, 1982.
- [202] Yurtsever, E., *Franck-Condon integrals over a Sturmian basis. An application to photoelectron spectra of molecular hydrogen and molecular nitrogen*, Chem. Phys. Lett., **91** 21-26, 1982.
- [Cario and Franck, 1923] G. Cario and J. Frank, Z. Phys. **17**, 202, (1923)
- [Avery, Bay and Szent-Györgyi, 1961] J.S. Avery, Z. Bay and A. Szent-Györgyi, *On the Energy Transfer in Biological Systems*, Proc. Nat. Acad. Sci. (US), **47**, 1742-1744, (1961)
- [Avery, 1966] J.S. Avery, *Resonance energy transfer and spontaneous photon emission*, Proc. Phys. Soc. (London) **88**, Part 1, (1966)
- [Avery, 1984] J.S. Avery, *Use of the S-Matrix in the Relativistic Treatment of Resonance Energy Reansfer*. Int. J. Quantum Chemistry, **XXV**, 79-96, (1984)
- [Gomberoff and Power, 1966] L. Gomberoff and E.A. Power, *The resonance transfer of excitation*, Proc. Pys. Soc. **88**, 281. (1966)
- [Scholes, 2003] G.D. Scholes, *Long-Range Resonance Energy Transfer in Molecular Systems*, Annu. Rev. Phys. Chem. **54**, 57-87, (2003)
- [Chou and Dennis, 2015] K.F. Chou and A.M. Dennis, *Förster Resonance Energy Transfer Between Quantum Dot Donors and Quantum Dot Acceptors*, Sensors, **2015**, 15, 13288-13325, (2015)

- [Wegner, 2015] K-D. Wegner, *Förster Resonance Energy Transfer From Terbium Complexes to Quantum Dots for Multiplexed Immunoassays and Molecular Rulers*, Doctoral Thesis, Université Paris-Sud, (2015)
- [Kusku and Akan, 2011] M. Kusku and O.B. Akan, *A Nanoscale Communication Channel with Fluorescence Resonance Energy Transfer (FRET)*, 1st IEEE International Workshop on Molecular and Nano Scale Communication (MoNaCom), Published in: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), (2011)
- [Pearlstein, 2002] R.M. Pearlstein, *Photosynthetic Exciton Transfer in the 1960's*, Photosynthesis Research **73**, 119-126, (2002)
- [Andrews and Demidov, 1999] D.L. Andrews and A.A. Demidov, *Resonance Energy Transfer*, (Wiley, 1999)
- [Juzeliunas and Andrews, 2000] G. Juzeliunas and D.L. Andrews, *Quantum Electrodynamics of Resonance Energy Transfer*, Adv. Chem. Physics, **112**, (2000)
- [Anderson, 2008] D.L. Anderson, Editor, *Encyclopedia of Applied Spectroscopy*, Wiley, (2008)
- [Daniels et al., 2003] G.J. Daniels, R.G. Jenkins, D.S. Bradshaw, and D.L. Andrews, *Resonance Energy Transfer: The Unified Theory Revisited*, J. Chem. Phys. **119**, 2264-2274, (2003)
- [Bradshaw and Andrews, 2008] D.S. Bradshaw and D.L. Andrews, *Optically controlled resonance energy transfer: Mechanisms and configuration for all-optical switching*, J. Chem. Phys. **128**, 144506, (2008)
- [Wheeler and Feynman, 1949] J.A. Wheeler and R.P. Feynman, *Classical Electrodynamics in Terms of Direct Interparticle Interaction*, Rev. Mod. Phys. **21**, 3, 425-433, (1949)
- [Kastner, 2015a] R.E. Kastner, *Haag's Theorem as a Reason to Reconsider Direct-Action Theories*, International Journal of Quantum Foundations, **1**, (2), 56-64, (2015) (<https://arxiv.org/pdf/1502.03814.pdf>)
- [Kastner, 2015b] R.E. Kastner, *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility* (Cambridge University Press, 2012)
- [Grundler, pdf] G. Gründler, *Remarks on Wheeler-Feynman Absorber Theory* (<https://arxiv.org/pdf/1501.03516.pdf>)
- [Lazarovici, preprint, pdf] D. Lazarovici, *Against Fields*, (preprint, pdf available online),

- [Klaczynski, 2015] L. Klaczynski, *Haag's Theorem in Renormalisable Quantum Field Theories*, Doctoral Thesis at Humboldt-Universität zu Berlin, (2015), (available on line at <http://www2.mathematik.hu-berlin.de/kreimer/wp-content/uploads/PhDLutz.pdf>)
- [Thompson et al., 2018] D. Thompson et al. *Hundred-fold enhancement in far-field radiative heat transfer over the blackbody limit*, Nature, 3 September, 2018.
- [203] R. Haag, *On quantum field theories*, Matematisk-fysiske Meddelelser, **29**, 12, 1955.
- [204] A.I. Akhiezer and V.B. Berestetski, *Quantum Electrodynamics*, Interscience, New York, 1965.
- [205] R.S. Swathi and K.L. Sebastian, *Distance dependance of fluorescence resonance energy transfer*, J. Chem. Soc. **121**, 777-787, (2009)
- [206] R.S. Swathi and K.L. Sebastian, *Long range resonance energy transfer from a dye molecule to graphene has (distance)⁻⁴ dependence*, J. Chem Phys. **130**, 086101, (2009).
- [207] L.I. Rangel-Zamudio, D.S. Rushforth, and R.P. Van Duyne, *Long-Range Resonance Energy Transfer from Aromatic Hydrocarbons to the Anion Radical of Tetracyanoquinodimethane*, J. Phys. Chem. **90**, 807, (1986)
- [208] R. Deshmukh et al., *Long-Range Resonant Energy Transfer Using Optical Topological Transitions in Metamaterials*, ACS Photonics, **5**, 2737-2741, (2018)
- [209] K. Akulov et al., *Long-Distance Resonant Energy Transfer Mediated by Hybrid Plasmonic-Photonic Modes*, J. Phys. Chem. C., just-accepted manuscript, 18 June, (2018).
- [210] D. Martin-Cano et al., *Resonance Energy Transfer and Superradiance Mediated by Plasmonic Nanowaveguides*, Nano Lett. **10** (8), 3129-3134, (2010).
- [211] City College of New York, Team study breaks Forster resonant energy transfer (FRET) distance limit, Physics, Optics and Photonics, Oct. 25, (2018).

Index

- Ab-initio calculations, 101
- Abelian groups, 252
- Absolute velocity, 155
- Accelerated development, 221
- Accelerators, 20
- Accidental degeneracies, 254
- Active site, 125
- Active transport, 223
- ADA programming language, 203
- Adler, Ellen, 23
- Agriculture, 225
- Aiken, Howard, 206, 210
- Aircraft engines, 223
- Albert Einstein, 257
- Alchemy, 12, 20
- Allowed energy bands, 102, 214
- Allowed orbits, 19, 31
- Alpha particles, 13, 20
- Alpha rays, 12
- America, 259
- Amino acids, 125
- Analytical engine, 202
- Analytical machine, 202
- Anderson, Carl David, 46
- Angular momentum, 27, 43
- Anode, 119
- Anti-war manifesto, 260
- Antibonding orbitals, 82
- Anticommutation, 158
- Antimatter, 46
- Antisymmetric functions, 282
- Antisymmetry requirement, 76
- Approximate N -electron Schrödinger equation, 287, 290
- Approximate Schrödinger equation, 346
- Arms race, danger of, 38
- Artificial intelligence, 199, 226
- Artificial transmutation of elements, 20
- Artistic creativity, 226
- As big as a house, 101
- ASCC calculator, 206, 210
- Aslos mission, 38
- Atom, model of, 16, 26
- Atomic bomb, 260
- Atomic numbers, 31
- Atomic orbitals, 67, 72
- Atomic spectra, 28
- Atomic spin-orbitals, 301
- Atomic units, 157, 158, 279, 291
- Augusta Ada, 203
- Austria, 43
- Autoassembly, 218
- Automatic warehouses, 224
- Automation, 222, 225
- Autoradiography, 129
- Axons, 226
- B-like isoelectronic series, 293
- Babbage, Charles, 199, 201, 202
- Bacterial cell wall, 125
- Ballhausen, Carl J., 222
- Band structure of crystals, 102, 214
- Banking, 224
- Bardeen, John, 102, 213, 215
- Bare nucleus, 293
- Base pairs, 128
- Basis functions, 242, 252, 305
- Bateson, Gregory, 210, 213
- Be-like isoelectronic series, 293
- Beck, Guido, 32
- Becquerel, Henri, 11

- Bell Telephone Laboratories, 102, 207, 209, 213
- Beltrami, Eugenio, 267
- Bending of a light ray, 261
- Berkeley, 80
- Bernal, J.D., 125
- Beryllium, 73
- Beta rays, 12
- Binary digit, 211
- Binary numbers, 104, 207, 218
- Binding energies, 19, 31
- Binomial coefficient, 301
- Biosemitotics, 211
- Biosynthesis of proteins, 129
- Bits, 104, 218
- Bits per unit area, 218
- Black body radiation, 26
- Bloch, Felix, 32
- Bloch, Felix, 46
- Block-diagonal form, 242, 247
- Bohr contacts Roosevelt and Churchill, 38
- Bohr effect, 23
- Bohr's atomic theory, 31, 43
- Bohr's escape to Sweden, 33
- Bohr, Aage, 23, 33, 38, 39
- Bohr, Christian, 23
- Bohr, Harold, 23
- Bohr, Margrethe, 23
- Bohr, Niels, 16, 20, 23, 26, 33, 38
- Bohr, Thomas, 23
- Bohr, Vilhelm, 23
- Boltzman's kinetic theory, 259
- Bombyx mori, 211
- Bonding orbitals, 82
- Born, Max, 44
- Boron, 73
- Boundary conditions, 166
- Box normalization, 166
- Bragg angle, 121
- Bragg, William Henry, 19, 121
- Bragg, William Lawrence, 19, 121
- Brain, mechanism of, 226
- Brattain, Walter, 102, 213, 215
- Brogie, Louis de, 43, 46
- Brownian motion, 258, 259
- Bush, Vannevar, 207
- Bytes, 104, 218
- C-like isoelectronic series, 293
- Cajal, Ramòn y, 226
- Calculating box, 199
- Calculations took hundreds of hours, 101
- Cambridge University, 11, 44, 76, 121
- Canada, 12
- Canonically conjugate momentum, 148
- Carbon, 73
- Carbon-like ground state, 302
- Cartesian coordinates, 263
- Cathode rays, 119
- Cavendish Laboratory, 11
- Central processing unit, 103, 214
- Central processing units, 82, 218
- Chain of point masses, 149
- Chains of subgroups, 278
- Character tables, 252
- Characters, 242, 247, 251
- Charge acceptors, 82
- Charge conservation, 157
- Charge density, 155
- Charge donors, 82
- Chemical bonds, 44, 72
- Chemical industries, 223
- Chemistry, 19, 222
- Cherwell, Lord, 38
- Chess-playing program, 227
- Chips, 104, 217
- Choice of V_0 , 300
- Christoffel symbol, 272
- Chromatography, 129
- Church-Turing hypothesis, 207
- Churchill, Winston, 38
- Classes, 248
- Classification of basis functions, 252
- Classified basis functions, 253
- Clebsch-Gordan coefficients, 160
- Clementi, Enrico, 97, 101, 222

- Cockcroft, Sir John, 20
 Collapse of Rutherford's atom, 27
 Collection of harmonic oscillators, 168
 Collective consciousness, 222
 Colossus, 207
 Commutation, 254
 Complementarity, 129
 Completeness, 282
 Computer disc storage, 83, 218
 Computer memories, 104, 218
 Computer networks, 83, 104, 218, 224
 Computer-assisted design, 223
 Computerization of commerce, 223
 Computers, 199
 Conduction band, 103, 214
 Conduction bands, 102, 214
 Conductor, 102, 214
 Configuration Interaction, 285
 Configurations, 332
 Confluent hypergeometric function, 64, 161, 273
 Conjugate eigenvalue problem, 273
 Conjugate function, 290
 Conjugate transpose, 244
 Consciousness, 211
 Container ships, 225
 Continuous media, 163
 Continuous systems, 164
 Contraction of tensors, 264
 Contravariant metric tensor, 263, 268
 Contravariant tensors, 263
 Control and communication, 209
 Coordinate transformations, 240
 Copenhagen, 16, 31
 Coulomb Sturmian basis functions, 277
 Coulomb Sturmiants, 273, 319, 320
 Coulson, Charles, 80
 Countess of Lovelace, 203
 Covariant metric tensor, 263, 267
 Covariant tensors, 263
 Creation and annihilation operators, 168
 Crookes' tubes, 11
 Crookes, Sir William, 119
 Crystal fields, 311
 Crystallography, 222
 Crystals, 17, 102, 213
 Cubic field, 313
 Cultural evolution, 199, 221, 226
 Current density 3-vector, 155
 Current density 4-vector, 156, 157
 Curvature of space, 262
 Curvature, intrinsic, 271
 Curved spaces, 272
 Curvilinear coordinate transformation, 266
 Cybernetics, 207, 209, 213
 Cytoplasm, 129
 D'Alembertian operator, 156
 Danish resistance movement, 33
 Darwin, Sir Charles Galton, 17
 Data banks, 222
 Debye, Peter, 46
 Definition of a finite group, 239
 Degeneracy, 254, 278
 Degenerate configurations, 346
 Dendrites, 226
 Denmark's Jewish community escapes, 33
 Densities, 335
 Derivation of the secular equations, 287
 Detection of radio waves, 11
 Diagnosis, 223
 Diatomic molecules, 322, 333
 Differential and integral calculus, 201
 Diffraction grating, 17
 Diffraction of X-rays, 17
 Diffractometers, 222
 Digital computers, 76
 Dirac alpha matrices, 158
 Dirac equation, 157, 158
 Dirac equation for a hydrogenlike atom, 159
 Dirac gamma matrices, 158
 Dirac Hamiltonian, 158
 Dirac notation, 254
 Dirac spinors, 158
 Dirac's relativistic wave equation, 46
 Dirac, P.A.M, 202

- Dirac, P.A.M., 44, 157
Discharge tube, 119
DNA, 129
Doping, 103, 214, 217
Dorothy Crowfoot Hodgkin, 125
Double-stranded DNA, 128
Drum parity, 101
Due degeneracies, 254, 346
Dynamical symmetry, 278
Dynamics of reactions, 223
- Eastern Europe, 101
Effective charge, 286
Ehrenfest, Paul, 27
Eigenvalues of L_z , 306
Eigenvalues of S_z , 306
Eigenvalues of the interelectron repulsion matrix, 305
Einstein convention, 263, 267
Einstein in Italy, 258
Einstein's family, 258
Einstein's letter to Freud, 262
Einstein, A., 155
Einstein, Albert, 16, 26, 27, 43
Ekert, J.P., 207
Elector of Hanover, 201
Electric field vector, 156
Electric organs, 211
Electrodynamics of Moving Bodies, 259
Electromagnetic fields, 164
Electromagnetic potential 4-vector, 156, 157
Electromagnetic potentials, 165
Electromagnetic radiation, 26
Electromagnetic theory, 155
Electromagnetic waves, 119
Electromagnetism, 259
Electromechanical calculators, 207
Electron exchange, 282
Electron microscopy, 129
Electron rest mass, 158
Electron spin resonance, 129
Electron-hole pair, 46
Electronic data, 224
Electronic digital computers, 207
Electronic valves, 103, 214
Electrons, 27
Electrophoresis, 129
Electrostatic forces, 125
Electrostatic potential, 155
Elements, 12, 19
Embedded computer, 224
Emmeche, Claus, 211
Energy-independent matrix, 290
England, 13
ENIAC digital computer, 207
Entropy, 209
Equilibrium position, 147
Equivalence of inertial frames, 155
Equivalent representations, 242
Etched away by acid, 103, 217
Ether, 259
Ethics, 38
Ethology, 211
ETO's, 319
Euclidean geometry, 257
Euclidian space, 262
Euler-Lagrange equations, 266, 272
Excited configurations, 306
Excited states, 305
Expansion of products, 335
Exponential factor, 321
Exponential-Type Orbitals, 319
External electromagnetic potential, 157, 158
- F-like isoelectronic series, 302
Fairchild Semiconductor, 215
Feedback in biology, 209
Few-electron atoms, 305
Fiber optics, 83, 221
Field-point, 157
Fields, 163
Film animation, 225
Financial book-keeping, 224
Fine structure constant, 158
Finite-dimensional groups, 239
Fisher, R.A., 209

- Flat organic molecules, 82
Floating-point operations, 104, 217
Flops, 82, 104, 217, 218
Fluorine, 73
Fock projection, 277, 278, 326, 337
Fock's transformation, 69
Fock's treatment, 68
Fock, V., 97, 222, 277, 278, 326
Fock, Vladimir A., 73
Foerster, Heinz von, 210
Forbidden energy bands, 102, 214
Force constant of a spring, 150
Four-vectors, 155
Fourier coefficients, 166
Fourier convolution theorem, 281
Fourier transform representation, 335
Fourier transformed hydrogenlike orbitals, 71
Fourier transforms, 277, 278, 280
Franck, James, 32
Franklin, Rosalind, 129
Free electromagnetic field, 168
Frequency distribution, 26
Frequency spectrum, 150
Frisch, Otto, 32
- Göttingen, 44
Gamma rays, 12
Gasaneo, Gustavo, 300
Gaussian basis functions, 319
Gaussian expansions, 339
Gegenbauer polynomials, 278, 327
Geiger, Hans, 13
Geiger-Marsden experiment, 13
General relativity, 272
Generalized angular integral, 62
Generalized Shibuya-Wulfman matrix, 333
Generalized Sturmian basis set, 279
Generalized Sturmian Method, 285, 305, 346
Generalized Sturmian secular equations, 287, 290
Generalized Sturmians, 278–281
Generalized Wigner coefficients, 328
Genetic evolution, 221
Genetic material, 128
Geodesics, 272
German nuclear program, 38
Germanium, 103, 214
Germany, 17, 119
Gigaflop 11, 104, 217
Gold medal competition, 23
Goscinski, O., 278, 286
Goscinskian configurations, 286, 287, 293, 306, 311
Goscinskian configurations, advantages and disadvantages, 299
Graphics chips, 225
Gravitation, 260
Great orthogonality theorem, 244, 245, 247, 249, 254
Grossman, Marcel, 258, 262
Group elements, 251
Group representations, 240
Group-theoretical projection, 249
Groves, General L., 39
- Hückel, Erich, 82
Hamiltonian, 290
Hamiltonian density, 165
Hansen, H.M., 28
Hardware, 226
Harmonic normal modes, 166
Harmonic oscillator Lagrangians, 150, 167
Harmonics, 43
Hartree potential, 300
Hartree, D.R., 222
Hartree, Douglas, 73
Hartree, Douglas R., 97
Hartree, William, 97
Hartree-Fock-Roothaan equations, 285
Hartree-Fock-Roothaan Method, 76
Hartrees, 291
Hawking, Steven, 202
Heisenberg, Werner, 33, 44, 46
Heitler, Walter, 80
Heitler-London theory, 80
Helium, 12, 73

- Hemoglobin, 23, 125
Hermann Minkowski, 258
Hermitian adjoint, 244
Herschbach, Dudley, 223
Hevesy, George de, 32, 38
Highest filled orbitals, 82
Hilbert space, 252, 253
Hilbert, David, 44
Hippel, Arthur von, 32
Hodgkin, Alan, 226
Hodgkin, Dorothy, 222
Hodgkin, Dorothy Crowfoot, 125
Hoffmeyer, Jesper, 211
Hollerith, Herman, 206
Homeostasis, 209
Homogeneous functions, 58
Homogeneous wave equation, 166
Hospitals, 223
Hundreds of times faster and cheaper, 102
Hyperspherical coordinates, 267
Huxley, Andrew, 226
Hydrogen, 73
Hydrogen bonds, 128
Hydrogen molecule, 82
Hydrogen spectrum, 28
Hydrogenlike orbitals, 273, 321
Hydrogenlike Schrödinger equation, 273, 321
Hydrogenlike spin-orbitals, 286, 287, 300
Hydrogenlike wave equation, 286
Hydrophilic groups, 125
Hydrophobic groups, 125
Hyperangular integration, 328
Hypergeometric functions, 161
Hypersphere, 277, 326
Hyperspherical harmonics, 59, 60, 62, 66, 278, 326, 327, 335, 336
IBM Corporation, 104, 206, 217
IBM machine, 101
Idempotents, 252
Identity element, 239, 251
Imperial College, 101
Impurities, 103, 217
Income policies, 225
Increasingly paranoid, 214
Independent electrons, 293
Inertial frames, 155
Infinitesimal element of length, 271
Infinitesimal length, 263
Informality, enthusiasm and speed, 16
Information, 104, 218
Information accumulation, 221
Information explosion, 221
Information technology, 199
Inner product, 280
Institute for Theoretical Physics, 31
Insulator, 102, 214
Insulin, 125
Integral equation, 281
Integrated circuits, 103, 217
Integration by parts, 164
Intel, 218
Interactive calculations, 104, 218
Interelectron repulsion, 288, 293, 300, 346
Interelectron repulsion integrals, 335, 336
Interelectron repulsion matrix, 293, 302, 305
Internet, 222
INTERNIST-1, 223
Intrinsic curvature, 271
Invariance, 164
Invariant blocks, 300, 306
Invariant subsets, 300, 306, 346
Invariants, 248
Invention of computers, 199
Invention of transistors, 102
Inventory data base, 224
Inverse, 239, 244
Inverse transform, 277
Inversions, 300
Ionization, 12
Ions, 73
Ions in gases, 11
Irreducible representations, 244, 252
Isoenergetic basis sets, 278
Isoenergetic configurations, 332, 346

- Jacobi, Carl Gustav, 265
Jacobian, 265
Jacobians, 267
Jacquard's loom, 203
Jacquard, Joseph Marie, 203
Jordan, Pascal, 44
- Kaiser Wilhelm Institute, 260
Kendrew, J.C, 125
Kendrew, J.C., 125
Kinetic energy density, 163
Kinetic energy operator, 287
Kinetic energy term, 279, 280
King George I, 201
Klein-Gordon equation, 46
Koga, T., 321
Krogh, August, 23
Kronecker delta, 263
Kull, K., 211
- Löwdin, Per-Olov, 222
Ladik, Janos, 101
Lagrangian densities, 163
Lagrangian density, 164
Laplace-Beltrami operator, 266
Laplacian operator, 156, 157
Large and small components, 160
Large-Z approximation, 293, 346
Laser, 224
Laue, Max von, 17
Lawrence, E.O., 20
LCAO approximation, 82
LCAO SCF Method, 76
Learning by computers, 226
Lee, Yuan, 223
Legendre polynomials, 278, 327
Leibniz, Gottfried Wilhelm, 199, 201
Leontief, Wassily W., 225
Levi, Hilde, 32
Levi-Civita tensor, 265
Levi-Civita, Tulio, 265
Lewin, Kurt, 210
Lewis structures, 80
- Lewis, G.N., 80
Lie detector test, 215
Ligands, 311
Linear coordinate transformations, 240
Linear field, 313
Linus Pauling, 82
Literature searches, 223
Lithium, 73
Lithium-like ground state, 301
Lithiumlike isoelectronic series, 293
Logic density, 218
London, Fritz, 80
Lord and Lady Byron, 203
Lord Moulton, 202
Lorentz gauge, 157
Lorentz invariance, 155, 163, 165, 259
Lorentz transformation, 155, 165
Lorentz, H.A., 259
Lorenz, Konrad, 211
Los Alamos, 38
Lovelace, Countess of, 203
Lowest empty orbitals, 82
Lysozyme, 125
- Machine instructions, 104, 217
Macy Conferences, 213
Macy Foundation, 210
Magnetic disc memories, 104, 218
Magnetic field vector, 156
Magnetic ink, 224
Magnetic moment of electrons, 73
Magnetic splitting of levels, 73
Manchester, 20, 31
Manchester University, 13
Many-center Coulomb Sturmians, 319
Many-center potentials, 321
Many-particle problems, 278
Marcel Grossman, 258, 262
Marsden, Ernest, 13
Mass defect, 260
Mass increases with velocity, 260
Mass-weighted coordinates, 147
Mass-weighted potential energy matrix, 148

- Mathematica, 337
Matrix mechanics, 44
Matrix representation of operators, 253
Matter waves, 43
Mauchley, J.W., 207
Maurice Wilkins, 129
Maxwell's equations, 16, 26, 157, 259, 260
McCulloch, Warren, 210
McGill University, 12
MCIR system, 224
Mead, Margaret, 210, 213
Mechanization of agriculture, 225
Medicine, 223
Meitner, Lise, 32
Memories, 104, 218
Mercury computer, 101
Messenger RNA (mRNA), 129
Metric tensors, 262
Michaelson-Morley experiment, 155
Michelson, A.A., 259
Michelson-Morley experiment, 259
Microelectronics, 103, 213, 217
Microminiaturization, 101
Microprocessors, 82, 104, 217, 218
Microscope, 103, 217, 226
Military mentality, 258
Miniaturization, 82, 103, 217, 218
Minicomputer, 104, 217
Minimal geodesics, 272
Minkowski space, 155, 157, 260
Minkowski, Hermann, 260
MIPS, 104, 217
Mixed tensors, 263
Molecular biology, 129
Molecular calculations, 332
Molecular dynamics, 223
Molecular orbital theory, 82
Molecular orbitals, 319
Molecular structure, 19
Molecular Sturmians, 335
Molecular vibrations and rotations, 150
Molecules, 320
Momentum space, 277, 326, 335
Momentum space orthonormality, 280
Momentum-space, 69, 277
Momentum-space orthonormality, 280, 281
Moore's law, 218
Moore's law, 218, 221
Moore, Gordon, 218
Moore, Gordon C., 215
Moore, Gordon E., 218
Morley, E.W., 259
Moseley's law, 19
Moseley, Harry, 17
Moseley, Henry, 31
Motion of the nucleus, 287
Mottelson, Ben Roy, 38
Mullikin, R.S., 222
Multiplication, 199
Multiplication and division, 201
Multiplication of group elements, 239
Multiplication table, 241, 251
Music synthesizers, 225
Myoglobin, 125

N-like isoelectronic series, 302
Nature of the Chemical Bond, 80
Nazi occupation of Denmark, 33
Ne-like isoelectronic series, 302
Negative energy states, 46
Negative feedback, 211
Neon, 73
Nerve cells, 226
Nervous system, 226
Networks, 104, 218
Neumann, John von, 207, 209
Neural networks, 226
Neurophysiology, 209
New Zealand, 12
Newman, M.H.A., 207
Newton's equations of motion, 43, 260
Newton, Isaac, 201
Newtonian mechanics, 259, 261
Niels Bohr Institute, 16, 38
Nitrogen, 73
Nobel Prize in Chemistry, 80

- Non-Abelian groups, 252
 Non-Euclidean geometry, 262
 Non-relativistic wave equation, 288
 Nonequivalent irreducible representations, 249
 Nonhomogeneous polynomials, 62
 Normal coordinates, 148, 150
 Normal modes, 147, 151, 163, 166
 Normalization, 289
 Normalizing constant, 278, 327
 Nuclear arms race, 38
 Nuclear attraction, 288
 Nuclear attraction integrals, 322
 Nuclear attraction matrix, 290
 Nuclear attraction matrix diagonal and energy
 independent, 305
 Nuclear charge, 13, 20
 Nuclear charge, weighted, 286
 Nuclear charges, 311
 Nuclear magnetic resonance, 129
 Nucleic acids, 129
 Nucleus, 13
 Null matrix, 245

 O-like isoelectronic series, 302
 One-center expansions, 311
 Optical storage devices, 104, 218
 Optimal basis set, 305
 Orbital angular momentum, 301
 Orbital hybridization, 80
 Order of a group, 239
 Orthonormal sets, 278
 Orthonormality, 287
 Orthonormality relations, 167
 Oxygen, 73

 Pacifism, 262
 Palade, George Emil, 129
 Pale-faced teenagers, 101
 Parallel-processing, 104, 217
 Parallelization, 82, 218
 Pascal, Blaise, 199
 Pauli exclusion principle, 72, 80, 82
 Pauli principle, 301
 Pauli spin matrices, 158
 Pauli, Wolfgang, 46, 73, 76, 80
 Pauli-allowed states, 306
 Pauling, Linus, 80, 82, 222
 PC hard-drive capacity, 218
 Peirce, Charles Sanders, 211
 Periodic system, 19
 Periodic table, 44, 72
 Personal memories of early computers, 100
 Perturbation, 301
 Perutz, Max, 125
 Phillips, D.C., 125
 Photoelectric effect, 27, 258
 Photon annihilation operators, 169
 Photon creation operators, 169
 Photoresist, 103, 217
 Photosynthesis, 223
 Physics and Beyond, 33
 Pitts, Walter, 210
 Placzek, George, 32
 Planck's constant, 26, 158
 Planck's quantum hypothesis, 26, 259, 260
 Planck, Max, 16, 26, 260
 Planck-Einstein formula, 19, 28, 31
 Plane wave, Sturmian expansion, 282
 Point charges, 311
 Poison gas, 260
 Polanyi, J., 223
 Polarized light, 211
 Polyatomic molecules, 333
 Polynucleotides, 129
 Polypeptides and polynucleotides, 101
 Population explosion, 221
 Position vector, 155
 Positron, 46
 Potential, 288
 Potential energy density, 163
 Potential-weighted Orthonormality, 287
 Potential-weighted orthonormality, 275, 279,
 280, 289, 335
 Pre-calculated integrals, 337
 Pre-calculated matrices, 336
 Principal quantum number, 287

- Principal quantum numbers, 300
Principle of Equivalence, 261
Product of characters, 249
Products of Coulomb Sturmians, 335
Projection operator, 282
Projection operators, 249, 252
Protein chain, 129
Protein structure, 125
Protein synthesis, 129
Protons, 20
Pseudo-Euclidean space-time continuum, 155
Pseudo-rotation, 155
Pullman, Alberte, 82, 223
Pullman, Bernard, 82, 223
Punched cards, 101, 203
Pythagoras, 43
- Quanta, 26
Quantization of angular momentum, 27, 43
Quantum Biochemistry, 82
Quantum biochemistry, 223
Quantum chemical calculations, 101
Quantum chemistry, 222
Quantum dot technology, 218
Quantum dots, 218
Quantum hypothesis, 26
Quantum numbers, 19, 31, 287
Quantum theory, 44, 102, 213, 221
Quick Medical Reference, 223
- R-blocks, 300, 306
Rabinowitch, Eugene, 32
Radial functions, 273, 320
Radial functions, relativistic, 160
Radial wave functions, 65
Radioactive tracer techniques, 129
Radioactive tracers, 129
Radioactive transmutation, 12
Radioactivity, 20, 121, 260
Radium, 20, 260
Raleigh, Lord, 23
Rank of a tensor, 263
Rapidly rotating drum, 101
Rayleigh, Lord, 28
Reactive scattering, 223
Reciprocal space, 326
Red shift, 261
Reducibility, 242
Reducible representations, 247
Reflections, 300
Regular representation, 251
Relativistic 1-electron energies, 161
Relativistic effects, 155
Relativistic electrodynamics, 156
Relativistic hydrogenlike orbitals, 160
Relativistic wave equation, 157
Relativity theory, 258
Removal of degeneracy, 301
Repeated eigenvalues, 245
Replacement orders, 224
Representations of groups, 240
Resonance, 80
Rest energy, 158
Retailing, 224
Ribonucleic acid, 129
Ribosomes, 129
RNA, 129
RNA and ribosomes, 129
Robots, 223, 226
Rockefeller Institute, 129
Roentgen, Wilhelm Konrad, 119
Rolls Royce Ltd., 223
Roosevelt, Franklin, 38
Roothaan, Clemens C.J., 97
Roothaan, Clemens, C.J., 76, 82
Rosalind Franklin, 129
Rosenbluth, Arturo, 209
Rotations, 150, 239, 300
Royal Institution, 121
Royal Institution, London, 125
Royal Society, 202
Rozental, Stefan, 32
Russell-Saunders states, 301
Rutherford's model of the atom, 13, 16, 26
Rutherford, Lord, 11, 20, 31
Rydberg, Johannes, 19, 31

- Scalar density, 266
 Scalar products, 247
 Scalars, 264
 Scaling of the coordinate system, 333
 Scaling parameter, 280, 287, 305
 Scattering, 13
 SCF calculations, 285
 Scheraga, Herald, 223
 Schneider, Erich Ernst, 32
 Schrödinger equation, 279
 Schrödinger equation in reciprocal space, 68
 Schrödinger's wave equation, 73
 Schrödinger, Erwin, 209
 Schrödinger equation, 43
 Schrödinger, Erwin, 43, 222
 Schur's lemma, 244, 245, 252
 Second Industrial Revolution, 225
 Secular equation, 319
 Secular equations, 148, 287, 290, 333
 Self-Consistent Field method, 97
 Self-consistent field method, 76
 Self-reinforcing accumulation, 221
 Semiconductors, 103, 214
 Semiempirical approximations, 82
 Semiotics, 211
 Sensation, 213
 Sensors, 223
 Shannon, Claude, 209
 Shells, 301
 Shibuya-Wulfman integrals, 326, 327
 Shift operator, 254
 Shockley, William, 102, 213, 214
 Sigmund Freud, 262
 Sign systems, 211
 Silicon, 103, 214
 Silicon Valley, 214
 Similarity transformation, 244
 Similarity transformations, 241
 Simple harmonic oscillators, 149
 Single-stranded DNA, 128
 Sixth-order polynomials, 202
 Slater determinants, 285, 286, 332
 Slater exponents, 305
 Slater, J.C., 222
 Slater-Condon rules, 287
 Sleeping on three chairs, 101
 Small displacements from equilibrium, 147
 Sobolev spaces, 281, 282
 Social communication, 209
 Socially beneficial projects, 226
 Soddy, Frederick, 12
 Sodium, 73
 Software, 226
 Soho, 101
 Solid angle element, 69
 Sorbonne, 43
 Source-point, 157
 Space exploration, 103, 217
 Space-component, 156
 Space-time continuum, 259, 272
 Space-time symmetry, 155–157, 259, 260
 Spain, 226
 Sparse matrices, 336
 Special relativity, 258, 259
 Special theory of relativity, 155
 Specific heats, 27
 Spectra of X-rays, 17
 Spectral lines, 28
 Spectrum of energies, 305
 Speed of computers, 104, 217
 Speed of light, 103, 214, 259, 260
 Spherical harmonics, 160, 273, 286, 320, 327
 Spherical polar coordinates, 262
 Spherical spinor, 160
 Spin, 73, 82
 Spin-orbit coupling, 287
 Spin-spin coupling, 287
 Spinors, 4-component, 158
 Spray-painting, 223
 Square matrix, 245
 Square-planar field, 313
 Stability of atoms, 16, 26
 Standard irreducible representation, 254
 Standard set, 278
 Standard tree, 267
 Standard unitary representation, 249

- Standardization, 224, 225
Stibitz, George R., 207
Stockmarket, 224
Storage density, 83, 218
Strong projection operators, 252, 253
Structure of DNA, 128
Structure of proteins, 125
Sturmian basis sets, 278, 280
Sturmian expansion of a plane wave, 282
Sturmian overlap integrals, 326
Sturmian radial function, 335
Sturmian secular equations, 280, 287, 290, 321
Sturmians, 319
Subcellular particles, 129
Subgroups, 278
Subshell sets, 306
Subspaces, 252
Sugar-phosphate backbone, 128, 129
Supermarkets, 224
Surface tension measurement, 23
Swiss Patent Office, 258
Switzerland, 27, 43
Symmetry group, 346
Symmetry of molecules, 239
Symmetry-adapted basis functions, 346
Symmetry-adapted basis sets, 253, 306
System of N electrons, 282
Szent-Györgyi, Albert, 82
Szent-Györgyi, Albert, 223
Szilard, Leo, 209
- Tables of quadratic functions, 202
Taylor series expansion, 147
Technological unemployment, 225
Teller, Edward, 32
Templates, 128
Tensor analysis, 262
Tensor capacities, 266
Tensor densities, 266
Tensors, 263
Tetrahedral field, 313
Theory of Games, 210
- Thin films, 103, 217
Thomson's model of the atom, 13
Thomson, J.J., 11, 31
Thought experiment, 261
Thymine, 129
Time-component, 156
Time-independent Dirac equation, 158
Time-independent solution, 156
Tisvilde, 33
Total angular momentum, 160
Total internal reflection, 83, 221
Total potential, 333
Total spin operator, 301
Trace, 247
Traitorous eight, 214
Transfer RNA, 129
Transformation laws, 263
Transistors, 104, 217, 221
Transistors, invention of, 102
Translations, 150
Transmutation of elements, 12
Transportation, 224
Trial Lagrangian density, 166
Trigonal field, 313
Triplet states, 302
Truman, Harry, 39
Tubes sometimes failed, 101
Turing, Alan, 207
Two-center densities, 339
- Uexküll, Jakob von, 211
Ultracentrifugation, 129
Ultraviolet light, 211
Ultraviolet spectroscopy, 129
Umwelt, 211
Unemployment, 226
Uniform motion, 155
Unit matrix, 252, 293
Unitarity, 241
Unitary representations, 247
Universal coefficients, 339
Universal computer, 207
Universal integrals, 337

- Universal product code, 224
- Universal programmable computer, 203
- University of Adelaide, 121
- University of California, 80
- University of Copenhagen, 23
- Uracil, 129
- Uranium, 11, 260

- Vacuum tubes, 207
- Vacuum tubes required cooling, 101
- Valance band, 103, 214
- Valence, 80
- Valence bands, 102, 214
- Valence bond theory, 80
- Valence electrons, 301
- Van de Graff, J.H., 20
- Van Vleck, J.H., 222
- Variational optimization, 305
- Variational principles, 164
- Vector potential, 155
- Velocity of light, 158
- Vibrations, 150
- Virial theorem, 287
- Vitamin B12, 125
- Volume element, 264
- Von Laue, Max, 121

- Wafers, 103, 217
- Waggle dance, 211
- Watson-Crick model, 128
- Wave equation, 166, 222
- Wave equations, 163
- Wave lengths, 163
- Wave mechanics, 43
- Wave theory of light, 17, 119, 259
- Weak projection operators, 252
- Weighted nuclear charge, 285, 286
- Weighted potential, 278, 280, 346
- Weighting factors, 300
- Weisskopf, Victor, 32
- Wholesaling, 224
- Why War?, 262
- Wiener, Norbert, 207, 209, 213

- Wilkins, Maurice, 129
- Wilson, E. Bright, 80, 222
- Word-processors, 225
- World War I, 19, 31
- Writing, 226
- Wulfman, C.E., 321

- X-ray crystallography, 19, 121
- X-ray diffraction, 129, 222
- X-rays, 11, 119
- X-rays, spectra of, 17

- Zürich Polytechnic Institute, 258
- Zürich, 43
- Zero-frequency modes, 151
- Zuse, Konrad, 207