Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

R. Bergström (203), Uppsala University, Dept. of Quantum Chemistry, P.O. Box 518, Uppsala S-75120, Sweden

Manuel Berrondo (107), Brigham Young Univ., Dept. of Physics, 268 FB, Provo, UT 84602

Erkki J. Brändas (121), Uppsala University, Dept. of Quantum Chemistry, P.O. Box 518, Uppsala S-75120, Sweden

Sylvio Canuto (161), Instituto de Fisica, Universidade de São Paulo, Caixa Postal 66318, São Paulo 05315-970, Brazil

Kaline Coutinho (161), C/O Dr. Sylvio Canuto, Instituto de Fisica, Universidade de São Paulo, Caixa Postal 66318, São Paulo 05315-970, Brazil

Piotr Froelich (185), Uppsala University, Dept. of Quantum Chemistry, P.O. Box 518, Uppsala S-75120, Sweden

Osvaldo Goscinski (51), Uppsala University, Dept. of Quantum Chemistry, P.O. Box 518, Uppsala S-75120, Sweden

Göran Karlsson (1), Department of Mechanics, Kungl. Tekniska Högskolan – KTH, (Royal Institute of Technology), S-100 44 Stockholm, Sweden

Sven Larsson (9), Chalmers University of Technology, Dept. of Physical Chemistry, S-41296 Göteborg, Sweden

Jan Linderberg (xi, 87), Aarhus University, Dept. of Chemistry, Langelandsgade 140, Aarhus C. DK 8000, Denmark

Sten Lunell (203), Uppsala University, Dept. of Quantum Chemistry, P.O. Box 518, Uppsala S-75120, Sweden

David A. Micha (139), Quantum Theory Project, University of Florida, P.O. Box 118435, Gainesville, FL 32611-8435

Yngve Ohrn (xi, 35), Quantum Theory Project, University of Florida, P.O. Box 118435, Gainesville, FL 32611-8435

Lars Ojamäe (203), Dept. of Chemistry, Kemi–IFM, Linköpings Universitet, S-581 83 Linköping, Sweden

P. Persson (203), Uppsala University, Dept. of Quantum Chemistry, P.O. Box 518, Uppsala S-75120, Sweden

John Sabin (xi), Quantum Theory Project, University of Florida, P.O. Box 118435, Gainesville, FL 32611-8435

Daniel Trzesniak (161), C/O Dr. Sylvio Canuto, Instituto de Fisica, Universidade de São Paulo, C.P. 66318, São Paulo 05315-970, Brazil

Advancing Quantum Chemistry: Per-Olov Löwdin 1916-2000

Uppsala and Gainesville were the nodal points in Per-Olov Löwdin's World Wide Web of quantum chemistry. He was a scientist who introduced new and lasting methods and concepts in his field, while maintaining an unusual zeal in spreading the good word though extensive travel and indefatigable concern about teaching quantum chemistry through summer schools, winter institutes, and lecture tours. He took new initiatives for making scientists from Latin America, Africa, and Asia members of his community.

Per-Olov was born in Uppsala on October 28, 1916 as the son of the musician Erik Wilhelm Löwdin and his wife Eva Kristina, née Östgren. He showed mathematical proficiency early in his school work, and when he entered Uppsala University in 1935 his plans were to major in mathematical physics. His first scientific paper in 1939 on the Lorentz-transformation and the kinematical principle of relativity was published in Swedish in the journal *Elementa*. The years of war that followed meant interruptions in communications and research opportunities and Per-Olov, like most young Swedish men, spent time in the military defending his country.

The new quantum electrodynamics presented serious challenges to a theoretician and Per-Olov's research notes from the 1940's on this topic show that he spent much time and effort working on these problems. An extended stay with Wolfgang Pauli at Zürich in 1946 became a turning point in Per-Olov's scientific interests. At about this time, he started work in solid state physics and began his dissertation research with Professor Ivar Waller at Uppsala. The topic of his thesis work was a quantum mechanical study of ionic crystals and the examination of the Cauchy relations. These relations arise from the assumption of pair-wise interactions between spherical ions, but were not satisfied in real crystals such as sodium chloride. This pioneering application of quantum mechanics, before the advent of electronic computers, was accomplished with the aid of a cadre of doctoral students using desk calculators and ingenious numerical algorithms designed by Per-Olov.

A Theoretical Investigation into Some Properties of Ionic Crystals. A Quantum Mechanical Treatment of the Cohesive energy, the Interionic Distance, the Elastic Constants, and the Compression at High Pressures with Numerical Ap-

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

plications to some Alkali Halides. (Almqvist & Wiksell, Uppsala, 1948) is the full title of Per-Olov Löwdin's dissertation for the degree of Doctor of Philosophy. The preface is marked May 1, 1948, the successful defense took place some three weeks later, and the ring, laurel, and diploma as insignia of the doctorate were awarded at the commencement ceremony on May 31.

It is interesting to read Löwdin's discussion of his method of investigation and the numerical results from Chapter XI. He specifies the fixed nuclei Hamiltonian in the Schrödinger form where no relativistic effects are considered. He specifies the one-electron approximation with the free ion orbitals and shows that polarization and van der Waals effects are small compared to the terms from exchange and the *S-energy.* This latter concept was to remain a key feature in Per-Olov's future work. He considered all overlap integrals, denoted $S_{\mu\nu}$ and chose to evaluate the formal energy expression to second order. Only nearest neighbor overlap was included since he found the others to be smaller than the general tolerance level he adopted. The detailed numerical calculations were based on the so-called α-function expansions. Orbitals on one atomic center were expanded in terms of spherical harmonics on a neighboring center. The resulting radial factors are the Löwdin α-functions.

The failure of the Cauchy relations derives from the three- and four-body interactions, which stem from the overlap terms. The description of the properties of ionic crystals was brought to a new and improved level by Per-Olov's thesis and he developed an arsenal of tools, which were sharpened and extended, throughout his career.

A five month stay with Neville Mott at Bristol in 1948, and extended periods with Robert Mulliken at Chicago, and with the group of Hertha Sponer at Duke University in the early 1950's set the tone for life filled with international travel. Particularly significant was a stay with the Solid State and Molecular Theory Group of John C. Slater at the Massachusetts Institute of Technology. There developed a close association between Slater and Löwdin which was to last until Slater's death in 1976.

Ram's Head Inn on Shelter Island, in Gardiners Bay at the Eastern end of Long Island, was the venue for a conference that was considered epoch-making for the emerging field of quantum chemistry by the participants. The National Academy of Sciences sponsored the *Conference on Quantum-Mechanical Methods in Valence Theory* with financial support of the Office of Naval Research.

It was held September 8 to 10, 1951, and was attended by the leading figures in the field.[1] Emphasis was given to numerical work. Slater and Ufford reported on the use of electronic computers for self-consistent field calculations, Coulson and Barnett were implementing the ζ-function technique in London, Roothaan, Ruedenberg, and Shull at Chicago and Ames were pursuing diatomic integrals, Kotani's table project was well under way, and Löwdin could present experiences with direct numerical integration methods, which he was developing together with Stig Lundqvist. Per-Olov was now an established scientist in the new field of quantum chemistry.

Three publications are particularly noteworthy in defining the scientist Per-Olov Löwdin in the first half of the 1950's. His concern with overlap led him to formulate *symmetric orthogonalization* as a means of forming an orthonormal basis from an overlapping one.[2] This paper may still be among his most cited works[3] and has been an essential element in the justification of *neglect of differential overlap*.[4] The trilogy *Quantum Theory of Many-Particle Systems*[5] was worked out during Löwdin's stay with Slater's group, and contains a detailed analysis of the configuration space method for dealing with the correlation problem. Density matrices were used as the principle vehicle for the advancement of interpretation of many-electron wave functions. Detailed prescriptions were offered for the evaluation of density matrices from general wave functions in the form of superposition of configurations. General orbital basis sets with overlaps were used. The *natural spin orbitals* were defined and the first efforts to formulate *extended Hartree-Fock* methods appeared here. Löwdin also initiated the use of projection operators in order to handle degenerate states. Spin multiplets were of primary concern and the foundation was laid for the later development of the *alternant molecular orbital* approach.[6] *Quantum Theory of Cohesive Properties of Solids*[7] exhibits, in its 172 pages, the structure of the theoretical approach towards the description of electronic properties of matter

[1] R. G. Parr and B. L. Crawford, Jr., *Proc. Nat. Acad. Sci.* 38 (1952) 547.

[2] P. O. Löwdin, *J. Chem. Phys.* 18 (1950) 365.

[3] R. Manne in *Quantum Science. Methods and Structure* (Plenum Press, New York 1976) Eds. J.-L. Calais, O. Goscinski, J. Linderberg, and Y. Öhrn, p. 25

[4] R. G. Parr, *The Quantum Theory of Molecular Electronic Structure* (W. A. Benjamin, Inc. 1963) p. 51; and *J. Chem. Phys.* 33 (1960) 1184.

[5] P. O. Löwdin, *Phys. Rev.* 97 (1955), 1474, 1490, 1509.

[6] J.-L. Calais, *Abstracts of Uppsala Dissertations in Science* 52 (1965), R. Pauncz, *Alternant Molecular Orbital Method* (W. B. Saunders, Philadelphia 1967)

[7] P. O. Löwdin, *Adv. Phys.* 5 (1956) 1.

as Per-Olov saw it at the time. The numerical procedures, density matrices, natural orbitals, and projection operators were developed in these papers and a thorough review of the literature was included.

Per-Olov received substantial grants that enabled the creation of the Quantum Chemistry Group at Uppsala University in 1955. The *King Gustaf VI Adolf's 70-years Fund for Swedish Culture* and the *Knut and Alice Wallenberg's Foundation* provided funding for the project, while Per-Olov's position as research professor was supported by the *Swedish Natural Sciences Research Council.* Together with Inga Fischer-Hjalmars, Per-Olov arranged the first symposium on quantum chemistry in Sweden at Uppsala and Stockholm in 1955 with participation from the leading practitioners.[8]

Extension of the group was made possible by a research grant starting in 1957 from *Aerospace Research Laboratories, OAR,* through the *European Office of Aerospace Research (OAR), United States Air Force.* An agreement with the Swedish agencies opened the possibility for purchasing a state-of-the-art electronic computer, the *ALWAC IIIE.* This was the first device of its kind at Uppsala, and was formally inaugurated in conjunction with the group's new offices at Rundelsgränd on April 23, 1958, the centennial of the birth of Max Planck.

The electronic correlation problem was a primary research theme in the Quantum Chemistry Group when Jan Linderberg and Yngve Öhrn joined the group in 1957 and 1958 respectively. The review[9] and Yoshizumi's bibliography[10] had been submitted, natural orbitals had been determined,[11] and the Quantum Chemistry Group was emerging as a center for quantum chemical investigation. Further enhancement came through the Summer Schools. Vålådalen in the mountains of Sweden, was the location for the first Summer School, held in August 1958. Löwdin managed to gather several prominent lecturers such as Robert Mulliken, Linus Pauling, Kenneth Hedberg, F. A. Matsen and others for the final symposium week. Ruben Pauncz, a participant then, returned as a valued lecturer for more than twenty years in these annual Summer Schools.

Quantum Chemistry became recognized as an academic discipline in its own

[8] I. Fischer-Hjalmars and P. O. Löwdin, *Sv. Kem. Tidskrift* **67** (1955) 365.

[9] P. O. Löwdin, *Adv. Chem. Phys.* **2** (1959), 207.

[10] H. Yoshizumi, *Adv. Chem. Phys.* **2** (1959) 323.

[11] P. O. Löwdin and H. Shull, *Phys. Rev.* **101** (1956), 1730; J. O. Hirschfelder and P. O. Löwdin, *Mol. Phys.* **2** (1959), 229.

right at Uppsala University with the establishment of a chair in the subject. Per-Olov was appointed to the chair in 1960, the same year that he founded an interdiscplinary research institute, the Quantum Theory Project for Research in Atomic, Molecular, and Solid-State Theory (informally known as QTP)), at the University of Florida. Dean Linton E. Grinter and the heads of the departments of chemistry and physics at UF, Professor Harry S. Sisler and Professor Stanley S. Ballard, respectively, provided the local support for Löwdin and a small contingent of young Swedish scientists[12] as the original staff of the new QTP. With academic homes on both sides of the Atlantic the number of quantum chemists connected to Per-Olov grew and was sustained by the Scandinavian Summer Schools and Florida Winter Schools in Quantum Chemistry. It is estimated that about 4,000 scientists have attended a summer school, winter institute, or conference organized by Per-Olov and his colleagues.

There were certain areas of theory that were not embraced by Per-Olov in the early years. Group theory was not given a satisfactory form to his liking until John Coleman had lectured on group algebra and Per-Olov could cast this in the form of linear algebra.[13] Second quantization had been familiar to him since the 1940's when he worked with quantum electrodynamics and he wanted to connect the use of the Fock space in many-electron theories with the configuration space formulation. In spite of asking his junior colleagues to study the early papers of V. Fock, and others on the subject, no concrete project arose from these efforts and and Per-Olov seemed to share Slater's feelings[14], that field operators were of little use in the study of electronic systems. It was the influence of Stig Lundqvist that inspired the use of field theoretical methods in the Uppsala and Florida groups.

Through 1960, Per-Olov had authored and coauthored some fifty papers. He added nearly two hundred in the following thirty-five years. These covered topics in quantum genetics, proton tunneling, and science in society in addition to further pursuits in quantum theory. The series *Studies in Perturbation Theory* I-XIV demonstrated a search for economy and elegance in presentation, which was important to Per-Olov and became one of his trademarks as a scientist. Per-Olov frequently referred to "the economy of thinking" as shorthand for elegance

[12] P. O. Löwdin, in *Partners in Progress* (Kastrup, A.; Olsson, N. W. Eds. Swedish Council of America, 1977) p. 255.

[13] P. O. Löwdin, *Rev. Mod. Phys.* **39 (1967), 259.**

[14] J. C. Slater, *Am. J. Phys.* **36** (1968), 69.

and compactness of a mathematical derivation. This characteristic manifests itself in his book *Linear Algebra for Quantum Theory* (J. Wiley & Sons, New York 1998), which summarizes over five decades of Per-Olov's teaching efforts.

Even after his official retirement form the University of Florida faculty in 1992, he and his wife Karin continued to travel between Uppsala and Gainesville. Their Florida stays coincided with the annual Sanibel Symposia, where Per-Olov always attended all sessions and from his first row seat continued his habit of making insightful and often complimentary remarks after many of the plenary presentations.

Uppsala University bestows a *Jubilee doctorate* on those alumni who have survived fifty years after their commencement. Thus it was that Per-Olov received a new laurel and an additional diploma on May 29, 1998 as well as a two shot salute by the artillery cannon outside the university aula. The president of the student union saluted the five recipients with a speech that was answered by Per-Olov in his characteristically youthful style. Five months later he had a serious heart operation. He never recovered fully. His health deteriorated, and he passed away quietly on October 6, 2000. A memorial symposium was arranged at Uppsala on the day before the funeral, on October 26. Many of Per-Olov's colleagues and friends also gathered on the University of Florida campus to honor his memory with eulogies by Per-Olov's long time collaborator and friend Harrison Shull and by Robert A. Bryan, who held many top administrative positions at UF, including that of President, during most of Per-Olov's tenure at Florida. The 42'nd Sanibel Symposium held in 2002 was dedicated to his memory.

We remember him as a teacher, an indefatigable lecturer, a helpful colleague, and a friendly competitor. Perhaps his most important characteristic is that of a true internationalist. The Scandinavian Summer Schools, the Florida Winter Schools, and the Sanibel Symposia have, through the vision of Per-Olov Löwdin, furthered international cooperation and friendship between scientists from all continents.

Jan Linderberg,
Yngve Öhrn,
John R. Sabin.

The Scientific, Educational and Internationalization Influence of Per-Olov Löwdin

Göran Karlsson, Department of Mechanics,
KTH, SE-100 44 STOCKHOLM, Sweden

Abstract

The influence of Per-Olov Löwdin on the author's scientific, educational and international activities are described.

Contents

1 Introduction

I came to the Dept. of Quantum Chemistry and the Quantum Chemistry Group at Rundelsgränd (where the department was situated) as a Ph.D. student in the summer of 1965. During my physics studies I had become interested in understanding more of Quantum Mechanics and I was also interested to understand how molecules were formed. I left the department with my Ph.D. dissertation in late spring 1970. First afterwards I really came to understand

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

how special the time at Rundelsgränd was: Per-Olov Löwdin had created an international and creative environment which was far ahead of what now is not only on the research but also on the education agenda. Most of those receiving their Ph.D. from Per continued in their field at universities in or outside Sweden. Even if I myself turned away from Quantum Chemistry I adopted to Per's ideas of the importance of international contacts and of "baby swim".

Immediately after I had decided to start my Ph.D. studies, Per suggested that I should participate in the Summer School 1965 in Abisko and Uppsala. There I was brought into contact with Per's enormous activity and energy: An intensive lecture programme, mountain climbing, ascending of Kebnekaise from Tarfala valley, rushing up Nuolja, night swims in Abisko canyon, social activities every evening. It was intensive. And the activities at the department from the fall were the same: lectures and seminars every morning 8.30-10 (to get people out of bed). And here and in what to study Per had a pedagogical approach which can be called "baby swim" (just throw the baby into the water and he/she is forced to swim without knowing how to do it) but today would referred to as constructivism. And then soccer every Saturday afternoon and social activities.

In pedagogy there is today a strong tendency to require learners to take accountability for their own learning. Piaget, Dewey and Lewin [1] laid the foundation for constructivism and experience-based learning. Bringuier [2] and Papert [3, 4] have described the process of constructive learning as opposed to instructivism: The basic idea of constructivism is the own construction of knowledge by the learner. And in fact it was the Belarusian Lev Semenovich Vygotsky (1896-1934) who introduced the concept. But not before 1962 his book Thought and Language [5] was translated into English. During the last decade constructivism has been supported by Problem-Based Learning (PBL) where real world problems are used to support critical thinking and problem solving skills.

This was Per's educational concept: To give the student immediately the research literature and then support this with advanced lectures and seminars and forcing the student himself/herself to find out the necessary mathematics and numerical methods needed. The concept we today try to use in project education and problem based learning. And it functioned. Today many educational systems try to introduce this educational scheme to get the learner to be accountable for his/her own learning.

More pronounced, visible and well-known was Per's international activity. Not only at the summer schools but he also created a creative international

environment at Rundelsgränd. Having many small children I for many years I suffered during near 10 years of not being able to go abroad, but later at the end of the 70's I started to take international contacts first within Western Europe participating in EU project and during the 80's and before Sweden had joined EU pushing for Swedish participation in EU programs. That was not easy, but then Sweden joined EEA and later EU and that simplified things.

Per also led me into the world of computers (mostly the main frames ones with punched cards but also a table one: Olivetti). At the end of the 60's this was natural, else the discipline should never have evolved in the way it has. But without that introduction I am quite sure I should not be involved in the activities I am today. But even here he used the "baby swim" strategy. No elementary teaching, just consider it as self-evident that computers should be used; the programming and the numerical methods one had to learn oneself. But once he gave an 8.30 seminar series in numerical analysis and I can still inside myself see him lecturing and writing Gramm-Schmidt's orthogonalization procedure, which then, of course, was substituted by symmetric orthogonalization. This was one of his own research concepts: Starting to lecture about something known, and using that to develop a new, pedagogically mathematically feasible formalism. He had an inherent feeling and skill for developing easily handleable mathematical formalisms. Not to speak about his beautiful and easily read calligraphic handwriting. Anyone who has tried to read a Mathematics or Physics textbook in German using old Gothic types appreciates as simple notation as possible; it must be possible to tacitly read the symbols.

2 Florida off-springs

To the Dept. of Mechanics at KTH I came in the fall of 1973 after half a year 1971 at the Department of Chemistry in Guelph, Ont., Canada with Mike Zerner, half a year at the Dept. of Theoretical Physics with Mechanics in Lund, Sweden and one year as a high school lecturer in Mathematics in Köping, Sweden. When I and my wife left Canada for Lund we went by car down from Ontario to Florida (visiting Paul Seybold, another "Uppsala" Quantum Chemist in Dayton, Ohio on our way down). In Gainesville we visited Peter Lindner with family. On Christmas Day we were invited to Per and Karin and there I met Rune "Tex" Lindgren for the first time. He still spends 2/3 of the year in Gainesville. But one of the first persons I ran into when I took up my job at Mechanics at KTH was Tex. 1/3 of the year (3 summer months + 1 fall month) he nowadays spends at KTH. And the last 5-6 years we have shared office, which is not particularly disturbing since about

half of these 4 months I have vacation or is abroad, and the other months we have very interesting conversations. In this way I still get up-to-date reports from Gainesville twice a year.

There is a peculiar similarity between Per and Tex; those being their American vocative names, while both in Sweden got more extended vocative names: Pelle and Texas.

3 The PLATO and CAI time

Quite early in the beginnings of 80's I became interested in the use of Computer Aided Instruction. I was lead into the field by Stig Björklund and we cooperated very well for many years until his retirement. Especially he had introduced the main frame computer (nobody uses this terminology any longer) system PLATO into KTH. PLATO used its own operating system on the Control Data Corporation (CDC) Cyber computers. This was a system far ahead of its time using touch sensitive screens on the terminals, advanced graphics and advanced graphics mail and computer conferencing linking all over the world (communicating with these other places which had the system). The mail and conferencing system was implemented by CDC, but its graphics facilities were invented by Björklund and coworkers at KTH [6]. The main draw backs of PLATO were: (i) ahead of time; (ii) main frame computer with terminals; (iii) own protocol (there was no TCP/IP for generic use), (iv) due to failure of instructional technology CAI did not have a good reputation. But this instructional technology approach is now often used today in web based education (neglecting the possibilities of a constructivistic approach and with too little faculty support). And in fact PLATO has survived from its start 1963 with Don Bitzer at University of Illinois, and later Bruce Sherwood and Control Data and it has now landed at PLATO Learning, Inc. in Bloomington, Minnesota and on the web at http://www.plato.com. I mention this extensively since I found Per's way of operating very inspiring for many of my activities and my work with networking (of people and computers), international contacts, computer use in education and distance education where I also adopted thoughts and ideas from PLATO and Björklund.

From the meetings and conferences organized by SEFI (Societé Européenne pour la Formations des Ingéniurs - European Society for Engineering Education) I had come to know lots of people in the Scandinavian countries and it was then natural to organize the First Conference for Computer Aided Learning in University and University College Education [7] at KTH in Stockholm 1987. As a consequence of this conference, Nordic Forum for Computer Aided

Higher Education was formed 1988 and I was its chairman from the interimistic board of 1988 until its Helsinki conference in August 1991. I had for two years — with no success since it was not politically correct at that time — tried to open the association for universities (only institutional membership was allowed) from the Baltic states. At the Helsinki conference we had, however, many Baltic participants (and also Russians and Belarusians) and they were very concerned about what was going on in their home countries those days. And the Finnish radio and TV did not provide much information. But CNN was a good source of information.

4 Eastern and Central Europe

Directly after the Nordic Forum for Computer Aided Higher Education was formed, distance education started to benefit from computer networking and video conferencing. An application to the EU COMETT program was successful (DECAL — Methodology of Open Learning and Distance Education in Technology and Engineering). This was a start for cooperation with the Open University in UK, Aarhus University, Aalborg University, University of Trondheim and several companies. This lead to further cooperation with Romania and Moldova, and during the latest year initial cooperation within the Baltech University Consortium in Science and Technology with Kaunas Technical University in Lithuania and Riga Technical University. The award "KTH Internationalization Price" 1990 lead to several visits to and initiatives from Stig Hagström at Stanford University and Alfred Bork at University of California at Irvine.

When I 1992 during a conference in Portsmouth, UK was invited to join the program committee of CAEE'93 (Computer Aided Engineering Education) to be held in Bucharest I could not resist. I had very limited experiences of visits to Eastern Europe and of Bucharest I only had film memories from the 2nd World War and I was fascinated of getting a chance to see a new world. I was overwhelmed by the Romanian hospitality, but later I came to learn about the professional sides which were not so easy to handle (color of stamps, signatures on papers, a sometimes somewhat strange concept of research, authoritarian education systems, little understanding of financing aspects, frequent sliding back to plan economy thinking, expectations of functioning as a travel agency for visitors to Sweden *etc.*). In any case with some troubles I was successful with the help of money and practical assistance from the Open Society Institute in New York belonging to the family of Soros' foundations in implementing VSAT Internet connectivity for Bucharest, Cluj-Napoca, Timisoara, and Iasi and Chisinau (in Moldova) from 1995-96 [8,9]. Later an extensive

project with University of Craiova. And now when this is written I have just started together with Mihai Nicolescu (professor of Production Engineering at KTH but of Romanian origin) a project with the Institute of Solid Mechanics of the Romanian Academy in Bucharest on *Active and Passive Damping with Applications in Manufacturing* on how to avoid shaping defects through vibrations in manufacturing machines. It involves theoretical studies, modelling and analyzing methodologies, and practical applications of damping materials, such as viscoelastic and piezoelectrical ones to increase dynamic stability. Besides the still somewhat difficult psychological pressure in cooperation with Romania, I think the country is developing to the better. It is a tragedy that the situation is much worse in Moldova.

Hopefully the cooperation with the Baltic states will not be limited to distance education programs but also to consider the needed change to introduce more activating education methods through a constructivistic approach with problem oriented projects in groups. One has to consider the rôle of engineering education for the society, the tools and methods used in design, construction and production and also the change of earlier knowledge (especially in mathematics and physics) of the student population entering the university. Then it might not be necessary (nor possible) to cover all parts of classical mechanics but instead to train the use of advanced mathematical tools for *e.g.* transformations of kinematic quantities for multibody systems as interconnected rigid bodies in robots and for solving dynamical equations for such systems which has been developed by Lesser [10] with the use of a computer algebra programs as Maple. Such packages can be used as black boxes, especially if only kinematical problems are considered. The algebra is quite cumbersome both from the notation point of view (requiring notation for which frame of reference which is in use) and from the algebraic rules in use (for triads and dyads; scalar, vector and dyadic products; rotation matrices). Dankowicz [11] has succeeded to simply this somewhat for kinematical problems, and I had hoped it should be possible in the spirit of Per to find a formalism still more transparent, but it has not been possible. Might it be that Per was the man who would have found a formalism giving such a pedagogical simplification.

Final remarks

Noone more than I myself knows how much Per's way of attacking scientific problems and of creating a creative and inspiring environment, his approach to education and his international activities meant to me. In fact it formed many of my thoughts, actions and activities. And I am very thankful to him that I still have a community of friends in Uppsala and in other parts of the

world from that time. It is a pity that besides Per some of them are not with us any more.

It has was a honor to me know Per.

5 Acknowledgments

I want to express my gratitude to John R. Sabin for giving me the opportunity to write this paper and in this way giving me the possibility to express how much Per meant to me.

6 References

1. J. Piaget, **The Construction of Reality in the Child**, Basic Books, New York, NY, (1954).

2. J.C. Bringuier, **Conversations with Jean Piaget**, The University of Chicago Press (1980).

3. S. Papert, **The Children's Machine; Rethinking School in the Age of the Computer**, Harvester Wheatsheaf, Hetfordshire (1993).

4. S. Papert, **The Connected Family - Bridging the Digital Generation Gap**, Atlanta, Georgia: Longstreet Press (1996). Companion CD-ROM and web site http://www.connectedFamily.com/.

5. L.S. Vygotsky, **Thought and Language**, MIT Press, Cambridge, Ma. (1962).

6. G. Karlsson, in **UNESCO-SEFI International Symposium on Innovative Technological Education**, Université Paris-Sud XI, Orsay, p. 120 (1987).

7. K. Askeland (ed.), **Dokument från Nordisk konferens för datorstödd utbildning i universitets- och högskoleutbildningen**, Inst för mekanik, KTH, Stockholm (1988), ISBN 91-7170-937-1.

8. G. Karlsson, in **Proceedings 3rd International Conference on Computer Aided Engineering Education**, J. Breza, D. Donoval, R. Redhammer, editors, FEI STU, Bratislava, Slovakia, p.329 (1995).

9. R. Dumbraveanu, Flash Informatique, Sommaire du numéro spécial du 3 septembre 1996 - Terra informatica, sawww.epfl.ch/SIC/SA/publications/FI96/fi-sp-96/sp-96-page18.html.

10. M. Lesser, **The Analysis of Complex Nonlinear Mechanical Systems. A Computer Algebra Approach**, World Scientif Publishing Co. (1995), ISBN 981-02-2209-2.

11. H.J. Dankowicz, **Mechanics for Computer Scientists — Language, Method, Perspective**, Harry Dankowicz, Dept. of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksbur, Virgina 24061 (2000).

Partitioning technique in electron transfer and excitation energy transfer

Sven Larsson

Department of Physical Chemistry
Chalmers University of Technology
S-41296 Göteborg, Sweden

Abstract

The partitioning technique of Löwdin is used to derive practical and useful models for the mediation properties of matter in electron transfer (ET) and excitation energy transfer (EET). Partitioning technique is first used on the many-electron level for EET, where direct as well as mediated Förster and Dexter coupling is treated. Corresponding equations for ET, referred to as "superexchange" models, are derived. ET is also treated on the one-electron level, where the connection between ET capability and structure is discussed. A practically useful procedure to calculate the electronic factor on a PC screen in large systems is described. Finally the usefulness of models derived from partitioning technique is critically assessed.

Contents

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

1. Introduction

A number of phenomena in physics, chemistry, and biology [1] are related to tunnelling of nucleons, nuclei, protons, or electrons. The tunnelling model of the text-book was invented by Gamow and first used to explain radioactive decay [2]. The nuclear particle is confined to a region of space by a barrier whose potential energy V is greater than its kinetic energy E. The barrier is thus impenetrable in classical mechanics. In wave mechanics, however, where the wave function is required to have a derivative at the border of the energetically forbidden region, the probability for penetration is not zero but decreases exponentially with distance R. The probability to pass a barrier of width R and constant height V-E is is given by:

$$P = C\exp\left[-\sqrt{2(V-E)/27.21}\cdot 2(R-R_0)\cdot 1.89\right] = P_0\exp\left[-\beta(R-R_0)\right] \qquad (1)$$

where P_0 is the probability at distance R_0. V and E (V>E) are expressed in eV and R and R_0 in Å. For V-E=2 eV we obtain β=1.45. For example in the case of electron transfer (ET) in proteins [3,4], $\beta \approx 0.9\text{-}1.4$, depending on the detailed structure. β also depends, of course, on where R is measured (for example edge-to-edge or center-to-center in the case of π systems). This rather slow decrease of probability with distance compared to empty space permits biological tunnelling over some 15-20 Å at average time constants in the millisecond range.

The theory to be derived here covers a great number of cases where "particles" pass a forbidden region, for example excitation energy transfer (EET), of great importance in the antenna systems of photosynthesis. The decrease of probability with distance in the case of singlet EET is weaker than exponential [5]. Another case is transfer of vibrational energy through a large molecule. Naturally, the more pronounced the wave nature of the transferred energy, the less surprising is it that the "particle" passes through a forbidden region, since tunnelling is essentially a wave resonance phenomenon.

It is difficult theoretically to calculate a tunnelling barrier of Gamow type (not the same as activation barrier) from first principles. This paper summarizes methods that employ molecular orbital (MO) approaches to calculate the "electronic factor" which accounts for the transmittance in the forbidden region. The electronic factor will be structure dependent and there is no reason why it should be isotropic in space.

The rate of a given reaction depends on the thermal activation conditions of the particle in donor and acceptor, factors which are accounted for in the Marcus model [6,7] or models where the vibrational wave functions are included [8-10]. The reaction rate is derived in rather much the same way as for ordinary chemical reactions, using the concept of potential energy surfaces (PES's) [6]. The electronic factor is introduced either as a matrix element H_{12} or as an

energy gap Δ at an avoided crossing between two PES's [11]. For several reasons it is simpler to use an "effective" coupling $H_{12}=\Delta/2$.

As far as I know he first attempt to calculate an electronic factor using molecular orbital (MO) methods was done by Halpern and Orgel [12]. Soon afterwards McConnell [13] treated a system consisting of identical aromatic groups as donor and acceptor, connected by a bridge consisting of a saturated alkane chain (fig.1). To describe ET between the two aromatic systems, he assumed the presence of carbon 3d orbitals to mediate the coupling. An important conclusion is that if the interaction between the local orbitals is small, the coupling should decrease exponentially, in agreement with the Gamow model [2].

The McConnell model did not answer the question about the ET capability properties of the medium, since he assumed that transfer was promoted by some rather mysterious carbon 3d orbitals. The first attempt to calculate the rate of an ET reaction, using established quantum chemical methods, was done by Newton [14]. Newton assumed open space between the metal ions and it was still unclear whether the medium between the metal ions enhances the tunnelling rate. Before 1980 the little that was known on the connection between electron structure and mediation properties was actually discussed in connection with X-ray photo-electron spectra (XPS). Hoffman [15] and Heilbronner [16] introduced "through space" and "through bond" interactions. Symmetrically placed lone pair electrons correspond to a double peak in XPS, indicating interactions. The energy splitting between the components, equivalent to the electronic factor, is the result of superposition of through space and through bond interaction [15,16].

In 1980 I used partitioning technique [17,18] to derive an equation for the electronic factor in terms of MO theory, including the medium between donor and acceptor [19]. My knowledge of partitioning technique goes back to lectures on perturbation theory given by Löwdin in April 1964. These lectures reiterated the already published partitioning technique [17,18] and continued to the subject of wave and reaction operators of the tenth paper in the series "Studies in perturbation theory" [20]. To save time, Löwdin used to combine lecturing with research for his next paper, and he did so this time too. Probably he wrote down his new discoveries immediately after the lecture, while we, his students, enjoyed "morgonkaffe och wienerbröd" in the soft chairs at Ofvandahls Hofkonditori, sometimes discussing what use there might be of this formidable theory. Obviously partitioning technique is more general than perturbation theory since it can treat the case of near degeneracy in a pedagogically attractive way, but quite few applications were at hand in 1964.

In long distance ET, the presence of a wave function in the medium between donor and acceptor is the only means for donor and acceptor to communicate. The nature of this connection is expected to influence the reaction rates for ET or EET between the subsystems. Partitioning technique together with the Marcus-Hush model [6,7] may be viewed as an adaptation to practical chemistry of a full quantum mechanical treatment [21], where nuclei and electrons are treated as equal partners. In particular the influence on ET from the medium between the redox centres is formalized.

The present paper provides a summary of many rather different applications of partitioning technique in the field of ET and EET since 1980. Particularly the symmetric case will be treated here, since this case is simple but still sufficiently general to describe applications of partitioning technique. Two things distinguish ET reactions from other chemical reactions. (1) The reaction path consists usually of rather small motions of the nuclei away from the equilibrium positions. Hence the potential energy surface (PES) is well approximated by harmonic potential wells (parabolas), for which the force constants may be known, except possibly in the intersection region, which is an "avoided crossing". (2) At the avoided crossing, corrections to the Born-Oppenheimer-approximation may become important. The latter may be calculated [22], but a far more used way has been the Landau-Zener approximation [23,24]. The nuclei move classically and the electrons time dependently. The probability for jump between two PES's is given as a function of the energy difference between PES at the avoided crossing [11,25]. At small energy differences the rate of ET is reduced. This happens particularly if the distance between donor and acceptor is large. For example in the case of biological ET between distant metal centers, the avoided crossing energy splitting is extremely small, but still sufficiently large to accomplish ET in about 10^{-6} s.

In the case of two-electron transfer the mediating interaction is of a completely different type [26-28]. The direct interaction between the wave functions, corresponding to two electrons at either of adjacent sites, is very small even if the sites are close. An effective mediating wave function is one where only one of the two electrons has transferred to the acceptor. In the case of superconductivity this means that an antiferromagnetic wave function may be mediator in electron pair transfer [27]. There is increasing evidence that this is actually the case. A small modification of a superconducting system often turns it into an antiferromagnetic system, which just shows that the mediator state may become the ground state.

2. Partitioning of Hamiltonian matrix

The configuration interaction (CI) method is applied and the eigenvalue problem written as [19]:

$$\begin{pmatrix} H_{dd} & H_{da} & H_{db} \\ H_{ad} & H_{aa} & H_{ab} \\ H_{bd} & H_{ba} & H_{bb} \end{pmatrix} \begin{pmatrix} C_d \\ C_a \\ C_b \end{pmatrix} = \begin{pmatrix} E & 0 & 0 \\ 0 & E & 0 \\ 0 & 0 & E \end{pmatrix} \begin{pmatrix} C_d \\ C_a \\ C_b \end{pmatrix} \tag{2}$$

H_{xy} and C_x are matrices in general. Index d, a, and b denote donor, acceptor, and bridge, respectively. One example may be an interaction leading to transfer of an excitation from donor to acceptor. The rate depends on a CI matrix element between an excitation on the donor and an excitation on the acceptor [5]. The b states are excitations on the bridge, higher in energy than the d and a states. In the ET case, the interaction is between two states where the electron is either on the donor or on the acceptor. The b states are states higher in energy with the electron located on the bridge. In both cases there is very little mixing into the ground state of the whole system, as is typical for tunnelling. It is not always clear whether all configurations can be strictly partitioned the ways mentioned, but we assume for the moment that this can be done. The d and a configurations involved may be referred to as diabatic states.

C_b is calculated from the third equation of the eigenvalue problem, eq.(2), with the result:

$$C_b = -(H_{bb} - E)^{-1}(H_{bd}C_d + H_{ba}C_a). \tag{3}$$

Eq.(3) is inserted in the first and second equation to obtain:

$$\overline{H}\begin{pmatrix} C_d \\ C_a \end{pmatrix} = \begin{pmatrix} H_{dd} - H_{db}(H_{bb} - E)^{-1}H_{bd} & H_{da} - H_{db}(H_{bb} - E)^{-1}H_{ba} \\ H_{ad} - H_{ab}(H_{bb} - E)^{-1}H_{bd} & H_{aa} - H_{ab}(H_{bb} - E)^{-1}H_{ba} \end{pmatrix} \begin{pmatrix} C_d \\ C_a \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix} \begin{pmatrix} C_d \\ C_a \end{pmatrix} \tag{4}$$

If the Hamiltonian is a one-electron Hamiltonian, for example the Fock operator, the partitioning is done by basis functions, since the latter are usually centered on the atomic nuclei, which belong to donor (d), bridge (b) or acceptor (a). In the Hartree-Fock case, the total wave function is a Slater determinant. There may be problems with symmetry breaking in the symmetric case. CI that includes the two localized solutions can solve this problem [29-31]. The problem is that the Hartree-Fock method gives energy advantage to a localized state, which holds true also in the unsymmetric case.

In the Marcus model the important part of the system is donor and acceptor and the process takes place on the potential energy surface of the ground state of the total system, while the first excited state corresponds to the remaining, upper parts of the Marcus' parabolas. If eq.(4) is solved, a number of excited states correspond to excitations of the bridge. It is interesting to

solve eq.(4) at the activated, avoided crossing point, where the ground state and first, excited state are very close. Eq.(4) is in principle solved by iterations [17,18], but to get the two lowest solutions, it likely suffices with one iteration. This is possible since the (H_{bb}-E)-matrix, that determines the distance between the two lowest states at the avoided crossing is not strongly depending on the value of E in eq.(4). Hence one may choose an average of the upper and lower state. The latter equation is an ordinary matrix eigenvalue problem for the donor-acceptor interaction. The diagonal correction terms are overshadowed by fluctuations in the environment and will not be important. The non-diagonal correction terms involving the bridge, on the other hand, may become much more important than the direct interaction terms of H_{ad}.

3. Marcus model

The idea of a coupling between nuclear and electronic motion goes back to Landau. In physics the Holstein model for small polaron motion may be considered the most succinct formulation of electron conductivity in a one-dimensional metal or a semi-conductor [32]. To simplify, Holstein considers each site to be a diatomic molecule with variable interatomic distance R_i. The equilibrium distances are R_i^0, R_i^+ and R_i^- for neutral, positive ion and negative ion, respectively. The coupling between the sites is assumed to be small, implying that ionic charges or excitations are localized on single sites. However, if the ratio reorganization energy (λ) to electronic factor (Δ) decreases the system may become delocalized [33].

The Holstein model was preceded by the Pekar model [34] and in chemistry by the Marcus model [6]. In chemistry donor-acceptor systems are more frequent objects of study than conducting wires but the coupling between electronic and nuclear motion of similar nature. For example if the coupling is large a small nuclear displacement is sufficient to change the wave function much and in a way which corresponds to ET or EET. We use the effective, many-electron Hamiltonian $\overline{H}$ of eq.(4) and assume that it is solved for donor and acceptor, giving the energies H_{aa} and H_{dd}, respectively. We use the new nuclear coordinates:

$$Q_1 = (R_1 - R_2)/\sqrt{2}$$
$$Q_2 = (R_1 + R_2)/\sqrt{2}\,. \qquad (5)$$

The Marcus diabatic PES's may be represented as [6,22,34]:

$$H_{aa} = \frac{k}{2}(Q_1 + Q_{10})^2 + \frac{k}{2}(Q_2 + Q_{20})^2$$

$$H_{dd} = \frac{k}{2}(Q_1 - Q_{10})^2 + \frac{k}{2}(Q_2 - Q_{20})^2 \qquad (6)$$

The coordinate Q_1 corresponds to a bond or angle in donor or acceptor, which is sensitive to the presence of the electron on either donor or acceptor. The coordinate Q_2 corresponds to the non-important, perpendicular coordinate, not coupled to ET. A good example is ET between two metal complexes in different oxidation states, which is the problem the Marcus model was originally designed to solve. The lower oxidation state of the metal ion has larger metal-ligand bond lengths, since the additional electron, compared to the oxidized state, occupies an anti-bonding MO and since the ionic bonding also decreases due to increased screening of the metal atom nucleus. Q_1 represents simultaneous shortening of the bonds at the reduced metal ion and lengthening of the bonds at the oxidized metal ion. Q_2 on the other hand represents simultaneous shortening (or lengthening) at the two centers. The equilibrium positions are $-Q_{10}$ and $-Q_{20}$ in the acceptor state and Q_{10} and Q_{20} in the donor state. We consider the reaction path where $Q_2=0$ and $Q_{20}=0$, and set $Q=Q_1$ and $Q_0=Q_{10}$. The parabolic expressions used in eq.(6) correspond, then, to the Marcus model:

$$H_{aa} = \frac{k}{2}(Q + Q_0)^2$$

$$H_{dd} = \frac{k}{2}(Q - Q_0)^2 \qquad (7)$$

The reorganization energy is [35]:

$$\lambda = H_{aa}(Q_0, 0) = 2kQ_0^2 \qquad (8)$$

and the distance between the minima along the reaction coordinate:

$$2Q_0 = \sqrt{2\lambda / k} \qquad (9)$$

We may write the matrix eigenvalue problem [eq.(4)] as:

$$\begin{pmatrix} H_{aa} - E & H_{ad} \\ H_{ad} & H_{dd} - E \end{pmatrix} \begin{pmatrix} \cos\vartheta \\ \sin\vartheta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad (10)$$

where H_{ad} includes direct as well as indirect matrix elements. The energy eigenvalues

$$E_{\pm} = \frac{H_{aa} + H_{dd}}{2} \pm \sqrt{\left(\frac{H_{aa} - H_{dd}}{2}\right)^2 + H_{ad}^2} = \frac{k}{2}\left(Q^2 + Q_0^2\right)^2 \pm \left(k^2 Q_0^2 Q^2 + H_{ad}^2\right)^{1/2} \qquad (11)$$

define the PES. $\vartheta=45^0$ ($Q=0$) defines the top of the barrier, where

$$E_{\pm} = \frac{k}{2}Q_0^2 \pm H_{ad}; \qquad \Delta = 2|H_{ad}| \tag{12}$$

The height of the barier is:

$$E_a = \frac{(\Delta - \lambda)^2}{4\lambda} \tag{13}$$

The eigenvectors are positive and negative combinations of Ψ_d and Ψ_a, respectively. Ψ_d (Ψ_a) is a wave function where the mobile electron is at the donor (acceptor). In a time dependent theory, the wave function is set equal to Ψ_d at t=0. The electronic wave packet moves across the barrier to the acceptor. Landau-Zener theory [23,24] provides the probability for crossing to the upper surface at the avoided crossing:

$$P = \exp\left(-\pi\Delta \Big/ 2\hbar \left|\frac{d}{dt}(H_{11} - H_{22})\right|\right) \tag{14}$$

Δ (>0) is the electronic factor. 1-P is the probability for continuing on the lower PES, which corresponds to ET. If the barrier disappears, the Landau-Zener model should not be used and it may be necessary to include the nuclear coordinates in a wave packet model.

The success of the Marcus model is connected to the consistent description of self-exchange reactions and later to ET reactions with non-zero free energy. Using the easily measured free energy of reaction ($-\Delta G^{\Theta}$) in the PES diagram, gives the Arrhenius rate:

$$k = A\exp\left(-\frac{E_a}{k_B T}\right); \text{ where } E_a = \frac{\lambda}{4}\left(1 + \frac{\Delta G^0}{\lambda}\right)^2 \tag{15}$$

If we plot log k as a function of $-\Delta G_0$ we obtain a parabolic curve. In particular the part where E_a increases for $-\Delta G^0 > \lambda$ is called the inverted region [36].

Particularly in the case of protein ET, it became clear that there must be a factor in eq.(15) which accounts for the dependence on distance between the two subsystems, for example two metal centres exchanging electrons in proteins. For this reason eq.(15) is supplemented by the factor κ:

$$k = \nu_n \kappa \exp[-c/k_B T] \tag{16}$$

ν_n is the average frequency for nuclear fluctuation which take the system across the barrier. κ can be calculated with the help of Sutin's equation [11,25]:

$$\kappa = \frac{2[1 - \exp(-\nu_{el}/2\nu_n)]}{2 - \exp(-\nu_{el}/2\nu_n)} \tag{17}$$

where

$$\nu_{el} = \frac{2\pi}{\hbar}\frac{\Delta^2}{2}\left(\frac{1}{4\pi\lambda k_B T}\right)^{1/2} \tag{18}$$

For $\nu_{el} \ll \nu_n$, usually due to a large distance between D and A, we obtain that $\kappa = \nu_{el} / \nu_n$ and in this case we may rewrite eq.(16) as:

$$k = \nu_{el} \exp[-c / k_B T]. \tag{19}$$

From eqs.(18) and (19) follows that for small electronic coupling the rate is proportional to Δ^2. This is also the case in the theories of Jortner et al., [10,37-39] but there are exceptions when a linear dependence is more appropriate [40].

4. Electronic factor

For simplicity we will continue to assume that the reaction path is mapped out by a single coordinate Q. In the Landau-Zener approximation [23,24] we need the distance between the upper and lower PES. This energy may be calculated variationally as an excitation energy at the top of the barrier when the symmetric system has full symmetry. In the case of an asymmetric system, the transition state occurs when the relevant MO's are localized half on donor and half on acceptor.

We first consider CI as described in section 2 above. In particular we are interested in the admixture of excited states of the bridge at the avoided crossing. We have diagonalized donor, acceptor and bridge matrices separately. To simplify the notation we include the donor and acceptor part of the matrix in H_0. Let the unitary matrix U_b diagonalize H_b and rewrite eq.(2) as:

$$\begin{pmatrix} 1 & 0 \\ 0 & U_b^\dagger \end{pmatrix}\begin{pmatrix} H_{00} & H_{0b} \\ H_{b0} & H_{bb} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & U_b \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & U_b^\dagger \end{pmatrix}\begin{pmatrix} C_0 \\ C_b \end{pmatrix} = E\begin{pmatrix} 1 & 0 \\ 0 & U_b^\dagger \end{pmatrix}\begin{pmatrix} C_0 \\ C_b \end{pmatrix} \tag{20}$$

Eq.(15) is equivalent to:

$$\begin{pmatrix} H_0 & H_{0b}U_b \\ U_b^\dagger H_{b0} & U_b^\dagger H_{bb} U_b \end{pmatrix}\begin{pmatrix} C_0 \\ U_b^\dagger C_b \end{pmatrix} = E\begin{pmatrix} C_0 \\ U_b^\dagger C_b \end{pmatrix} \tag{21}$$

After diagonalization of the bridge matrix we may make the following replacements in eq.(4):

$$H_{db} \rightarrow H_{db}U_b = \eta; \quad H_{ab} \rightarrow H_{ab}U_b = \vartheta$$
$$H_{bb} \rightarrow U_b^\dagger H_{bb} U_b \tag{22}$$

The off-diagonal matrix element in eq.(4) may be written:

$$\overline{H}_{da} = H_{da} - H_{db}U_b\left(U_b^{\dagger}H_{bb}U_b - E\right)^{-1}U_b^{\dagger}H_{ba} = H_{da} - \eta\left(U_b^{\dagger}H_{bb}U_b - E\right)^{-1}\vartheta^{\dagger} \qquad (23)$$

$U_b^{\dagger}H_{bb}U_b - E$ is a diagonal matrix. The matrices η and ϑ refer to interactions between donor and acceptor and the bridge. If we assume that donor and acceptor also have been diagonalized and the ground state and first excited state chosen, we may write down the matrix element for interaction between state i on the donor and state j on the acceptor:

$$\overline{H}_{ad} = H_{ad} - \sum_k \frac{\eta_{aik}\vartheta_{kd}^{\dagger}}{(E - E_k)} \qquad (24)$$

An advantage with this equation is that it may be carried out on donor, acceptor, and bridge separately. Alternatively experimental data may be used.

5. Excitation energy transfer (EET)

If we want to study the probability for an excitation to move from donor to acceptor, the simplest possible approximation gives the direct matrix element (Förster coupling [5]) as [41]:

$$H_{12} = \left\langle {}^1\Phi_i^a \middle| H \middle| {}^1\Phi_j^b \right\rangle = 2\left(\phi_a\phi_i|\phi_j\phi_b\right) - \left(\phi_a\phi_b|\phi_i\phi_j\right) \qquad (25)$$

where H_{12} is either a matrix element between donor and acceptor or between bridge and donor or acceptor. The Mulliken notation is used in the right member:

$$\left(\phi_a\phi_i|\phi_j\phi_b\right) = \int\phi_a^*(\underline{r}_1)\phi_i^*(\underline{r}_2)\frac{1}{r_{12}}\phi_j(\underline{r}_2)\phi_b(\underline{r}_2)dv_1dv_2 \qquad (26)$$

If we assume that acceptor and donor have low lying excited states the probability for singlet EET is proportional to H_{da}^2. The first part of the right member of eq.(25) is the Förster part [5] and the second part the Dexter part [42]. For increasing distance R between donor and acceptor the Förster part [eq.(26)] converges to a dipole interaction between two transition moments, tending to zero as a dipole – dipole interaction, i e as $1/R^3$. The Dexter part tends to zero exponentially with R, since $\phi_a\cdot\phi_b$ and $\phi_i\cdot\phi_j$ both are exponentially decreasing with distance R. The Förster part is thus dominating over the Dexter part at a large enough distance. This *direct* Förster part leads to slow $1/R^6$ decrease of the rate with R and permits singlet EET over large distances.

If an organic molecule, for example an aromatic molecule with a π system and with low-lying excited states is situated between donor and acceptor, there may be a contribution to the coupling via this bridge according to the summation in eq.(24) which we may term

Förster-Förster coupling, since η couples donor and bridge via a Förster term and 9 couples bridge to acceptor via another Förster term [43,44]. Under particular circumstances the Förster-Förster coupling may be larger than the direct coupling.

In the case of triplet EET, the direct matrix element is

$$H_{da} = \left\langle {}^3\Phi_i^a \left| H \right| {}^3\Phi_j^b \right\rangle = -\left(\phi_a \phi_b \middle| \phi_i \phi_j\right) \tag{27}$$

Since only the exponentially decreasing Dexter matrix element appears, triplet excitations cannot be transferred over large distances. The most important means of triplet EET is sensitation [42,45], i e direct interaction between the donor and acceptor chromophores. Indirect coupling via a bridge in many cases gives a larger contribution than via a solvent. Since the coupling matrix element in the triplet case decreases exponentially with distance, one expects the same rules as for ET, possibly with a larger β (see below).

6. Superexchange

Eq.(24) is seldom used in ET since it is often hard to partition the configurations according to donor, acceptor, and bridge as would be necessary. Furthermore in practical calculations one has to use the same nuclear geometry for all bridge and donor and acceptor states. In that case the bridge states do not have the meaning of actual states of the bridge, since in the latter the geometry is often much modified. One important exception is the superexchange model, where normally the lowest state of the bridge is included.

The superexchange model has its origin in the description of magnetic interactions between paramagnetic or antiferromagnetic sites in crystals by Kramers [46] and later by P.W. Anderson [47]. In these descriptions the nuclear motion is not considered as in the Marcus model and the electronic coupling of a completely different nature than the coupling in the case of ET [48]. There may be a number of intermediate states of the type D^+B^-A interacting with D^*BA, but we are interested in particularly one of them at a low energy. The coupling is expressed as [49-51]:

$$V_{super} = \frac{V_{db} V_{ba}}{\Delta E_{db}} \tag{28}$$

The advantage with this model is that one may study in great detail what happens if the energy of the lowest state of B^- becomes so low that the electron is delocalized to the bridge and temporally occupying an orbital on B [50,51].

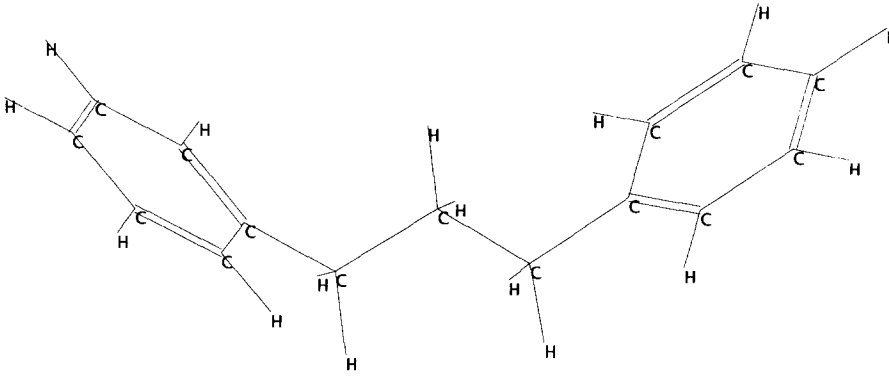


Figure 1. Typical donor-bridge-acceptor system.

A superexchange model including many CI states is possible by simple adaption of a calculational model such as CNDO/S to the excited states in a case of photo-induced ET [52]. The advantage is that the treatment is more general than the simplified superexchange model of eq.(28).

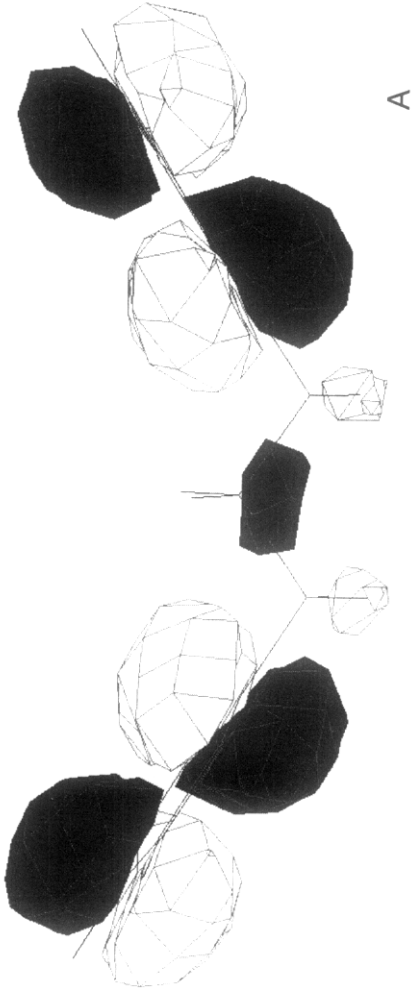
In the case of the primary charge separation in bacterial photosynthesis, it was for a long time believed that the lowest state of the negative ion of an accessory bacteriochlorophyll, in between the special pair, that is excited by the antenna system, and a bacteriopheophytin molecule, that is the end station of the electron in the primary charge separation, acts as a superexchange state. Calculations for the whole system revealed, however, that superexchange via the accessory chlorophyll is much less important than via a side group, which is placed in a more strategic position [54,55]. Both contributions are too small to explain the fast ET measured. Experiments [56] also show that the negative ion of the accessory bacteriochlorophyll is used as a real intermediate state. The mechanism for charge separation between the special pair and bacteriopheophytin is thus step-wise hopping rather than tunnelling.

7. Orbital interactions

A quite simple picture is obtained in the ET case if we go from many-particle theory to orbitals via Koopmans' theorem. This treatment is correct in the limit case when donor and acceptor exchange electrons using well-separated MO's on donor and acceptor. The simplest example is the case studied by McConnell [7] (fig.1). Two identical π systems are connected via a hydrocarbon chain, which acts as a bridge. ET is possible in an open shell system. We may assume that either the whole ET system is either a negative ion or a positive ion and that the corresponding neutral molecule has closed shells. In the former case the electron occupies π LUMO. In the latter case there is a "hole" in π HOMO.

In the case of ET, the reaction path starts with the electron in LUMO of the donor. Along the reaction path, the orbital shifts character, from donor to acceptor orbital. At the transition state LUMO and LUMO+1 for the full system are symmetric and antisymmetric combinations of the two local LUMO's (fig.2). Since for any of the two orbitals, the negative of the orbital energy approximates the electron affinity according to Koopmans' theorem, the following implication holds:

$$A_a = E^a_{N-1} - E_N = -\varepsilon_a \qquad \Rightarrow \qquad E^a_{N-1} - E^b_{N-1} = \varepsilon_b - \varepsilon_a \tag{29}$$


A

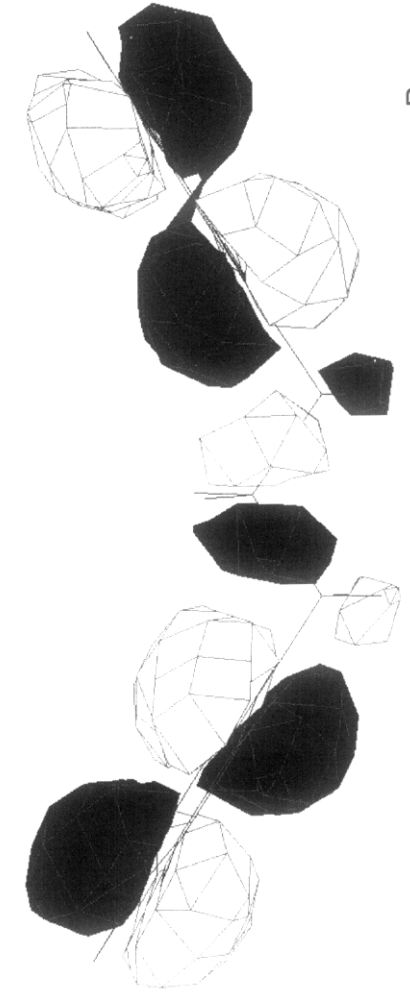


Figure 2. Symmetric (A) and antisymmetric (B) combination of LUMO orbitals in the system shown in figure 1.

The corresponding procedure is carried out in the case of hole transfer, now with use of the ionization energies of the symmetric and antisymmeric linear combination of HOMO's. From eq.(29) follows that the total energy difference at the avoided crossing, equal to the electronic factor, can be obtained as an orbital energy difference between two MO's which are sum and difference of the electron exchanging donor and acceptor MO's ϕ_d and ϕ_a, respectively.

$$(\phi_d + \phi_a)/\sqrt{2} \qquad \text{and} \qquad (\phi_d - \phi_a)/\sqrt{2} \tag{30}$$

with

$$\Delta = E^-_{N-1} - E^+_{N-1} = \varepsilon_+ - \varepsilon_- \tag{31}$$

Please notice that the orbital energies must be calculated for the system with one electron removed in the case of electron transfer and for the ET system with one electron added in the case of hole transfer. In both cases the orbital energies ε_+ and ε_- are calculated for the neutral, closed shell molecule.

In the asymmmetric case the 50/50 composition of eq.(30) happens at the transition state for ET. This point on the reaction surface is more complicated to find than in the symmetric case, particularly if Δ is small [56]. It can be found instructively on the screen using, for example, *HyperChem* [57]. The general procedure is to first find a sensitive bond length, or possibly a charge, which represents the solvent. Changing the bond length or moving the charge represents a move on the PES. When equal distribution is obtained as in eq.(30), the transition state has been reached and Δ can be read off as an orbital energy difference for the neutral system. Of course Δ may vary slightly with position on the transition state and depends slightly on the geometry difference between neutral system and ion. The correct way is to optimize the structure for the ion at the symmetric transition state and apply Koopmans' theorem and calculate the orbital energies for the neutral, closed shells system.

The influence of the intervening matter is formalized by use of eq.(4). For simplicity we include only ϕ_d and ϕ_a on donor and acceptor, respectively. The H_{bb} matrix is diagonalized. In the new basis, the interaction matrix element may be written as in eq.(23):

$$\overline{H}_{da} = H_{da} - H_{db}(H_{bb} - E)^{-1}H_{ba} = H_{da} - \eta\left(U_b^{\dagger}H_{bb}U_b - E\right)^{-1}\vartheta \tag{32}$$

H_{ad} contains the direct interaction which decreases exponentially with the distance R between donor and acceptor. It is reasonable to expect that the second term of the right member may give a larger contribution at a large enough distance between donor and

acceptor, since the direct term H_{da} has negligible matrix elements due to exponentially decreasing overlap with large β. We assume that the unitary transformation is chosen to diagonalize H_{bb}. The eigenvalues of H_{bb} are called b_v. We obtain:

$$(U_b^{\dagger}H_{bb}U_b - E)^{-1} = \begin{pmatrix} b_1 - E & 0 & . \\ 0 & b_2 - E & . \\ . & . & . \end{pmatrix}^{-1} = \begin{pmatrix} (b_1 - E)^{-1} & 0 & . \\ 0 & (b_2 - E)^{-1} & . \\ . & . & . \end{pmatrix} \qquad (33)$$

A common case is that the matrix element (=coupling, η_1) is non-zero between donor and bridge at a single atomic orbital which we give index 1. The acceptor is similarly in contact with the last orbital of the bridge, with index n (coupling=ϑ). Most matrices of eq.(27) are thus sparse and it is easy to derive the effective interaction matrix element [19]. The energy difference Δ between ϕ_d and ϕ_a may be written as:

$$\Delta = 2\eta_1\eta_2 \sum_i \frac{c_{1i}c_{ni}}{b_i - E} \qquad (34)$$

c_{1i} and c_{ni} are the coefficients bridge MO i at the contact atomic orbitals 1 and n, respectively. The summation runs over the MO's of the bridge. E is the energy of the donor-acceptor MO. Eq.(34) has also been called the Green function method [58]. Similar equations have also been given in references [59-71].

8. Molecular wire

In many cases the transferring electron has its energy E in the quite wide HOMO-LUMO gap. The absolute values of the diagonal matrix elements of H_{bb}–E are likely to be much larger than the non-diagonal matrix elements. One may use eq.(34) for the case of tunnelling through a wire and let E be the energy of the electron being transmitted [43,72-75]. The results of [43] shows that a straight π wire is better than a σ wire. A real wire should be able to accept electrons, of course, and transfer them in its orbitals. The detailed solution of that problem is very extensive and includes the so called Coulomb blockade. The latter problem means simply that the number of electrons on the bridge must be limited and there must be redox centres to transmit the electrons, or the bridge should have delocalized electrons. In each step ET takes place.

A simplified model may still be developed from eq.(34). As the bias voltage E is increased, one of the denominators tends to zero and consequently Δ to infinity. The model does not hold close to an orbital energy of the bridge and the molecular bridge must be

recalculated with an increased number of electrons. The new bridge orbital HOMO-LUMO gap is bracketing E, since another electron is present. The new electron is soon transferred to the other electrode. There will be a stepwise increase of the current. Further increase of E leads to other possible reductions of the bridge followed by oxidation, leading to a real current through the bridge. Since the Hartree-Fock orbital energies are strongly sensitive to the number of electrons, the Coulomb blockade will be visible in the IV diagram.

9. Pathway model

Additional insight may then be obtained if instead of diagonalizing the matrix H_{bb}–E we write:

$$H_{bb} - E = D + B \tag{35}$$

where D is a diagonal matrix and B an interaction matrix with zeroes in the diagonal. By repeated use of the matrix identity [18,19]:

$$(D+B)^{-1}=D^{-1}-(D+B)^{-1}BD^{-1} \tag{36}$$

we obtain, as shown by Löwdin [18]:

$$(D+B)^{-1}=D^{-1}-D^{-1}BD^{-1}+D^{-1}BD^{-1}BD^{-1}-\ldots \tag{37}$$

In eq.(32) the first term in the donor-acceptor interaction is the direct matrix H_{ad}. A matrix element $H_{\mu\nu}$ is the resonance integral between MO μ on the donor and MO ν on the acceptor.

The next contribution to the matrix element $\overline{H}_{\mu\nu}$ of eq.(32), is from D^{-1} of eq.(37):

$$-H_{\mu i}D_{ii}^{-1}H_{i\nu} = -\sum_i \frac{H_{\mu i}H_{i\nu}}{E_i - E} \tag{38}$$

The terms in this matrix element correspond to paths from μ on the donor, to atomic orbital in the medium, and from i in the medium to ν on the acceptor.

The following terms in the expression for the coupling between donor and acceptor is obtained by taking the higher terms in the eq.(37). The term $D^{-1}BD^{-1}$ gives for example:

$$H_{db}D^{-1}BD^{-1}H_{ba} = \frac{H_{\mu i}H_{ij}H_{j\nu}}{(E_i - E)(E_j - E)} \tag{39}$$

and finally:

$$\overline{H}_{da} = H_{\mu\nu} - \sum_i \frac{H_{\mu i}H_{i\nu}}{E_i - E} + \sum_{i,j} \frac{H_{\mu i}H_{ij}H_{j\nu}}{(E_i - E)(E_j - E)} - \ldots \tag{40}$$

Finally we obtain pathways connecting μ and ν in all possible ways. In the case when few sites are involved in the medium, the large distances between the sites make all terms

small. On the other hand, when many sites are involved, the great number of factors in each term leads also to a vanishing contribution. The large contributions are obtained when the connection pathway through the medium is the shortest possible. Thus one expects the shortest pathways between μ and ν to give the largest contributions.

This theory provides some sort of justification of the very simplified methods invented by Beratan and Onuchich that appears to have been very useful in understanding structural dependencies in proteins [66,67].

10. Summary. Why is partitioning technique useful?

It is possible, of course, to calculate the electronic factor without any reference to partitioning technique. Eq.(34) is useful in the case of wires, as pointed out above, but generally partitioning technique is useful only to the extent that closed expressions can be derived from it, which shows the connection between molecular structure and electronic structure in simple equations. Thus it is easy to show that the exponential decrease is obtained in the case the bridge consists of a number of independent solvent molecules [64]. It is thus correct to consider the electron transfer phenomenon as a form of tunnelling in this case, as has been pointed out early by Davydov in a simplified model [71]. It is also simple to show that if the energy levels of the bridge come closer to the energy of the electron to be transferred as the length of the bridge is extended, the transfer capability does not decrease exponentially [65]. This problem was further discussed in references [68,69].

The partitioning technique also nicely shows why a bridge, consisting of a molecule or a solvent, is a better mediator than empty space. In empty space the direct matrix element H_{da} is the only coupling between donor and acceptor. This matrix element decreases with $\beta \approx 4$, which effectively forbids ET distances larger than 4-5 Å.

In the paper by Halpern and Orgel [12], the electronic coupling was calculated for a very simplified system with a two-electron bridge. The expression obtained was generalized and used for planar π bridges and it was found that the mobile bond order [76] is important for the transfer capability of the bridge. In fact the final equation has a clear resemblance to eq.(34). The proportionality to bond order applies only if severe approximations are made, however. A superexchange contribution was also noted. It is clear from the form of eq.(34), however, that the superexchange contribution is included already at the orbital

level and hence one should not include both. Unfortunately, since the Halpern and Orgel paper only treats π-electrons, it has unintentionally given the impression that only π-electrons are capable of mediating electrons, and this was the prevailing opinion in 1980.

The method of McConnell [13] assumes the presence of a set of virtual orbitals all with higher energy than the electron to be transferred, in the form of Hückel combinations of C 3d orbitals. This somewhat odd assumption leads to exponential decrease. Incorrectly one gets the impression that only unoccupied MO's are useful for ET. The virtual orbitals of the kind used by McConnell, are nowadays standard in any reasonable basis set for ab initio calculations, but it is fair to say that the electronic factor is not much dependent on whether they are included or not. Ordinary valence MO's of the bridge, occupied or unoccupied are of a much greater importance for ET. The McConnell method [13] is therefore mainly of historical interest.

Partitioning technique may also be used to treat EET [43,44]. It is easy to show that triplet transfer, rapidly decreasing exponentially with distance in empty space, decreases with a smaller β because of mediating properties of the bridge [41]. EET for singlets decreases slowly with distance, as $1/R^6$, and the transfer properties may also be improved further by a bridge in fortunate cases.

It is interesting to discuss to what extent the detailed study of ET systems has changed our views on ET and EET since 1980, not only what concerns partitioning technique but actual calculations. In 1980 it was shown that the XPS energy splitting between symmetrically placed MO's may be mediated by bonds [15,16]. Ingenious systems have since been synthesized to see how this interaction decreases with distance [61-63,77]. Miller and Closs made systems which demonstrate internal ET [77]. Theoretical calculations on the ab initio level showed clear agreement with these experimental results [62,63,78].

In proteins Dutton et al.[4] concluded by studying an example from photosynthetic reaction centres, that a protein acts by mediating ET in a fixed exponential way without much possibility for Nature to improve the ET properties in evolution, other than choosing the optimum distance between the active ET centres. It has been shown both experimentally and theoretically that π systems are better than σ systems in transmitting electrons, if the orientation of the former system is favourable. Thus the statement of Dutton et al. [4] could be rephrased by saying that π systems or other favourable ET systems cannot be incorporated in a favourable position in a protein, by evolutional

pressure. Gray, Winkler and others opposed this view [3]. It is quite possible, however, that the differences between different forms of protein structure are surprisingly small. On the other hand it is clear from the above theory that orbitals directly influence the directional and distance dependence. The results of Dutton et al. are partly explained by the fact that they studied well functioning system with rather small modifications between them. If the functioning half (L) of the bacterial photosynthetic system is compared to the unused half (M), it turns out that the functioning part has a favourably placed π system for ET [79,80]. This supports the idea that evolutional pressure can lead to structures that are favourable for a large electronic factor. In this sense partitioning technique is of great importance also in biophysics and biochemistry.

Acknowledgement: I thank the Swedish Natural Research Council for support.

References

1. B. Chance, D.C. DeVault, H. Frauenfelder, R.A. Marcus, J.R. Schrieffer, N. Sutin, eds. *Tunneling in Biological Systems*; (Academic Press, New York, 1979).
2. G. Gamow, *Z. Phys.* **43**, 204 (1928).
3. H.B. Gray, D.N. Beratan, J.N. Onuchic and J.R. Winkler, *Science* **258**, 1740-1741 (1992); H.B. Gray and J.R. Winkler, *Annu. Rev. Biochem.* **65**, 537-561 (1996); F.A. Tezcan, B.R. Crane, J.R. Winkler and H.B. Gray, *Proc. Natl. Acad. Science (USA)* **98**, 5002-5006 (2001).
4. C.C. Moser, J.M. Keske, K. Warncke, R.S. Farid, and P.L. Dutton, Nature **355**, 796-802 (1992); R.S. Farid, C.C. Moser, P.L. Dutton, *Curr. Opin. Struct. Biol.* **3**, 225 (1993); C.C. Page, C.C. Moser, X. Chen, and P.L. Dutton, *Nature* **402**, 47-52 (1999).
5. T. Förster, *Naturwiss.* 33,166 (1954); T. Förster in *Modern Quantum Chemistry*, Ed. O. Sininoğlu (Academic Press, 1965), pp 93-137.
6. R.A. Marcus, *J. Chem. Phys.* **24**, 966, 979 (1956); **43**, 679 (1965); *Discuss Faraday Soc.* **29**, 21 (1960); *Ann. Rev. Phys. Chem.* **15**, 155-163 (1964).
7. N.S. Hush, *Trans. Faraday Soc.* **57**, 155 (1961).
8. V.G. Levich, *Adv. Electrochem. Electrochem. Eng.* **4**, 249 (1966).
9. R.R. Dogonadze and A.A. Kornyshev, *Phys. Stat. Solidi* **53**, 439 (1972).
10. N. Kestner, J. Logan, and J. Jortner, J. Phys. Chem. **78**, 2148-2166 (1974).
11. N. Sutin, ref. 1, page 201; *Annu. Rev. Nucl. Sci.* **12**, 285 (1962).
12. J. Halpern and L.E. Orgel, *Discuss. Faraday Soc.* **29**, 32-41 (1960).
13. H.M. McConnell, *J. Chem. Phys.* **35**, 508-515 (1961).
14. M.D. Newton, *Int. J. Quant. Chem. S* **14**, 363-391 (1980).
15. R. Hoffmann, A. Imamura, and W.J. Hehre, *J. Am. Chem. Soc.* **90**, 1499 (1968); R. Hoffmann, *Accounts Chem. Res.* **4**, 1 (1971).
16. E. Heilbronner and K.A. Muszkat, *J. Am. Chem. Soc.* **92**, 3818 (1974); E. Heilbronner and J.P. Maier, *Helv. Chim. Acta* 57, 151 (1974).
17. P.-O. Löwdin, *J. Molecular Spectr.* **10**, 12-33 (1963).
18. P.-O. Löwdin, *J. Math. Phys.* **3**, 969-982 (1962).
19. S. Larsson, *J. Am. Chem. Soc.* **103**, 4034-4040 (1981).
20. P.-O. Löwdin, *Phys. Rev.* **139**, A357-A372 (1965).
21. E. Deumens, A. Diz, R. Longo, Y. Öhrn, *Rev. Modern Phys.* **66**, 917 (1994).
22. A. Klimkāns and S. Larsson, *Int. J. Quant. Chem.* **77**, 211-220 (2000).
23. L. Landau, *Phys. Z. Sow.* **2**, 46 (1932).

24. C. Zener, *Proc. Roy. Soc. A* **137**, 696 (1932).
25. B.S. Brunschwig, J. Logan, M.D. Newton, and N. Sutin, *J. Am. Chem. Soc.* **102**, 5798-5809 (1980); B. Brunschwig, C. Creutz, H. Macartney, T.-K. Sham, N. Sutin, *Faraday Discuss. Chem. Soc.* **74**, 113 (1982).
26. K. Prassides and P. Day, *J. Chem. Soc., Faraday Trans. 2* **80**, 85 (1984).
27. S. Larsson, *Chem. Phys. Letters* **157**, 403-408 (1989); S. Larsson, *Chem. Phys.* **236**, 135-150 (1998).
28. L.D. Zusman ans D.N. Beratan, *J. Phys. Chem. A* **101**, 4136 (1997); *J. Chem. Phys.* **105**, 165 (1996).
29. H.B. Broer-Braam, PhD Thesis, University of Groningen, The Netherlands (1981); R. Broer and W.C. Nieuwpoort, *Chem. Phys.* **54**, 291 (1981); *Theor. Chim. Acta* **73**, 405 (1988).
30. P.Å. Malmqvist, *Int. J. Quant. Chem.* **30**, 479 (1986); P.-Å. Malmqvist and B.O. Roos, *Chem. Phys. Letters* **155**, 189 (1989).
31. A. Broo and S. Larsson, *Chem. Phys.* **148**, 103-115 (1990).
32. T. Holstein, *Ann. Phys. (N.Y.)* **8**, 343-342, 343-389 (1959).
33. S. Larsson and A. Klimkāns, *Molecular Crystals and Liquid Crystals* **355**, 217-229 (2001); *Int. J. Quant. Chem.* **80**, 713-720 (2000).
34. S.I. Pekar, *Untersuchungen über die Elektronentheorie der Kristallektronentheorie der Kristalle*, Akademie-Verlag, Berlin, 1954.
35. P. George, J.S. Griffith, in Eds.: P.D. Boyer, H. Hardy, N. Myrbäck, *The Enzymes*, Vol. 1; (Academic Press, N.Y., 1955), p. 347.
36. A.I. Burshtein, *J. Chem. Phys.* **103**, 7927-7933 (1995).
37. J. Jortner, *Biochim. Biophys. Acta* **594**, 193 (1980; *J. Chem. Phys.* **64**, 4860 (1976).
38. M. Bixon and J. Jortner, *Adv. Chem. Phys.* **106**, 35 (1999); *J. Phys. Chem. B* **104**, 3906-3913 (2000);
39. J. Ulstrup and J. Jortner, *J. Chem. Phys.* **63**, 4358-4368 (1975).
40. J.R. Reimers and N.S. Hush, *Chem. Phys.* **134**, 323 (1989).
41. S. Eriksson, B. Källebring, S. Larsson, J. Mårtensson, O. Wennerström, *Chem. Phys.* **146**, 165 (1965).
42. D.L. Dexter, *J. Chem. Phys.* **21**, 836-850 (1953).
43. S. Larsson and A. Klimkāns, *J. Mol. Structure (Theochem)* **464**, 59-65 (1999).
44. K. Kilså, S. Larsson, and B. Albinsson, in preparation.
45. H. Kallman and F. London, *Z. Physik.Chem.* **B2**, 207 (1929).

46. H.A. Kramers, *Physica* **1**, 182-193 (1934).
47. P.W. Anderson, *Phys. Rev.* **109**, 1492 (1958).
48. P. Bertrand, *Chem. Phys. Letters* **113**, 104-107 (1985).
49. M. Bixon, M.E. Michel-Beyerle, and J. Jortner, *Isr. J. Chem.* **28**, 155-168 (1988); M. Bixon, J. Jortner, M.E. Michel-Beyerle, and A. Ogrodnik, *Biochim. Biophys. Acta* **977**, 273-286 (1989).
50. Y.I. Kharkats and J. Ulstrup, *Chem. Phys. Lett.* **182**, 81 (1991); A.M. Kuznetsov and J. Ulstrup, *J. Chem. Phys.*. **157**, 25 (1991); Y.I. Kharkats, A.M. Kuznetsov, and J. Ulstrup, *J. Phys. Chem.* **99**, 1345, 1355 (1995).
51. H. Sumi and T. Kakitani, *Chem. Phys. Lett.* **252**, 85 (1996).
52. S. Larsson and A. Volosov, *J. Chem. Phys.* 85, 2548-2554 (1986) Erratum: *J. Chem. Phys.* 86, 5223 (1987); S. Larsson and A. Volosov, *J. Chem. Phys.* 87, 6623-6625 (1987).
53. Källebring, B.; Larsson, S. Paper VI in Källebring, B., *Thesis*, Department of Biochemistry and Biophysics, Chalmers University of Technology, 1990; Larsson, S.; Broo, A.; Källebring, B.; Volosov, A. *Int. J. Quant. Chem. QBS* 1988, *15*, 1-22; Larsson, S.; Braga, M.; Broo, A.; Källebring, B. *Int. J. Quant. Chem. QBS* 1991, *18*, 99.
54. Ivashin, N.; Källebring, B.; Larsson, S.; Hansson, Ö. *J. Phys. Chem. B* 1998, *102*, 5017; Larsson, S.; Ivashin, N.V. *J. Appl. Spectr.* 1999, *66*, 539-543 (1999).
55. V.A. Shuvalov, A.V. Klevanik, A.V. Sharkov, Y.A. Matveetz, and P.G. Kryukov, *FEBS Lett.* **91**, 135 (1978); W. Holzapfel, U. Finkele, W. Kaiser, D. Oesterhelt, H. Scheer, H.U. Stilz, and W. Zinth, *Chem. Phys. Lett.* **160**, 1 (1989).
56. S. Larsson *J. Chem. Soc. Faraday Trans. 2* **79**, 1375-1388 (1983).
57. *HyperChem,* Molecular Visualization and Simulation, Hypercube, Inc..
58. A.A.S. Da Gama, *Theor. Chim. Acta* **68**, 159 (1985). A.A.S. Da Gama, *Quim. Nova* **11**, 76 (1989).
59. J.W. Verhoeven and P. Pasman, *Tetrahedron* **37**, 943-947 (1981).
60. I. Lee, *Tetrahedron* **39**, 2409-2416 (1983).
61. M.N. Paddon-Row and S.S. Wong, *Chem. Phys Letters* **167**, 432 (1990).
62. K.D. Jordan and M.N. Paddon-Row, *J. Phys. Chem.* **96**, 1188 (1992).
63. K.D. Jordan and M.N. Paddon-Row, *Chem. Rev.* **92**, 395 (1992).
64. S. Larsson *J. Phys. Chem.* **88**, 1321-1323 (1984).
65. S. Larsson *Chem. Phys. Letters* **90**, 136-139 (1982).
66. D.N. Beratan, J.N. Onuchic, *J. Photosynth. Res.* **22**, 173-186 (1989).
67. D.N. Beratan, J.N. Betts, and J.N. Onuchic, *J. Phys. Chem.* **96**, 2852-2855 (1992).

68. J.R. Reimers and N.S. Hush, *Chem. Phys.* **146**, 89 (1990); J.R. Reimers and N.S. Hush, *J. Photochem. Photobiol. A (Chem)* **82**, 31 (1994).

69. J.W. Evenson and M. Karplus, *J. Chem. Phys.* **96**, 5272 (1992).

70. A.A. Stuchebrukhov, *J. Chem. Phys.* **104**, 8424 (1996); **105**, 10819 (1996); **107**, 6495 (1997); **108**, 8499 (1998); *Int. J. Quant. Chem.* **77**, 16-26 (2000).

71. A.S. Davydov *phys. stat. sol. (b)* **90**, 457-464 (1978).

72. R. Landauer, *IBM J. Res. Dev.***32**, 306 (1992); *Phys. Scripta* **T42**, 110 (1992).

73. S. Datta, *Electron Transport in Mesoscopic Systems*, Cambridge University Press, London, New York, 1995.

74. M.P. Samantha, W. Tian, S. Datta, J.I. Henderson, C.P. Kubiak, *Phys. Rev. B* **53**, R7626 (1996).

75. V. Mujica, A. Nitzan, Y. Mao, W. Davis, M. Kemp, A. Roitberg, and M.A. Ratner, *Adv. Chem. Phys.* **107**, 403 (1999).

76. C. Coulson and H.C. Longuet-Higgins, *Proc. Roy. Soc. A* **191**, 39 (1947).

77. L.T. Calcaterra, G.L. Closs, J.R. Miller, *J. Am. Chem. Soc.* **105**, 670 (1983); J.R. Miller, L.T. Calcaterra, G.L. Closs, *J. Am. Chem. Soc.* **106**, 3047 (1984); G.L. Closs, L.T. Calcaterra, N.J. Green, K.W. Penfield, and J.R. Miller, *J. Phys. Chem.* **90**, 3673 (1986); G.L. Closs and J.R. Miller, *Science* **240**, 440 (1988);

78. M. Braga and S. Larsson, *J. Phys. Chem.* **97**, 8929-8936 (1993); M. Braga and S. Larsson, *Chem. Phys. Letters* **213**, 217-223 (1993).

79. Hasegawa, J.; Ohkawa, K., Nakatsuji, H., *J. Phys. Chem B* **1998**, *102*, 10410; Hasegawa, J, Nakatsuji, H. *J. Phys. Chem. B* **1998**, *102*, 10420-10430; Nakatsuji, H.; Hasegawa, J.; Ohkawa, K. *Chem. Phys. Lett.* **1998**, *296*, 499-504.

80. Larsson, S.; Ivashin, N.V. *J. Appl. Spectr.* **1999**, *66*, 539-543 (1999).

Density, Density Matrix, or Propagator

Yngve Öhrn
Quantum Theory Project
University of Florida
Gainesville, Florida 32611–8435

ABSTRACT

In spite of the current popularity of density functional methods and the many efforts to construct functionals that accurately describe molecular electronic properties, an exact exchange-correlation potential expressed only in terms of the electron density is elusive. This paper presents some thoughts around this problem that try to reflect some of the concerns about density functionals expressed by Per-Olov Lowdin on many occasions.

CONTENTS

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

1. Introduction

Computations in electronic structure theory for molecular systems, which is based on the Hamiltonian

$$\sum_{i=1}^{N} h(x_i) + \frac{1}{2} \sum_{i \neq j=1}^{N} V(x_i, x_j), \tag{1}$$

with $h = -\frac{1}{2}\nabla^2 - \sum_j Z_j/|\vec{r} - \vec{R}_j|$, and $V = 1/|\vec{r}_i - \vec{r}_j|$, are dominated by various forms of density functional approaches. These methods offer possibilities to compute primarily ground state properties and energetics with much less cost than do the quantum chemistry theory that uses many-particle wavefunctions. The numerical results obtained with the most popular density functionals are in many cases in good agreement with experimental numbers, a fact that coupled with their relatively low demands on compute power accounts for the growth in popularity of these methods. There is also the basic attractiveness of being able to describe chemistry in terms of the electron density; a much simpler and more intuitive quantity than a many-electron wavefunction [1].

The ideas of an energy functional in terms of the electron density $n(\mathbf{r})$ have a long history starting with the Thomas-Fermi [2, 3] and the Fermi-Thomas-Dirac [4] models for atoms. In this approach a statistical treatment of the free electron gas yields a total energy expression involving fractional powers of the electron density. The total kinetic energy term is proportional to a volume integral over $n^{5/3}(\mathbf{r})$. The electron-nuclear attraction energy is linear, while the electron-electron repulsion energy is quadratic in $n(\mathbf{r})$. An exchange energy term proportional to a volume integral over $n^{4/3}(\mathbf{r})$ in the Fermi-Thomas-Dirac model subtracts out some of the unphysical self-interaction. This model has obviously served as an inspiration to seek a universal energy functional for general electronic systems in terms of the electron density. It also justifies a "local density" approximation [5] as a resonable first step towards such a universal direct energy functional.

Further developments of the Thomas-Fermi and Thomas-Fermi-Dirac models are many and varied. It was recognized early by Slater [6], Gaspar [7], and by Kohn and Sham [5] that a proper quantum mechanical treatment of the kinetic energy uses the reduced one-electron density matrix and the Laplacian rather than the electron density. Such functionals are then not direct functionals of the electron density, but uses the much richer underlying structure of the Kohn-Sham determinant. Applications to molecules exposed the inadequacy of these models and augmentations in terms of so called gradient corrections, *i.e.* energy terms dependent on $\nabla n(\mathbf{r})$, $\nabla^2 n(\mathbf{r})$, and $(\nabla n(\mathbf{r}))^2$ were introduced [8, 9] to account for some of the inhomogeneity of the electron density.

Löwdin, who contributed in no small measure to the development of formal many-electron theory through his seminal work on electron correlation, reduced density matrices, perturbation theory, etc. many times expressed his concerns about the theoretical aspects of density functional approaches. This short review of the interconnected features of formal many-electron theory in terms of propagators, reduced pure state density matrices, and density functionals is dedicated to the memory of Per-Olov Löwdin.

2. Density functionals

The density functional methods were, in the views of many, legitimized by the introduction of the first Hohenberg-Kohn theorem [10]. The consequence of this celebrated theorem is that for a non-degenerate ground state and a given external potential, $v(\mathbf{r})$, the electronic ground state energy can be expressed as

$$E[n] = \int v(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + F[n], \tag{2}$$

where $F[n]$ is a universal functional of the electron density, and $E[n]$ is minimized by the exact density. The external potential $v(\mathbf{r})$, that one generally thinks about is the potential from the Coulomb potential from all the participating atomic nuclei.

It is often stated that this theorem only states the existence of a universal energy functional of electron density and provides no guidance for the construction of such a functional. It has also been challenged as not being all that profound given the general meaning of the term "functional". Maybe it only states the operational meaning introduced by E. Bright Wilson[1], namely knowing all the positions and charges of the atomic nuclei, and the charge and mass of an electron one can write down the electronic Hamiltonian, solve the corresponding Schrödinger equation and from the many-electron ground state wavefunction calculate the electron density, which then defines the ground state energy.

It is clear that the various density functional schemes for molecular applications rely on physical arguments pertaining to specific systems, such as an electron gas, and fitting of parameters to produce energy functionals, which are certainly not universal. By focusing on the energy functional one has given up the connection to established quantum mechanics, which employs Hamiltonians and Hilbert spaces. One has then also abandoned the tradition of quantum chemistry of the development of hierarchies of approximations, which allows for step-wise systematic improvements of the description of electronic properties.

Many theoretical approaches employing quantum mechanics for the study of molecular electronic structure considers the Hartree-Fock single determinantal wavefunction as the starting point for a hierarchy of approximations that permit the introduction of some measure of electron correlation. Already at the Hartree-Fock level one finds that the antisymmetry of the wavefunction under permutation of electron labels imposed by the Fermion nature of the electron introduces the exchange energy. This results in the non-local exchange potential in the single-electron Hartree-Fock equations that determine the "best" spin orbitals that make up the many-electron wavefunction. This non-local potential is essential for the description of bonding in even the simplest of molecules. Applications of density functional methods may quote the

[1] Comment made in discussions about the Hohenberg-Kohn theorem at the Sanibel Symposium, March, 1980.

Hohenberg-Kohn theorems, but use the Kohn-Sham construction and local approximations to such non-local potentials and often lump together the exchange and the correlation energies into an exchange-correlation energy $E_{xc}[n]$. This yields a local exchange-correlation potential $v_{xc}(\mathbf{r})$ in the Kohn-Sham equations that determine the Kohn-Sham spin orbitals $\{\phi_j\}$, *i.e.*

$$[-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})]\phi_j = \epsilon_j \phi_j \tag{3}$$

with $v_{eff}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r})$, which like the Hartree-Fock equations are solved self-consistently.

The Kohn-Sham electron density is then

$$n(\mathbf{r}) = \sum_{j=1}^{N} \sum_{\zeta} |\phi_j(\mathbf{r}, \zeta)|^2, \tag{4}$$

expressed in terms of the orthonormal Kohn-Sham spin orbitals and the total energy can be obtained as

$$E_0(N) = \sum_{j=1}^{N} \epsilon_j - \frac{1}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[n] - \int v_{xc}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \tag{5}$$

In this manner local density functional methods can be thought of as a Hartree-Fock-like theory with a different potential.

The density functional approaches include electron correlation. This means that the electron density used must be one corresponding to a rather sophisticated many-electron ground state wavefunction. This is the well-known *N*-representability problem. This problem is solved by constructing the density in terms of the Kohn-Sham orbitals, making the density correspond to a single-determinantal wavefunction. It is claimed [11] that when the *N* orbitals are allowed to vary over the space of functions that are continuous with a finite kinetic energy, and are square integrable then the density covers all *N*-representable densities. A constrained-search algorithm due to Levy [12, 13] makes possible to ascertain that the density that minimizes $E[n]$ in Eq. (2)

is N-representable and, thus, corresponds to a properly antisymmetric ground state wavefunction. The construction of such a wavefunction from a given density is of course a different matter.

The construction of the exchange-correlation potential in terms of the electron density is the crucial step. This is not easy, since the exact exchange correlation potential is most certainly non-local and even energy-dependent. The Hartree-Fock exchange potential has been approximated by a local potential with some success. For instance, the Xα method introduced by Slater used a local exchange potential proportional to $n^{1/3}(\mathbf{r})$, which gave approximate Hartree-Fock results and this multiplicative local potential simplified the calculations in comparison to those using the correct Hartree-Fock non-local exchange.

The construction of exchange correlation potentials and energies becomes a task for which not much guidance can be obtained from fundamental theory. The form of dependence on the electron density is generally not known and can only to a limited extent be obtained from theoretical considerations. The best one can do is to assume some functional dependence on the density with parameters to satisfy some consistency criteria and to fit calculated results to some model systems for which applications of proper quantum mechanical theory can be used as comparisons. At best this results in some form of ad-hoc semi-empirical method, which may be used with success for simulations of molecular ground state properties, but is certainly not universal.

Clever use of this approach and choice of exchange correlation potential has produced methods for calculating molecular ground state properties and energetics in reasonable agreement with experimental numbers for an impressive array of systems. The fact that these methods are computationally inexpensive make them applicable to larger systems than can routinely be studied with proper quantum mechanical approaches. This makes them useful simulation tools.

3. Density matrices

The N-electron problem in terms of reduced density matrices was considered early by Husimi, [14], and Löwdin [15]. The energy functional

$$E[\Gamma] = \int \left[Nh(x) + \binom{N}{2} V(x,y) \right] \Gamma(x,y|x,y) \mathrm{d}x\mathrm{d}y / \int \Gamma(x,y|x,y) \mathrm{d}x\mathrm{d}y, \tag{6}$$

where Γ is the two-electron reduced density matrix, which can be obtained from an N-electron ground state wavefunction Ψ as

$$\Gamma(x,y|x',y') = \binom{N}{2} \int \Psi(x,y,x_3,\cdots x_N) \Psi^*(x',y',x_3,\cdots x_N) dx_3 \cdots dx_N, \tag{7}$$

where Ψ is normalized to unity.

The use of this expression for a variational determination of Γ is a complex problem because of the N-representability requirement [15, 16, 17]. Nevertheless, there is a renewed interest in this problem and a number of methods, including so called cumulant-based approximations [18, 19] are being put forth as solutions to the representability problem. Although some advances can be obtained for special cases there appears to be no systematic scheme of approximating the density matrix with a well-defined measure of the N-representability error. Obviously, the variational determination of density matrices that are not guaranteed to correspond to an antisymmetric electronic wavefunction can lead to non-physical results.

The one-electron reduced density matrix is

$$\gamma(x|x') = \frac{2}{N-1} \int \Gamma(x,y|x',y) \mathrm{d}y, \tag{8}$$

the diagonal of which yields the electron density by summing over spin variables, *i.e.* $n(\vec{r}) = \sum_\zeta \gamma(x|x)$. The one-electron density matrix can be expressed in terms of the Löwdin natural spin orbitals $\{\chi_k\}$ as

$$\gamma(x|x') = \sum\nolimits_k n_k \chi_k(x) \chi_k^*(x'), \tag{9}$$

with the occupation numbers satisfying $0 \leq n_k \leq 1$.

The two-electron reduced density matrix is a considerably simpler quantity than the N-electron wavefunction and again, if the N-representability problem could be solved in a simple and systematic manner the "two-matrix" would offer possibilities for accurate treatment of very large systems. The natural expansion may be compared in form to the expansion of the electron density in terms of Kohn-Sham spin orbitals and it raises the question of the connection between the spin orbital space and the N-electron space when working with reduced quantities, such as density matrices and the electron density.

It would be very useful if there were an equation of motion for the electron density or for the reduced density matrices corresponding to a pure state of a many-electron system from which these quantities could directly be determined. Unfortunately this is not the case. The quantity closest to the "one-matrix" that has an equation of motion from which it can be determined by well-defined approximations is the one-electron Green's function or electron propagator. We explore its connection to the electron density in the next section.

4. Propagators

The single particle propagator [20] $G(x, x'; E)$ is an energy-dependent (or time-dependent) function of two compound space spin variables. It satisfies the Dyson-like equation

$$h(x)G(x, x'; E) + \int \Sigma(x, y; E)G(y, x'; E)\mathrm{d}y = \delta(x - y), \tag{10}$$

where $h(x)$ is the same as in Eq. (1). The non-local, complex, and energy-dependent operator $\Sigma(x, y; E)$ is called the self-energy or mass operator and describes exchange and correlation effects. The propagator satisfies the formal spectral expansion in terms of the Dyson amplitudes

$\{f_m\}$ and $\{g_m\}$, *i.e.*

$$G(x,x';E) = \sum_m \Big[\frac{f_m(x)f_m^*(x')}{E + E_0(N) - E_m(N+1)} + \frac{g_m(x)g_m^*(x')}{E - E_0(N) + E_m(N-1)}\Big], \tag{11}$$

with the energy eigenvalues of the stationary states of the (N+1)-electron and the (N-1)-electron systems and for a general complex energy (frequency) parameter E. The one-electron reduced density matrix for the N-electron ground state can be expressed as

$$\begin{aligned}\gamma(x|x') &= (2\pi i)^{-1}\int_C G(x,x';E)\mathrm{d}E \\ &= \sum_m g_m(x)g_m^*(x'),\end{aligned} \tag{12}$$

where the contour C consists of the real energy axis and an infinite semicircle in the upper half energy plane.

The amplitudes $\{g_m\}$ satisfy the equation

$$h(x)g_m(x) + \int \Sigma(x,x';E)g_m(x')\mathrm{d}x' = Eg_m(x). \tag{13}$$

These amplitudes are nonorthogonal and even linearly dependent and when orthogonalized by diagonalizing their metric by a unitary transformation **V**, we can write the natural orbitals as [21]

$$\chi_k(x) = \sum_m g_m(x)V_{mk}n_k^{-\frac{1}{2}}, \tag{14}$$

where n_k are the nonvanishing eigenvalues of the metric.

In comparing Eq. (13) to the Kohn-Sham equations Eq. (3) one concludes that $\Sigma(x,x';E)$, since it is derived from exact many-electron theory [22], is the exact Coulomb (direct) plus exchange-correlation potential. It is non-local and also energy-dependent. In view of this it is hard to see how the various forms of constructed local exchange correlation potentials that are in use today can ever capture the full details of the correlation problem.

One way to proceed would be to device a procedure for obtaining systematic local approximations to $\Sigma(x,x';E)$. This could possibly be obtained from known perturbation expansions

of the self-energy. The details of such a development falls outside the scope of this paper. It might be more straightforward to start from a general form of the two-matrix for a correlated many-electron wavefunction, such as the antisymmetrized geminal power (AGP) wavefunction [23, 24, 25]. Then the procedure advocated by Löwdin [15] to approximate the Hartree-Fock exchange potential by a "best" local potential could be adapted to find a local exchange-correlation potential. The AGP form of ground state wavefunction has been shown to be a consistent ground state for the random phase approximation (RPA) [26, 27].

Starting from a set of $2s$ spin orbitals $\{\phi_i\}$ for a $2N$-electron system one forms the antisymmetric geminal (two-electron function)

$$g(x_1, x_2) \equiv g(1,2) = \sum_{i=1}^{s} g_i |\phi_i(1)\phi_{i+s}(2)\rangle, \tag{15}$$

where the coefficients $\{g_i\}$ are in general complex. The one-electron reduced density matrix $D^1(g)$ corresponding to the geminal $g(1,2)$ has eigenvalues $n_j = n_{j+s} = |g_j|^2$ displaying at least double degeneracy. The corresponding natural orbitals [15] are $\{\chi_j\}_1^{2s}$, and we write

$$D^1(g) = \sum_{j=1}^{s} n_j \left[\chi_j \chi_j^* + \chi_{j+s}\chi_{j+s}^*\right]. \tag{16}$$

The many-electron wavefunction is constructed from the single geminal as

$$|g^N\rangle = S_N^{-1/2} O_{AS} \prod_{i=1}^{N} g(2i-1, 2i), \tag{17}$$

where O_{AS} is the antisymmetric projector. The normalization

$$S_N = \sum_{1 \le j_1 \le \cdots \le j_N \le s} n_{j_1} n_{j_2} \cdots n_{j_N} \tag{18}$$

is a symmetric function of order N in terms of the s eigenvalues $n_i = |g_i|^2$. This wavefunction is called an antisymmetric geminal product (AGP) and is the famous BCS-state of superconductivity projected onto the $2N$-electron space.

The total energy can be expressed succinctly as [25, 28]

$$E(g) = \mathrm{Tr}\left\{hD^1\left(g^N\right)\right\} + \mathrm{Tr}\{VD^2\left(g^N\right)\}, \tag{19}$$

where $D^1\left(g^N\right)$ and $D^2\left(g^N\right)$ are the one- and two-electron reduced density matrices of the AGP. It has been demonstrated by applications to small molecules that the AGP wavefunction is capable of satisfactorily describing electron correlation for various molecular geometries [25, 29, 30, 31].

The natural orbitals of the AGP are the same as those of the geminal, but the eigenvalues of $D^1\left(g^N\right)$ are of course different. We can write

$$\begin{aligned} D^1\left(g^N\right) &= \sum_{j=1}^{s} N_j\left[\chi_j\chi_j^* + \chi_{j+s}\chi_{j+s}^*\right] \\ &= \sum_{k=1}^{2s} N_k\chi_k(1)\chi_k^*(1') = \gamma(1|1') \end{aligned} \tag{20}$$

with

$$N_j = S_N^{-1} n_j \frac{\partial S_N}{\partial n_j}. \tag{21}$$

The two-electron reduced density matrix of the AGP can be brought to an interesting form, called, "Box and Tail"

$$D^2\left(g^N\right) = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix} \tag{22}$$

with **B** ann $s\times s$ matrix and **T** a diagonal matrix of dimension $2s(s - 1)$ achieved with a particular ordering of the natural orbitals. We write

$$\begin{aligned} D^2\left(g^N\right) &= \sum_{i=1}^{s} B_{i,i+s;i,i+s}|\chi_i\chi_{i+s}\rangle\langle\chi_i\chi_{i+s}| \\ &+ \sum_{i<j}^{s} B_{i,i+s;j,j+s}|\chi_i\chi_{i+s}\rangle\langle\chi_j\chi_{j+s}| \\ &+ \sum_{\substack{i<j \\ j\neq i+s}}^{2s} T_{ij;ij}|\chi_i\chi_j\rangle\langle\chi_i\chi_j| \\ &= \sum_{k<l} \Gamma_{kl;kl}|\chi_k\chi_l\rangle\langle\chi_k\chi_l| \end{aligned} \tag{23}$$

where

$$B_{i,i+s;i,i+s} = S_N^{-1} n_i \frac{\partial \ln S_N}{\partial n_i}, \tag{24}$$

$$B_{i,i+s;j,j+s} = S_N^{-1} g_i^* g_j \frac{\partial^2 S_{N+1}}{\partial n_i \partial n_j}, \tag{25}$$

and

$$T_{ij;ij} = S_N^{-1} n_i n_j \frac{\partial^2 S_N}{\partial n_i \partial n_j}. \tag{26}$$

Note that

$$\sum_{j(i<j)} T_{ij;ij} = (N-1)N_i \tag{27}$$

so that Eq. (8) is satisfied.

If one has a non-local one-electron potential $K(1)$ with a kernel $k(1,1')$, such that

$$K(1)\chi(1) = \int k(1,1')\chi(1')\mathrm{d}(1') \tag{28}$$

then one would like to minimize the difference

$$\int K(1)\gamma(1|1')\mathrm{d}(1) - \int v_{xc}(1)\gamma(1|1)\mathrm{d}(1), \tag{29}$$

where v_{xc} is the "best" local exchange-correlation potential. Let

$$k(1,1') = \int \frac{\Gamma(1,2|1',2)}{r_{12}}\mathrm{d}(2) \tag{30}$$

and then minimize [15]

$$\sum_k N_k |K(1)\chi_k(1) - v_{xc}(1)\chi_k(1)|^2 \tag{31}$$

with the result that the "best" exchange-correlation potential is

$$\begin{aligned} v_{xc}(1) &= \frac{\sum_k N_k \chi_k^*(1) K(1)\chi_k(1)}{\sum_k N_k \chi_k^*(1)\chi_k(1)} \\ &= \sum_{k=1}^{2s} N_k \chi_k^*(1) \frac{\int \Gamma(1,2|1',2)}{r_{12}\gamma(1|1)}\mathrm{d}(2)\chi_k(1')\mathrm{d}(1'). \end{aligned} \tag{32}$$

We can now write

$$v_{xc}(1) = \sum_k \sum_j \frac{N_k \Gamma_{kj;kj}}{\gamma(1|1)} \times \left[\chi_k^*(1)\chi_k(1) \int \frac{\chi_j^*(2)\chi_j(2)}{r_{12}} d(2) - \chi_k^*(1)\chi_j(1) \int \frac{\chi_j^*(2)\chi_k(2)}{r_{12}} d(2)\right] \tag{33}$$

and use the AGP form of density matrices. These forms are parametrized by the complex geminal coefficients $\{g_i\}$ and by the geminal occupation numbers $n_i = |g_i|^2$. The dependence is highly nonlinear and will demand some careful analysis to find stable and fast optimization algorithms [32].

The problem is that this exchange correlation potential does not in a simple fashion depend on only the density or the diagonal part of the one-matrix. The antisymmetry property of the two-matrix, which yields the last term in Eq. (33) as well as the apperance of the coefficents $\Gamma_{kj;kj}$, prevent this. Only some further (ad hoc) assumptions will allow an exchange-correlation potential that is simply a function of the electron density. This can be achieved in a number of ways. For a full discussion on this subject see [11].

5. Conclusion

Density functional approaches to molecular electronic structure rely on the existence theorem [10] of a universal functional of the electron density. Since this theorem does not provide any direction as to how such a functional should be constructed, the functionals in existence are obtained by relying on various physical models, such as the uniform electron gas and others. In particular, the construction of an exchange-correlation potential that depends on the electron density only locally seems impossible without some approximations. Such approximate exchange-correlation potentials have been derived and applied with some success for the description of molecular electronic ground states and their properties. However, there is no credible evidence that such simple constructions can lead to either systematic approximate treatments , or an exact description of molecular electronic properties. The exact functional that seems to

elude the density functional methods is most likely the operational construction mentioned by E. B. Wilson.

References

[1] Y. Öhrn, Propagator theory of atoms and molecules, in *The New World of Quantum Chemistry*, edited by B. Pullman and R. G. Parr, Proceedings of the second international congress of quantum c hemistry held in New Orleans, April 19-24, 1976, pages 57–78, Reidel, 1976.

[2] L. H. Thomas, Proc. Camb. Phil. Soc. **23**, 542 (1927).

[3] E. Fermi, Z. Physik **48**, 73 (1928).

[4] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).

[5] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

[6] J. C. Slater, Phys. Rev. **81**, 385 (1951).

[7] R. Gaspar, Acta Physica Hungarica **3**, 263 (1954).

[8] D. A. Kirzhnits, Sov. Phys.-JETP **5**, 64 (1957).

[9] C. H. Hodges, Can. J. Phys. **51**, 1428 (1973).

[10]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[11]R. G. Parr and W. Yang, *Density-Funtional Theory of Atoms and Molecules*, International Series of Monographs on Chemistry, Oxford University Press, Clarendon Press, New York, Oxford, 1989.

[12]M. Levy, Phys. Rev. A **26**, 1200 (1982).

[13]M. Levy and J. P. Perdew, The constrained search formulation of density functional theo ry, in *Density functional methods in physics*, edited by R. M. Dreizler and J. da Providencia, pages 11–30, Plenum, 1985.

[14]K. Husimi, Proc. Phys. Math. Soc. Japan **22**, 264 (1940).

[15]P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[16]A. J. Coleman, Rev. Mod. Phys. **36**, 668 (1963).

[17]H. Nakatsuji, Theor. Chem. Acc. **102**, 97 (1999).

[18]D. A. Mazziotti, Chm. Phys. Lett. **326**, 212 (2000).

[19]D. Mukherjee and W. Kutzelnigg, J. Chem. Phys. **114**, 8226 (2001).

[20]J. Linderberg and Y. Öhrn, *Propagators in Quantum Chemistry*, Academic, New York, 1973.

[21]O. Goscinski and P. Lindner, J. Math. Phys. **11**, 1313 (1970).

[22]P. Nozière, *Theory of Interacting Fermi Systems*, W. A. Benjamin Inc., New York, 1964.

[23]A. J. Coleman, J. Math. Phys. **6**, 1425 (1965).

[24]J. Linderberg, Israel J. Chem. **19**, 93 (1980).

[25]J. V. Ortiz, B. Weiner, and Y. Öhrn, Int. J. Quantum Chem. **S15**, 113 (1981).

[26]J. Linderberg and Y. Öhrn, Int. J. Quantum Chem. **12**, 161 (1977).

[27]Y. Öhrn and J. Linderberg, Int. J. Quantum Chem. **15**, 343 (1979).

[28]E. Deumens, B. Weiner, and Y. Öhrn, Nucl. Phys. **A466**, 85 (1987).

[29]B. Weiner, H. J. Jensen, and YÖhrn, J. Chem. Phys. **80**, 2009 (1984).

[30]Y. Öhrn, Int. J. Quantum Chem.: Quant. Chem. Symp. **19**, 39 (1986).

[31]E. Sangfelt, R. Roychowdhury, B. Weiner, and YÖhrn, J. Chem. Phys. **86**, 4523 (1987).

[32]H. J. Jensen, B. Weiner, and Y.Öhrn, Int. J. Quantum Chem. **S16**, 615 (1982).

CONJUGATE EIGENVALUE PROBLEMS AND GENERALIZED STURMIANS

Osvaldo Goscinski
Department of Quantum Chemistry
The Ångström Laboratory
Box 518 SE-75120 Uppsala Sweden

Abstract

A short presentation of the unpublished Research Report No 217, by Osvaldo Goscinski, from the Quantum Chemistry Group at Uppsala University is made. The Report has the title "Conjugate Eigenvalue Problems and the Theory of Upper and Lower Bounds". Some justification of its *verbatim* inclusion in this Volume in honour of Per-Olov Löwdin is made. It is essentially motivated by the attention that the theory presented there has received in the field of Generalized and Molecular Sturmians. Current work by John Avery and collaborators is alluded to. It is included as an Appendix.

Contents

1 Introduction

The Preliminary Research Report No. 217, June 17 1968, was not really completely finished at the specified date. Perusal indicates that two sections, Abstract and Acknowledgements, are missing. A change of civil status took place at the end of that month followed by Summer vacation.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

Urgency in publishing the report did not become apparent at that time. At the beginning I thought about it sporadically. The unpublished note was part though of my list of publications. It dawned upon me gradually though through the daunting efforts and interest from John Avery that it should become available to others. It has existed as typewritten manuscript and thanks to the effort of John Avery in a Latex version which is included below as an Appendix. Publication without any changes is a challenge. It provides though a view of the style and problem solving that pervaded at the Quantum Chemistry Group under the enthusiastic leadership of Per-Olov Löwdin, unique mentor and adviser, to whom this contribution is dedicated. It is particularly frustrating that we never had the opportunity to discuss this article.

2 The conjugate eigenvalue problem

In the Introduction of the Technical Note it is stated that "the purpose of this paper is to consider splittings of a Hamiltonian which is assumed to be bounded from below where V is not necessarily non-negative. In fact an explicit treatment can be given for the part of V which is negative definite. The technique employed is an eigenvalue transformation in which the energy is a fixed parameter and the "strength" of the perturbation, i. e. the numerical coefficient multiplying V, has to be determined".
This is what is meant by the conjugate eigenvalue problem. The perturbation involved is the simple Coulomb operator, or parts of it, appearing in molecular physics.

The theory of upper and lower bounds is used in order to get estimates of the resolvent in the one-electron as well as in the many-electron case. The issue of completeness of the basis sets employed is central.

3 Generalized Sturmians

The results obtained by a number of authors, though specially by John Avery and collaborators in the field of generalized Sturmians, illustrate the power

of the methods involved.It is appropriate, though with some hesitation, to cite, *in extenso*, the Introduction of the book by John Avery /1/:

"This book explores the connections between the theory of hyperspherical harmonics, momentum-space quantum theory, and generalized Sturmian basis functions; and introduces methods which may be used to solve many-electron problems directly, without the use of the self-consistent-field approximation."

...

"Sturmian basis functions have an independent history, although they are closely related to the topics which we have just been discussing. Very early in the history of quantum theory, it was thought that hydrogenlike wave functions could be used as building blocks to construct the wave functions of more complicated systems- for example many-electron atoms or molecules. However it was soon realized that unless the continuum is included, a set of hydrogenlike orbitals is not complete. To remedy this defect Shull and Löwdin /2)/ introduced sets of radial functions which could be expressed in terms of Laguerre polynomials multiplied by exponential factors. The sets were constructed in such a way as to be complete, i.e., any radial function obeying the appropriate boundary conditions could be expanded in terms of the Shull-Löwdin basis set. Later Rotenberg /3, 4/ gave the name "Sturmian" to basis sets of this type in order to emphasize their connection with Sturm-Liouville theory. There is a large and rapidly growing literature on Sturmian basis functions, and selections from this literature are cited in the bibliography."

4 Molecular Sturmians

"In 1968, Goscinski /5/completed a study of Sturmian basis sets, formulating the problem in such a way as to make generalization of the concept very easy. In the present text we shall follow Goscinski's easily generalizable definition of Sturmians. After this presentation the book continues with an exhaustive presentation of Sturmians and their use. "

This is actually the work in question. The generalization to a many-electron conjugate eigenvalue problem was done in 1968 but the actual use of the name Sturmian was not done there.

A new development is the notion of Molecular Sturmians, recently published /6/. It is stated there that:

"In 1968 Goscinski introduced a powerful generalization of the Sturmian concept. He defined a Sturmian basis set a a solution of the Schrödinger equation with a weighted potential, the weighting factor being chosen for each basis function in such a way that all the members of the set correspond to the same value of the energy. In his 1968 work, explicitly considered many-electron Sturmian basis sets. Later, two-electron Sturmians were used by Gazeau and Maquet while Bang and Vaagen and their coworkers introduced many-particle Sturmians into nuclear physics.

Herschbach, Avery and Antonsen introduced and studied many particle Sturmian basis sets where the a basis potential Vo (**x**) was taken to be the potential of a *d*-dimensional hydrogenlike atom but in most applications convergence with this basis proved to be slow. In the book by Avery (1) a detailed description is given of the notion of many-electron Sturmian basis sets where the potential used in the conjugate eigenvalue problem , Vo (x) is the positive Coulomb potential of the nucleus, as discussed by Aquilanti and Avery , generalized to molecules by Avery and Sauer where it is the potential due to the nuclei. Convergence of these expansions was very rapid which makes them very interesting from the computational point of view. For a good account and references see the book by Avery /1/.

An excellent review of the state of the art was given recently in the case when the actual external potential experienced by an N-electron system is given by the attractive potential of the nuclei. This leads to accurate calculations including explicitly including correlation effects /7 /

It is the Technical Note from 1968, alluded to as reference /5/ which is included here as an Appendix.

5 Acknowledgements

This brings me to some reflections about the activities carried out at the Quantum Chemistry Group under the guidance of Per-Olov Löwdin to whom this work is dedicated. In a way the Technical Note, now included as an Appendix, is still unfinished because it does not have an abstract nor acknowledgments to the various members of the group with which I discussed the subject matter.

The author is grateful to Professsor John Avery for his interest in Sturmian functions and in the unpublished report from 1968. He is further indebted to him for a Latex translation of the original manuscript where the conjugate eigenvalue problem as well Sturmians are introduced.

6 References

1. J. Avery, "Hyperspherical Harmonics and Generalized Sturmians" Kluwer Academic Publishers Dordrecht (2000)

2. H. Shull and P-O Löwdin, J. Chem Phys **30**, 617 (1959).

3. M. Rotenberg, Ann. Phys. (New York) **19**, 262 (1962).

4. M. Rotenberg, Adv. At. Mol. Phys.. **6**, 233 (1970).

5. O. Goscinski, Preliminary Research Report No. 217, Quantum Chemistry Group, Uppsala University, (1968)

6. J. Avery, Int. J. Quantum Chem. **83** p 1 (2001)

7. J. Avery, Adv. Quantum Chem. **31**, 201 (1999.)

APPENDIX

QUANTUM CHEMISTRY GROUP
FOR RESEARCH IN ATOMIC, MOLECULAR AND SOLID-STATE THEORY
UPPSALA UNIVERSITY, UPPSALA, SWEDEN

CONJUGATE EIGENVALUE PROBLEMS AND THE THEORY OF UPPER AND LOWER BOUNDS

by

Osvaldo Goscinski

PRELIMINARY RESEARCH REPORT NO. 217
June 17, 1968

Sponsored in part by the
SWEDISH NATIONAL RESEARCH COUNCIL
and in part by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (OSR)
THROUGH THE EUROPEAN OFFICE OF AEROSPACE RESEARCH (OAR)
UNITED STATES AIR FORCE
under Grant AF EOAR 67-50

INTRODUCTION

It is a truism that the Schrödinger equation very seldom may be solved exactly. It is then desirable to obtain error estimates, namely, upper and lower bounds to the eigenvalues and accuracy criteria for the approximate eigenfunctions.

The main approximation methods belong to two categories: variation and perturbation type.

The former yield upper bounds to the eigenvalues through the solution of secular equations.[1] It is possible to obtain lower bounds from variational solutions by additional computation and additional information: in Temple's method[2] the expectation value of $\mathcal{H}^2$ and the first excited eigenvalue are needed in order to compute a lower bound to the ground state energy. Lower bounds from variational methods can be constructed by the technique of intermediate problems, involving truncations and projections of operators, essentially due to Weinstein, Aronszajn, Bazley and Fox.[3]

The partitioning technique developed by Löwdin provides a synthesis of different perturbation schemes, and by the use of a bracketing theorem upper and lower bounds can be obtained.[4]

A feature of the intermediate problem approach, is that the construction of comparison operators involves the splitting of the Hamiltonian, $\mathcal{H}$, into two parts: $\mathcal{H}^0$ and V, as in perturbation theory, but in such a way that lower bounds to only those eigenvalues of $\mathcal{H}$ that lie below the first limit point of the spectrum of $\mathcal{H}^0$ can be obtained, with the concomitant requirement that V is non-negative definite.[5] In the partitioning technique this is reflected by the same condition on V, and in addition, if $E_0 < E_p^0$ where E_0 and E_p^0 are the ground state of $\mathcal{H}$ and the p-th state of $\mathcal{H}^0$ respectively, the explicit inclusion of the first p eigenfunctions of $\mathcal{H}^0$ in the lower bound formula is necessary.[6] These requisites are often not compatible with each other. An obvious V is the interelectronic repulsion which leaves as $\mathcal{H}^0$ a sum of independent particle operators. As it is well known*), the ground state of lithium lies above the onset of the continuous

*) Pointed out to us by Sven G. Larsson

spectrum of $\mathcal{H}^0$. It is therefore not tractable by the methods currently available. The spectrum of the unperturbed Hamiltonian or at least its lower part, i. e., the first eigenvalues and eigenfunctions has to be known, in order to construct a tractable reference function in the partitioning method.[4] The spectrum is not actually needed to compute the reduced resolvent T_0 , and the first part of the spectrum is needed in the method of truncations. At any rate, an $\mathcal{H}^0$ with a partially discrete spectrum whose first accumulation point lies above the ground state of $\mathcal{H}$ is not compatible with a positive V .

If in atoms it is natural to split the Hamiltonian into an unperturbed part with the desired characteristics and a non-negative definite perturbation V , in molecules, if one considers an unperturbed part for which the lower part of the spectrum is known, the perturbation is not positive. If one chooses the perturbation in the same way as in atoms i. e. the interelectronic repulsion, the unperturbed spectrum has no known spectrum. This dilemma was faced by Bazley and Fox[5], who suggested a method of truncations combined with a splitting of the unperturbed Hamiltonian into parts corresponding to the different nuclei. Coulson and Johnson[7] carried out a calculation for H_2^+ with this scheme, and the results were not very good, the reason being that the truncations involved threw away the significant continuum contribution.

The purpose of this paper is to consider splittings of a Hamiltonian which is assumed bounded from below, where V is not necessarily non negative. In fact, an explicit treatment can be given for V , or . the part of V which is negative definite.

The technique employed is an eigenvalue transformation in which the energy is a fixed parameter and the "strength" of the perturbation, i. e., the numerical coefficient multiplying V , has to be determined These conjugate eigenvalue problems are of course, well known in the theory of integral equations. Transformations of this type were introduced in connection with one-electron problems by Schrödinger[8] and by Hylleraas, [9] and it was recognized that the conjugate problem could have a complete discrete spectrum, in counter distinction with the original eigenvalue problem. Walmsley exploited this fact in calculations on H_2^+ , truncating the conjugating problem, and thus avoiding the continuum problem. [10]

A preliminary account of this approach, which emphasizes applications to many-electron systems was made, where the fact that the conjugate problem is associated to a non-negative operator, was exploited to obtain upper and lower bounds to the eigenvalues by the projection technique.[11]

The very interesting work of Joseph[12] on the determination of the exact number of bound states of a given potential uses the conjugate eigenvalue problem for arbitrary one particle, N-dimensional potentials. It turns out that the conjugate problem is exactly soluble in several cases of interest, but for an arbitrary problem the techniques discussed in this paper, with approximate solutions, are needed.

CONJUGATE EIGENVALUE PROBLEMS

Let $\mathcal{H}$ be the Hamiltonian of the system; T the kinetic energy operator, U and V potential energy operators which are non-negative definite such that:

$$\mathcal{H} = T + U - \mu V \tag{1}$$

For an atom we could choose $\mu = Z$ and

$$U = \sum_{i<j}^{N} r_{ij}^{-1} \; ; \; V = \sum_{i=1}^{N} r_i^{-1}$$

though this is not the only possible choice.

The Schrödinger equation can be written in the following way:

$$[T + U - \mathcal{E}]\, \underline{\Psi}_n(\mathcal{E}) = \mu_n(\mathcal{E})\, V\, \underline{\Psi}_n(\mathcal{E}) \tag{2}$$

The bound states of the system have square integrable eigenfunctions

and the corresponding eigenvalues are negative. The parameter μ is given. Suppose now that instead of (2) we consider

$$[T+U-\varepsilon]\bar{\Psi}_n(\varepsilon) = \mu_n(\varepsilon) V \bar{\Psi}_n(\varepsilon) \tag{3}$$

where ε is a parameter at our disposal and the $\mu_n(\varepsilon)$ and $\bar{\Psi}_n(\varepsilon)$ the new eigenvalues and eigenfunctions. Following Joseph[12] we shall call (3)' the conjugate problem.

It is a Sturm-Liouville type of problem, and its properties are well known. The functions $\bar{\Psi}_n(\varepsilon)$ for real ε are orthogonal with the metric V. Define

$$\bar{\Phi}_n(\varepsilon) = V^{1/2}\bar{\Psi}_n(\varepsilon) \tag{4}$$

then (3) becomes

$$V^{-1/2}(T+U-\varepsilon)V^{-1/2}\bar{\Phi}_n(\varepsilon) = \mu_n(\varepsilon)\bar{\Phi}_n(\varepsilon) \tag{5}$$

The operator

$$\Omega(\varepsilon) \equiv V^{-1/2}(T+U-\varepsilon)V^{-1/2} \tag{6}$$

is Hermitean, its eigenvalues, if any, are real, its eigenfunctions, if any, are orthogonal, and if square-integrable can be assumed, normalizable. From

$$\delta_{nm} = \langle \bar{\Phi}_n(\varepsilon) | \bar{\Phi}_m(\varepsilon) \rangle = \langle \bar{\Psi}_n(\varepsilon) | V | \bar{\Psi}_m(\varepsilon) \rangle$$

follows the statement about orthogonality of the $\Phi_n(\varepsilon)$. In addition, for $\varepsilon < 0$, $\Omega(\varepsilon)$ is a positive definite operator, with positive eigenvalues. $\Omega(\varepsilon)$ will in general have a partially discrete spectrum, of positive eigenvalues, and a continuum. In some exceptional cases the spectrum will be discrete.

The eigenvalues $E_n(\varepsilon)$ and $\mu_n(\varepsilon)$ can be related through an application of the Hellmann-Feynman theorem.[13] For any eigenvalue of $\Omega(\varepsilon)$, $\varepsilon < 0$ we have:

$$\frac{d\mu_n(\varepsilon)}{d\varepsilon} = \frac{\langle \bar{\Psi}_n(\varepsilon) | \bar{\Phi}_n(\varepsilon) \rangle}{\langle \bar{\Phi}_n(\varepsilon) | V | \bar{\Psi}_n(\varepsilon) \rangle} (-1) < 0 \tag{7}$$

assuming square integrability of both Φ_n and $\bar{\Psi}_n$.

This is not at all surprising, since in a similar way we may obtain the familiar result

$$\frac{dE_n(\mu)}{d\mu} < 0 \tag{8}$$

which indicates, i. e. that to increase the nuclear charge in an isoelectronic series decreases the energy.

A digression can be made at this point. Consider a many-body problem

$$(T + \lambda W)\Psi_n = E_n \Psi_n \,, \quad \lambda > 0$$

where the potential W is not necessarily definite. For bound state problems Ψ_n can be assumed normalized and $E_n < 0$:

$$E_n = \langle \varphi_n | T | \varphi_n \rangle + \lambda \langle \varphi_n | W | \varphi_n \rangle$$

Since T is positive definite, $\langle \varphi_n | W | \varphi_n \rangle < 0$ even though W is not negative definite. Then:

$$\frac{dE_n}{d\lambda} < 0$$

which indicates that in an arbitrary system, the bound-state energies decrease when the "strength" of the potential increases. This result was proven previously for local or non-local perturbations, irrespective of their negative character, by studying the integral equation associated to one spinless particle.[14] The present proof, of deceptive simplicity, lacks the proof of the continuity with respect to $\mathcal{E}$ of the $\mu_n(\mathcal{E})$.

The functions $\mu(\mathcal{E})$ and $\mathcal{E}(\mu)$ are interrelated. In certain problems one can be obtained from the other. In general, it is unwieldy to obtain $\mathcal{E} = \mathcal{E}(\mu)$. That is the problem of perturbation theory. Sometimes it is possible to obtain the dependence of the interaction parameter μ on the energy $\mathcal{E}$, and this is the basis for iteration-variation methods.[15] It is by no means necessary that the solution of (2) implies the solution of (3), or vice versa, yet for some potentials this is indeed the case, and it can be used to count the number of bound states of a particle in certain potentials.[12]

Before discussing the nature of the spectrum of the conjugate problem we shall relate the solutions of (2) and (3) in a more specific way.

Suppose that by some method we are able to provide upper and lower bounds to $\mu_n(\mathcal{E})$ i. e. that for an arbitrary $\mathcal{E}$ (which should certainly be negative) we can obtain $\mu_n^<(\mathcal{E})$ and $\mu_n^>(\mathcal{E})$ such that

$$\mu_n^<(\mathcal{E}) \leq \mu_n(\mathcal{E}) \leq \mu_n^>(\mathcal{E}) \qquad (9)$$

We are assuming continuity with respect to $\mathcal{E}$ of the $\mu_n(\mathcal{E})$ associated with the bound states. The important result in this connection is that $\mathcal{E}$ for which

$$\mu_n^{>}(\mathcal{E}) \leq \mu(\mathcal{E}) \rightarrow E_n(\mu) \leq \mathcal{E} \tag{10}$$

is an upper bound to the energy of the state under consideration. And also for the lower bounds:

$$\mu_n^{<}(\mathcal{E}) \geqslant \mu(\mathcal{E}) \rightarrow E_n(\mu) \geqslant \mathcal{E} \tag{11}$$

The proof follows from the fact that if in (3) we set $\mathcal{E} = E_n(\mu)$ it has the n-th. solution $\mu_n(E_n(\mu)) = \mu$, $\Xi_n(E_n(\mu)) = \Xi_n(\mu)$. In other words, the operator $\Omega(\mathcal{E})$ for $\mathcal{E} = E_n(\mu)$ has as n-th. eigenvalue $\mu_n(E_n(\mu)) = \mu$ and as n-th. eigenfunction $\Xi_n(E_n(\mu))$.
Notice that this does not mean that the set $\Phi_n(E_n(\mu))$ belongs to the same eigenvalue $\mathcal{Z}$. They are eigenfunctions of different operators $\Omega(E_n(\mu))$.
Then it follows that:

$$\mu_n^{>}(\mathcal{E}) \leq \mu \doteq \mu_n(E_n(\mu)) \tag{12}$$

and this implies in turn that $E_n(\mu) \leq \mathcal{E}$ on account of the negative slope of $\mu_n(\mathcal{E})$. See (4). Also

$$\mu_n^{<}(\mathcal{E}) \geqslant \mu = \mu_n(E_n(\mu)) \tag{13}$$

implies (11). Therefore upper bounds to $\mu_n(\mathcal{E})$ are related

to upper bounds to $E_n(\mu)$, and the same applies to the respective lower bounds. This treatment is completely general and is applicable to any negative part of the Hamiltonian.

So far nothing has been done towards solving (2), since (3) or (5) are just as difficult. But $\Omega(\varepsilon)$ is positive by construction, in fact any part of $\Omega(\varepsilon)$ is positive, and therefore it is possible to carry inner projections of $\Omega(\varepsilon)$ or of parts of $\Omega(\varepsilon)$ which are a convenient tool in obtaining the $\mu_n(\varepsilon)$ [4]. In addition it is possible to split $\Omega(\varepsilon)$ into an unperturbed part, with known spectrum and a positive perturbation. In the case of one-electron systems, the spectrum is discrete. For many-electron systems it is partially discrete, as we shall see in the next section.

NATURE OF THE SPECTRUM OF THE CONJUGATED PROBLEM

This is indeed not an easy question, but considerable insight can be gained by studying first a non-interacting system. Consider then, as an unperturbed problem that $U = 0$ and that

$$T = \sum_{i=1}^{N} T(i) \; ; \quad V = \sum_{i=1}^{N} V(i) \tag{14}$$

Equation (3) becomes then:

$$\left[\sum T(i) - \varepsilon\right] \overline{\Psi}_n^0(\varepsilon) = \mu_n^0(\varepsilon) \sum V(i) \overline{\Psi}_n^0(\varepsilon) \tag{15}$$

The problem is to find solutions $\mu_n^0(\varepsilon)$ and $\overline{\Psi}_n^0(\varepsilon)$ for a given $\varepsilon < 0$. As antisymmetry is a condition on the sol-

utions of (2) we shall impose the same requirement on $\bar{\Psi}_n^0(\varepsilon)$. We purposely will ignore all other questions of symmetry of the solutions of Eq. (15). As the operators involved are a sum of one electron operators, we may seek solutions in the form of antisymmetized products of spin orbitals, i. e., construct determinants, take linear combinations of them in order to construct eigenfunctions of L^2, S^2, etc. But for our purposes it is enough to assume that all that can be done following the treatment given in the book by Slater,[10] or by using a projection operator formalism.

Let $\mathcal{A}$ be the antisymmetizer operator. We seek solutions of (15) of the form:

$$\bar{\Psi}_n^0(\varepsilon) = \mathcal{A} \prod_{j=1}^{N} u_j(j) \tag{16}$$

Replacing (16) in (15) we obtain:

$$\sum_{i=1}^{N} \prod_{j \neq i} u_j(j) \left[T(i) - \mu_n^0(\varepsilon) V(i) \right] u_i(i) = \varepsilon \bar{\Psi}_n^0(\varepsilon) \tag{17}$$

Let us set $\varepsilon = \sum_{i=1}^{N} \varepsilon_i$ where the ε_i are arbitrary. Then (17) becomes:

$$\mathcal{A} \sum_{i=1}^{N} \prod_{j \neq i} u_j(j) \left[T(i) - \mu_n^0(\varepsilon) V(i) - \varepsilon_i \right] u_i(i) = 0 \tag{18}$$

This allows us to assert that (16) is a solution of (15) provided the following conditions are fulfilled:

$$\text{(i)} \quad \mathcal{E} = \sum_{i=1}^{N} \varepsilon_i < 0$$

$$\text{(ii)} \quad \left[T(1) - \mu_i(\varepsilon_i) V(1) - \varepsilon_i\right] u_i(1) = 0 \qquad (19)$$

$$\text{(iii)} \quad \mu_i(\varepsilon_i) = \mu_n^0(\varepsilon) \quad \text{for all } i .$$

Consider condition (ii). If $V(1) = r_1^{-1}$ it is entirely analogous to the hydrogen atom eigenvalue problem, but with the energy as a fixed arbitrary parameter and the nuclear charge as the eigenvalue to be determined. If one uses the method of Frobenius to solve (ii) for this potential one finds the same terminating condition for the series as in the hydrogen atom, namely that

$$\mu_i(\varepsilon_i) = n_i \sqrt{-2\varepsilon_i}, \quad n_i = 1, 2, \ldots \qquad (20)$$

The eigenfunctions are hydrogen-like functions, but with no dependence on the quantum number n in the argument:

$$u_i(1) = u_{n_i \ell_i m_i \sigma_i}(1) = R_{n_i \ell_i}(\rho) Y_{\ell_i m_i}(\theta, \varphi) \Theta_{\sigma_i} \qquad (21)$$

with $\rho = 2\sqrt{-2\varepsilon_i}\, r$

These functions were already known by Schrödinger.[18] The fact that they form a complete discrete set was used by him and by others.[19] There is a discussion by Hylleraas[20] of a transformation of the Schrödinger equation for the S-states of the hydrogen atom which yields the same solutions as (20) and (21), and where he proves the completeness of the discrete spectrum of (ii) for $\varepsilon_i < 0$ Clearly, they are orthogonal with weight factor $V(1) = r_1^{-1}$ In the present context the significant point is that for a one-electron Kepler problem the transformation being discussed yields a complete discrete spectrum. Unfortunately this does not follow for many electron problems. Consider condition (i) of (19). The sum of "one-electron energies" can be negative even if one, or at the most, N-1 of them

are positive. Therefore, there are solutions of Eq. (15) which correspond to positive $\mathcal{E}_i$ and these form a continuum.

Let us consider the discrete part of the spectrum and examine the solutions in more detail. We have to impose condition (iii). Let the $\mathcal{E}_i$ satisfy:

$$\mathcal{E}_i = \epsilon p_i^2 ; \quad \epsilon < 0 \tag{22}$$

Then (20) becomes

$$\mu_i(\mathcal{E}_i) = n_i p_i \sqrt{-2\epsilon} \tag{23}$$

The p_i's are integers at our disposal. Let's denote a set of p_i's by $p = \{p_i\}$. Then, given a set p, we impose a restriction on the associated n_i's:

$$n_i p_i = n M(p) ; \quad n = 1, 2, \ldots, \tag{24}$$

where M(p) is the lowest common multiple of the set p. In such a case Eq. (23) becomes

$$\mu_i(\mathcal{E}_i) = \mu^0_{n,p} = n M(p) \sqrt{-2\epsilon}, \text{ for all } i. \tag{25}$$

We see then that the set p determines a "configuration", and for each configuration there is a set of eigenvalues $\mu^0_{n,p}$ and a set of eigenfunctions, antisymmetized products of spin orbitals, where the radial functions are given by (21) provided the condition (24)

is imposed on the indices n_i . The eigenfunctions are given by

$$\underline{\Psi}^0(\varepsilon) = \underline{\Psi}^0_{np}(\varepsilon) = \mathcal{A} \prod_i R_{n_i l_i m_i}(\rho_i) Y_{l_i m_i}(\theta_i, \varphi_i) \begin{Bmatrix} \alpha_i \\ \beta_i \end{Bmatrix} \quad (26)$$

where

$$\rho_i = 2 p_i \sqrt{-2\epsilon} \; ; \; n_i = M(p)/p_i \quad (27)$$

We should notice that this shows explicitly how each one-electron function which appears in the solution of (15) depends on the quantum numbers of all the other spin orbitals as suggested by Lundqvist.[21] It should be emphasized that not all choices of the set p are independent. For a given ε , there is a restriction on any set p' :

$$\varepsilon = \epsilon \sum p_i^2 = \epsilon' \sum p_i'^2 \quad (28)$$

or

$$\epsilon' = \sum p_i^2 / \sum p_i'^2 \quad (29)$$

Therefore, for any set p' the eigenvalues $\lambda^0_{n,p'}$ are related to the eigenvalues $\lambda^0_{n,p}$ by:

$$M^0_{n,p'} = n M(p') \sqrt{-2\epsilon'} = n M(p) \sqrt{-2\epsilon} \frac{M(p')}{M(p')} \sqrt{\frac{\sum p_i^2}{\sum p_i'^2}} \quad (30)$$

Any choice p' such that $p' = \alpha p_i$ should be discarded because, by Eq. (30) does not give anything new. If $p_i = \alpha p_i$; $M(p') = \alpha M(p)$ and $\mu^0_{n,p'} = \mu^0_{n,p}$.

For the two-electron problem the choices of "configurations" p are: (1, 1). (1, 2), (1, 3), . . . ; (2, 3), (2, 5), (2, 7), . . . ; (3, 4), (3, 5), (3, 7), . . . , etc. Each configuration will lead to a set of solutions by letting n assume the values $1, 2, 3, \ldots, \infty$.

The discussion of the operator $\Omega(\epsilon)$ should have the preceding as a starting point, but for our purposes it is sufficient to have found an unperturbed operator with a partially discrete spectrum, with known solutions and therefore tractable by methods of the partitioning technique.

The many center problem would require a similar discussion, but it is clear from the one-center case that for many-electron problems each unperturbed "state" is described by a product wavefunction in which each "orbital" or one-electron function depends explicitly on the quantum numbers of all the other electrons. Lundqvist[21] seems to have been the first to have noticed that a purely discrete spectrum is not obtained by the sort of transformation being considered. Yet, for a one-electron problem he obtained an expansion of the Green's function which apparently is not very well known. Due to the relation of Green's functions and the reduced resolvents of perturbation theory we present in the next section some expressions for R_0 , the reduced resolvent operator used in perturbation theory derived by the method of this paper.

RESOLVENTS FOR ONE-ELECTRON SYSTEMS

There are a considerable number of one-electron problems which are still of interest: perturbation treatments of Hartree-Fock equations or of atomic systems in the Coulomb approximation, to mention just two. In this connection the explicit solution of the inhomogeneous equation has proven to be very effective.[22] This method is of course equivalent to the use of the reduced resolvent operator R_0 [23]

$$R_0 = \sum_{K\neq 1}{}' \frac{|\varphi^0_K\rangle\langle\varphi^0_K|}{E^0_1 - E^0_K} \tag{31}$$

where

$$\mathcal{H}^0 \psi^0_K = E^0_K \varphi^0_K \tag{32}$$

When $\mathcal{H}^0$ is a hydrogenic Hamiltonian (31) is not very useful on account of the integration over the continuum spectrum involved in (31). Expansions in terms of complete discrete sets have been introduced several times.[24] and as pointed out by Hirschfelder et al in Appendix B of their review[22], the second order energy involves the inversion of a matrix and a double summation when an arbitrary set is employed in the computation. We wish to show that a simple expression for R_0 can be given, analogous to (31), involving a complete discrete set.

Consider (2) and (3) for a hydrogen-like ion:

$$U=0\,;\quad E^0_n = -Z^2/2n^2\,,\quad \mu = Z\,;\quad V = r^{-1}$$

Then

$$(\mathcal{E} - \mathcal{H}^0)^{-1} = V^{-1/2} [Z - \Omega^0(\mathcal{E})]^{-1} V^{-1/2} \tag{33}$$

In the previous section we discussed the Hermitean operator $\Omega^0(\mathcal{E})$. It is just a one particle operator with a complete, discrete spectrum, and the following relations hold, for $\mathcal{E} < 0$:

$$\Omega^0(\mathcal{E})\, \phi^0_n(\mathcal{E}) = z_n(\mathcal{E})\, \phi^0_n(\mathcal{E}) \tag{34a}$$

$$\hat{1} = \sum_{n=1}^{\infty} |\phi^0_n(\mathcal{E})\rangle\langle\phi^0_n(\mathcal{E})| \tag{34b}$$

$$\Omega^0(\mathcal{E}) = \sum_{n=1}^{\infty} |\phi^0_n(\mathcal{E})\rangle z_n(\mathcal{E}) \langle\phi^0_n(\mathcal{E})| \tag{34c}$$

$$[Z - \Omega^0(\mathcal{E})]^{-1} = \sum_{n=1}^{\infty} \frac{|\phi^0_n(\mathcal{E})\rangle\langle\phi^0_n(\mathcal{E})|}{Z - z_n(\mathcal{E})} \tag{34d}$$

The normalized eigenvalues are given by $z_n(\mathcal{E}) = n\sqrt{-2\mathcal{E}}$ and the normalized eigenfunctions are given by $\phi^0_n(\mathcal{E}) = V^{1/2} \bar{\varphi}^0_n(\mathcal{E})$ where $\bar{\varphi}^0_n(\mathcal{E})$ is explicitly defined by (21) with

$(n_i, \ell_i, m_i, \mathcal{E}_i) = (n, \ell, m, \mathcal{E})$.

Then it follows from (33) and (34d) that:

$$(\mathcal{E}-\mathcal{H}^0)^{-1} = \sum_{n=1}^{\infty} \frac{|\bar{\varphi}^0_n(\mathcal{E})\rangle\langle\bar{\varphi}^0_n(\mathcal{E})|V|\bar{\psi}^0_n(\mathcal{E})\rangle^{-1}\langle\bar{\varphi}^0_n(\mathcal{E})|}{Z - n\sqrt{-2\mathcal{E}}} \tag{35}$$

An expression of a similar nature was obtained by Lundqvist.[21] As it stands, (35) does not exist for $\mathcal{E}=E^0_k$. The reduced resolvent (31) can be written in a formal way:[4]

$$R_0 = P(E^0_1 - \mathcal{H}^0)^{-1}P; \quad P = \hat{1} - |\psi^0_1\rangle\langle\varphi^0_1| \tag{36}$$

and in this case, $\varphi^0_1 = \psi^0_1(z)$; $\mathcal{E} = E^0_1 = -z^2/2$:

$$R_0 = \sum_{n\neq 1} \frac{P|\bar{\varphi}^0_n(E^0_1)\rangle\langle\bar{\varphi}^0_n(E^0_1)|\frac{1}{2}|\bar{\varphi}^0_n(E^0_1)\rangle^{-1}\langle\bar{\varphi}^0_n(E^0_1)|P}{Z(1-n)} \tag{37}$$

This explicit form of the resolvent, of remarkable simplicity, leads to single sums of simple matrix elements, with no matrix inversions, for the second order energy associated with a perturbation V

$$\epsilon_2 = \langle\varphi^0_1(z)|V R_0 V|\varphi^0_1(z)\rangle \tag{38}$$

This explicit solution is particularly useful, since the integrals involved are in general not difficult to obtain. As an example we may point out that Linderberg[25] treated by the expansion method the Z^{-1} expression of the Hartree-Fock energies using as basis first order Laguerre functions $\{L_K^1(2\alpha r)e^{-\alpha r}\}$ which are essentially the functions given in (21) with $\ell=0$. It is not difficult to see that his results follow immediately from (38). Another example, of trivial nature, is the Stark effect, where inspection shows that there are only two nonvanishing terms remaining in (38), to yield of course the exact result, though it has been thought that the set (21) involving the Laguerre functions $\{L_{n+\ell}^{2\ell+1}(2\sqrt{2\epsilon}r)$ would not give a simple result.[26]

We are currently employing a modification of (37), appropriate to the treatment of H_2^+ by perturbation theory,[27] which avoids both matrix inversions and explicit solutions of the inhomogeneous equation.

This simple form of the resolvent does not hold for many-electron systems, but in the same way as in variational calculations[20] the set (21) simplifies very much the matrix of the resolvent if ϵ_ν is the same for all ν.

THE CONJUGATE PROBLEM AND THE SPECTRUM

We shall quote first two lemmas which are useful in what follows:[28]

Lemma 1: Let $\mathcal{H}$ be any self adjoint operator with the domain $\mathfrak{D}_{\mathcal{H}}$ and let $E(\lambda)$ be the corresponding resolution of the identity. If there is an N - dimensional manifold $\mathfrak{M}_N \subseteq \mathfrak{D}_{\mathcal{H}}$ such that $\langle f|\mathcal{H}|f\rangle \leq \mu \langle f|f\rangle$ for every $f \in \mathfrak{M}_N$, then dim $E(\mu) \geq N$

Lemma 2: Let $\mathcal{H}$ and $\mathcal{H}'$ be two self-adjoint operators with the domains $\mathfrak{D}_{\mathcal{H}}$ and $\mathfrak{D}_{\mathcal{H}'}$ and let $E(\lambda)$ and $E'(\lambda)$ be the corresponding resolutions of the identity. If $\mathfrak{D}_{\mathcal{H}} \subseteq \mathfrak{D}_{\mathcal{H}'}$ and $\langle f|\mathcal{H}|f\rangle \geq \langle f|\mathcal{H}'|f\rangle$ for all $f \in \mathfrak{D}_{\mathcal{H}}$ then dim $E(\lambda) \leq \dim E'(\lambda)$ for every λ

We denote the spectral resolutions as

$$\mathcal{H} = \int_{-\infty}^{\infty} \lambda \, dE(\lambda) \tag{39}$$

and dim $E(\mu)$ is the dimension number of the range of the projection $E(\mu)$. $\mathfrak{D}_{\mathcal{H}}$ is the domain of $\mathcal{H}$. These two lemmas encompass the variation principle as well as the ordering theorem in operator inequalities:[29]

Let Θ be an arbitrary projection operator:

$$\Theta^2 = \Theta; \quad \Theta^+ = \Theta; \quad Tr(\Theta) = N \tag{40}$$

for instance, the projection operator associated to a linear manifold $\mathfrak{f} = (f_1, f_2, \ldots, f_N)$:

$$\Theta = |\mathfrak{f}\rangle\langle\mathfrak{f}|\mathfrak{f}\rangle^{-1}\langle\mathfrak{f}| \tag{41}$$

The "outer projection" $\bar{A}$, of a self-adjoint operator A with respect to the manifold $\mathcal{M}$ is given by

$$\bar{A} = \Theta A \Theta \; ; \; \bar{A}\bar{u}_\kappa = \bar{a}_\kappa \bar{u}_\kappa \tag{42}$$

and the eigenvalues of $\bar{A}$ are upper bounds to the eigenvalues of A in order:

$$a_\kappa \leq \bar{a}_\kappa \tag{43}$$

Furthermore using the notation $A \leq B$ to denote that $\langle f|A|f\rangle \leq \langle f|B|f\rangle$ for $f \in \mathcal{D}_B$ and $\mathcal{D}_B \subseteq \mathcal{D}_A$ one has the inequality for the eigenvalues:

$$a_\kappa \leq b_\kappa \tag{44}$$

in order.

Joseph[12] made an efficient use of lemma 2 after explicitly solving the eigenvalue problem of the conjugate problem associated to a class of one-body potentials. This allowed the determination of the exact number of bound states of the system. The reasoning involved is quite simple. Consider (3), and suppose that for a given $\mathcal{E}$:

$$\mu_1(\mathcal{E}) \leq \mu_2(\mathcal{E}) \leq \ldots . \leq \mu_p(\mathcal{E}) < \mu < \mu_{p+1}(\mathcal{E}) \tag{45}$$

Then, according to (10) and (11) the eigenvalue problem (2) has the solutions

$$E_1(\mu) \leq E_2(\mu) \leq \ldots . \leq E_p(\mu) < \mathcal{E} < E_{p+1}(\mu) \tag{46}$$

Notice also that $\mu = \mu(E_p(\mu)) = \mu_{p+1}(E_{p+1}(\mu))$.
Then

$$\mu_p(\varepsilon) < \mu = \mu_p(E_p(\mu))$$

implies, on account of (4), that:

$$E_p(\mu) < \varepsilon \tag{47}$$

and also

$$\mu = \mu_{p+1}(E_{p+1}(\mu)) < \mu_{p+1}(\varepsilon)$$

leads to

$$\varepsilon < E_{p+1}(\mu) \tag{48}$$

Of course this assumes continuity of $\mu_n(\varepsilon)$.

In the case of one-body potentials, one can set $\varepsilon = 0$ and thus count the number of bound states, given by p provided one can solve (3) exactly. We refer to Joseph for an excellent account.[12]

If (3) cannot be solved exactly we would be at a loss unless approximate solutions do help. But we saw in (10) and (11) that upper and lower bounds to the $\mu_n(\varepsilon)$, the $\mu^{>}_n(\varepsilon)$ and $\mu^{<}_n(\varepsilon)$, yielded upper and lower bounds to the $E_n(\mu)$. In addition we may state that if the set $\{\mu^{>}_1(\varepsilon), \ldots, \mu^{>}_q(\varepsilon)\}$ is a set of upper bounds to the $\mu_n(\varepsilon)$ obtained from an outer projection on an arbitrary manifold, and which satisfies

$$\mu^{>}_1(\varepsilon) \leq \mu^{>}_2(\varepsilon) \leq \ldots \leq \mu^{>}_q(\varepsilon) < \mu. \tag{49}$$

it follows that $q \leq p$, i. e. it is a lower bound to the number of

states whose energies satisfy $E_1(\mu) \leq \ldots \leq E_p(\mu) < \mathcal{E}$.
Also, if a set of lower bounds could be obtained, $\{\mu_1^<(\mathcal{E}), \ldots, \mu_2^<(\mathcal{E})\}$ with the property

$$\mu_1^<(\mathcal{E}) \leq \mu_2^<(\mathcal{E}) \leq \ldots \leq \mu_{2-1}^<(\mathcal{E}) < \mu < \mu_2^<(\mathcal{E}) \quad (50)$$

the two aforesaid lemmas imply that r is an upper bound to the number of states with energy less than $\mathcal{E}$:

$$q \leq p \leq r \quad (51)$$

Clearly, the whole approach hinges on both the existence of solutions, the continuity of the $\mathcal{E}$- dependence and on the possibility of computing the bounds to the $\mu_n(\mathcal{E})$. But this last part, at least formally, is perhaps simpler than to compute bounds to the $\mu_n(\mathcal{E})$. We shall draw special advantage from the positive character of $\Omega(\mathcal{E})$. We first review the basic properties of inner-projections.

A general discussion of the subject is available.[3] We follow here the treatment of Löwdin,[4, 6, 29] and we note that the infinite set $\{f_1, f_2, \ldots\}$ is complete if, and only if, the associated projection operator, (41), is the identity operator. Furthermore the projection operator Θ satisfies the operator inequality.

$$\hat{0} \leq \Theta \leq \hat{1} \quad (52)$$

Consider now a self-adjoint positive-definite operator A; $A > 0$. Its eigenvalues are all positive, and the positive square root can be defined in terms of the spectral resolution:

$$A^{1/2} = \sum_\kappa |u_\kappa\rangle a_\kappa^{1/2} \langle u_\kappa|$$

Which is a symbolic way to denote

$$A^{1/2} = \int_{-\infty}^{\infty} x^{1/2} dE(x)$$

If A is just a positive function the square root $A^{\frac{1}{2}}$ exists trivially and in either case it follows from (52) by multiplying on both sides by $A^{\frac{1}{2}}$ that:

$$\hat{0} \leq A^{1/2}\Theta A^{1/2} \leq A \tag{53}$$

The operator $A' = A^{1/2}\Theta A^{1/2}$ is called the "inner projection" by A by Θ .[29] Since Θ is <u>any</u> projection operator in the space being considered we may write for an arbitrary linear manifold:

$$A \geqslant A' = A^{1/2}|f\rangle\langle f|f\rangle^{-1}\langle f|A^{1/2} \geqslant 0 \tag{54}$$

which is not a very useful form. But by using the manifolds g and h characterized by $f = A^{1/2}g$ and $f = A^{-1/2}h$ respectively one obtains:

$$A \geqslant A' = A|g\rangle\langle g|A|g\rangle^{-1}\langle g|A \geqslant 0 \tag{55}$$

i. e. the so called Aronszajn projection and

$$A \geqslant A' = |h\rangle\langle h|A^{-1}|h\rangle^{-1}\langle h| \tag{56}$$

the Bazley projection.[29] The three forms (54) (55) and (56) are equivalent except for questions of the domains of the operators involved. The last two are particularly convenient for computation purposes. In the partitioning technique these inequalities are used to construct lower bounds to the reaction operator when it is positive definite. If we denote

the eigenvalues of A by a_K, the lemmas from the previous section and in particular the inequality (44), together with (54) - (56) imply that

$$a_K' \leq a_K \tag{57}$$

in order. Notice now that we could not try to apply this technique directly to the Hamiltonian (2), since it is not positive definite. We could subtract a lower bound to it, in order to construct a positive operator. On the other hand, for $\varepsilon < 0$, the operator $\Omega(\varepsilon)$, given in (6) is positive definite. We can apply to it the intermediate problems approach [3] or the partitioning technique [4] and this is perhaps the most interesting aspect of the conjugate eigenvalue problem.

UPPER AND LOWER BOUNDS

Our purpose here is to show how upper and lower bounds to the $\mu_n(\varepsilon)$ can be obtained. Consider the eigenvalue problem associated with $\Omega(\varepsilon)$

$$\Omega(\varepsilon)\,\phi_n(\varepsilon) = \mu_n(\varepsilon)\,\phi_n(\varepsilon) \tag{58}$$

where $\Omega(\varepsilon)$ is given by (6) and $\Omega(\varepsilon) > 0$ for $\varepsilon < 0$. Regardless of the nature of the problem, approximate upper bounds, follow from the variation principle, i. e. from outer projections. So, introducing the projection operator (41) we define:

$$\bar{\Omega}(\varepsilon) \equiv |f\rangle\langle f|f\rangle^{-1}\langle f|\,\Omega(\varepsilon)\,|f\rangle\langle f|f\rangle^{-1}\langle f| \tag{59}$$

The eigenvalues of $\bar{\Omega}(\varepsilon)$, the $\mu_n^0(\varepsilon)$ are obtained from solving a secular equation:

$$\left| \langle \phi | \Omega(\varepsilon) | \phi \rangle - \mu^{\circ}(\varepsilon) \langle \phi | \phi \rangle \right| = 0 \tag{60}$$

Since the set ϕ is arbitrary we can introduce the transformation $V^{-1/2}\phi = i$ and from (6) it follows that the eigenvalues are given by the roots of:

$$\left| \langle i | T + U - \varepsilon | i' \rangle - \mu^{\circ}(\varepsilon) \langle i | V | i' \rangle \right| = 0 \tag{61}$$

or, using (1):

$$\left| \langle i | \mathcal{H} - \varepsilon + (\mu - \mu^{\circ}(\varepsilon)) V | i' \rangle \right| = 0 \tag{62}$$

where it should be remembered that ε is a given parameter. The similarity with the conventional secular equation for the eigenvalue problem of $\mathcal{H}$ is apparent.

On the other hand lower bounds to the $\mu_n(\varepsilon)$ are more difficult to obtain, as it could be expected. One could try to take advantage of the fact that $\Omega(\varepsilon)$ is positive definite and apply the projection technique to it, but this must be done with some care.

Let us write (6) in the form:

$$\Omega(\varepsilon) = V^{-1/2}\, \omega(\varepsilon)\, V^{-1/2} \tag{63}$$

$\omega(\varepsilon)$ is of course positive definite. Then:

$$\Omega(\varepsilon) \geqslant \Omega'(\varepsilon) \equiv V^{-1/2} \omega^{1/2}(\varepsilon)\, \theta\, \omega^{1/2}(\varepsilon) V^{-1/2} \geqslant 0 \tag{64}$$

The eigenvalues of $\Omega'(\varepsilon)$, the $M_n^L(\varepsilon)$ are lower bounds, in order, to the eigenvalues of $\Omega(\varepsilon)$ The eigenvalue problem of $\Omega'(\varepsilon)$ is not easily solvable in general. Consider yet the transformation $V^{-1/2}\omega(\varepsilon)^{1/2}\Phi = \mathcal{F}$ Then $\Omega'(\varepsilon)$ can be written in the form:

$$\Omega'(\varepsilon) = |\mathcal{F}\rangle\langle\mathcal{F}| V^{1/2}\omega(\varepsilon)^{-1}V^{1/2}|\mathcal{F}\rangle^{-1}\langle\mathcal{F}|$$

or simply:

$$\Omega'(\varepsilon) = |\mathcal{F}\rangle\langle\mathcal{F}|\Omega^{-1}(\varepsilon)|\mathcal{F}\rangle^{-1}\langle\mathcal{F}| \tag{64}$$

and the eigenvalue problem is solvable:

$$|\langle\mathcal{F}|\Omega'(\varepsilon)|\mathcal{F}\rangle^{-1} - M^L(\varepsilon)\langle\mathcal{F}|\mathcal{F}\rangle| = 0 \tag{65}$$

provided the matrix elements can be evaluated. This is by no means a trivial point, but it is interesting nevertheless to compare, at least formally the secular equations for the upper and the lower bounds, given by (60) and (65) respectively.

Instead of proceeding this way, a division of $\Omega(\varepsilon)$ into two parts:

$$\Omega(\varepsilon) = \Omega^0(\varepsilon) + \Omega^1(\varepsilon) \tag{66}$$

where $\Omega^0(\varepsilon)$ has a known spectrum and $\Omega^1(\varepsilon)$ is positive would allow the use of the conventional techniques of truncations and projections as in the method of intermediate operators[3, 30] or of the partitioning technique,[4] The operator $\Omega^0(\varepsilon)$ can be defined by (15) and $\Omega^1(\varepsilon)$ chosen to be the interelectronic repulsion U which is positive definite.

The work by Walmsley and Coulson[10] on H_2^+ indicates that the conjugate transformation increases the transformation increases the convergence properties of the truncation method in molecular systems[5] applied in its original form by Johnson and Coulson.[7]

The use of the resolvents (35), (37) can make the treatment more efficient.[27] For many-particle problems the eigenvalue transformation discussed can lead to useful results. The Sturmian functions used by Rotenberg[31] in the treatment of electron-hydrogen scattering are clearly related to the solutions of (19).

DISCUSSION

We have reviewed and derived some properties of a class of conjugate problems arising in the theory of upper and lower bounds to eigenvalues of atomic and molecular systems.

The procedure employed, namely to use the "strength" of a perturbation as eigenvalue, goes back to Hylleraas,[9] The treatment of the change as the eigenvalue has recently been discussed by Bazley and Fox,[32] who pointed out that the transformation to a conjugate problem does not lead to a purely discrete spectrum in many atomic or molecular systems. Lundqvist[21] had examined the possibility of finding a transformation that yielded a purely discrete spectrum for the Green's function and foresaw the difficulties arising in many electron systems.

We have explicitly considered the many-electron conjugate problem and verified the complex nature of the solutions. Lundqvist's results are rederived as well as expressions for the reduced and the full resolvents, R_0 and $(\mathcal{E}-\mathcal{H}^0)^{-1}$ in terms of discrete sums. The second order energy, when $\mathcal{H}^0$ is a hydrogen-like Hamiltonian involves then a single sum without an integration over the continuum. In some cases this leads to an alternative way of solving the inhomogeneous equations of perturbation theory. Our formulae lead to Linderberg's[25] results for the $(Z)^{-1}$ expansion of Hartree-Fock energies.

In combining the projection technique for upper and lower bounds with the conjugate eigenvalue problems, we are able to extend some of the results of Joseph,[12] on the counting of the eigenvalues of given one-particle potentials, to arbitrary problems. Approximate techniques are of course needed.

We expect to present applications of these formulae shortly, but some of the present results are perhaps interesting as such, e. g. the fact that in an arbitrary system the bound state energies decrease when the "strength" of the potential increases.

BIBLIOGRAPHY

1. E. A. Hylleraas and B. Undheim, Z. Phys. 65, 759 (1930)
   J. K. McDonald, Phys. Rev. 43, 830 (1933)
2. G. Temple, Proc. Roy. Soc. (London) A119, 276 (1928)
3. For an excellent account and bibliography see S. H. Gould, "Variational methods for eigenvalue problems" (University of Toronto Press, Toronto, 1966)
4. A review by P. O. Löwdin in "Perturbation Theory and its Applications to Quantum Mechanics" (Ed. C. H. Wilcox, John Wiley and Sons, New York, 1966)
5. N. W. Bazley and D. W. Fox, J. Math. Phys. 4, 1147 (1963)
6. See discussion in P. O. Löwdin, J. Chem. Phys. 43, S175 (1965)
7. B. P. Johnson and C. A. Coulson, Proc. Phys. Soc. 84, 263 (1964)
8. E. Schrödinger, Ann. Phys. 80, 437 (1926)
9. E. A. Hylleraas, Z. Phys. 48, 469 (1928)
10. M. Walmsley and C. A. Coulson, Proc. Camb. Phil. Soc. 62, 769 (1966)
11. O. Goscinski, University of Florida Quantum Chemistry Group, Technical Note 82, April 1, 1966 (unpublished)
12. A. Joseph, Int. J. Quantum Chemistry, 1, 615 (1967)
13. H. Hellman, "Quantenchemie" (Deuticke Co. Leipzig, 1937) p. 285. P. R. Feynman, Phys. Rev. 56, 340 (1939).
14. G. C. Ghirardi and A. Rimini, J. Math. Phys. 6, 40 (1965)
15. P. M. Morse and H. Feshbach, "Methods of Theoretical Physics" (McGraw Hill Co., New York, 1953). Vol. 2, Chapter 9.
16. J. C. Slater, "Quantum Theory of Atomic Structure (McGraw Hill Co., New York 1960), Vol. 2.
17. E. Schrödinger, Ann. Phys. 79, 361 (1962).
    E. A. Hylleraas, Z. Physik 48, 469 (1928); L. Pauling and J. J. Beach, Phys. Rev. 47, 686 (1935).
    E. A. Hylleraas, Institute for Theoretical Physics, University of Oslo, Preprint No. 1, 1961. (unpublished)

21. S. O. Lundqvist, Univ. of Uppsala Quantum Chemistry Group, Technical Note 49, April 15, 1960. (unpublished)

22. J. C. Slater and Kirkwood, Phys. Rev. A. Dalgarno and J. T. Lewis, Proc. Roy. Soc. (London) A233, 70 (1955); for a good review of this technique see J. O. Hirschfelder, W. B. Brown and S. T. Epstein in "Advances in Quantum Chemistry". (Academic Press Inc. New York 1964), Vol. 1

23. P. O. Löwdin, J. Math. Phys. 6, 1341 (1965).

24. P. Epstein, Phys. Rev. 28, 695 (1926); J. E. Lennard Jones, Proc. Roy. Soc. (London) A129, 598 (1930); P. O. Löwdin and H. Shull, J. Chem. Phys. 23, 1362 (1955).

25. J. Linderberg, Phys. Rev. 121, 816 (1961)

26. E. Steiner, J. Chem. Phys. 37, 454 (1962)

27. E. Brändas and O. Goscinski, to be published.

28. T. Kato, Trans. Am. Math. Soc. 70, 195 (1951); for a useful and compact review of definitions and properties of Hamiltonian operators see T. Kato, Suppl. Prog. Theor. Phys. 40, 3 (1967)

29. P. O. Löwdin, Phys. Rev. 139, A357 (1965).

30. N. Bazley and D. W. Fox, "Comparison Operators for Lower Bounds to Eigenvalues." Technical Report, July 1963, Batelle Memorial Institute, Geneva.
M. Rotenberg, An Phys. 19, 262 (1962)
N. Bazley and D. W. Fox, "Lower Bounds to Eigenvalues of Quantum Mechanical Systems", Mathematics Report No. 13, Battelle Inst. Geneva 1967.

Approximate propagator for the radial Dirac problem

Jan Linderberg
Department of Chemistry
Aarhus Universitet
DK-8000 Aarhus C, Denmark

ABSTRACT

Two coupled first order differential equations derived for the atomic central field problem within the relativistic framework are transformed to integral equations through the use of approximate Wentzel-Kramers-Brillouin solutions. It is shown that a finite charge density can be derived for a relativistic form of the Fermi-Thomas atomic model by appropriate attention to the boundary conditions. A numerical solution for the effective nuclear charge in the Xenon atom is calculated and fitted to a rational expression.

CONTENTS

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

1. INTRODUCTION

Per-Olov Löwdin had a long and lasting interest in the analytical methods of quantum mechanics and my tribute to his legacy involves an application of the Wentzel-Kramers-Brillouin (WKB) asymptotic approximation method. It was the subject of a contribution(1) by Löwdin to the Solid State and Molecular Theory Group created by John C. Slater at the Massachusetts Institute of Technology.

Our esteemed teacher looked with skepticism on the efforts by Yngve Öhrn and myself to make use of propagators but learned to accept and appreciate our results. We offered a derivation of the Fermi-Thomas central field model for atoms through a WKB-approach to the radial propagator in our book(2). I developed an integral equation for the general multichannel scattering problem(3) from the WKB solutions and have nourished a feeling that a similar attack could be used for the relativistic, two component, radial Dirac central field case. This paper delineates such an approach.

The next section establishes the basic formulation and is followed by the transformation to an integral equation form by means of the WKB-type solutions where all terms are retained. It is demonstrated that the quantization condition derives from the reciprocal kernel of the integral equation in the fourth part. There follows a derivation of the electron and energy densities for the relativistic central field model where it is shown that a Fermi-Thomas type screening function can be obtained without the disturbing singularity at the origin by observing proper boundary conditions. An example solution is computed for the Xenon atom and presented in the sixth section. A short review of the attempts towards a relativistic Fermi-Thomas theory concludes this paper.

2. PRELIMINARIES

Condon and Shortley wrote *the* book(4) on atomic theory and I will use their equations for the radial components of the relativistic wave function for an electron in a central field(5). Thus I quote

$$\begin{aligned} -\mu c^2 F(r) + c\hbar\left(\frac{dG(r)}{dr} + \frac{k}{r}G(r)\right) + U(r)F(r) &= WF(r); \\ +\mu c^2 G(r) - c\hbar\left(\frac{dF(r)}{dr} - \frac{k}{r}F(r)\right) + U(r)G(r) &= WG(r); \end{aligned} \tag{1}$$

The quantum number k takes positive and negative integer values, the appropriate electron mass parameter is given by μ, the energy parameter is W, and the central field potential energy is

$$U(r) = -\frac{e^2 Z(r)}{r} \tag{2}$$

where a general screening function is used instead of the pure Coulomb field from Condon and Shortley. The radial amplitudes, F and G, are the "small" and "large" components respectively. They define the radial density distribution as

$$\rho(r)dr = \left(|F(r)|^2 + |G(r)|^2\right)dr \tag{3}$$

and a state has the degeneracy is $2|k| = 2j + 1$.

Early applications of WKB approximations to the Coulomb problem in Schrödinger theory demonstrated the necessity and expediency of the Kramers modification(6):

$$\ell(\ell+1) \Rightarrow \left(\ell + \frac{1}{2}\right)^2 \tag{4}$$

It holds that

$$\ell = \begin{cases} k, & k > 0 \\ -k-1 & k < 0 \end{cases} \tag{5}$$

so that

$$\ell(\ell+1) = k(k+1) \Rightarrow \left(k+\tfrac{1}{2}\right)^2 \tag{6}$$

is the corresponding replacement for the Dirac case. Thus it is useful to introduce an auxiliary variable through the replacement

$$k = r\kappa(r) - \tfrac{1}{2} \tag{7}$$

Another substitution serves to simplify the equations. The combination W-U defines an additional function:

$$h(r) = \frac{1}{c}\big(W - U(r)\big) = \frac{1}{c}\left(W + \frac{e^2 Z(r)}{r}\right) \tag{8}$$

Atomic units, where μ, $\hbar$, and e equal unity, will be used. The basic equations will be used in the inhomogeneous form in order to prepare for the Green function formulation:

$$\begin{aligned} &\big[h(r)+c\big]F(r) - \frac{dG(r)}{dr} - \kappa(r)G(r) = \Phi(r) - \frac{G(r)}{2r}; \\ &\frac{dF(r)}{dr} - \kappa(r)F(r) + \big[h(r)-c\big]G(r) = -\frac{F(r)}{2r} + \Gamma(r); \end{aligned} \tag{9}$$

Notations Φ and Γ are used for the source functions. These coupled first order differential equations will be transformed into integral equations with a WKB approach.

3.TOWARDS INTEGRAL EQUATIONS

Approximate solutions to the homogeneous equations on the left-hand side of Eq. (9) are constructed in the WKB manner as

$$\begin{Bmatrix} f^{+}(r) = \dfrac{1-i}{2}\left[\dfrac{\kappa(r)+ip(r)}{\sqrt{p(r)\left(h(r)+c\right)}}\right]\exp\left[is(r)\right] \\ g^{+}(r) = \dfrac{1-i}{2}\left[\dfrac{h(r)+c}{\sqrt{p(r)\left(h(r)+c\right)}}\right]\exp\left[is(r)\right] \end{Bmatrix}$$
$$\begin{Bmatrix} f^{-}(r) = \dfrac{1-i}{2}\left[\dfrac{\kappa(r)-ip(r)}{\sqrt{p(r)\left(h(r)+c\right)}}\right]\exp\left[-is(r)\right] \\ g^{-}(r) = \dfrac{1-i}{2}\left[\dfrac{h(r)+c}{\sqrt{p(r)\left(h(r)+c\right)}}\right]\exp\left[-is(r)\right] \end{Bmatrix} \tag{10}$$

The phase integral $s(r)$ has the derivative $p(r)$ and it satisfies the secular equation

$$\left|\begin{Bmatrix} h(r)+c & -\kappa(r)-ip(r) \\ -\kappa(r)+ip(r) & h(r)-c \end{Bmatrix}\right| = h^2(r)-c^2-\kappa^2(r)-p^2(r)=0 \tag{11}$$

General complex energy parameters W will be considered and it is useful to choose the branch of the function $p(r)$ in the upper part of the complex plane. This means that the phase integral $s(r)$ has an imaginary part which is increasing with r. Accordingly it holds that the upper set of functions of Eq. (10) vanish at large values of r while the lower set go to zero at the origin. The limiting form of $p(r)$ at small r-values is seen to be

$$p(r) \xrightarrow[r\to 0]{} \frac{i}{r}\sqrt{\left(k+\tfrac{1}{2}\right)^2-\left[Z(0)/c\right]^2} \tag{12}$$

and we conclude that

$$\left|k+\tfrac{1}{2}\right| \geq Z(0)/c \tag{13}$$

This relation will prove to be significant later.

Two auxiliary constructs are defined as

$$\begin{aligned} X(r) &= g^{+}(r)F(r) - f^{+}(r)G(r); \\ Y(r) &= g^{-}(r)F(r) - f^{-}(r)G(r); \end{aligned} \tag{14}$$

The inverse relation is

$$\begin{aligned} F(r) &= f^{+}(r)Y(r) - f^{-}(r)X(r); \\ G(r) &= g^{+}(r)Y(r) - g^{-}(r)X(r); \end{aligned} \tag{15}$$

and the reason for the rather arbitrary complex factor in Eq. (10) becomes clear. Differentiation of the functions (15) and usage of the basic equations (9) gives, after some algebra with primes indicating derivatives, that

$$\begin{aligned} \frac{dX(r)}{dr} &= g^{+}(r)\left\{\frac{h'(r)}{2[h(r)+c]} - \frac{p'(r)}{2p(r)} - \frac{1}{2r}\right\}F(r) \\ &+ f^{+}(r)\left\{\frac{h'(r)}{2[h(r)+c]} - \frac{p(r)+i\kappa(r)}{p(r)-i\kappa(r)}\left[\frac{p'(r)}{2p(r)} + \frac{1}{2r}\right]\right\}G(r) \\ &+ f^{+}(r)\Phi(r) + g^{+}(r)\Gamma(r); \\ \frac{dY(r)}{dr} &= g^{-}(r)\left\{\frac{h'(r)}{2[h(r)+c]} - \frac{p'(r)}{2p(r)} - \frac{1}{2r}\right\}F(r) \\ &+ f^{-}(r)\left\{\frac{h'(r)}{2[h(r)+c]} - \frac{p(r)-i\kappa(r)}{p(r)+i\kappa(r)}\left[\frac{p'(r)}{2p(r)} + \frac{1}{2r}\right]\right\}G(r) \\ &+ f^{-}(r)\Phi(r) + g^{-}(r)\Gamma(r); \end{aligned} \tag{16}$$

The significance of the Kramers correction is evident from the form with the logarithmic derivative of $p(r)$. Its singularity at the origin is cancelled by the Kramers term. The logarithmic derivative of the function $c+h(r)$ is small except at the origin and its effect will be neglected here. Essential singularities occur for real energy parameter values at the so-called classical turning points where $p(r)$ goes through zero. Their treatment lead to the Born-Sommerfeld quantization condition, the subject of the next section.

Boundary conditions requires that

$$X(r) = -\int_r^\infty dr' X'(r'); \quad Y(r) = \int_0^r dr' Y'(r'); \tag{17}$$

and combination with Eq. (15) provides the integral equation form

$$\begin{bmatrix} F(r) \\ G(r) \end{bmatrix} = \int dr' \begin{bmatrix} f^+(r_>) f^-(r_<) & f^+(r_>) g^-(r_<) \\ g^+(r_>) f^-(r_<) & g^+(r_>) g^-(r_<) \end{bmatrix} \begin{bmatrix} \Phi(r') \\ \Gamma(r') \end{bmatrix}$$
$$+ \int dr' \begin{bmatrix} f^+(r_>) f^-(r_<) & f^+(r_>) g^-(r_<) \\ g^+(r_>) f^-(r_<) & g^+(r_>) g^-(r_<) \end{bmatrix} \begin{bmatrix} 0 & \tilde{n}(r') \\ n(r') & 0 \end{bmatrix} \begin{bmatrix} F(r') \\ G(r') \end{bmatrix}$$
$$n(r) = \frac{h'(r)}{2[h(r)+c]} - \frac{p'(r)}{2p(r)} - \frac{1}{2r} \tag{18}$$
$$r_> = \max(r, r')$$
$$r_< = \min(r, r')$$

Equation (16) contains the information to relate the function $\tilde{n}(r)$ to $n(r)$ and it involves a factor, which changes when r' equals r. There seems to be no convenient way to express this in a compact way. Its importance arises in connection with the derivation of the Bohr-Sommerfeld quantization rule.

The matrix kernel in the integral equation (18) defines the zeroth order propagator or Green function for the central field Dirac problem:

$$\mathbf{G}_0(r, r'; W) = \frac{1}{c} \begin{bmatrix} f^+(r_>) f^-(r_<) & f^+(r_>) g^-(r_<) \\ g^+(r_>) f^-(r_<) & g^+(r_>) g^-(r_<) \end{bmatrix} \tag{19}$$

The energy parameter W is indicated explicitly and the notation for the smaller and larger of the radial arguments is used.

4. BOHR-SOMMERFELD QUANTIZATION

The zeroth order propagator has no isolated poles but exhibits a cut along the real energy axis for radial values where $p(r)$ takes real values. A pole like structure is obtained from a Fredholm analysis of the integral equation(7). The only terms of concern arise from the logarithmic derivative of $p(r)$. It is assumed that there are two values of r where $p(r)$ has zeroes for a given real energy. The addition of a small imaginary part to the energy causes small displacements of the zeroes and it is concluded that the forms

$$p(r) \cong \begin{cases} \sqrt{\alpha(r-a+i\eta)}, & |r-a| \text{ small} \\ \sqrt{\beta(b+i\eta'-r)}, & |b-r| \text{ small} \end{cases} \tag{20}$$

apply. The η's are positive when the energy parameter is in the upper half of the complex plane and negative otherwise.

The relevant matrix kernel from the second term on the right hand side of Eq. (18) is expressed in terms of a couple of modified functions:

$$\begin{aligned} \tilde{f}^{+}(r) &= -\frac{1-i}{2}\left[\frac{\kappa(r)-ip(r)}{\sqrt{p(r)(h(r)+c)}}\right]\exp[is(r)] \\ \tilde{f}^{-}(r) &= -\frac{1-i}{2}\left[\frac{\kappa(r)+ip(r)}{\sqrt{p(r)(h(r)+c)}}\right]\exp[-is(r)] \end{aligned} \tag{21}$$

Thus we define

$$\mathbf{Q}(r,r') = \begin{cases} \begin{bmatrix} f^{+}(r)g^{-}(r') & f^{+}(r)\tilde{f}^{-}(r') \\ g^{+}(r)g^{-}(r') & g^{+}(r)\tilde{f}^{-}(r') \end{bmatrix} & r > r' \\ \begin{bmatrix} f^{-}(r)g^{+}(r') & f^{-}(r)\tilde{f}^{+}(r') \\ g^{-}(r)g^{+}(r') & g^{-}(r)\tilde{f}^{+}(r') \end{bmatrix} & r < r' \end{cases} \tag{22}$$

and consider the homogeneous integral equation

$$\begin{bmatrix} F(r) \\ G(r) \end{bmatrix} = \int dr' \mathbf{Q}(r,r') \begin{bmatrix} -\frac{p'(r')}{2p(r')} \end{bmatrix} \begin{bmatrix} F(r') \\ G(r') \end{bmatrix} \tag{23}$$

It should be appreciated here that the singularities at $r'=a$ and $r'=b$ makes the kernel of this equation practically separable of rank two. Fredholm theory shows that a solution to the homogeneous equation (23) requires a singular reciprocal kernel. Thus it holds that

$$\begin{aligned} 0 = 1 &- \int dr \left[Tr\mathbf{Q}(r,r) \frac{p'(r)}{2p(r)} \right] \\ &+ \frac{1}{2} \int dr dr' \left[Tr\mathbf{Q}(r,r) \frac{p'(r)}{2p(r)} \right] \left[Tr\mathbf{Q}(r',r') \frac{p'(r')}{2p(r')} \right] \\ &- \frac{1}{2} \int dr dr' \left[\frac{p'(r)}{2p(r)} Tr\mathbf{Q}(r,r') \mathbf{Q}(r',r) \frac{p'(r')}{2p(r')} \right] + \ldots \end{aligned} \tag{24}$$

The second and the third term on the right hand side equal zero since the trace of the matrix kernel $\mathbf{Q}(r,r)$ vanishes. The next term reduces when the explicit form is used,

$$Tr\mathbf{Q}(r,r')\mathbf{Q}(r',r) = -\exp\left[2is(r_>) - 2is(r_<)\right] \tag{25}$$

so that the essential condition is

$$0 = 1 + \int_0^\infty dr \exp\left[2is(r)\right] \frac{p'(r)}{2p(r)} \int_0^r dr' \exp\left[-2is(r')\right] \frac{p'(r')}{2p(r')} + \ldots \tag{26}$$

Approximate evaluation of the integrals by means of the saddle point method8 will be used. The integration over a small interval around the singular point at $r=a$ comes out as follows(3):

$$\begin{aligned} \int_{r\approx a} dr \exp\left[-2is(r)\right] \frac{p'(r)}{2p(r)} &= \int_{r\approx a} dr \exp\left[-2is(a) - \tfrac{4}{3} i \sqrt{\alpha (r-a+i\eta)^3} \right] \frac{1}{4(r-a+i\eta)} \\ &= -i \operatorname{sgn}(\eta) \sqrt{\frac{\pi}{12}} \exp\left[\tfrac{2}{3} - 2is(a)\right] \end{aligned} \tag{27}$$

A similar result for the other singular point gives the result that

$$\begin{aligned} 0 &= 1+\exp\left[2is(b)-2is(a)+\frac{4}{3}-\ln\left(\frac{\pi}{12}\right)\right]+\ldots \\ &= 1+\exp\left[2is(b)-2is(a)-0.0068434\ldots\right]+\ldots \end{aligned} \tag{28}$$

and we learn that the Fredholm determinant is small when the Bohr-Sommerfeld quantization condition is satisfied:

$$2s(b)-2s(a)=2\int_a^b drp(r)=(2\nu+1)\pi \tag{29}$$

Our analysis demonstrates that the integral equation approach offers an alternative to the differential equation methods for connection formulas across the turning points. The degeneracy between states with quantum number k and those with $-k-1$ should be noticed here. The soluble case with a pure Coulomb field will not give the accurate energy eigenvalues here as is the case in the Schrödinger form. The quantization condition (29) gives the result

$$\begin{aligned} W &= c^2\left\{1+\frac{(Z/c)^2}{\left[\nu+\frac{1}{2}+\sqrt{\left(k+\frac{1}{2}\right)^2-(Z/c)^2}\right]^2}\right\}^{-\frac{1}{2}} \\ &= c^2-\frac{Z^2}{2n^2}-\frac{Z^4}{2c^2n^4}\left(\frac{n}{\left|k+\frac{1}{2}\right|}-\frac{3}{4}\right)\ldots \\ n &= \nu+\frac{1}{2}+\left|k+\frac{1}{2}\right| \end{aligned} \tag{30}$$

where the second order term differs from the correct result on account of the one half from the Kramers modification.

5. ELECTRON DENSITY DISTRIBUTION

Propagator theory(2) provides an expression for the electron density in terms of an integral over the residues of the Green function at the poles corresponding to the occupied levels. Coulson suggested the equivalent procedure(9) of a contour integration in the complex energy plane. The present case calls for a contour that encloses the part of the spectrum that corresponds to bound electron states so it should include the part of the real axis between 0 and c^2. States with quantum number k are degenerate and thus the total radial density emanating from these states equals

$$\begin{aligned}\rho_k(r) &= \frac{2|k|}{2\pi i}\oint dW Tr\mathbf{G}_0(r,r;W) \\ &= \frac{2|k|}{2\pi c i}\oint dW\left[f^+(r)f^-(r)+g^+(r)g^-(r)\right] \\ &= -\frac{|k|}{\pi c}\oint dW\frac{h(r)}{p(r)} = -\frac{|k|}{\pi c}\oint dp\end{aligned} \tag{31}$$

An illustration of the path in the complex energy plane and its mapping onto the complex p-plane is offered in Fig. 1.

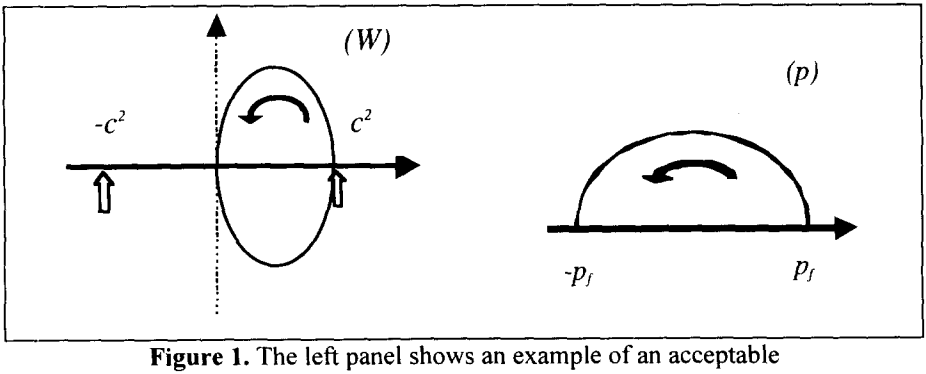


Figure 1. The left panel shows an example of an acceptable contour in the energy plane and the right panel gives its mapping onto the p-plane.

The integration gives the result

$$\rho_k(r) = \frac{2|k|}{\pi}\sqrt{\left[c + \frac{Z(r)}{cr}\right]^2 - c^2 - \left(\frac{k+1/2}{r}\right)^2} \tag{32}$$

The total radial density is obtained from a summation over quantum numbers k and its approximation with an integral:

$$\begin{aligned}
\rho(r) &= \Sigma_k \rho_k(r) = \Sigma_j\left[\rho_{-j-1/2}(r) + \rho_{j+1/2}(r)\right] \\
&= \Sigma_j \frac{2j+1}{\pi}\sqrt{\left[c + \frac{Z(r)}{cr}\right]^2 - c^2 - \left[\frac{j}{r}\right]^2} \\
&+ \Sigma_j \frac{2j+1}{\pi}\sqrt{\left[c + \frac{Z(r)}{cr}\right]^2 - c^2 - \left[\frac{j+1}{r}\right]^2} \\
&\cong \Sigma_j \frac{4j+2}{\pi}\sqrt{\left[c + \frac{Z(r)}{cr}\right]^2 - c^2 - \left[\frac{j+1/2}{r}\right]^2} \\
&= -\Sigma_j \frac{4r^2}{3\pi}\frac{\partial}{\partial j}\left\{\left[c + \frac{Z(r)}{cr}\right]^2 - c^2 - \left[\frac{j+1/2}{r}\right]^2\right\}^{3/2} \\
&\cong \frac{4r^2}{3\pi}\left\{\left[c + \frac{Z(r)}{cr}\right]^2 - c^2 - \left[\frac{Z(0)}{cr}\right]^2\right\}^{3/2}
\end{aligned} \tag{33}$$

The lower limit of the integration is determined by the condition (13) and it secures that a self-consistent field can be sought from Poisson's equation,

$$\nabla^2 U(r) = -\frac{\rho(r)}{r^2} \tag{34}$$

which gives the equation for the screening function as

$$\frac{d^2Z(r)}{dr^2} = \frac{4}{3\pi\sqrt{r}}\left[2Z(r) - \frac{Z^2(0) - Z^2(r)}{c^2 r}\right]^{3/2} \tag{35}$$

This simplifies to the regular Fermi-Thomas equation for infinitely large values of c and reduces to a well-known difficult form when the term in $Z(0)$ is neglected(10).

The propagator gives a value for the sum of the energies of the one-electron states through the integral

$$\begin{aligned}
\varepsilon_k(r) &= \frac{2|k|}{2\pi i}\oint W dW\left[Tr\mathbf{G}_0(r,r;W)\right] = -\frac{|k|}{\pi}\oint W dp \\
&= \frac{|k|}{\pi}\oint\left[\frac{Z(r)}{r} - c\sqrt{p^2 + c^2 + \kappa^2}\right]dp \\
&= -\frac{Z(r)\rho_k(r)}{r} + \frac{|k|c}{\pi}\left[c + \frac{Z(r)}{cr}\right]^2\left\{\xi_k(r) + \tfrac{1}{2}\left[1 - \xi_k^2(r)\right]\ln\frac{1+\xi_k(r)}{1-\xi_k(r)}\right\}
\end{aligned} \tag{36}$$

$$\xi_k(r) \equiv \sqrt{1 - \frac{c^4 r^2 + c^2\left(k + \tfrac{1}{2}\right)^2}{\left[Z(r) + c^2 r\right]^2}}$$

It is important to choose the branch of the square root with a positive real part in the integrations above. Summation over quantum numbers k proceeds as before with the result for the energy density that

$$\begin{aligned}
\varepsilon(r) &= -\frac{Z(r)\rho(r)}{r} + \frac{cr^2}{2\pi}\left[c^2 + \frac{Z^2(0)}{c^2 r^2}\right]^2\left\{\frac{\xi(r) + \xi^3(r)}{\left[1 - \xi^2(r)\right]^2} - \tfrac{1}{2}\ln\frac{1+\xi(r)}{1-\xi(r)}\right\} \\
\xi(r) &\equiv \sqrt{1 - \frac{c^4 r^2 + Z^2(0)}{\left[Z(r) + c^2 r\right]^2}}
\end{aligned} \tag{37}$$

The first term in the energy density expression is the expectation value of the potential energy and it is concluded that the second one includes the kinetic energy

as well as the rest energy. The auxiliary variable $\xi(r)$ can be expressed in terms of the density rather than the potential in the form

$$\xi(r) = \sqrt[3]{\frac{3\pi\rho(r)}{4c^3r^2}} \Bigg/ \sqrt{1 + \frac{Z^2(0)}{c^4r^2} + \left[\sqrt[3]{\frac{3\pi\rho(r)}{4c^3r^2}}\right]^2} \tag{38}$$

so that the energy may related directly to the density. It is also seen that this variable is small even for rather small values of the radius. An expansion of the energy density in inverse powers of c gives the result that

$$\begin{aligned} \varepsilon(r) &= -\frac{Z(r)\rho(r)}{r} \\ &+c^2\rho(r)\left\{1 + \frac{3}{10c^2}\left[\frac{3\pi\rho(r)}{4r^2}\right]^{2/3} - \frac{1}{c^4}\left[\frac{Z^2(0)}{r^2} + \frac{3}{56}\left[\frac{3\pi\rho(r)}{4r^2}\right]^{4/3}\right] + \ldots\right\} \end{aligned} \tag{39}$$

where the rest energy and the classical Fermi-Thomas kinetic energy density are readily recognized. The expansion is not valid near the origin and the last explicit term is not integrable over the range of radii.

6. THE CENTRAL FIELD

Equation (35) does not exhibit the singularity at the origin that mars the form where $Z(0)$ is omitted. A solution is not readily computed because a new condition appears. The argument of the root on the right hand side needs to be non-negative, that is

$$2Z(r)-\frac{Z^2(0)-Z^2(r)}{c^2 r}\geq 0 \quad \Rightarrow$$
$$Z(r)+c^2 r-\sqrt{c^4 r^2+Z^2(0)}=Z(r)-\frac{Z^2(0)}{c^2 r+\sqrt{c^4 r^2+Z^2(0)}}\geq 0; \tag{40}$$

and thus it seems as if the screening function goes to an asymptotic inverse distance behavior. An expansion can be generated in the form

$$Z(r)=\frac{Z^2(0)}{2c^2 r}\left[1+a_1 r^{-2/3}+a_2 r^{-4/3}+\ldots\right] \tag{41}$$

for the large radius regime. The coefficients are rapidly increasing,

$$a_1=\left[\frac{3\pi c}{4Z(0)}\right]^{2/3};\quad a_2=\frac{40}{27}\left[\frac{3\pi c}{4Z(0)}\right]^{4/3}; \tag{42}$$

and the series is of little use presently. This asymptotic form ensures that the number of electrons in the system equals the nuclear charge.

Small radii suggest a series of the form

$$Z(r)=Z(0)\left[1+b_1 r+b_2 r^{3/2}+\ldots\right]$$
$$b_2=\frac{32\sqrt{2Z(0)}}{9\pi}\left[1+\frac{Z(0)b_1}{c^2}\right]^{3/2} \tag{43}$$

where b_1 is negative and will be determined by the appropriate asymptotic behavior. An accurate representation of the non-relativistic Fermi-Thomas function can be found as a rational function of the variable $\sqrt{r}$. The present case, with the different form for large radii, does not seem to offer a similar approach.

Direct numerical integration is approached through a couple of transformations. The radius variable and the screening function are expressed in the traditional Fermi-Thomas scaled ones:

$$r = xb; \quad b = \sqrt[3]{\frac{9\pi^2}{128Z(0)}}; \quad Z(r) = Z(0)\chi(x); \tag{44}$$

Thus we find from Eq. (34) that

$$\sqrt{x}\,\frac{d^2\chi(x)}{dx^2} = \left\{\chi(x) - \frac{\beta}{x}\left[1 - \chi^2(x)\right]\right\}^{3/2}; \quad \beta = \frac{Z(0)}{2bc^2}; \tag{45}$$

and the atomic number occurs only in the parameter β. It has been convenient to replace Eq. (45) by a system of coupled first order differential equations with yet another independent variable:

$$\begin{aligned} x &= \xi^2; \\ \frac{d\chi(\xi)}{d\xi} &= \xi\varphi(\xi); \\ \frac{d\varphi(\xi)}{d\xi} &= 4\left\{\chi(\xi) - \frac{\beta}{\xi^2}\left[1 - \chi^2(\xi)\right]\right\}^{3/2}; \end{aligned} \tag{46}$$

A satisfactory solution requires that $\varphi(\xi)$is negative and that its derivative is positive. Integration by means of the numerical Runge-Kutta(11) stepping procedure starts from values $\varphi(0)$ and $\chi(0)$ = 1and ends when either of the two conditions is violated. A solution has been sought such that the screening function is as small as possible while still having a negative derivative. This will result in a finite radius for the charge distribution, a consequence that seems to be unavoidable in the present approximation. A similar situation occurs when Dirac exchange is included with the non-relativistic version(10).

An example calculation for $Z=54$, the Xenon atom, gives the result in Table 1. The "edge" of the charge distribution appears at 6.5 Bohr radii, well outside the normally accepted atomic radius. The screening function is well represented by the rational form

$$\chi(\xi) = \frac{1 + \sum_{j=1}^{5} a_j \xi^j}{1 + \sum_{j=1}^{7} c_j \xi^j} \tag{47}$$

with the coefficients from Table 2. There are no poles on the real, positive ξ-axis and the zeroes occur outside the range where the numerical solution is acceptable. The initial value $\varphi(0)$ was chosen as -3.12851 which gives $d\chi(0)/dx$ =-1.564255, slightly less than the regular value for the Fermi-Thomas function, -1.588041. The two screening functions are close and differ significantly only at larger distances from the atomic nucleus.

Table 1. Result of integration of Eq. (46) for the Xenon atom where $Z=54$, β=0.0061383, b=0.936928

ξ	r	$\chi(\xi)$	$\varphi(\xi)$
0.000	0.000000	1.000000	-3.128510
0.088	0.007256	0.988767	-2.788210
0.176	0.029022	0.958524	-2.458086
0.264	0.065300	0.914188	-2.146406
0.352	0.116089	0.860156	-1.858689
0.440	0.181389	0.800175	-1.598044
0.528	0.261201	0.737311	-1.365625
0.616	0.355523	0.673965	-1.161112
0.704	0.464356	0.611934	-0.983153
0.792	0.587701	0.552496	-0.829748
0.880	0.725557	0.496492	-0.698544
0.968	0.877924	0.444422	-0.587063
1.056	1.044802	0.396515	-0.492859
1.144	1.226191	0.352804	-0.413621
1.232	1.422092	0.313179	-0.347230
1.320	1.632503	0.277432	-0.291792

1.408	1.857426	0.245293	-0.245642
1.496	2.096860	0.216454	-0.207338
1.584	2.350805	0.190592	-0.175648
1.672	2.619261	0.167377	-0.149525
1.760	2.902228	0.146483	-0.128092
1.848	3.199706	0.127595	-0.110615
1.936	3.511696	0.110406	-0.096486
2.024	3.838196	0.094623	-0.085203
2.112	4.179208	0.079962	-0.076351
2.200	4.534731	0.066153	-0.069587
2.288	4.904765	0.052932	-0.064626
2.376	5.289310	0.040046	-0.061219
2.464	5.688367	0.027254	-0.059136
2.552	6.101934	0.014330	-0.058140
2.640	6.530013	0.001084	-0.057917

Table 2. Coefficients in the rational approximation for the screening function $\chi(\xi)$ of Eq. (47).

j	a_j	c_j
1	0.606287	0.606287
2	-1.020611	0.543644
3	0.397868	0.0509944
4	-0.0577654	-0.0031777
5	0.0000232	-0.0216453
6	0	-0.0092576
7	0	0.0037789

7. SOME COMMENTS

Several suggestions have been put forth for a relativistic generalization of the traditional Fermi-Thomas atomic model. It seems that Vallarta and Rosen(12) initiated this quest. They arrived at the expression for the electron density from the homogenous electron gas and could not then obtain the boundary condition (13), which was essential here to eliminate the singularity at the origin(13). Rudkjøbing(14) considered the spherical case and approximated the radial equations in a shell and obtained a density distribution dependent on the gradient of the potential in addition to the potential itself. His purpose was a study of relativistically degenerate stars and there seems to have been no atomic applications. Weizsäcker's(15) correction, based on the density gradient, has been introduced by Tomishima(16) in an effort to remove the singularity at the origin.

There are a couple of attempts to use the WKB approximation to the two component radial Dirac problem for screened Coulomb fields. Kosaka and Yonei(17) review these and apply the form offered by Goldberg and Pratt(18) for calculations of bound state orbitals in the potentials suggested by Tomishima(16). Their procedure is rather more involved than the present one.

The current effort demonstrates the necessity to include a proper coupling between the angular and radial behavior in order to obtain a reasonable charge density distribution in the Dirac case. This coupling manifests itself through the boundary condition on the Green function at the atomic nucleus as given by relation (13). A less satisfactory situation occurs far from the nucleus. It appears unavoidable to get a density with a finite radius. The effective nuclear charge will thus be zero outside this radius.

The demonstration here of the efficacy of the integral formulation and the propagator techniques may be used in further attempts at approximating the more awkward terms in the electron-electron interactions, such as exchange and Breit terms(19), albeit giving rise to rather involved integrals.

REFERENCES:

(1) Löwdin, P. O. Quarterly Progress Report SSMTG, MIT, January 15, 1952, 12.
(2) Linderberg, J.; Öhrn, Y. "Propagators in Quantum Chemistry"; Academic Press: London 1973; 36.
(3) Linderberg, J. *Int. J. Quant. Chem.* **1981,** *S15*, 559.
(4) Condon, E. U.; Shortley, G. H. "The Theory of Atomic Spectra"; University Press: Cambridge 1963.
(5) Ref. 4, p. 128.
(6) Kemble, E. C. "The Fundamental Principles of Quantum Mechanics with Elementary Applications", Dover Publications: New York 1958, 107
(7) Courant, R.; Hilbert, D. "Methoden der Mathematischen Physik", Springer-Verlag: Berlin 1968, 3. ed. Vol. 1, 121.
(8) Ref. 7, p.455.
(9) Coulson, C. A. *Proc. Camb. Phil. Soc.* **1940,** *40*, 201.
(10) March, N. H. "Self-Consistent Fields in Atoms. Hartree and Thomas-Fermi Atoms"; Pergamon: Oxford 1975,131.
(11) Ince, E. L. "Ordinary Differential Equations"; Dover Publications: New York 1956, Appendix B, 540.
(12) Vallarta, M. S.; Rosen, N. *Phys. Rev.* **1932,** *41*, 708.
(13) Jensen, H. *Z. Physik* **1933,** *82*, 794.
(14) Rudkjøbing, M. *Kgl. Danske Vid. Selskab, Mat.-fys. Medd.* **1952**, *27,* no.5.
(15) Weizsäcker, C. F. v. *Z. Physik* **1935,** *96*, 431.
(16) Tomishima, Y. *Progr. Theor. Phys.* **1969,** *42*, 437.
(17) Kosaka, K.; Yonei, K. *J. Phys. Soc. Japan* **1991,** *60,* 850.
(18) Goldberg, I. B.; Pratt, R. H. *J. Math. Phys.* **1987,** *28,* 1351. Also Goldberg, I. B.; Stein, J.; Ron, A.; Pratt, R. H. *Phys. Rev.* **1989,** *A39,* 506.
(19) Bethe, H. A.; Salpeter, and E. E. "Quantum Mechanics of One- and Two-Electron Atoms"; Springer-Verlag: Berlin 1957, 195.

Learning Differential Calculus without Limits

M. Berrondo
Department of Physics and Astronomy,
Brigham Young University,
Provo, UT 84097

Abstract

The concept of limit seems to be essential in the understanding and the present teaching of Calculus. In this article, however, we show how to structure and use differential calculus without introducing this concept. The crucial idea in this development is to use Leibniz rule for the derivative of a product of two functions as one of the postulates, rather than as a derived theorem. Within this approach, the idea of limit could be introduced belatedly and only in order to define concepts such as continuity and differentiability in a more rigorous fashion.

Contents

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

1 *In Memoriam*

Per Olov Löwdin had two love affairs: Teaching and Algebra. They were crowned by his passion to teach algebra. I would like to dedicate the present work to my Teacher, Mentor, and Friend in this memorial volume. I always thought of Per Olov as immortal, so it came to me as a huge shock when I heard the news of his untimely death. Although we can no longer listen to his inspiring lectures or to his wise advise, Per Olov's spirit will always remain with us, always smiling, always optimistic, always enthusiastic. I am absolutely convinced that Per Olov would have greatly enjoyed my novel way of presenting and teaching Calculus, since it is so close to his beloved Algebra. He also would have undoubtedly improved on it after his first reading.

2 Introduction

It is an understatement to say that Science and Engineering have benefited enormously from the concept of the derivative of a function. During the last three centuries, differential and integral calculus have become the main quantitative language of applied mathematics in science. Differential calculus takes simple ideas and theorems from elementary geometry and extends them to more general situations. For instance, it expands the concept of the slope of a straight line and applies it to a curve at each point of the graph. It generalizes the idea of a tangent to the circle obtained as the straight line perpendicular to the radius, to the tangent line of any smooth curve [1]. In physics, it allows for the calculation of instantaneous velocity and acceleration, instead of being limited to the corresponding average quantities. Actually, we can say that Calculus began with the question "what is velocity and how can we calculate it?" [2]. The derivative, in essence, encapsulates the ideas of slope and of rate of change and applies them to a myriad of practical problems.

The concept of derivative is presently taught introducing the idea of a limit foremost [1, 3]. We often think of differential calculus as a difficult subject, until we realize that it can be presented in a much simpler and interesting fashion than the way it has traditionally been done. An introduction to Calculus should present the subject in a more intuitive way and relate it to geometry, algebra, and science in general. Beginning students should not be expected to learn all the subtleties of Mathematical Analysis needed to cover the exceptional cases. Rigorous definitions of continuity and differentiability of functions should only come later [3, 4]. Even Cauchy, the father of rigor, missed the difference between these two concepts at first! Instead, the students should concentrate on learning the basics of the subject, as well as comprehend the important concepts and master the applications and techniques, based on

their grasp of algebra and geometry.

The rule to calculate the derivative of a product of two functions was first introduced by Leibniz [5, 6]. The crux of our presentation is to take the multiplication rule as an *initial postulate*, rather than as a derived result. Leibniz rule for the derivative of a product of functions is not privy of calculus. It also appears when calculating commutators of matrices or linear operators: $[..., BC] = B[..., C] + [..., B]C$. There is no need to invoke the concept of limit in this case, or when dealing with Lie brackets, or other *derivations*. The ultimate justification for this choice of initial postulate is given *a posteriori* in terms of the logarithmic function [7].

In this work we shall consider continuous functions of a single real variable. We will take the concept of continuity of a function in the simple geometrical sense, namely, that we can draw the graph of the function without lifting up the pencil. Likewise, we shall say that a function is differentiable wherever the slope of its tangent line is well defined.

Caveat: The purpose of this paper is to introduce the principles on how to teach differential calculus without the idea of a limit. A more pedagogical version is being prepared.

3 The Three Canonical Rules

Consider a real function $y = f(x)$ of a real variable x. By this we mean a mapping of the real number x to a unique real number y, given by the rule f. Furthermore, let us assume it to be continuous. We will introduce the concept of the derivative of $f(x)$ with respect to x in terms of the slope of the tangent line at the point $(x, f(x))$. In order to do this, we need to consider three simple constructive rules using the *slope of a straight line* as our starting point, and Leibniz rule as our keystone. The slope is calculated as a ratio of two displacements: "rise over run". Hence, we define the derivative of y with respect to x as a quotient of the two corresponding differentials, denoted by dy (the "rise") and dx (the "run"):

$$\text{derivative of } y \text{ w.r. to } x = \frac{dy}{dx}. \tag{1}$$

Rule 1: Unit slope and zero slope.

The derivative of the function $y = x$ is given by

$$y = x$$

$$\frac{dx}{dx} = 1, \tag{2}$$

corresponding to a unit slope. A constant function $y = c$ has a vanishing slope:

$$y = c$$

$$\frac{dc}{dx} = 0. \tag{3}$$

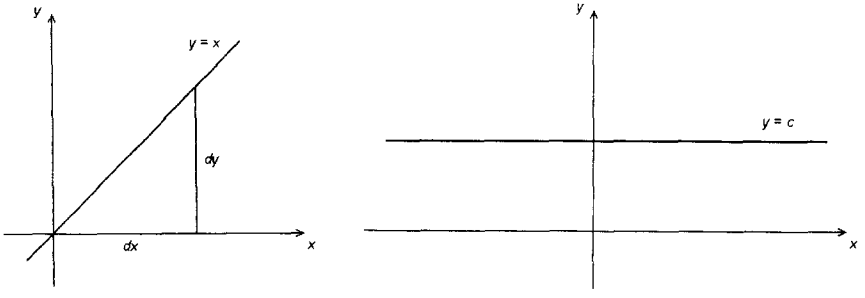


Figure 1 Unit slope and zero slope

Rule 2: Derivative of the sum.

Given two functions $f(x)$ and $g(x)$,

$$\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}. \tag{4}$$

Rule 3: Leibniz rule for the product.

Given two functions $f(x)$ and $g(x)$,

$$\frac{d(fg)}{dx} = f\frac{dg}{dx} + \frac{df}{dx}g. \tag{5}$$

This rule is the **core of differential calculus**. We will actually prove that it allows us to interpret the derivative of a function at a point (x, y) as the slope of the tangent line touching it. We shall henceforth refer to the latter simply as the *slope of the curve.*

These three rules define the differential calculus without involving the concept of limit. The derivative of positive powers and of polynomials follow directly from a straightforward application of these rules. The case of the

reciprocal of a function, and the quotient rule are also easy to derive. The geometric interpretation of the derivative of a function as the slope of its tangent line is also a consequence of Leibniz rule. This is a remarkable result! Another satisfying feature of this procedure is the fact that the second and third rules refer to the two elementary operations of addition and multiplication of real numbers.

As a first example, consider the usual equation for a straight line $y = mx+b$, where m and b are constant. We proceed to calculate its derivative using the three canonical rules as follows:

$$y = mx + b,$$

$$\frac{dy}{dx} = \frac{d(mx)}{dx} + \frac{db}{dx} = m\frac{dx}{dx} + \frac{dm}{dx}x = m, \tag{6}$$

corresponding to the slope of the line $y(x)$.

In the same vein as Eq. (6), Rule 2 can be generalized to a linearity condition, namely:

Rule 2': Linearity of the derivative.

The derivative of a linear combination of functions is the linear combination of the corresponding derivatives.

4 The Parabola

The first nontrivial example is that of a quadratic function. We start with the equation for a simple parabola $y = x^2$. Its graph is a continuous smooth function and we can evaluate its derivative using the third rule:

$$y = x^2,$$

$$\frac{dy}{dx} = \frac{d(x^2)}{dx} = x\frac{dx}{dx} + \frac{dx}{dx}x = 2x. \tag{7}$$

We notice immediately that there are no limits involved in this calculation, there are no terms to be neglected, and there is no undefined "0/0" to begin with. This simple use of Leibniz rule illustrates its power and scope. The generalization to any power of x should be easy to visualize.

The key point here is to realize that Eq. (7) represents the slope of the parabola at point x. In order to this, we first need to know how the tangent line can be defined in algebraic terms.

Tangent line: The tangent to a smooth function $f(x)$ at a given point $x = x_o$ is defined as the straight line $y = mx + b$ such that the difference function:

$$f(x) - mx - b \tag{8}$$

has a double root at the point $x = x_o$.

Applying this criterion to the parabola $y = x^2$, we look for the double zero of the quadratic $x^2 - mx - b$. This corresponds to a vanishing value of the discriminant, and the result $x = m/2$ is independent of b. Rewriting this as

$$m = 2x \tag{9}$$

shows explicitly that the slope of the parabola is given by its derivative as calculated in Eq. (7).

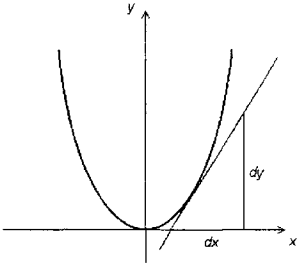


Figure 2 Parabola and its tangent line

The derivative of the more general quadratic function $y = ax^2 + bx + c$, with a, b, c constant is again readily calculated using Rule 2' with the result:

$$y = ax^2 + bx + c,$$

$$\frac{dy}{dx} = 2ax + b. \tag{10}$$

5 Powers, Polynomials, and Reciprocals

The derivative of a positive integer power follows from induction. Before we show this, let us work out explicitly the case $y = x^3$:

$$y = x^3,$$

$$\frac{dy}{dx} = \frac{dx^3}{dx} = x\frac{dx^2}{dx} + \frac{dx}{dx}x^2 = 3x^2 \tag{11}$$

where we have used Eq. (7).

The same procedure is used for the n-th power. For

$$y = x^n,$$

$$\frac{dy}{dx} = nx^{n-1}. \tag{12}$$

In order to prove Eq. (12), let us assume that it is valid for $n-1$. Then we show that it is valid for n:

$$\frac{dy}{dx} = \frac{dx^n}{dx} = x\frac{dx^{n-1}}{dx} + \frac{dx}{dx}x^{n-1} = (n-1)x^{n-1} + x^{n-1} = nx^{n-1}. \tag{13}$$

Since, by definition, it is also valid for $n = 1$ from Rule 1, the proof by induction is complete.

With the aid of Rule 2', we can obviously calculate the derivative of any polynomial.

Let us now turn our attention to the case of the reciprocal $g^{-1}(x)$ of the (nonvanishing) function $g(x)$. If we take the derivative of the identity $g^{-1}g = 1$, we can evaluate the derivative of the reciprocal function as follows:

$$\frac{d(g^{-1}g)}{dx} = g^{-1}\frac{dg}{dx} + \frac{dg^{-1}}{dx}g = 0,$$

$$\frac{d(1/g)}{dx} = -\frac{1}{g^2}\frac{dg}{dx}. \tag{14}$$

whenever $g(x) \neq 0$.

This allows us to extend Eq. (13) to the case of negative exponents:

$$y = x^{-m} = \frac{1}{x^m},$$

$$\frac{dy}{dx} = -\frac{mx^{m-1}}{x^{2m}} = -mx^{-m-1} = -\frac{m}{x^{m+1}}. \tag{15}$$

The corresponding rule for the quotient of two functions $f(x)$ and $g(x)$ obtains from Rule 2 as:

$$\frac{d(f/g)}{dx} = \frac{1}{g^2}\left(\frac{df}{dx}g - f\frac{dg}{dx}\right). \tag{16}$$

This case includes, of course, rational functions *i.e.* the ratio of two polynomials in x.

When the exponent refers to a root rather than a power *i.e.* $y = x^{1/q}$, we use the fact that the derivative is defined in terms of a ratio and consider the root as the inverse function of the power y^q. Let us illustrate this for the square root function $y = \sqrt{x}$, or $x = y^2$. We calculate first:

$$\frac{dx}{dy} = 2y = 2\sqrt{x} \tag{17}$$

and then take the reciprocal dy/dx:

$$y = \sqrt{x} = x^{1/2},$$

$$\frac{dy}{dx} = \frac{1}{dx/dy} = \frac{1}{2\sqrt{x}} = \frac{1}{2}x^{-1/2} \tag{18}$$

using Eq. (17).

In summary, we can generalize Eq. (13) to the case of any rational power:

$$y = x^{\alpha},$$

$$\frac{dy}{dx} = \alpha x^{\alpha-1} \tag{19}$$

The case of α irrational is deferred until later, although the result reads the same as in Eq. (19).

6 The Chain Rule and the Exponential Function

We have dealt with the sum and the product of functions, as well as the quotient. In order to calculate the derivative of the composition of two functions, we take advantage of the fact that the derivative is defined as the quotient of two differentials. In the case in which f is a function of u, in turn a function of x, the chain rule will tell us how to calculate the derivative of $f(u(x))$ with respect to x simply as:

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}. \tag{20}$$

A prominent application of this rule is the case of a parametric dependence. In finding the derivative of trigonometric functions, the angle plays the role of a parameter and we can use Eq. (20) to calculate them.

The more usual applications of the chain rule imply breaking up the function whose derivative we want to compute into smaller portions, where each piece is itself a function of the independent variable. As a simple example, let us calculate the derivative of a rational function such as:

$$y = \frac{1-x^2}{1+x^2} = \frac{1-u}{1+u}$$

$$\frac{dy}{dx} = 2x\frac{-2}{(1+u)^2} = -\frac{4x}{(1+x^2)^2}. \tag{21}$$

The chain rule is an indispensable tool in differential calculus. It allows for the simplification of derivatives of composite functions.

We next turn our attention to the exponential function, $E(x)$, *defined* as the function whose derivative is the same function. In order to choose the base in a unique way, we impose the condition that $E(0) = 1$. Thus,

$$\frac{dE(x)}{dx} = E(x), \qquad E(0) = 1, \tag{22}$$

or, also, using the chain rule for $E(\alpha x)$

$$\frac{dE(\alpha x)}{dx} = \alpha E(\alpha x) \qquad E(0) = 1. \tag{23}$$

for constant α. We next want to prove that this function does indeed behave like an exponential function, *i.e.* that it fulfills the law of multiplication $E(a+b) = E(a)\,E(b)$. To this end, we apply Rule 3 and Eq. (23) to the product:

$$y(x) = E(\alpha x)\,E(\beta x),$$

$$\frac{dy}{dx} = \alpha E(\alpha x)\,E(\beta x) + \beta E(\alpha x)\,E(\beta x) = \alpha y + \beta y = (\alpha+\beta)y \qquad y(0) = 1. \tag{24}$$

Hence, using Eq. (23) once more, $y(x) = E((\alpha+\beta)x)$. We will henceforth write $E(x)$ as e^x, so that Eqs. (22) and (23) may be rewritten as:

$$\frac{de^x}{dx} = e^x \tag{25}$$

$$\frac{de^{\alpha x}}{dx} = \alpha e^{\alpha x}, \tag{26}$$

with the property $e^0 = 1$ now being obvious.

The irrational number e is to calculus what π is to geometry. The approximate value of the transcendental number e, corresponding to e^1,

$$e = 2.7182818... \tag{27}$$

can be calculated as a series, as will be shown in a later section.

7 Euler's Formula and Trigonometric Functions

The derivatives of the trigonometric functions are obtained more easily through the use of Euler's formula. This identity relates the exponential function to the trigonometric functions through the imaginary number $i = \sqrt{-1}$. It can be understood geometrically using Pythagoras theorem:

$$\sin^2\theta + \cos^2\theta = 1. \tag{28}$$

Considering θ as variable, this can be read off as the parametric equation of a **unit circle**, whose Cartesian coordinates are given by the usual projections:

$$x = \cos\theta, \qquad y = \sin\theta, \tag{29}$$

actually defining these functions.

A different perspective of this same geometry obtains by factorizing Eq. (28) as:

$$(\cos\theta + i\sin\theta)(\cos\theta - i\sin\theta) = 1, \tag{30}$$

where the two factors are complex conjugate of each other. The function $\cos\theta + i\sin\theta$ spans the unit circle in the complex plane as we vary the polar angle θ from 0 to 2π, with real and imaginary parts given by Eq. (29). We now want to identify this function with $e^{i\theta}$ since the latter also represents the unit circle in the complex plane . This can be easily achieved using Eq. (26) with $\alpha = i$, *i.e.*

$$\frac{de^{i\theta}}{d\theta} = ie^{i\theta}. \tag{31}$$

The left-hand side of this equation is the slope of $e^{i\theta}$, while the right-hand side represents the same function rotated by 90 degrees, so the tangent line turns out to be perpendicular to the radius vector, therefore forming a circle in the complex plane. Furthermore, since the modulus of $e^{i\theta}$ is always unity, the corresponding circle is the **unit circle**. Euler's formula then asserts that:

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{32}$$

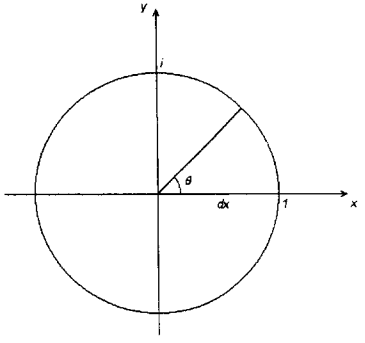


Figure 3 Unit circle in the complex plane

The derivative of the sine and the cosine functions now follow directly taking the real and the imaginary parts, using (31) and (32):

$$y = \sin\theta$$

$$\frac{d\sin\theta}{d\theta} = \cos\theta, \tag{33}$$

$$x = \cos\theta$$

$$\frac{d\cos\theta}{d\theta} = -\sin\theta. \tag{34}$$

In this case, the origin of the minus sign in Eq. (34) is easily understood as the factor i^2.

8 Inverse Functions. Power series

We have already seen an example of how to calculate the derivative of inverse functions when dealing with fractional powers. In the case of the exponential function the corresponding inverse is the logarithmic function $y = \ln(x)$. Its derivative follows from Eq. (25):

$$y = \ln(x), \qquad x = e^y,$$

$$\frac{dy}{dx} = \frac{1}{dx/dy} = \frac{1}{x} \qquad \text{for } x \neq 0. \tag{35}$$

This equation can be used to obtain the derivative of any power of x, including the case of irrational exponents:

$$y = x^{\alpha}, \qquad \ln y = \alpha \ln x,$$

$$\frac{dy}{dx} = y\,\frac{d\ln y}{dx} = x^{\alpha}\,\frac{\alpha}{x} = \alpha x^{\alpha-1}. \tag{36}$$

Derivatives of inverse trigonometric functions can be obtained in a similar fashion. For instance, the derivative of

$$\theta = \sin^{-1} y, \qquad y = \sin\theta$$

is given by:

$$\frac{d\theta}{dy} = \frac{1}{dy/d\theta} = \frac{1}{\cos\theta} = \frac{1}{\sqrt{1-y^2}}. \tag{37}$$

In order to estimate the transcendental number e, we will expand the exponential function e^x in a power series using a simple iterative procedure starting from its definition Eq. (25) together with Eq. (12). As a prelude, we first find the power series expansion of the geometric series $y = 1/(1+x)$, iterating the equivalent expression:

$$y = 1 - xy$$

$$\frac{1}{1+x} = 1 - x(1-xy) = 1 - x + x^2 y = 1 - x + x^2 - x^3 + \ldots + (-)^n x^n y. \tag{38}$$

However, given the definition of the exponential e^x, we need to start with an expression involving derivatives. The iteration follows by defining the function $z = e^x - 1$, whose derivative is given by:

$$\frac{dz}{dx} = 1 + z. \tag{39}$$

Starting from $dz/dx \simeq 1$ and using the inverse of Eq. (2), we find the first approximation, $z_1 = x$. Substituting this in the right hand side of Eq. (39), the next approximation will be given by $z_2 = x + x^2/2$, recalling Eqs. (2) and (7). Iterating this procedure, with the aid of Eq. (12), we end up with:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \tag{40}$$

with a corresponding expansion:

$$e = 2 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ldots \tag{41}$$

A more accurate approximation can be obtained by using a fraction $x = 1/q$, for instance, in Eq. (40) and writing $e = (e^{1/q})^q$. For example, using the first five terms in Eq. (40) for $x = 1/16$, we get the value:

$$e = 2.7182818...$$

In order to attain this same accuracy, we would need at least ten terms in the series (41).

9 Discussion

We have identified the derivative of a function with its slope, or more properly, the slope of its tangent line through various examples. It is now time to prove this statement in general.

Theorem: The slope of the tangent line of a function $f(x)$ at a point x_o is given by its derivative at that point.

This is a generalization of the case of a parabola, Eq. (9). It turns out to be a simple consequence of our three canonical rules. Using the definition of the tangent line in terms of a double root x_o, we can write:

$$f(x) - mx - b = (x - x_o)^2 g(x), \tag{42}$$

where $g(x)$ and dg/dx evaluated at x_o have to be finite. Using our three rules to take the derivative of Eq. (42),

$$\frac{df}{dx} - m = 2(x - x_o)g(x) + (x - x_o)^2 \frac{dg}{dx}.$$

Evaluating this expression at $x = x_o$, we get the desired relationship between the derivative of the function and its slope:

$$\frac{df}{dx}(x_o) = m. \tag{43}$$

We obtain a new insight of Leibniz rule considering the properties of the logarithmic function, namely, that the logarithm of a product of two functions equals the sum of the logarithms of each function, $\ln(f(x)g(x)) = \ln(f(x)) + \ln(g(x))$. Taking the derivative of this relationship and using Eq. (35), together with the chain rule, we obtain:

$$\frac{1}{fg}\frac{d(fg)}{dx} = \frac{1}{f}\frac{df}{dx} + \frac{1}{g}\frac{dg}{dx}. \tag{44}$$

Multiplying Eq. (44) by fg, we obtain Rule 3 as a special case of the linearity assumption, Rule 2!

The idea of the limit of a real function gives us a rigorous way to define *continuity* by comparing the approach to a point from the right with the one from the left. In the usual Calculus textbooks [1, 4], based upon Fermat's definition of a tangent line as the limit of chords [5], the derivative of a function is defined as the limit of a quotient. Hence the same limiting procedure gives us a rigorous definition of *differentiability*. However, once we encounter a discontinuous function, the method of limits does not tell us how to calculate the derivative. Instead, it is much more fruitful to introduce the idea of distributions and define the Dirac delta function, so useful in applied mathematics, especially in physics and engineering.

The procedure based on the multiplication rule as an initial postulate can also be generalized to the derivative of complex functions of a complex variable. Vector calculus would benefit from this approach, since the ∇ operator also obeys Leibniz rule. In both of these cases we would have to generalize the basic Rule 1 and make it consistent with the corresponding case.

In closing, we would like to point out that our viewpoint on differential calculus of real functions is consistent with the modern view of *derivations* found in differential geometry, in which the derivative is understood as a mapping to the tangent space. Restricted to a planar curve, this is nothing but the tangent line.

References

1. G.J. Etgen, **Salas and Hille's Calculus: One and Several Variables**, Wiley, New York, (1999).

2. I. Newton, **Principia**, reprinted by Prometheus, Amherst, NY, (1995)

3. R. Beatley, Amer. Math. Monthly, **57**, 369 (1950).

4. M. Kline, **Calculus, An Intuitive and Physical Approach**, reprinted by Dover, Mineola, NY, (1998).

5. C.H. Edwards, **The Historical Development of the Calculus**, Springer Verlag, New York, (1979).

6. D.J. Struik, **A Source Book in Mathematics**, Harvard University Press, Cambridge, MS, (1969), Chap. V.

7. M. Berrondo, to be published.

Dissipative Systems and Microscopic Selforganization

Erkki J. Brändas
Department of Quantum Chemistry, Uppsala University
Box 518, S-751 20 Uppsala, Sweden

Abstract
In this tribute and memorial to Per-Olov Löwdin we discuss and review the extension of Quantum Mechanics to so-called open dissipative systems via complex deformation techniques of both Hamiltonian and Liouvillian dynamics. The review also covers briefly the emergence of time scales, the definition of the quasibosonic pair entropy as well as the precise quantization relation between the temperature and the phenomenological relaxation time. The issue of microscopic selforganization is approached through the formation of certain units identified as classical Jordan blocks appearing naturally in the generalised dynamical picture.

CONTENTS

1. INTRODUCTION

The international field of Quantum Chemistry was defined and established [1, 2] through the pioneering work of Per-Olov Löwdin. In his 1948-thesis *Some Ionic Properties of Ionic Crystals* - reprinted in this volume of the Advances of Quantum Chemistry - he carried out the first *ab initio* calculation in the solid state domain using a "computer" consisting of graduate students and assistants equipped with the best mechanical devices available in Uppsala at the time. Initially Quantum Chemistry acquired special connections with the Uppsala University and the University of Florida in Gainesville, where POL founded two active research groups, which still, after more than 40 years, continue to influence the area through

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

the organisation of annual Sanibel Symposia, various congresses and workshops not to mention editoring the *International Journal of Quantum Chemistry* and the book series the *Advances in Quantum Chemistry* created through his boundless productivity and efficiency.

The subject of Quantum Chemistry is today represented at every major University and the activities mentioned above have resulted in strong international and global networks thanks to the leadership of Per-Olov Löwdin. Through outstanding activities during his lifetime he has influenced generations of physicists and chemists and established important links between the fields of physics, chemistry and biology.

In this contribution, dedicated to the memory of Per-Olov Löwdin, we will show how to extend applications in Quantum Chemistry, based on standard Quantum Mechanics for isolated systems, to so-called dissipative or open systems in contact with its environment, see e.g. [3] for some recent discussions. This route goes via the formation of resonances both in the Hamiltonian as well as in the Liouvillian picture [4]. In the most general formulation in terms of reduced density matrices it is demonstrated how particular selforganisational units emerges from a quantization condition relating the temperature with the phenomenological relaxation time scale leading to increasing order during a specific number of life times [4, 5]. These units are identified as certain finite dimensional Jordan blocks appearing through the generalised dynamics of the extended resonance picture previously established. There exist several applications [4-10] of this development and in the concluding statement we summarise the formulation in terms of a time super operator which bears the main characteristics of the transition from the microscopic level to a higher order phenomenological level [4].

2. FORMATION OF RESONANCES

The extension of quantum theory to incorporate the resonance picture of unstable states are often used via so-called dilation analytic techniques [11] or through precise semigroup constructions, see e.g. [12]. Depending on the physical situation one or the other is to be preferred. For instance, one might find that direct analytic continuation is possible via naturally occurring parameters, cf. the dilation group, but at the same time that it is impossible to prove the existence of an underlying semigroup [13]. Alternatively one might realise cases where the emergence of a singularly continuous spectrum makes the former approach impossible. We will not dwell further on these aspects except pointing out that the analytic approach appears more convenient in most of the applications so far addressed in quantum chemistry and chemical physics.

For convenience we will make a simple demonstration of how to transform a 2x2 matrix problem to complex symmetric form. In so doing we will also recognise the appearence of a Jordan block off the real axis as an immediate consequence of the generalisation. The example referred to is treated in some detail in Ref. [15], where in addition to the presence of complex eigenvalues one also demonstrates the crossing relations on and off the real axis. The Hamiltonian

$$H = H_0 + aV = -\tfrac{1}{2}\frac{d^2}{dr^2} - \frac{1}{r} + a\frac{e^2}{4} r^2 e^{-r} \tag{1}$$

describing the hydrogen atom perturbed by a barrier potential modelled by the parameter a displays how the Coulomb eigenvalues change when the parameter a varies from small to large values. As expected the well-known non-crossing rule is naturally obeyed. Complex eigenvalues, see below, may occur when either or both, the scaling factor η defined by

$$r \to \eta r;\ \eta = e^{i\vartheta};\ |\vartheta| > \tfrac{1}{2}|\arg(E_{res})| \tag{2}$$

where ϑ is a real angle (sufficiently large to uncover the complex energy) and the parameter a is made complex. For a general spectral theory behind the occurrence of non-real eigenvalues we refer to [11], but see also [16]. The prescription for the energy $E(a)$ and the matrix elements $H_{ij}(a)$; $i,j = 1,2$ are simply given as

$$E(a) = \frac{\langle \Psi(a^*)|H|\Psi(a)\rangle}{\langle \Psi(a^*)|\Psi(a)\rangle};\ \langle \Psi(a^*)|\Psi(a)\rangle \neq 0 \tag{3}$$

where $\Psi(a) = c_1\psi_1 + c_2\psi_2$ is obtained from the matrix, see below, via a complex symmetric similarity transformation yielding the canonical form

$$\begin{pmatrix} H_{11}(a) & H_{12}(a) \\ H_{21}(a) & H_{22}(a) \end{pmatrix} \to \begin{pmatrix} E_1(a) & 0 \\ 0 & E_2(a) \end{pmatrix} \tag{4}$$

with

$$H_{ij}(a) = \langle \psi_i(a^*)|H|\psi_j(a)\rangle;\ \langle \psi_i(a^*)|\psi_j(a)\rangle = \delta_{ij}. \tag{5}$$

In Eqs. (3-5) above we note that a^* occurs in the bra-position in order to produce matrix elements that are analytic in the parameter a. Eq. (4), however, displays the hope or wish that the matrix can be diagonalised and self-orthogonality avoided.

This is easily proved for Hermitean or self-adjoint problems, see e.g. POL's elegant linear algebra lectures that he finally collected into the excellent monograph *Linear Algebra for Quantum Theory* [17]. However, even if our matrix problem emanates from a so-called self-adjoint analytic family of operators, there is no guarantee that we will find simple eigenvalues, i.e. with Segrè characteristics equal to one. For more on this and its relation to complex symmetric forms, see e.g. Ref. [18]. Rather than discussing this important aspect we will demonstrate the dilemma as follows. Noting that we have a complex symmetric problem at hand for our generalised dynamical picture we consider the non-Hermitean case, writing (H_{11}, H_{22}, and υ are here chosen to be real)

$$H_{12} = H_{21} = \iota\upsilon;\ \upsilon \in \Re;\ \iota = \pm i \tag{6}$$

where the indicator ι is introduced for convenience, see further below. The result is somewhat surprising in that the non-Hermitean eigenvalue, sometimes referred to as "off the real axis", turns out to be real!

To see this in more detail we will attempt to diagonalise the matrix defined in Eq. (4) finding the surprising result that

$$\begin{pmatrix} H_{11} & \iota\upsilon \\ \iota\upsilon & H_{22} \end{pmatrix} \rightarrow \begin{pmatrix} E & 1 \\ 0 & E \end{pmatrix} \tag{7}$$

with eigenvalues given by

$$\lambda_{\pm} = \tfrac{1}{2}(H_{11} + H_{22}) \pm \sqrt{(H_{11} - H_{22})^2 - 4\upsilon^2} \tag{8}$$

exhibiting a degeneracy (on the real axis) when

$$H_{22} = H_{11} \pm 2\upsilon;\ \lambda_+ = \lambda_- = E = \tfrac{1}{2}(H_{11} + H_{22}) = H_{11} \pm \upsilon\,. \tag{9}$$

Note that any given value of υ (or prescribed difference between the two diagonal elements) results in two different sets of degenerate solutions in Eq. (9), one of which should be unphysical (for more on this see Ref. [14]).

Thus we learn three things: 1) the non-crossing rule is not obeyed in the present picture of unstable resonance states, 2) complex resonances may appear on the real axis and 3) unphysical states may appear as solutions to the secular equation. Thus avoided crossings in standard molecular dynamics are accompanied by branch points in the complex plane corresponding to Jordan blocks in the classical canonical form of the associated matrix representation of the actual operator.

To belabour the point on the occurrence of unphysical solutions made above, we will briefly refer to the issue of the Lamb shift in anti hydrogen, i.e.

$$\bar{p} + e^{+} \tag{10}$$

and the possibility of breaking the fundamental *CPT* invariance. The zero order problem is obtained from the Dirac Hamiltonian for $p + e^{-}$

$$\begin{aligned} H_{11}(E) &= \langle \psi_1 | e\phi + mc^2 + c^2(\sigma \cdot \pi)(-e\phi + mc^2 + E)^{-1}(\sigma \cdot \pi) | \psi_1 \rangle \\ H_{22}(E) &= \langle \psi_2 | e\phi + mc^2 + c^2(\sigma \cdot \pi)(-e\phi + mc^2 + E)^{-1}(\sigma \cdot \pi) | \psi_2 \rangle \end{aligned} \tag{11}$$

with $H_{12}(E) = H_{21}(E) = 0$. In Eq. (11) ψ_1 and ψ_2 correspond to the large and small components respectively, ϕ, σ and π have their usual meanings and the interaction with (10) in zero order is obtained by standard partitioning technique. Let us now try to model the fine structure and vacuum fluctuations by a small complex symmetric perturbation, thereby avoiding the usual renormalisation problem [14]. From Eqs. (7-9) we identify as the physical solution

$$\lambda = E = H_{11} + \upsilon \approx E^0 + \upsilon; \; \upsilon > 0 . \tag{12}$$

The model is oversimplified in the sense that we have not attempted to specify what effects are incorporated in υ. We will, however, consider the main effects from the vacuum fluctuations as well as other possible perturbations needed to produce the degeneracy above as well as, if necessary, considering the weak energy dependence in the Hamiltonian referred to in Eq. (11). To see how the *CPT* theorem affects our formulation we note that our zero order problem is an irreducible representation of

$$C \otimes \tau L_0 \tag{13}$$

where C carry the unit and charge conjugation and τL_0 contain translations in the Minkowski space (L_0 is the proper orthochronous inhomogeneous Lorentz group). Adding strong inversion, i.e. parity and time inversion, in (13) we obtain the full inhomogeneous Lorentz group τL. Consider now the following transformation

$$\langle \psi_i(x_\mu) | H | \psi_i(x_\mu) \rangle \rightarrow \langle \psi_j(-x_\mu) | \bar{H} | \psi_j(-x_\mu) \rangle; \; i,j = 1,2 \tag{14}$$

where $x_\mu = (x_1, x_2, x_3, ict)$ and $\bar{H}$ is the *CPT* transformed Hamiltonian. Assuming *CPT* symmetry we also deduce that

$$-E = \overline{H_{22}} - \upsilon \approx -E^0 - \upsilon \tag{12'}$$

and that the solutions $E^0 - \upsilon$ and $-E^0 + \upsilon$ are unphysical. However, if *CPT* symmetry is broken the presumed "unphysical solution" for the $\bar{p} + e^+$ system may in fact be the physical one. Even if this may not occur in the celebrated Lamb shift (rigorous measurements should, however, settle the question beyond any reasonable doubts) the present scenario is perceivable when degeneracies creating Jordan blocks appear. In the next section we will proceed to further discuss the consequences of emerging Jordan blocks on the associated dynamics.

3. THE DYNAMICAL PICTURE

As we have seen in the previous section, Jordan blocks appear easily in an ever so slight non-Hermitean extension of Quantum Mechanics [19]. The success in both atomic as well as molecular applications has also been noted [20-22]. Non-trivial extensions from the Hamiltonian to the Liouville picture was moreover soon realised [23].

To give a simple demonstration of the consequences of an existing Jordan block in the associated dynamics we will consider the frequency operator

$$\hat{T} = \frac{\hat{P}}{\tau}; \; \hat{P} = (\omega_0 \tau - i)\hat{I} + \hat{J} \tag{15}$$

In Eq. (15) we have introduced a fast time scale or frequency ω_0 and a life time τ. $\hat{I}$ is the unit and $\hat{J}$ the operator corresponding to a Jordan block representation.

In the Liouville formulation one obtains the well-known connection between τ and the imaginary part Γ of the complex resonance eigenvalue

$$\tau = \frac{h}{2\pi\,\Gamma} \tag{16}$$

Note that the appearance of a generic time scale is a characteristic property of a dissipative system and $\hat{T}$ generates its time evolution in scaled time units. Such time operators are strictly speaking forbidden in standard Quantum Mechanics, see Ref. [24] for further aspects on the problem, however, in open systems far from equilibrium they do not only exist but might also be useful in many applications, see below and [4-10, 13-15]. The form (15) has been investigated and obtained

from general principles, see Refs. [4] and [5]. In an r-dimensional representation

$$\mathbf{J} = \begin{pmatrix} 0 & 1 & 0 & . & 0 \\ 0 & 0 & 1 & . & . \\ . & . & . & . & 0 \\ . & . & . & . & 1 \\ 0 & . & . & . & 0 \end{pmatrix} \tag{17}$$

which leads to the nilpotent relations $\mathbf{J}^r = \mathbf{0}$; $\mathbf{J}^{r-1} \neq \mathbf{0}$ and to the evolution law

$$e^{-i\frac{\hat{P}t}{\tau}} = e^{-i\omega_0 t}\, e^{-\frac{t}{\tau}} \sum_{k=0}^{r-1} \left(\frac{-it}{\tau}\right)^k \frac{1}{k!}\, \hat{J}^k \tag{18}$$

and the associated resolvent expression

$$(\omega\tau\,\hat{I} - \hat{P})^{-1} = \sum_{k=1}^{r} \left((\omega - \omega_0)\tau + i\right)^{-k}\, \hat{J}^{k-1} \tag{19}$$

From Eqs. (18-19) one first deduces that the evolution of a dissipative system generated by $\hat{T}$ no longer follows a pure exponential decay law, rather the exponential is multiplied by a polynomial (in the limit of an infinite dimensional Jordan block one might consider more general evolution laws). Second, since the resolvent contains multiple poles the result will be a spectral gain-loss function that goes beyond the simple Lorentzian (or distorted Lorentzian) shape resulting from conventional simple pole structures in the related Green's function. Although we have not yet specified any physical representation, it is easy to comprehend the consequences of the shift operator $\hat{J}$. From appropriate matrix elements in Eq. (18) (note that the Liouvillian super operator formulation with its immediate relation to probability interpretations via a carrier space of density operators or density matrices still needs modification in a consistent resonance picture [9]) one may conclude that the conventional decay law

$$N(t) \propto e^{-\frac{t}{\tau}} \tag{20}$$

i. e. that the time dependent probability $N(t)$ of finding the system at time t in its

initial configuration obey the exponential behaviour (20) with the well-known decay rule

$$dN = -\frac{1}{\tau} N(t)dt \tag{21}$$

is modified to

$$N(t) \propto t^{r-1} e^{-\frac{t}{\tau}} \tag{22}$$

and

$$dN = t^{r-2}\left(r - 1 - \frac{t}{\tau}\right) N(t)dt \tag{22'}$$

for the highest power in the prepolynomial in Eq.(18). While we find $dN(t) < 0$ for all times in (21), the law deduced in (22) and (22') yields

$$dN(t) > 0;\ t < (\mathrm{r}-1)\tau\,. \tag{23}$$

Instead of disintegration of the initial state we find that the law in Eq.(23) leads to increasing (self)-organization during a finite number of life times τ. It is in this context that we will speak of microscopic selforganization in what follows below. As we need a more detailed description we will proceed to discuss the density operator concept in general and the reduced density matrices in particular.

4. THE DENSITY MATRIX

The density operator is the key quantity in quantum statistical physics and as the approach from the quantum world to the classical domain can conveniently be formulated by the Liouville equation

$$i\frac{\partial \rho}{\partial t} = \hat{\mathrm{L}}\rho \tag{24}$$

we first focus on ρ. The realization that the exact energy of a system of identical particles can be expressed by the second order reduced density matrix goes back to Husimi [25], for a recent review on reduced density matrices and the famous N-representability problem, see Coleman and Yukalov [26].

The N particle representable density matrix Γ is an idempotent projector

$$\Gamma = |\Psi(x_1...x_N)\rangle\langle\Psi(x_1...x_N)|;\ \mathrm{Tr}\{\Gamma\} = 1;\ \Gamma^2 = \Gamma;\ \Gamma^\dagger = \Gamma \tag{25}$$

on the wave function $\Psi(x_1...x_N)$ representing the many-body quantum mechanical system. Quantum statistical formulations concern traditionally two types of particles, i. e. bosons and fermions and since quantum chemical applications often center on the electronic structure of atoms and molecules it is customary to investigate in particular the characteristics of fermion reduced density density matrices. Moreover, the electron correlation properties are known to be contained in the two matrix and it has therefore been (and still is) a much wanted goal to find the precise conditions under which a reduced density matrix actually corresponds to a wave function or at least an ensemble, see [26]. It is convenient to define the reduced matrices for fermions [2]

$$\Gamma^{(p)}(x_1...x_p|x_1'...x_p') = \binom{N}{p}\int\Psi^*(x_1...x_p,x_{p+1}...x_N)\Psi(x_1'...x_p',x_{p+1}...x_N)dx_{p+1}...dx_N \tag{26}$$

so that, e. g. for $p=2$ it is normalised to he number of pairings of N fermions. Other choices are that of Coleman [27] and Yang [28], see below:

$$\begin{aligned}&\text{Löwdin: } \Gamma^{(p)};\ \mathrm{Tr}\{\Gamma^{(p)}\} = \binom{N}{p}\\ &\text{Coleman: } D^{(p)};\ \mathrm{Tr}\{D^{(p)}\} = 1\\ &\text{Yang: } \rho^{(p)};\ \mathrm{Tr}\{\rho^{(p)}\} = p!\binom{N}{p}.\end{aligned} \tag{27}$$

The work cited above are important in yet another very important aspect, namely they bear upon the possible sizes or limits that the corresponding eigenvalues may obtain ranging from the independent particle model to a fully pair coherent condensate. The celebrated concept of off-diagonal long-range order (*ODLRO*) introduced by Yang in connection with his proof of the largest bound for $\Gamma^{(2)}$ has been of central importance in super conductivity and -fluidity and it will also play a central role below. In connection with this development Coleman [27] identified the precise condition for the so-called extreme state which in certain cases could develop *ODLRO*. The observation that a simple AGP-function gave the largest possible eigenvalues of the 2-matrix was given by Sasaki [29] in a report that, according to Ref. [26], was delayed through a misplacement by the publisher. The clue to describe a convenient representation for open systems will be obtained via the extreme state as follows:

Starting from a set of localised pair functions or geminals $\mathbf{h} = (h_1, h_2, \ldots h_m)$ obtained from appropriate pairing of one particle basis spin functions, we will devise the following transformations, the motive to be explained below,

$$\mathbf{B} = \frac{1}{\sqrt{m}} \begin{pmatrix} 1 & \omega & \omega^2 & . & \omega^{m-1} \\ 1 & \omega^3 & \omega^6 & . & \omega^{3(m-1)} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & \omega^{2m-1} & \omega^{2(2m-1)} & . & \omega^{(m-1)(2m-1)} \end{pmatrix}; \omega = \mathrm{e}^{\frac{i\pi}{m}} \tag{28}$$

introducing a coherent and a correlated basis $\mathbf{g}$ and $\mathbf{f}$ respectively through

$$\begin{aligned} |\mathbf{h}\rangle\mathbf{B} &= |\mathbf{g}\rangle = |g_1, g_2, \ldots g_m\rangle \\ |\mathbf{h}\rangle\mathbf{B}^{-1} &= |\mathbf{f}\rangle = |f_1, f_2, \ldots f_m\rangle \end{aligned} \tag{29}$$

There are several inter-connections to be mentioned here. The first one concerns the extreme state. If $\mathbf{h}$ is the set of two particle determinants and the AGP wave function is constructed from g_1, see Coleman [27] for the exact condition for the extreme state, the two-matrix (save the "tail contribution" from the remaining pair configurations) can be expressed as

$$\Gamma^{(2)} = \Gamma_{\mathrm{L}}^{(2)} + \Gamma_{\mathrm{S}}^{(2)} = \lambda_{\mathrm{L}} |g_1\rangle\langle g_1| + \lambda_{\mathrm{S}} \sum_{k=2}^{m} |g_k\rangle\langle g_k| \tag{30}$$

with the eigenvalues given by

$$\lambda_{\mathrm{L}} = \frac{N}{2} - (m-1)\lambda_{\mathrm{S}}; \; \lambda_{\mathrm{S}} = \frac{N(N-2)}{4m(m-1)} \tag{31}$$

Note that we have m basis pairfunctions or geminals and $2m$ spin orbitals. The number of fermion pairings and pair configurations are

$$\binom{N}{2}; \binom{2m}{2}$$

respectively. The dimension of the "box contribution" defined in Eq. (30) is m and consequently the dimension of the "tail contribution", not explicitly written out in Eq. (30), is $2m(m-1)$.

Since the eigenvalue associated with the tail is the same as λ_S [27] we have for the extreme state one large eigenvalue λ_L and an $(m-1)(2m+1)$ degenerate eigenvalue λ_S. As have been shown previously analogous equations can be derived in a statistical framework for the case of localised fermions in a specific pairing mode and/or bosons subject to a quantum transport setting, see [4,5,7] for more details. The relevance of the basis **f** will be demonstrated in the following section

5. TEMPERATURE AND ENTROPY

It is a well-known trick to make time imaginary by formally including temperature in the formulation through the relation

$$t \to t - i\beta;\ \beta = \frac{1}{kT} \tag{32}$$

where k is Boltzmann's constant and T the absolute temperature in Kelvin. In the extended dynamical picture presented here this analogy is not exactly true. To embody the temperature in a non-contradictional way we can resort to the Bloch equation via

$$-\frac{\partial \rho}{\partial \beta} = \hat{L}_B \rho \tag{33}$$

where $\hat{L}_B$ is to be distinguished from the Liouville generator in Eq. (24), i.e.

$$\hat{L}_B = \frac{1}{2}\{H|\ \rangle\langle *| + |\ \rangle\langle *|H\} \tag{34}$$

Note the complex conjugate in the bra-position compared to

$$\hat{L} = H|\ \rangle\langle\ |-|\ \rangle\langle\ |H^{\dagger} \tag{35}$$

where we admit the hermitean conjugate in the commutator in Eq. (35). For more on this, see Ref. [10]. Applying now Eqs. (33-34) to

$$\Gamma_S^{(2)} = \lambda_S \sum_{k=2}^{m} |g_k\rangle\langle g_k| = \lambda_S \sum_{k=1}^{m}\sum_{l=1}^{m} |h_k\rangle(\delta_{kl} - \frac{1}{m})\langle h_l| \tag{36}$$

we obtain

$$e^{-\beta \hat{L}_B}\tilde{\Gamma}_S^{(2)} = \lambda_S e^{-\beta E}\sum_{k=1}^{m}\sum_{l=1}^{m} |h_k\rangle e^{i\beta\frac{1}{2}(\varepsilon_k+\varepsilon_l)}(\delta_{kl} - \frac{1}{m})\langle h_l *| \tag{37}$$

where E is the degenerate (real part) energy and ε_k the imaginary part of the energy related to the life time as (h is Planck's constant)

$$\varepsilon_k = \frac{\Gamma_k}{2} = \frac{h}{4\pi\tau_k} \tag{38}$$

The tilde over the two matrix indicates an extended representation according to the analyticity requirements given for the resonance formulation above.

To appreciate the formulas (36) and (37) we will remind the reader of the result quoted in Ref. [18], i.e.

$$\mathbf{Q} = \mathbf{B}^{-1}\mathbf{J}\mathbf{B} \tag{39}$$

where ($r = m$)

$$\mathbf{Q}_{kl} = (\delta_{kl} - \frac{1}{m})e^{i\frac{\pi}{m}(k+l-2)}; \; k,l = 1,2,..m \tag{40}$$

Hence, if we can find a relation between ε_k, β and r so that Eq. (37) assumes the Jordan block form (i.e. proportional to $\mathbf{Q}$ or $\mathbf{J}$) one of the consequences for the dynamics is the emergence of an dramatically increased lifetime as indicated in our discussion in section 3. Hence if (see [10] for more details)

$$\beta\varepsilon_l = 2\pi\frac{l}{r}; \; l = 1,2,...r \tag{41}$$

we find that the thermalised density matrix in Eq. (37) becomes proportional to $\mathbf{Q}$. A similar analysis with $m = 2$ obtains for $\Gamma_L^{(2)}$. Using (32), (38) and (41) one finds that the longest relaxation time τ_{rel} (for $l = 1$ above) compatible with the smallest dimension or "size" r_{min} is given by

$$\beta\varepsilon_1 = \frac{h}{4\pi kT\tau_{rel}} = \frac{2\pi}{r_{min}}. \tag{41'}$$

Thus we obtain a direct relation between the absolute temperature, the relaxation time and the "minimal size" of our open system

$$\frac{h}{2\pi}\tau_{rel}^{-1} = \frac{4\pi kT}{r_{min}} \tag{42}$$

where the relaxation time is given by

$$\tau_{rel} = r_{min}\tau_{lim};\ \tau_{lim} = \frac{h}{8\pi^2 kT} \tag{43}$$

Note that τ_{lim} is the shortest time commensurate with the uncertainty relation.

The theory above has been applied in a variety of realistic situations. The range includes ionic conductance in aqueous solutions and molten alkali chlorides, damped spin-wave behaviour in paramagnetic systems, stimulated emission of radiation in masers, the fractional quantum Hall effect and quantum correlations in high-T_C cuprates and other non-BCS superconductors [4, 5, 7, 8, 14, 30]. In the next section we will also make some comments on the problem of long-range transcorrelations of protons in DNA [31].

Before proceeding we comment on the relevance of the **f**-basis. It appears through Eq. (39) and the unitarity property of **B**, namely (37) and (41) which connects $\Gamma^{(2)}$ with

$$|\mathbf{h}\rangle\mathbf{Q}\langle\mathbf{h}| = |\mathbf{h}\rangle\mathbf{B}^{-1}\mathbf{J}\mathbf{B}\langle\mathbf{h}| = |\mathbf{f}\rangle\mathbf{J}\langle\mathbf{f}| = \sum_{k=1}^{r-1}|f_k\rangle\langle f_{k+1}| \tag{44}$$

The vectors of $\mathbf{B}^\dagger$ express **f** in the original basis **h**, see Eqs. (28-29). To prepare for a discussion of proton tunnelling in DNA we will briefly deliberate on the concept of the micro entropy. A direct generalisation of Gibbs entropy follows from the definition below (Coleman's normalisation [27])

$$S_G^{(2)} = -k\mathrm{Tr}\left\{D^{(2)}\ln(D^{(2)})\right\} \tag{45}$$

where the superscript (2) refers to the pair entropy. We note that Gibbs classical entropy formula is only meaningful for equilibrium situations. In the present situation we will apply it to the second order reduced density matrix (before thermalization etc.) and then draw some simple conclusions from it.

A subtle point is the non-linear nature of the reduced dynamics, which may lead to surprises if not properly understood. This may also be considered as a "close to equlibrium" situation in contrast to the extended dynamical picture where resonances and Jordan forms may appear.

Applying Eq. (45) to the Hartree-Fock independent particle model we immediately

obtain from the fact that all the eigenvalues (equal to the number of pairings) of $\Gamma^{(2)}$ are equal to unity, and that m equals the number of pairs that (note that $D^{(2)}$ is normalised to unity)

$$S_G^{(2)} = k\ln\left\{\binom{N}{2}\right\} \tag{46}$$

Applying the same argument to the extreme case, Eqs. (30) and (31), including also the "tail contribution", one finds the surprising result that

$$S_G^{(2)} \to \infty;\ m \to \infty \tag{47}$$

depending on the large degeneracy associated with the small eigenvalue λ_S, see [27]. One way to remedy the situation is to evaluate the entropy (45) over the "box" as defined explicitly in Eqs. (30) and (31). Properly trace-normalised to unity we obtain for the Hartree-Fock model and the extreme case, respectively

$$S_G^{(2)} = k\ln\left\{\frac{N}{2}\right\} \tag{46'}$$

$$S_G^{(2)} \to 0;\ m \to \infty \tag{47'}$$

where as expected the entropy goes to zero when the pair condensate is 100%. Other possibilities would be to refrain from taking the limit and let e. g. m be equal to the number of pairings or alternatively choose the quotient between N and m as constant (in Eq. (45)). In either case one would obtain expressions that contain contributions from the large eigenvalue (the rest being dominated by the enormous degeneracy for the small eigenvalues). Finally one could use Eqs. (30) and (31), Coleman normalised as $D^{(2)}$, and then take the limit $m \to \infty$ to obtain

$$S_{G,L}^{(2)} = k\frac{\ln(N-1)}{N-1} \tag{48}$$

This is an attractive pair entropy since it refers to the weight of physical pairs in all possible pairings.

6. SELFORGANIZATION

We will apply the results of the previous section to the question of long-range proton correlations in DNA referring to some of the early suggestions of P. O. Löwdin on genetic information and the problem of mutation rates and aging [32].

Most of these ideas did not evolve depending on limited experimental data at the time. Notwithstanding dramatic developments in molecular biology there still remain many fundamental questions that have not yet been elucidated by a rigorous quantum theoretic model. The screw like symmetry of the double helix and the number of "stairs" in a full rotation are complicated features to account for by quantum mechanical and molecular mechanical simulations as well as possibly relevant long-range thermal correlation properties of the smallest microscopic self-organizing units co-operating in the open *in vivo* system. To study this further we will first consider a double helix of identical $C-G$ base pairs, e. g. Cytosine and Guanine, which are known to form a lefthanded double helix with a 30 degree angle between two consecutive "stairs" forming a full rotation in 12 base-pair units.

Conceiving a corrrelated model of tunnelling protons we will first discuss the smallest case $N=4$. It follows immediately that the number of pairings is

$$r=\binom{N}{2}=\binom{4}{2}=6;\ \lambda^{(1)}=\frac{N}{2r}=\frac{1}{3} \tag{49}$$

where $\lambda^{(1)}$ is the probability of finding a proton pair in a particular correlated state. The entropy (48) for the pairing system is then simply given by

$$S_{\mathrm{G}}^{(2)}=-k\frac{1}{3}\ln\left(\frac{1}{3}\right) \tag{50}$$

in which we note that every unit, see below, is in contact with an environment of other similar or identical units. To see the latter we consider the vector in Eq. (28)

$$\mathbf{B}_{\mathbf{s}}^{-1}=(1,\omega^{2s-1},\omega^{2(2s-1)},\ldots\omega^{(m-1)(2s-1)})^{\dagger};\ \omega=e^{\frac{i\pi}{m}} \tag{51}$$

corresponding to f_s; $s\le m$ in the **h** basis. In a given model, it is easy to see that a general vector in Eq. (51), will contain repetitive blocks of sub-vectors of dimension 6, corresponding to $r=6$, if m is large enough and contains appropriate factors of 6. Note that the smallest non-trivial factor is of course 3. The vector (51) for $m=r=6$ implies that we need 6 correlated pairs (or base pairs) to "get around" 180 degrees in the vector space. In an equivalent cyclic model 12 pairs would be needed for a full rotation but since we "do not return" to the origin here we have instead a pseudo-cyclic "screw-like" symmetry reflected by (28) and (51). Since every column in Eq. (28), let us say for $s\le m/2$ in Eq. (51), is the complex conjugate of the column $m+1-s$, we have in fact a kind of "left -right" symmetry additionally built into the formulation.

One might speculate what would be the consequences of the appearance of a large prime number $m = p$, occurring in (51), since the only "repetitive vector" in (28) would then be the middle one containing sub-blocks of (1,-1).

Note that the Gibb's partition of the micro entropy above is not maximum for a near equilibrium situation. This follows from the fact that the function $-x\ln x;\ 0 \le x \le 1$, has a maximum for $x = 1/e$. Assuming that our correlation model in a more realistic dissipative system of varying occurrence of base pairs would "prefer" a "close to equilibrium" entropy, we find that a recalculation of our parameters above yields the trend

$$N-1 \to e; r = \binom{N}{r} \to \binom{e+1}{2} \approx 5.05 \tag{52}$$

This would indicate a less orderly double helix with about 10-11 base pairs in a full turn as well as a possible change in directions of the screw. One might further imagine that a coherent situation linking together very many units would cause a sudden drop in the pair entropy, corresponding to a large N in Eq. (48), resulting in macroscopic selforganization of all participating base pairs, thereby extending the characteristic lifetime for the relevant units of protonic pairs. A more detailed study would combine the microscopic theory in section 5 with the general dynamical picture in section 3.

7. CONCLUSIONS

These coincidences might of course be incidential and fortuitous. Nevertheless it is quite remarkable that so many complex features are reproduced by the same simple correlation matrix approach. We have seen how a slight extension of quantum mechanics to rigorously incorporate resonance processes leads to the formation of so-called Jordan blocks. Indeed the existence of such non-diagonal structures seems more to be the rule than the exception. The dynamics emerging from such a degenerate model with a nilpotent part added was demonstrated to have remarkable properties identified as microscopic self-organisation. The generating time operator was given by

$$\hat{T} = (\omega_0 - i\frac{1}{\tau})\hat{I} + \frac{1}{\tau}\hat{J} \tag{53}$$

where ω_0 is the given "pump frequence", τ the phenomenological lifetime and $\hat{I}$, $\hat{J}$ the identity- and the nilpotent operators respectively. We have further derived explicit expressions for this operator in various situations where the main features are the description of phenomena defined at a specific level (e.g. microscopic or mesoscopic) to a higher phenomenological level (mesoscopic or macroscopic).

The present perspective is reminiscent of the modern complexity bias in that certain generic forms recur on various organisation levels provoking comprehensive and debatable modes of understanding. However, we emphasis here that our derivations are firmly established in

1) quantum theory at the fundamental microscopic level
2) density matrix theory at the Liouville master equation level
3) dilation analytic (or other) theory for resonance formation
4) Jordan algebra's for the time evolution
5) non-equilibrium quantum statistical mechanics for the pair entropy

Per-Olov Löwdin contributed to all of these in addition to being the grand maestro of Quantum Chemistry. The field owes its rigor to him and yet he did not hesitate to take on new challenges, provoke new formulations, affirm personal views maintaining the highest quality and distinction in his unbiased judgment and assessment. Attuning to this leitmotif we dedicate this review to him.

8. REFERENCES

1. P. O. Löwdin, J. Phys. Chem. **61,** 55 (1957).
2. P. O. Löwdin, Phys Rev. **97,** 1474 (1955); *ibid.* 1490 (1955); 1509 (1955).
3. E. B. Karlsson and E. Brändas, Phys. Scripta, **T76,** 7 (1998).
4. E. Brändas, in *Dynamics during Spectroscopic Transitions,* eds. E. Lippert and J. D. Macomber, Springer Verlag, Berlin, p. 148 (1995).
5. E. Brändas, Adv. Chem. Phys. **99,** 211 (1997).
6. C. A. Chatzidimitriou-Dreismann, Adv. Chem. Phys. **80,** 201 (1991).
7. E. Brändas and C. A. Chatzidimitriou-Dreismann, Int. J. Quant. Chem. **40,** 649 (1991).
8. E. J. Brändas, L. J. Dunne and J. N. Murrell, in *New Trends in Quantum Systems in Chemistry and Physics,* Kluwer, Dordrecht, Vol. 2, p. 289 (2000).
9. E. Brändas and B. Hessmo, Lecture Notes in Physics **504,** 359 (1998).
10. E. Brändas and C. A. Chatzidimitriou-Dreismann, Lecture Notes in Physics Vol. **325,** 486, (1989).
11. E. Balslev and J. M. Combes, Commun. Math. Phys. **22,** 280 (1971); C. van Winter, J. Math. Anal. **47,** 633 (1974).
12 A. Bohm and M. Gadella, *Dirac Kets, Gamow vectors and Gel'fand Triplets,* Springer Verlag, Berlin, (1989).
13. J. Kumicák and E. Brändas, Int. J. Quant. Chem. **46,** 391 (1993).
14. E. J. Brändas in *Dynamics of Transport in Plasmas and Charged Beams,* eds. G. Maino and M. Ottaviani, World Scientific Publ. Comp., 160 (1996).
15. M. A. Natiello, E. J. Brändas and A. R. Engelmann, Int. J. Quant. Chem. **S21,** 555 (1987).

16. B. Simon, Ann. Math. **97,** 247 (1973).
17. P. O. Löwdin, *Linear Algebra for Quantum Theory* Wiley, New York, (1998).
18. C. E. Reid and E. J. Brändas, Lecture Notes in Physics, **325,** 475 (1989).
19. E. Brändas and P. Froelich, Int. J. Quant. Chem. **13,** 619 (1978)
20. W. P. Reinhardt, Ann. Rev. Phys. Chem. **33,** 223 (1982).
21. B. R. Junker, Adv. Atom. Mol. Phys. **18,** 207 (1982).
22. N. Moiseyev, Lecture Notes in Physics **211,** 235 (1984).
23. Ch. Obcemea and E. Brändas, Ann. Phys. **151,** 383 (1983).
24. I. Prigogine, *From Being To Becoming* W. H. Freeman and Company, San Francisco, (1980).
25. K. Husimi, Proc. Phys. Math. Society Japan, **22,** 264 (1940).
26. A. J. Coleman and V. I. Yukalov, *Reduced Density Matrices Coulson's Challenge,* Lecture Notes in Chemistry, **72,** Springer-Verlag, Berlin (2000).
27. A. J. Coleman, Rev. Mod. Phys. **35,** 668 (1963).
28. C. N. Yang, Rev. Mod. Phys. **34,** 694 (1962).
29. F. Sasaki, Technical Report **77,** Uppsala Quantum Chemistry Group (1962); Phys. Rev. **138B,** 1338 (1965).
30. E. J. Brändas, Ber. Bunsenges. Phys. Chem. **96,** 49 (1992).
31. E. J. Brändas, Int. J. Quant. Chem. **46,** 499 (1993).
32. P. O. Löwdin, Biopolymers Symposia 1, 161-181 (1964).

From Few-Atom to Many-Atom Quantum Dynamics

David A. Micha

Quantum Theory Project
Departments of Chemistry and of Physics
University of Florida
Gainesville, FL 32611-8435, USA

Contents

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

1 Introduction

A quantum mechanical theory is in principle needed to describe molecular phenomena in both few-atom and many-atom systems . In some cases a single electronic state is involved, and it is possible to gain valuable insight using only classical molecular dynamics, which can be relatively easy to apply even for a system of many atoms. A quantum mechanical description of molecular phenomena is however clearly needed for electronic states, insofar these have pronounced wavemechanical properties. The need for a quantum description of nuclear motions in molecular dynamics is less apparent, but it is required in some important situations. If we consider a generic interaction between two species $A(\alpha)$ and $B(\beta)$ leading to formation of two others, $C(\gamma)$ and $D(\delta)$, all of them in the specified quantum states, so that

$$A(\alpha) + B(\beta) \rightarrow C(\gamma) + D(\delta)$$

it is clear that one must describe rotational and vibrational intramolecular motions as quantized to begin and end up with, so that quantization at all times is called for. During the interaction, nuclear motions are coupled to quantized electronic states, and light nuclei such as protons can tunnel quantum mechanically. Resonance (or long-lived) molecular states may also form during interactions, with properties similar to quantized stationary states but with finite lifetimes. Therefore it is necessary in many cases to work with a quantum dynamics of molecular interactions.

The focus of this chapter is on molecular interactions leading to electronically diabatic state-to-state transitions relevant to chemical kinetics, brought about by collisions, and electronic transitions of interest in photochemistry, induced by absorption of light. Their description involve two challenges in quantum molecular dynamics. One of them is to properly describe the *couplings of electronic and nuclear motions*, which occur in different time scales (with fast electron transitions coupled to slow nuclear motions) and different quantum regimes (the electron dynamics is fully quantal while nuclear dynamics is frequently quasiclassical). The other challenge is to describe *many-atom systems*, with correspondingly many degrees of freedom for nuclear motions.

These challenges can be dealt with the powerful mathematical tools of quantum chemistry, as advocated by Per-Olov Lowdin.[1, 2, 3, 4] In our studies, linear algebras with matrices,[4] partitioning techniques,[3] operators and superoperators in Liouville space, and the Liouville-von Neumann

equation for the density operator,[5, 6, 7, 1, 2, 8] are all relevant and will be used in what follows. The material covered here is not a review of the subject, which involves many publications and has been periodically reviewed elsewhere,[9, 10, 11, 12, 13, 14, 15, 16, 17, 18] but it is instead an overview of research in our group, using the mathematical methods advocated by Lowdin. In particular, our work on molecular interactions has involved partitioning techniques,[19, 20, 21] wave-operator and many-body methods applied to atom-diatom and atom-polyatomic collisions,[22, 23, 24, 25, 26, 27] linear algebra methods in the solution of differential equations appearing in time-dependent many-electron theory,[28, 29] and more recently density matrix methods developed to treat dissipative dynamics in many-atom systems,[30, 31, 32, 33, 34] and applied to surface phenomena.[35, 36, 37, 38]

2 Few-atom systems

2.1 The Born-Oppenheimer description

We deal first with electronic-nuclear coupling in systems with few atoms, and therefore few degrees of freedom for nuclear motions, so that we can concentrate on the first mentioned challenge. The structure and properties of a molecule in stationary states are well described within the Born-Oppenheimer picture in which the disparity in masses of nuclei and the electron, with $m_n >> m_e$ and the similarity of Coulomb forces on nuclei and electrons, $F_n \approx F_e$, mean that within a short time interval Δt, changes in velocities satisfy $\Delta v_n = F_n/(m_n \Delta t) << \Delta v_e = F_n/(m_n \Delta t)$ so that an interaction involving small velocities to begin with, and lasting a short time, would be described by slow nuclei. The well known Born-Oppenheimer prescription is then to construct the electronic Hamiltonian $H_Q^{(el)}$ for fixed nuclear positions $Q = (\vec{R}_1, ..., \vec{R}_N)$, to calculate electronic states $\Phi_I^{(el)}(q;Q)$ for electron variables $q = (\vec{r}_1, \zeta_1, ..., \vec{r}_n, \zeta_n)$ and fixed Q, and the related potential energy surfaces (PES's) $E_I(Q) = \langle \Phi_I^{(el)} | H_Q^{(el)} | \Phi_I^{(el)} \rangle$. Molecular structures and interaction potentials between molecules are given by the topography of PES's: minima, saddle points, valleys, etc.; and properties for fixed conformations of a molecule can be obtained as expectation values of operators.

The dynamics of nuclei on PES's and the transitions $I \rightarrow I'$ between electronic states, can be described with either a time-independent or a time-

dependent approach. For the collision-induced and photo-induced phenomena of interest here, it is usually found that initial states are non-stationary, in which case the time-dependent description is a natural one, and is next followed.

Expanding the total molecular wavefunction $\Psi(q,Q,t)$ in a basis set of N_{st} electronic states $\{\Phi_I^{(el)}(q;Q), I = 1 - N_{st}\}$ one finds

$$\Psi(q,Q,t) = \sum_I \Phi_I^{(el)}(q;Q) F_I(Q,t) \tag{1}$$

where F_I is a nuclear motion wavefunction . Substitution in the time-dependent Schroedinger equation leads to

$$\sum_{I'}\{(2M)^{-1}[\delta_{II'}(\hbar/i)\partial/\partial Q + G_{II'}(Q)]^2 \tag{2}$$
$$+H_{II'}^{(el)}(Q)\}F_{I'}(Q,t) = i\hbar\partial F_I/\partial t$$

$$G_{II'}(Q) = -i\hbar \int dq\, \Phi_I^{(el)}(q;Q)^* \partial\, \Phi_{I'}^{(el)}(q;Q)/\partial Q \tag{3}$$

which displays the momentum couplings $G_{II'}$. The electronic basis functions can be chosen to generate the adiabatic representation where $H_{II'}^{(a)}(Q) = \delta_{II'}E_I(Q)$, or a diabatic representation, with $G_{II'}^{(d)} = 0$. When light is present, its electric field couples to the molecular electric dipole operator and this leads to the appearance of the transition dipole couplings $D_{I'I}(Q)$.

The coupling of nuclear and electronic motions in electronic transitions may then happen through $D_{II'}(Q)$ if induced by light, or through $H_{II'}^{(el)}$ and $G_{II'}$ if induced by nuclear displacements. The nuclear motion functions can be obtained by additional expansion in a basis of functions of nuclear coordinates, or by numerical solutions on a grid of points in the space of nuclear positions. The second approach is specially suitable for non-stationary states, and is briefly described.

For a molecular system with only a few active electronic states, such as $N_{st} \leq 10$, and for example two degrees of freedom $Q = (x,y)$ and mass M for nuclear motions, it is convenient to represent operators with matrices in the electronic basis, whose elements are yet operators but only on the nuclear degrees of freedom. The equation for nuclear motions in the diabatic representation is

$$\hat{\mathbf{H}}\mathbf{F}(t) = i\hbar\partial\mathbf{F}/\partial t$$
$$\hat{\mathbf{H}} = \hat{\mathbf{K}} + \mathbf{V}(x,y,t), \quad \hat{\mathbf{K}} = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\mathbf{I} \tag{4}$$

where $\mathbf{I}$ is the unit matrix.

The time-evolution generated by the time-dependent hamiltonian is given by a time-ordered exponential form,

$$\mathbf{F}(t) = T\ exp[-i \int dt'\ \hat{\mathbf{H}}(t')/\hbar]\mathbf{F}(0) \tag{5}$$

which can be numerically implemented with a variety of methods, one of which relies on the split operator propagator for a time step $\Delta t = t_1 - t_0$,

$$exp[-i\hat{\mathbf{H}}(t_{1/2})\Delta t/\hbar] = exp(-i\mathbf{V}_{1/2}\Delta t/2\hbar)exp(-i\hat{\mathbf{K}}\Delta t/\hbar) \tag{6}$$

$$\times exp(-i\mathbf{V}_{1/2}\Delta t/2\hbar) + \mathcal{O}[(\Delta t)^3] \tag{7}$$

with $\mathbf{V}_{1/2}$ evaluated at $t_{1/2} = t_0 + \Delta t/2$, and $exp(-i\hat{\mathbf{K}}\Delta t/\hbar)$ obtained from a fast Fourier transform between nuclear position and momentum grids.

While this approach is quite accurate and general, it can be very demanding of computing time and data storage when the number of degrees of freedom N_F increases, since the number of grid points grows geometrically with N_F. In addition it becomes increasingly difficult to parameterize PESs and to extract physical insight from the dynamics as N_F increases. An alternative which bypasses these two obstacles is to introduce a quasiclassical description of the nuclear motions and to directly generate the electronic states along nuclear trajectories, in a *first principles dynamics* treatment.

2.2 A first principles quantum dynamics

An approach briefly presented here is based on a combination of the eikonal (or short wavelength) approximation for nuclei, and time-dependent Hartree-Fock states for the many-electron system, in what we have called the *Eikonal/TDHF* approach.[13] A similar description can be obtained with narrow wavepackets for the nuclear motions. Several other approaches have recently been proposed for doing first principles dynamics, a very active area of current research.[39, 11, 15]

Our approach to the dynamics of complex electronic rearrangements, has been based on an *eikonal representation* of the molecular wavefunction. [40, 41, 42] In this representation, wavefunctions are written in the form $\chi(q.Q,t)exp[iS(Q,t)/\hbar]$, with a factorized exponential function of classical-like variables Q, where S is a classical-like mechanical action. It can be applied without detailed preliminary knowledge of electronic rearrangements,

such as may happen at potential energy crossings or avoided crossings. The usual eikonal approximation is obtained in the limit of short deBroglie wavelengths $\lambda_{dB} = P/h << a_B$, the Bohr radius, with P a local momentum for nuclei. The procedure provides a straightforward derivation of transition integrals and expectation values expressed as sums over initial values of trajectories and can be described as an initial value method.

Taking the nuclear coordinates Q to be classical-like, the eikonal representation gives the wavefunction $\Psi(q,Q,t)$ for an initial electronic state I as a superposition of functions, of the form

$$\Psi_I(q,Q,t) = \int d\Lambda\, a(\Lambda)\chi_I(q,Q,t;\Lambda)exp[iS(Q,t;\Lambda)/\hbar] \tag{8}$$

with parameters Λ (such as initial values of momenta and coordinates) and combination coefficients $a(\Lambda)$ chosen to construct the initial state, $\Psi^{(in)}(q,Q)$ from $\chi^{(in)}(q,Q;\Lambda)$ and $S^{(in)}(Q;\Lambda)$. Here the variables q refer to electronic coordinates and spin, but more generally they could also include the coordinates of protons if they must be treated as quantum variables, for example in studies of proton transfer requiring phase interference and tunnelling. The function $S(Q,t)$ is chosen to be real. The pre-exponential factor $\chi(q,Q,t)$ is however complex and has its own phase, dependent on the electronic state. Differential equations satisfied by these two functions can be obtained quite generally [31] from the Dirac-Frenkel time-dependent variational principle (TDVP). [43] The equation for S is, for given I,

$$\frac{1}{2M}(\frac{\partial S}{\partial Q})^2 + V_{qu}(Q,\frac{\partial S}{\partial Q},t) + \frac{\partial S}{\partial t} = 0 \tag{9}$$

and the one for χ is

$$[\frac{1}{2M}(\frac{\hbar}{i}\frac{\partial}{\partial Q} + \frac{\partial S}{\partial Q})^2 + \hat{H}_Q - (i\hbar\frac{\partial}{\partial t} - \frac{\partial S}{\partial t})]\chi(q,Q,t) = 0 \tag{10}$$

where $\hat{H}_Q$ is the electronic Hamiltonian operator obtained fixing nuclear coordinates Q. It follows from the first equality that S is a solution of a Hamilton-Jacobi equation and can be interpreted as a mechanical action governed by a quantum potential V_{qu} while χ satisfies a time dependent Schroedinger equation with shifted momenta and energy. The quantal potential is given, for a

system starting initially in state I, by $V_{qu,I} = V_I + V_I' + V_I''$, [40, 42] with

$$\begin{aligned} V_I &= \langle\chi_I|H_Q|\chi_I\rangle/\langle\chi_I|\chi_I\rangle \\ V_I' &= \frac{i\hbar}{2M}\frac{\partial S}{\partial Q}\cdot\left[\langle\frac{\partial\chi_I}{\partial Q}|\chi_I\rangle - \langle\chi_I|\frac{\partial\chi_I}{\partial Q}\rangle\right]/\langle\chi_I|\chi_I\rangle \\ V_I'' &= -\frac{\hbar^2}{2M}\frac{1}{2}\left[\langle\frac{\partial^2\chi_I}{\partial Q^2}|\chi_I\rangle + \langle\chi_I|\frac{\partial^2\chi_I}{\partial Q^2}\rangle\right]/\langle\chi_I|\chi_I\rangle \end{aligned} \tag{11}$$

where the first term is identified as the Ehrenfest potential with its meaning of a classical effective potential, and the following terms are quantal corrections. The bracket notation here and in what follows signifies an integration only over electronic variables, so that the brackets are yet functions of nuclear variables. The quantal corrections are small in regions where the function χ varies slowly over distances of the order of the deBroglie wavelength of the nuclear momenta. These equations are formally exact. In practice, the quantal corrections are neglected or included only as corrections for short de Broglie wavelengths, in improved eikonal approximations. The effective potential is then $V_{qu} = V_I$ or, if an average is taken over several initial states I with weights w_I, it is $\bar{V}_{qu} = \sum w_I V_I$.

The mechanical action can be constructed introducing the momenta $P = \partial S/\partial Q$, from solutions to the hamiltonian equations with the same potential,

$$\begin{aligned} \partial P/\partial t = -\partial H_{qu}/\partial Q\ ,\ \partial Q/\partial t = \partial H_{qu}/\partial P \\ H_{qu} = P^2/(2M) + V_{qu}(Q,P,t) \end{aligned} \tag{12}$$

where now P and Q can be interpreted as the position and momentum variables of trajectories originating at initial values P_{in} and Q_{in} and evolving as $P(t;P_{in},Q_{in})$ and $Q(t;P_{in},Q_{in})$. Some initial values of generalized momenta may be fixed by initial conserved quantities such as energy and angular momentum, while the others are available to combine the elementary solutions χ_I and construct the initial states, so that one can write that $a(\Lambda) = a(P_{in},Q_{in})$ and integrate over the relevant P_{in}. This leaves a dependence of the states Ψ on the Q_{in} , which can be used to calculate expectation values as follows. The expectation values of properties $\hat{A}(Q,Q',t)$, written here as operators on the electronic states and functions of the nuclear coordinates, are given by integrals

$$\langle A(t)\rangle = \int dQ \int dQ' \langle\Psi(Q,t)|\hat{A}(Q,Q',t)|\Psi(Q',t)\rangle \tag{13}$$

which can be transformed making a change of variables from the space coordinates Q to the initial values Q_{in} to obtain an initial value expression as integrals over the sets $\{P_{in}, Q_{in}\}$ and $\{P'_{in}, Q'_{in}\}$, written in terms of the Jacobians $J = \partial(Q)/\partial(Q_{in})$ of the transformations. This expression suggests a computational approach based on the discretization of the initial phase space of the classical variables, followed by integration over time of the coupled equations for the P, Q, and χ originating at each discrete set of values.

The advantage of using an eikonal description is that the electronic wavefunctions are needed only along nuclear trajectories. In the eikonal limit of short deBroglie wavelengths, the potential terms including wavefunction gradients can be ignored, and the equation for χ simplifies to

$$[\frac{\hbar}{i}\frac{P}{M}\frac{\partial}{\partial Q} + \hat{H}_Q - W(P,Q,t) - i\hbar(\frac{\partial}{\partial t})_{q,Q}]\chi(q,Q,t) = 0 \qquad (14)$$

with $W(P,Q,t) = V_{qu}(Q,P,t) + i\hbar(\partial P/\partial Q)/(2M)$. Letting P and Q change with time along trajectories, with $P = M\dot{Q}$ and $\dot{Q}\partial/\partial Q + (\partial/\partial t)_{q,Q} = (\partial/\partial t)_q$, and $W(P,Q,t)/\hbar = \omega(t)$, the replacement

$$\chi_I[q,Q(t),t] = \eta_I(q,t)exp[i\int_{t_{in}}^{t}\omega(t')dt'] \qquad (15)$$

leads to the usual time-dependent Schroedinger equation for the electronic wavefunction η_I,

$$[\hat{H}_{el}(t) - i\hbar(\frac{\partial}{\partial t})_q]\eta_I(q,t) = 0 \qquad (16)$$

where $\hat{H}_{el}(t) = \hat{H}_{Q(t)}$. The procedure in a first principles dynamics is then to simultaneously integrate Eqs.(12) and the above equation for all the electronic states I involved in the dynamics.

In a general case, the N-electron wavefunctions are combinations of Slater determinants $D_I(1,2,...,N,t) = (N!)^{-1/2}det[\psi_{i_m}(n,t)]$, constructed from N time-dependent molecular spin-orbitals (MSOs),$\{\psi_i(\vec{r},\zeta,t)\}$, and where n in the arguments is a short hand for all the coordinates (position and spin) of electron n, and $m = 1$ *to* N. In the simplest choice, useful in many applications, the electronic function is a single determinant and the variational procedure leads to the TDHF approximation. This is widely applicable since the MSOs can be chosen at the initial time so that electrons are localized at a given nucleus, unlike the situation when the MSOs are time-independent and

delocalized. Furthermore, the TDHF approximation includes some electron correlation. [44]

Instead of working with the collection $\{\eta_I\}$ of electronic wavefunctions, it is frequently more convenient to work with the density operator

$$\hat{\rho}_{el}(t) = \sum_I w_I |\eta_I(t)\rangle\langle\eta_I(t)| \tag{17}$$

which evolves in time as given by the Liouville-von Neumann (L-vN) equation,

$$\partial\hat{\rho}_{el}/\partial t = -(i/\hbar)[\hat{H}_{el}(t)\hat{\rho}_{el}(t) - \hat{\rho}_{el}(t)\hat{H}_{el}(t)] \tag{18}$$

This has the advantage of treating all relevant states I at once, and allows them to mix over time as determined by the L-vN equation. The density operator can be represented by a density matrix (DM) $\boldsymbol{\rho}(t)$ by expanding the states $\eta_I(q,t)$ in an electronic basis set $\{\Phi_K(q)\}$, in which case the matrix L-vN equation for the electronic DM must be solved coupled to the hamiltonian equations for the nuclear motions. Details of this procedure have been given in a previous publication.[32]

The standard (Born-Oppenheimer) treatment and the first principles treatment of quantum molecular dynamics have both advantages and disadvantages. The standard treatment, which separates from the outset electronic and nuclear motions, provides to begin with potential energy surfaces and their couplings for each molecular conformation. This allows for an efficient choice of atomic distances as needed to parametrize those functions. However, as the number of atoms grows the parametrization becomes complicated and it looses physical meaning. The following step is calculation of the dynamics, which can be done accurately with a quantum treatment based on a basis set expansion or numerical integration on a grid. A first principles treatment bypasses the need for potential energy surfaces and couplings, and generates them along trajectories. As a result there is no need for parametrizing complicated potential functions. But this comes at the expense of making many intermediate electronic structure calculations as time is advanced and atomic positions change. A compensating advantage is that couplings of electronic states by nuclear momenta or transition dipoles are generated only along relevant trajectories. But the calculation of properties or transition probabilities are done approximately, as sums over initial conditions, or sums over paths, or adding narrow wavepackets. Therefore the

choice of treatment is best done keeping in mind what properties are needed, and how accurately they must be known.

Working with the density operator is a convenient alternative to using wavefunctions when dealing with a few-atom, isolated molecular system, insofar it suggests more efficient computational procedures or more consistent approximations, but it is not stricktly needed. The density operator is however essential in treatments of a many-atom system, when this interacts with a medium which constrains thermodynamical properties such as temperature or pressure, because the density operator incorporates statistical averages which would not be included in a treatment based on wavefunctions.

3 Many-atom systems

3.1 Dissipative dynamics

A many-atom system may contain hundreds of atoms, as in clusters, or macroscopic amounts of matter, as in the cases of condensed matter solutions or solid surface phenomena. Mesoscopic systems and nanostructures fall in between those two extremes. These objects may be embedded in a medium in thermodynamical equilibrium, which imposes constrains of temperature, pressure, or chemical potentials. The medium may alternatively be excited and near equilibrium, or even far from it, in which cases it may strongly affect the time evolution of the object of interest. A unified treatment of these situations can be done with the density operator and its L-vN equation of motion.

Furthermore it is frequently found in molecular systems that excitation phenomena are localized. As examples, collisional excitation may involve a few atoms in a region of impact, or photo-excitation may involve a few atoms within weakly interacting chromophores. This leads to a localized dynamics which is however coupled to its surroundings so that fluctuation forces and dissipation of energy must be considered. The localized dynamics can be described by models where a primary (p-) region of interest is differentiated from a secondary (s-) region to which it couples. Mathematically, the two region can be defined by selecting p- and s- atomic degrees of freedom, to separate them into different hamiltonian terms and density operator factors. This naturally leads to a self-consistent field (SCF) description as a starting point. An alternative mathematical description introduces a basis set of

localized functions, and projection operators on p- and s-regions. This leads to an effective hamiltonian and density operator for the p-region. In either mathematical description, a new L-vN equation is derived, and contains non-hamiltonian terms describing dissipation and fluctuation phenomena of the p-region. These can not be treated in terms of a p-region hamiltonian, so that the L-vN equation with the new terms becomes the fundamental equation of motion. Dissipation and fluctuation can be studied with classical models when the system evolves in a single electronic state, but it must be treated within quantum mechanics when it is electronically excited. In what follows we present the two treatments of localized dissipative dynamics.

It is convenient to start with an isolated system which includes the object and medium, with hamiltonian $\hat{H}$. Its state is given by the density operator $\hat{\Gamma}(t)$ which satisfies the Liouville-von Neumann (L-vN) equation of motion

$$i\hbar\partial\hat{\Gamma}/\partial t = \hat{H}\hat{\Gamma}(t) - \hat{\Gamma}(t)\hat{H} = \hat{\mathcal{H}}\hat{\Gamma}(t) \tag{19}$$

using caligraphic symbols to indicate superoperators in the Liouville space of operators, with the notation $\hat{\mathcal{A}}\hat{B} = \hat{A}\hat{B} - \hat{B}\hat{A}$. This equation must be solved with the initial condition $\hat{\Gamma}(t_{in}) = \hat{\Gamma}_{in}$ and normalization $tr[\hat{\Gamma}(t)] = 1$.

Introducing a p-region containing the molecule M and adjacent bath atoms, and a s-region including the remaining bath, the hamiltonian terms are regrouped into the form

$$\hat{H} = \hat{H}_p(\vec{X}, \frac{\partial}{\partial \vec{X}}) + \hat{H}_s(\{\hat{B}\}) + \hat{H}_{ps}(\vec{X}, \frac{\partial}{\partial \vec{X}}, \{\hat{B}\}) \tag{20}$$

with $\vec{X}$ now indicating a set of degrees of freedom for the atoms in the p-region, and $\{\hat{B}\}$ a collection of boson-like operators (electron-hole pairs or phonons) for the s-region. Each of the hamiltonian terms is an operator on both electronic and nuclear variables. The density operator of the p-region is obtained from the trace of the full density operator over the s-region variables, $\hat{\Gamma}^p = tr_s(\hat{\Gamma})$. Its equation of motion follows from the trace of the previous equation of motion, which gives the contracted L-vN equation

$$\begin{aligned} i\hbar\partial\hat{\Gamma}^p/\partial t &= tr_s[\hat{\mathcal{H}}\hat{\Gamma}(t)] \\ &= \hat{\mathcal{H}}_p\hat{\Gamma}^p(t) + tr_s[\hat{\mathcal{H}}_{ps}\hat{\Gamma}(t)] \end{aligned} \tag{21}$$

A similar equation is obtained for the reduced density operator $\hat{\Gamma}^s = tr_p[\hat{\Gamma}]$ for the s-region. The right-hand side can not be written in general as a commutator between a hamiltonian and a density operator, so that in this case

a hamiltonian does not exist, and the density operator $\hat{\Gamma}^p$ must be obtained solving the new, contracted, L-vN equation instead of a Schroedinger-type equation.

An extended medium comprising the s-region can be described statistically with the probability distribution $P_s(\mathbf{b_0})$ of initial boson amplitudes. The stochastic assumption means that only averages $\overline{\hat{A}} = \int d\mathbf{b_0}\, \mathbf{P_s}(\mathbf{b_0})\hat{A}(\mathbf{b_0})$ of properties $\hat{A}$ over this distribution are needed. With this in mind, we look for statistical averages $\hat{\rho} = \overline{\hat{\Gamma}^p}$ of the p-region density operator, satisfying an averaged equation of motion,

$$\begin{aligned} \partial\hat{\rho}/\partial t &= -(i/\hbar)\overline{tr_s[\hat{\mathcal{H}}\hat{\Gamma}(t)]} = \hat{\mathcal{L}}_p\hat{\rho}(t) \\ \hat{\mathcal{L}}_p\hat{\rho}(t) &= -(i/\hbar)\hat{\mathcal{H}}_p\hat{\rho}(t) + \hat{\mathcal{L}}_p^{(D)}(t)\hat{\rho}(t) \\ \hat{\mathcal{L}}_p^{(D)}(t)\hat{\rho}(t) &= -(i/\hbar)\overline{tr_s[\hat{\mathcal{H}}_{ps}\hat{\Gamma}(t)]} \end{aligned} \tag{22}$$

which defines a dissipative Liouville superoperator $\hat{\mathcal{L}}_p^{(D)}(t)$, to be constructed in suitable approximations. For an operator $\hat{A}_p$ defined only in the p-region, its statistical average is $\overline{\langle A_p\rangle} = tr_p[\hat{\rho}(t)\hat{A}_p]$. In particular, probabilities of state survival and state-to-state transition probabilities follow from projection operators on the desired states.

3.2 Models of dissipative dynamics

3.2.1 Fast dissipation models

Models for the dissipative dynamics can frequently be based on the assumption of fast decay of memory effects, due to the presence of many degrees of freedom in the s-region. This is the usual Markoff assumption of instantaneous dissipation. Two such models give the Lindblad form of dissipative rates, and rates from dissipative potentials. The Lindblad-type expression was originally derived using semigroup properties of time-evolution operators in dissipative systems. [45, 46] It has been rederived in a variety of ways and implemented in applications. [47, 48] It is given in our notation by

$$\hat{\mathcal{L}}_p^{(D)}(t)\hat{\rho}(t) = \sum_L \{\hat{C}_p^{(L)}\hat{\rho}(t)\hat{C}_p^{(L)\dagger} - [\hat{C}_p^{(L)\dagger}\hat{C}_p^{(L)}, \hat{\rho}(t)]_+/2\} \tag{23}$$

where the $\hat{C}_p^{(L)}$ are operators in the p-region constructed from information about relaxation and decoherence times in the p-region. This form maintains complete positivity of the density operator $\hat{\rho}(t)$ and conserves its norm.

The operators $\hat{C}_p^{(L)}$ can be constructed as combinations of position and momentum operators in the p-region, [45] or from empirical transition rates $k_{\alpha' \leftarrow \alpha}$ between orthonormal eigenstates Φ_α^p and $\Phi_{\alpha'}^p$ of $\hat{H}_p$. [47] The index L then refers to a given transition $\alpha \rightarrow \alpha'$, and the corresponding operator is $\hat{C}_p^{(L)} = \sqrt{k_{\alpha' \leftarrow \alpha}} |\Phi_{\alpha'}^p\rangle\langle\Phi_\alpha^p|$. This simplifies the form of the Lindblad dissipative rate. The contracted L-vN equation with this dissipative rate expression can be transformed into a matrix equation introducing a basis set of quantum states, arranged as a row matrix $|\mathbf{\Phi}\rangle = [|\Phi\rangle_1, |\Phi\rangle_2, ...]$, taken here to be orthonormal, and giving the matrix representations $\hat{\rho} = |\mathbf{\Phi}\rangle\hat{\boldsymbol{\rho}}\langle\mathbf{\Phi}|$, where $\hat{\boldsymbol{\rho}}$ is yet an operator on the nuclear variables, and similarly for the amplitudes $\hat{\mathbf{C}}^{(L)}$. This equation is useful in semiempirical treatments where the transition rates are obtained from experiment or parametrized models.

Dissipative rates can also be obtained from a dissipative p-region hamiltonian derived from the full hamiltonian $\hat{H}$. We have done this in a self-consistent field approximation for coupled p- and s-regions, starting with a factorized density operator for a stochastic medium, $\hat{\Gamma} = \hat{\Gamma}^p \hat{\Gamma}^s = \hat{\rho}\, \hat{\Gamma}^s$.[33] Rewriting the L-vN equation to display fluctuation and dissipation terms, simplifying these after averaging over medium distributions, and taking its trace over s-region variables, the following contracted L-vN equation is found,

$$\begin{aligned} \partial\hat{\rho}/\partial t &= -(i/\hbar)\hat{\mathcal{F}}_p\hat{\rho}(t) - \hat{\mathcal{W}}_p(t)\hat{\rho}(t)/(2\hbar) \\ \hat{F}_p &= \hat{H}_p + tr_s[\hat{H}_{ps}\overline{\hat{\Gamma}^s}], \quad \hat{\mathcal{W}}_p(t) = \overline{tr_s[\hat{\mathcal{W}}(t)\hat{\Gamma}^s(t)]} \end{aligned} \tag{24}$$

where $\hat{\mathcal{W}}(t)$ is an instantaneous dissipation superoperator, bilinear in the coupling $\hat{H}_{ps}$.

We have implemented this approach in studies of photodesorption, deriving a dissipative hamiltonian $\hat{F}_p - i\hat{W}_p/2$, where the term $\hat{W}_p$ is a hermitian positive dissipative potential which gives relaxation rates $k_\alpha^p = \langle\Psi_\alpha^p|\hat{W}_p|\Psi_\alpha^p\rangle/\hbar$ of p-region decaying states Ψ_α^p. Defining a new superoperator

$$\hat{\mathcal{K}}_p\hat{\Gamma}^p = \sum_\alpha w_\alpha^p k_\alpha^p |\Psi_\alpha^p(t)\rangle\langle\Psi_\alpha^p(t)| \tag{25}$$

with statistical weights w_α^p for the decaying states, the dissipative Liouvillian term in the L-vN equation is given by

$$\hat{\mathcal{L}}_p^{(D)}\hat{\rho} = \hat{\mathcal{K}}_p\hat{\rho} - [\hat{W}_p, \hat{\rho}]_+/(2\hbar) \tag{26}$$

where the last term to the right is an anticommutator. The reduced L-vN equation with this dissipative term can also be transformed as explained above into a matrix equation with elements operating on nuclear coordinates.

3.2.2 The partitioning approach

To consider next the partitioning approach for many-atom systems, we start again with the full density operator $\hat{\Gamma}(t)$, describing both electrons and nuclei. We describe a procedure useful in a first principles dynamics where the nuclear motions are described in an eikonal approximation with $Q = Q(t)$, so that the electronic part of the problem involves the calculation of an electronic density operator $\hat{\Gamma}[Q(t), Q'(t), t] = \hat{\rho}_{el}(t)$. This satisfies an electronic L-vN equation

$$\partial\hat{\rho}_{el}/\partial t = -(i/\hbar)\hat{\mathcal{H}}_{el}(t)\hat{\rho}_{el} \tag{27}$$

where $\hat{H}_{el}$ is an electronic hamiltonian of the whole system. Introducing a basis set of electronic functions localized in the p- and s-regions, a partition of the density operator and the L-vN equation can be done using projection operators $\hat{O}_p$ and $\hat{O}_s$, with $\hat{O}_p + \hat{O}_s = \hat{I}$, the unit operator, and the orthonormality property $\hat{O}_p\hat{O}_s = 0$. We define the diagonal and non-diagonal parts of an operator $\hat{A}$ by means of

$$\begin{aligned} \hat{A}_D(t) &= \hat{O}_p\hat{A}\hat{O}_p + \hat{O}_s\hat{A}\hat{O}_s \\ \hat{A}_N(t) &= \hat{O}_p\hat{A}\hat{O}_s + \hat{O}_s\hat{A}\hat{O}_p \end{aligned} \tag{28}$$

so that

$$\hat{\rho}_{el}(t) = \hat{\rho}_D(t) + \hat{\rho}_N(t) \tag{29}$$

Next we derive a reduced differential equation for the diagonal part. The derivation is simplified using the relations

$$\begin{aligned} \left(\hat{A}\hat{B}\right)_D &= \hat{A}_D\hat{B}_D + \hat{A}_N\hat{B}_N \\ \left(\hat{A}\hat{B}\right)_N &= \hat{A}_D\hat{B}_N + \hat{A}_N\hat{B}_D \end{aligned} \tag{30}$$

for two operators $\hat{A}$ and $\hat{B}$. The equations take a simpler form when the projectors are time independent, and we treat this case in detail . Taking diagonal and non-diagonal parts of the L-vN equation for the density operator,

$$\begin{aligned} i\hbar\partial\hat{\rho}_D/\partial t &= \hat{\mathcal{H}}^{el}_D\hat{\rho}_D + \hat{\mathcal{H}}^{el}_N\hat{\rho}_N \\ i\hbar\partial\hat{\rho}_N/\partial t &= \hat{\mathcal{H}}^{el}_D\hat{\rho}_N + \hat{\mathcal{H}}^{el}_N\hat{\rho}_D \end{aligned} \tag{31}$$

Solving formally for $\hat{\rho}_N$ and replacing in the equation for $\hat{\rho}_D$, we have obtained the reduced L-vN equation [30]

$$\begin{aligned} i\hbar\partial\hat{\rho}_D/\partial t &= \hat{\mathcal{H}}_D^{el}\hat{\rho}_D + \int_{t_0}^{t} dt'\ \hat{\mathcal{M}}_D(t,t')\hat{\rho}_D(t') + \hat{w}_D(t) \\ \hat{\mathcal{M}}_D(t,t') &= (i\hbar)^{-1}\hat{\mathcal{H}}_N^{el}(t)\hat{\mathcal{U}}_D(t,t')\hat{\mathcal{H}}_N^{el}(t') \\ w_D(t) &= \hat{\mathcal{H}}_N^{el}(t)\hat{\mathcal{U}}_D(t,t_0)\hat{\rho}_N(t_0) \end{aligned} \tag{32}$$

where $\hat{\mathcal{U}}_D$ is a time evolution superoperator generated by $\hat{\mathcal{H}}_D^{el}$, $\hat{\mathcal{M}}_D$ is a delayed dissipative rate and $\hat{w}_D$ is a fluctuation energy. This is a quantal generalized Langevin equation for the reduced density operator which can be simplified when one averages over distributions of the stochastic medium. In this case the fluctuation rate vanishes and the dissipation is instantaneous, so that the reduced L-vN equation, written for the p-region density operator $\hat{\rho}_{pp}$ is

$$\begin{aligned} \partial\overline{\hat{\rho}}_{pp}/\partial t &= -(i/\hbar)\hat{\mathcal{H}}_{pp}^{el}\overline{\hat{\rho}}_{pp} + \hat{\mathcal{L}}_p^{(D)}(t)\overline{\hat{\rho}}_{pp}(t) \\ \hat{\mathcal{L}}_p^{(D)}(t)\overline{\hat{\rho}}_{pp}(t) &= (i/\hbar)\overline{\int_{t_0}^{t} dt'\ \hat{\mathcal{M}}_{pp}(t,t')\hat{\rho}_{pp}(t')} \end{aligned} \tag{33}$$

containing the dissipative rate superoperator $\hat{\mathcal{L}}_p^{(D)}(t)$. Details of its derivation from the hamiltonian $\hat{H}_{el}$ can be found in Ref. [30].

A simpler procedure has been implemented in applications to electron transfer in collisions of ions with metal surfaces.[35] Returning to Eq.(31), the second line can be interpreted as describing the rate of change of coherence of p- and s-regions over time. This is not likely to change much from its initial value for short collision times, so that the equation can be replaced on the average by a phenomenological rate equation,

$$\partial\overline{\hat{\rho}}_N/\partial t = -(\kappa + i\omega)\overline{\hat{\rho}}_N \tag{34}$$

over short times $t_0 \leq t \leq t_1$, with κ a rate of relaxation and ω a frequency of oscillation for decoherence. Replacing this in the first line of Eq(31) gives the inhomogeneous differential equation

$$i\hbar\partial\overline{\hat{\rho}}_D/\partial t = \hat{\mathcal{H}}_D^{el}\overline{\hat{\rho}}_D + \hat{\mathcal{H}}_N^{el}\overline{\hat{\rho}}_N(t_0)exp[-(\kappa + i\omega)(t - t_0)] \tag{35}$$

which can be transformed into a matrix equation in the p-region. This must be solved coupled to the equations for the trajectories. Numerical applications have been given in the cited papers.

4 Conclusion

The powerful mathematical tools of linear algebra and superoperators in Liouville space can be used to proceed from the identification of molecular phenomena, to modelling and calculation of physical properties to interpret or predict experimental results. The present overview of our work shows a possible approach to the dissipative dynamics of a many-atom system undergoing localized electronic transitions. The density operator and its Liouville-von Neumann equation play a central role in its mathematical treatments.

This overview has also considered the advantages and disadvantages of a description using first principles dynamics. Its application to many-atom systems undergoing electronic transitions is a very active and challenging subject of molecular quantum dynamics.

Acknowledgements

This contribution has been inspired by the work of Per-Olov Lowdin. Our work has been was partly supported by the National Science Foundation, whose funding is gratefully acknowledged.

References

[1] P. O. Lowdin. On operators, superoperators, hamiltonians and liouvillians. *Intern. J. Quantum Chem.*, QCS 16:485, 1982.

[2] P. O. Lowdin. Some aspects of the hamiltonian and liouvillian formalisms, the special propagator method, and the equation of motion approach. *Adv. Quantum Chem.*, 17:285, 1985.

[3] P. O. Lowdin. Partitioning technique, perturbation theory, and rational approximations. *Intern. J. Quantum Chem.*, 21:69, 1982.

[4] P. O. Lowdin. *Linear Algebra for Quantum Theory.* Wiley, New York, 1998.

[5] R. C. Tolman. *The Principles of Statistical Mechanics.* Clarendon Press, Oxford, England, 1938.

[6] J. von Neumann. *Mathematical Foundations of Quantum Mechanics.* Princeton Univ. Press, Princeton, New Jersey, 1955.

[7] U. Fano. Description of states in quantum mechanics by density matrix and operator techniques. *Rev. Modern Phys.*, 29:74, 1957.

[8] K. Blum. *Density Matrix Theory and Applications, 2nd. edition.* Plenum Press, New York, NY, 1996.

[9] J. M. Bowman and G. C. Schatz. Theoretical studies of polyatomic bimolecular reaction dynamics. *Ann. Rev. Phys. Chem.*, 46:169, 1995.

[10] R. Wyatt and J. Zhang. *Dynamics of Molecules and Chemical Reactions.* Dekker, New York, 1996.

[11] T. J. Martinez, M. Ben Nun, and R. D. Levine. Multielectronic molecular dynamics: A wavefunction approach with applications. *J. Phys. Chem.*, 100:7884, 1996.

[12] N. Makri. Time-dependent quantum methods for large systems. *Ann. Rev. Phys. Chem.*, 50:167, 1999.

[13] D. A. Micha. Time-dependent many-electron treatment of electronic energy and charge transfer in atomic collisions. *J. Phys. Chem.*, 103:7562, 1999.

[14] V. May and O. Kuhn. *Charge and Energy Transfer Dynamics in Molecular Systems.* Wiley-VCH, Berlin, 2000.

[15] M. D. Hack and D. G. Truhlar. Nonadiabatic trajectories at an exhibition. *J. Phys. Chem. A*, 104:7917, 2000.

[16] J. C. Tully. Chemical dynamics at metal surfaces. *Ann. Rev. Phys. Chem.*, 51:153, 2000.

[17] W. H. Miller. The semiclassical initial value representation: Adding quantum effects to classical molecular dynamics simulations. *J. Phys. Chem. A*, 105:2942, 2001.

[18] C. Y. Zhu, Y. Teranishi, and H. Nakamura. Non-adiabatic transitions due to curve crossings: complete solution. *Adv. Chem. Phys.*, 117:127, 2001.

[19] D A Micha. Long-lived states in atom-molecule collisions. *Accounts Cem. Res*, 6:138, 1973.

[20] D A Micha. Effective hamiltonian methods for molecular collisions. *Adv. Quantum Chem.*, 8:231, 1974.

[21] D. A. Micha. *Dynamics of Molecular Collisions, Vol. IA*, chapter Optical models in molecular collision theory, page 81. Plenum Press, New York, 1976.

[22] D A Micha. Quantum theory of reactive molecular collisions. *Adv. Chem. Phys.*, 30:7, 1975.

[23] D. A. Micha. *Potential Energy Surfaces and Dynamics Calculations*, chapter Overview of non-reactive scattering, page 685. Plenum Press, New York, 1981.

[24] D A Micha. Few-body processes in atom-diatom collisions. *Nuclear Phys.*, A353:309c, 1981.

[25] D. A. Micha. *Theory of Chemical Reaction Dynamics, Vol. II*, chapter Rearrangement in molecular collisions: A many-body approach, page 181. CRC Press, Boca Raton, Florida, 1985.

[26] D. Srivastava and D. A. Micha. Complex path integration for extended molecular systems. *Computer Phys. Comm.*, 63:331, 1991.

[27] D. A. Micha and E. F. Vilallonga. The collisional time-correlation function approach to molecular energy transfer. *Adv. Chem. Phys.*, 84:1, 1993.

[28] D. A. Micha and Keith Runge. Time-dependent many-electron approach to slow ion-atom collisions: The coupling of electronic and nuclear motions. *Phys. Rev. A*, 50:322, 1994.

[29] K. Runge and D. A. Micha. Time-dependent approach to slow ion-atom collisions for systems with one active electron. *Phys. Rev. A*, 53:1388, 1996.

[30] D. A. Micha. Temporal rearrangement of electronic densities in slow ion-atom collisions. *Intern. J. Quantum Chem.*, 51:499, 1994.

[31] D. A. Micha. Time-evolution of multiconfiguration density functions driven by nuclear motions. *Intern. J. Quantum Chem.*, 60:109, 1996.

[32] D. A. Micha. Density matrix treatment of electronic rearrangement. *Adv. Quantum Chem.*, 35:317, 1999.

[33] D. A. Micha. Density matrix theory and computational aspects of quantum dynamics in an active medium. *Intern. J. Quantum Chem.*, 80:394, 2000.

[34] D. A. Micha and B. Thorndyke. Dissipative dynamics in many-atom systems: a density matrix treatment. *Intern. J. Quantum Chem.*, vv:ppp, 2001. submitted.

[35] D. A. Micha and E. Q. Feng. The calculation of electron transfer probabilities in slow ion-metal surface collisions. *Computer Phys. Comm.*, 90:242, 1994.

[36] D. Beksic and D. A. Micha. Electronically diabatic quantum dynamics of molecular desorption. *J. Chem. Phys.*, 103:3795, 1995.

[37] Z. Yi, D. A. Micha, and J. Sund. Density matrix theory and calculations of nonlinear yields of CO photodesorbed from Cu(001) by light pulses. *J. Chem. Phys.*, 110:10562, 1999.

[38] D. A. Micha, A. Santana, and A. Salam. Nonlinear optical response and yield in the femtosecond photodesorption of CO from the Cu(001) surface: a density matrix treatment. *J. Chem. Phys.*, vvv:ppp, 2002. to appear.

[39] E. Deumens, A. Diz., R. Longo, and Y. Ohrn. Time-dependent theoretical treatments of the dynamics of electrons and nuclei in molecular systems. *Rev. Modern Phys.*, 96:917, 1992.

[40] D A Micha. A selfconsistent eikonal treatment of electronic transitions in molecular collisions. *J. Chem. Phys.*, 78:7138, 1983.

[41] C. D. Stodden and D. A. Micha. Generating wave functions from classical trajectory calculations: The divergence of streamlines. *Intern. J. Quantum Chem. Symposium*, 21:239, 1987.

[42] J. M. Cohen and D. A. Micha. Electronically diabatic atom-atom collisions: A self-consistent eikonal approximation. *J. Chem. Phys.*, 97:1038, 1992.

[43] J. Frenkel. *Wave Mechanics. Advanced General Theory.* Clarendon Press, Oxford, England, 1934.

[44] R. McWeeny. *Methods of Molecular Quantum Mechanics.* Academic Press, San Diego, CA, 2nd edition, 1989.

[45] G. Lindblad. On the generators of quantum dynamical semigroups. *Commun. math. Phys.*, 48:119, 1976.

[46] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of n-level systems. *J. Math. Phys.*, 17:821, 1976.

[47] P. Saalfrank and R. Kosloff. Quantum dynamics of bond breaking in a dissipative environment: Indirect and direct photodesorption of neutrals from metals. *J. Chem. Phys.*, 105:2441, 1996.

[48] D. A. Lidar, Z. Bihary, and K. B. Whaley. From completely positive maps to the quantum markovian semigroup master equation. *Chem. Phys.*, 268:35, 2001.

New Developments in Monte Carlo/Quantum Mechanics Methodology. The Solvatochromism of β-Carotene in Different Solvents

by Sylvio Canuto[1], Kaline Coutinho[2] and Daniel Trzesniak[1]

[1]Instituto de Física, Universidade de São Paulo, CP 66318
05315-970 São Paulo, SP, Brazil

[2]Universidade de Mogi das Cruzes/CCET, CP 411
08701-970, Mogi das Cruzes, SP, Brazil

Abstract

The solvatochromic shifts of the $\pi - \pi^*$ transition of all-*trans*-β-carotene in isopentane, acetone, methanol and acetonitrile are studied using a sequential Monte Carlo/quantum mechanics (S-MC/QM) methodology. These different solvents are examples of systems of varied nature, differing in dielectric constants and covering a wide range of polarities, and including also polar and non-polar solvents. In S-MC/QM we first generate the structure of the liquid using Metropolis MC simulation and then perform the QM calculations in statistically uncorrelated configurations. It is shown that, in these cases, including only 40 QM calculations gives statistically converged results. To deal with elongated solutes the box of the MC simulation has been extended to a large rectangular shape. Then, a nearest-neighbor distribution function has been developed and generalizes the concept of solvation shells for a solute of any arbitrary shape. The calculated results are converged with respect to the number of solvent molecules that are included according to the nearest-neighbor distribution function. The results are found to be in very good quantitative and qualitative agreement with experiment. The dipole moments of the ground and excited $\pi - \pi^*$ states of β-carotene are both zero and the transition shifts are thus dominated by the dispersive interaction. The inclusion of dispersion interaction in energy differences is then discussed.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

Contents

1 Introduction

The study of molecular systems in the liquid phase is important for understanding a great number of chemical, physical and biological processes[1]. The solvent interaction leads to changes in the molecular solute affecting its spectroscopic, structural and reactive properties. For this reason, the study of solvent effects has been a topic of increased interest[2, 3, 4]. In the theoretical front the basic ideas developed by Onsager[5] and Kirkwood[6] have led to sophisticated cavity theories, where the solute is enclosed in a cavity and the solvent is treated by a continuum polarizable dielectric medium. Tapia and Goscinski[7] have developed one of the first successful self-consistent reaction field (SCRF) theories that has been extended further by many others[8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. Present continuum models include sophisticated procedures, where the solute is treated with electron correlation effects[15, 16] leading to more accurate reaction fields, and variants such as the COSMO[12] methodology. Warshel and Levitt[18] have suggested a hybrid quantum mechanical-molecular mechanics (QM/MM) methodology, where the most important part of the system is treated by quantum mechanics and the rest by classical mechanics. Thus in solute-solvent interaction the chromophore, and perhaps a few other molecules, are treated by QM and the solvent is considered by classical point charges[19, 20, 21, 22, 23, 24, 25]. This idea was further developed by Blair and co-workers[25], by Gao[20] and by Zeng and co-workers[26] that considered that a liquid has not one but many structures at a certain temperature. A liquid is, indeed, statistical by nature and the liquid properties are, in fact, statistical averages. Thus they performed Molecular Dynamics and Monte Carlo simulations to generate the structure of the liquid. Gao has further developed this idea generating a successful Monte Carlo QM/MM method[20, 21, 22].

In Monte Carlo simulation of liquids the configurational space necessary for configurational averages is generated by Metropolis sampling technique and includes temperature effects. Although this is a more realistic representation of the liquid nature of the solvent, it has the concomitant disadvantage that several quantum mechanical calculations are necessary to obtain the proper statistical average. For instance, in studying solvatochromic shifts the transition energy has to be calculated several times for structures generated by the simulation, in order to obtain the average value that corresponds to the solvation shift. In many cases millions of calculations have been performed on these supermolecular systems composed of the solute treated by QM and the solvent as classical point charges. Furthermore, if the solvent is not explicitly treated by QM it is difficult to include dispersion interaction that, in fact arises from the reciprocal polarizations of the solute by the solvent, and the solvent by the solute.

We have extended this idea to a sequential Monte Carlo/Quantum mechanics (S-MC/QM)[27]. In this procedure we first generate structures of the liquid and only subsequently perform the QM calculations in those structures. The basic advantage is that opposite to conventional QM/MM, in the S-MC/QM, the solute and *all solvent molecules*, up to a certain solvation shell, are treated by quantum mechanics[28, 29]. The number of necessary solvation shells to be included can be systematically analyzed and converged results obtained[28]. As an important development, we have also shown that the drawback of having to perform a large number of quantum mechanical calculations to obtain the average of the property of interest may be strongly alleviated considering the statistical correlation between successive configurations[27, 29, 30]. As MC generates structures that belong to a markovian chain, the auto-correlation function of the energy gives important information on the relative statistical importance of the successive structures generated by the simulation. Of course, highly correlated structures will give very little new statistical information[30]. In other words, performing calculations on every structure generated is an enormous waste that gives no new statistical information. Reducing the number of quantum mechanical calculations is a great saving in computational resources that can be used to explicitly include the solvent molecules and thus leading to a more realistic treatment of the intermolecular solute-solvent interactions. As an example, the solvatochromic shift of pyrimidine in water[31] was successfully treated including all water molecules up to the third solvation shell. This required supermolecular quantum-mechanical calculations of the pyrimidine and 213 water molecules, an explicit 1734 all-valence electrons, properly anti-symmetrized. With this procedure specific interactions such as charge transfer and hydrogen bonds are naturally treated. Detailed analysis of the convergence of the average value with the number of configurations are made

elsewhere[30, 32], and shows that the average value is indeed converged.

Inclusion of dispersive interaction in solvent effects[33] has been a real challenge for present theoretical methodologies[2, 3]. If the solvent molecules are not explicitly included the polarization of the solute onto the solvent is not considered and dispersion is omitted. As dispersion is a double excitation, derived from single excitation in the solute and single excitation in the solvent, one possibility is to have previously calculated and separated the spectrum of the solvent molecule and try including this information in the calculation of the solvatochromic shift. This was attempted by Rösch and Zerner[34]. Another possibility is, of course, to explicitly include the solvent molecules in the supermolecular QM calculations. Our S-MC/QM procedure allows this to be done in a natural way[27, 28, 29]. In this connection it is very important to note that it is possible to include dispersive interaction in transition energy, using a singly-excited configuration interaction (CIS). It has been demonstrated[35] before that a configuration interaction electronic structure calculation on a supermolecule that contains only single excitations includes dispersion interactions between the two subsystems when energy differences are taken between the Hartree Fock (SCF) ground state and low energy excited states in which single excitations dominate. This theorem is proven up to second order in perturbation theory[35]. This has been used to calculate the solvatochromic shifts of benzene in different solvents (polar and non-polar)[29] with very good agreement with the experimental results. As the dipole moment of benzene is zero in the ground state and in the low-lying $\pi - \pi^*$ excited states the solvatochromic shifts in different solvents is basically given by dispersion (quadrupolar interaction is very small) and leads to a red shift, as described earlier by Liptay[36].

In short, our S-MC/QM methodology uses structures generated by MC simulation to perform QM supermolecular calculations of the solute and all the solvent molecules up to a certain solvation shell. As the wave-function is properly anti-symmetrized over the entire system, CIS calculations include the dispersive interaction[35]. The solvation shells are obtained from the MC simulation using the radial distribution function. This has been used to treat solvatochromic shifts of several systems, such as benzene in CCl_4, cyclohexane, water and liquid benzene[29, 37]; formaldehyde in water[28, 38]; pyrimidine in water and in CCl_4[31]; acetone in water[39]; methyl-acetamide in water[40] etc.

In this paper we address to the solvation of all-*trans*-β-carotene in different solvents. The solvent effects on the visible spectrum of β-carotene is a real challenge for theoretical methodologies for at least two aspects. First, the visible spectrum is characterized by a strong $\pi - \pi^*$ absorption transition in the region of 450 nm that suffers only slight shifts in different solvents[41, 42]. As the dipole moment is zero both in the ground and excited state, the shift is

dominated by dispersion interaction. This interaction is, in addition, small for different solvents. For instance, the shift of the $\pi - \pi^*$ absorption transition from isopentane (non-polar, non-protic, has small polarity and small dielectric constant) to methanol (polar, protic, has large polarity and large dielectric constant) is only 120 cm^{-1}[42]. Because of the low volatility of β-carotene the gas phase value of the absorption transition is not known experimentally and correlation of solvatochromic shifts in different solvents is very important. Second, the elongated shape of the molecule (see below) suggests the use of a non-spherical distribution of solvent molecules around the solute. This is parallel to the cavity-shape problem in SCRF methods[43]. For our purposes, the use of spherically defined solvation shells is not recommended. This is a delicate point as the concept of solvation shells is in essence, but not compulsory, related to a spherical distribution[44]. We will see that this is not only inconvenient but, to some extent, incorrect. Thus, we develop a nearest-neighbor distribution that follows the molecular shape and can be used for any molecule, no matter how elongated or distorted. The visible spectrum of β-carotene has been analyzed before by Applequist[45] using a cavity model where the chromophore was treated as classical point dipole oscillator. Myers and Birge[41] studied the change in oscillator strength of the absorption of β-carotene in different solvents and found that the results depend on the prolate cavity geometry. Zerner made an estimate of the shift in cyclohexane[46] using SCRF. Abe and co-workers[42] analyzed solvent effects in 51 different solvents and made an empirical analysis in terms of reaction field models. Here, we use our S-MC/QM methodology in a nearest-neighbor solvation shell to calculate the solvatochromic shift of β-carotene in four different solvents; namely, isopentane, acetonitrile, acetone and methanol. These four solvents are selected on the basis of their nature, exemplifying polar, non-polar, protic, non-protic, low polarity and large polarity solvents.

2 Monte Carlo Simulation

2.1 Rectangular Box and Computational Details

Monte Carlo (MC) simulations were carried out for all-*trans*-β-carotene in four solvents: acetone ($(CH_3)_2CO$), acetonitrile (CH_3CN), isopentane ($(CH_3)_2CHCH_2CH_3$) and methanol (CH_3OH). Standard procedures[47] were used, including the Metropolis sampling technique[48] in the canonical (NVT) ensemble and periodic boundary conditions using the image method. Because of the prolate shape of β-carotene we used a rectangular box instead of the more conventional cubic box. Therefore, a cuttoff radius was not used, but each molecule was restricted to interact either with a molecule or its respective

image, not simultaneously with both. Therefore, this system (1 solute + N solvent molecules) corresponds to an infinitely dilute solution. In figure 1, the solute (all-*trans*-β-carotene) in the smallest rectangular box used in our sim-

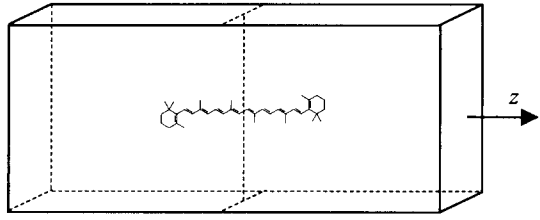


Figure 1: Illustration, in scale, of the all-*trans*-β-carotene in the smallest rectangular box (30x30x70.6)$\AA^3$ used in our simulations.

ulations is illustrated. Note that the β-carotene is almost a planar molecule with approximately 29 $\AA$ in the long axis and 6 $\AA$ in the small axis. Thus, even in the smallest box there was sufficient space to wrap the β-carotene in a bulk environment.

Table 1
Information of the simulated systems: the density, the box size, the dielectric constant and the polarity of the solvent.

Solvent	Density g/cm^3	Box Size (x, y, z) in $\AA$	Dielectric Constant (ϵ)	Normalized Polarity (E_T^N)
Isopentane	0.6001	(45.5, 45.5, 87.5)	1.828	0.006
Acetone	0.7682	(38.5, 38.5, 77.0)	21.36	0.355
Methanol	0.7676	(30.0, 30.0, 70.6)	32.66	0.762
Acetonitrile	0.7649	(33.5, 33.5, 72.5)	35.94	0.460

The four systems were simulated at T = 298K and were consisted of one β-carotene molecule and 900 solvent molecules in a rectangular box with linear dimensions, which correspond to the solvent densities[49]. The solvent density, the box size, the dielectric constant and the normalized E_T^N Reichardt polarity[1] are shown in table 1. Note the variation of the dielectric constants of the selected solvents. These solvents also exhibit large variations of the normalized solvent polarity, changing from 0.006 (isopentane) to 0.762 (methanol) with the intermediate values of 0.355 (acetone) and 0.460 (acetonitrile)[42].

The intermolecular interactions were described by the Lennard-Jones plus

Coulomb potential,

$$U_{ab} = \sum_{i}^{\text{on } a} \sum_{j}^{\text{on } b} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{e^2}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \tag{1}$$

where $\sum^a$ is the sum over the sites of molecule a, $\sum^b$ is the sum over the sites of molecule b, $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$, $\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$, $e^2/(4\pi\epsilon_o) = 331.9684$ Å kcal/mol and σ_i, ε_i and q_i are the parameters of the interacting sites. The potential parameters of the sites used in the our simulations were obtained in the OPLS force field[50] and are shown in table 2. The geometry of the β-carotene, shown in

Table 2
Potential parameters used in the Monte Carlo simulations (q_i in elementary charge unit, ε_i in kcal/mol and σ_i in Å).

Site	q_i	ε_i	σ_i
Isopentane			
CH_3	0.000	0.160	3.910
CH_2	0.000	0.118	3.905
CH	0.000	0.080	3.850
Acetone			
O	−0.424	0.210	2.960
C	0.300	0.105	3.750
CH_3	0.062	0.160	3.910
Methanol			
H	0.435	0.000	0.000
O	−0.700	0.170	3.070
CH_3	0.265	0.207	3.775
Acetonitrile			
N	−0.430	0.170	3.200
C	0.280	0.150	3.650
CH_3	0.150	0.207	3.775
β-Caroteno			
$C(sp^2)$	0.000	0.105	3.750
$C(sp^3)$	0.000	0.050	3.800
CH_3	0.000	0.175	3.905
CH_2	0.000	0.118	3.905
CH	0.000	0.115	3.800

figure 2, was obtained by gradient optimization, starting from the crystallographic experimental data[51], with the Becke three-parameter-functional[52]

and the Lee-Yang-Parr correlation[53], B3LYP/6-31G level of calculation. The

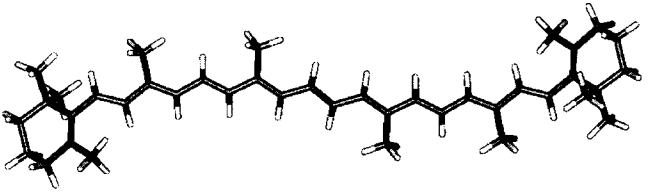


Figure 2: Geometry of the β-carotene obtained with B3LYP/6-31G optimization.

geometries of the solvents were obtained in the OPLS force field (acetone[54], acetonitrile[55], isopentane[56] and methanol[57]). All molecules were kept in the equilibrium geometry during the simulation. The initial configurations were generated randomly, considering the position and orientation of each molecule. A new MC step was generated by randomly selecting a solvent molecule, translating it randomly in all Cartesian directions and rotating it randomly about a randomly chosen axis. A new configuration was generated after 900 MC steps, i. e., after closing a loop over the solvent molecules. In this way, the number of configurations l generated here is equivalent to the number of configurations generated in a Molecular Dynamics simulation with an integration over l time steps. The acceptance of each random move was governed by the Metropolis sampling technique. The maximum displacement of the molecules was self adjusted after 50 MC steps to give an acceptance rate around 50%. The full simulation consisted of a thermalization stage of $4.5 \text{x} 10^6$ MC steps, which is not used in the statistics, followed by an averaging stage of $36 \text{x} 10^6$ MC steps. This is a long simulation by all present standards. Thus, the total number of configurations generated in each MC simulation was $l = 40000$. Instead of performing a quantum mechanical calculation on every configuration generated by the Monte Carlo simulation, we use the auto-correlation or statistical efficiency, to select the statistically relevant structures[27, 29, 30, 32]. In doing so, the subsequent quantum mechanical calculations are performed only on some uncorrelated structures. As in previous works[28, 29, 31, 38, 40] we fit the auto-correlation function of the energy to an exponentially decaying function and obtain the correlation step. This assures that the structures used in the quantum mechanical calculations are statistically (nearly) uncorrelated. As the total number of MC configurations generated in the simulation was 40000, the averages are then taken over only 40 configurations, separated by 1000 successive configurations. The convergence of the calculated values using this reduced number of uncorrelated configurations is discussed later in this paper. All simulation were performed with the DICE[58] Monte Carlo

statistical mechanics program.

To obtain the relative solvatochromic shifts, the excitation energies were calculated using the ZINDO program[59], within the INDO/CIS[60] approach. The quantum mechanical calculations were performed for the supermolecular clusters, generated by the MC simulations, composed of one β-carotene and all solvent molecules within a particular nearest-neighbor solvation shell. As the appropriate Boltzmann weights are included in the Metropolis Monte Carlo sampling technique, the average value of the solvatochromic shift is obtained from a simple average over a chain E_i of size L of uncorrelated configurations, where $L = 40$ for all the systems considered here and E_i corresponds to the excitation energy obtained for the supermolecular configuration i.

2.2 Nearest-Neighbor Solvation Shells

The molecular structure of liquids are best analyzed using the concept of the radial distribution function (RDF). This is of particular importance in solute-solvent structures as it defines the solvation shells around the solute molecule[44, 47]. The RDFs represent fluctuations in the local density due to structure in the liquid. Specifically, the average density of atoms of type Y around atoms of type X is $\rho_{X-Y}(r) = \rho_Y G_{X-Y}(r)$, where $\rho_Y = (N_Y/V)$ is the density of type Y atoms, r is the $X - Y$ separation and $G_{X-Y}(r)$ is the RDF between atoms of type X and Y. For a solute-solvent system, the first atom of the RDF (type X) belongs to the solute molecule and the second atom (type Y) belongs to the solvent molecule. In the simulation, the $G_{X-Y}(r)$ is obtained by accumulating and normalizing histograms with the total number of atom pairs $X - Y$ found in a distance between $r - \delta r/2$ and $r + \delta r/2$,

$$G_{X-Y}(r) = \frac{HISTOGRAM[r - \frac{\delta r}{2}, r + \frac{\delta r}{2}]}{ln_X n_Y N_X (\frac{4\pi\rho_Y}{3})[(r + \frac{\delta r}{2})^3 - (r - \frac{\delta r}{2})^3]} \tag{2}$$

where l is the number of MC configurations analyzed during the simulation, n_X is the number of type X atoms in the solute, n_Y is the number of type Y atoms in the solvent, N_X is the number of solute molecules, N_Y is the number of solvent molecules and δr is the width of each bin of the histogram. If a liquid is structureless, then $G_{X-Y}(r) = 1$.

For a system consisting of one β-carotene in solution the type X of the RDF could be defined as both carbon (C) or hydrogen (H) atoms. Although RDFs can be determined from diffraction experiments on liquid, such data are not currently available for β-carotene in solution. However, the calculated RDFs were used here to describe the distribution of the solvent molecules around the β-carotene and define its respective solvation shells. This structural analysis will be of great importance in the calculation of the absorption

spectrum of β-carotene in solution. Quantum mechanical calculation of the absorption spectrum of a supermolecular system, composed of 901 molecules, over 40000 MC configurations is of course not possible. The alternative we suggested[27, 29] was to perform quantum mechanical calculations after the simulation, but using only a few selected number of solvent molecules[31, 38] and a selected number of MC configurations[28, 29, 30, 31, 32]. The number of solvent molecules included in the calculation is obtained from the analysis of RDFs, using all molecules surrounding the solute up to a certain solvation shell. This number is still large enough to preclude sophisticated *ab initio* calculations but lies well within the range of semiempirical methods. The number l of necessary MC configurations for ensemble average has already been reduced dramatically from 40000 to only 40, as discussed in the previous section. We shall show that this small number gives indeed converged result, and is a consequence of the markovian chain generated by MC simulations, as documented before[28, 29, 30, 32]. In figure 3, the two RDFs between C and H of β-carotene and the central C atom of acetone are shown as an example of the liquid structure obtained in the simulations. Irrespective of the solvent

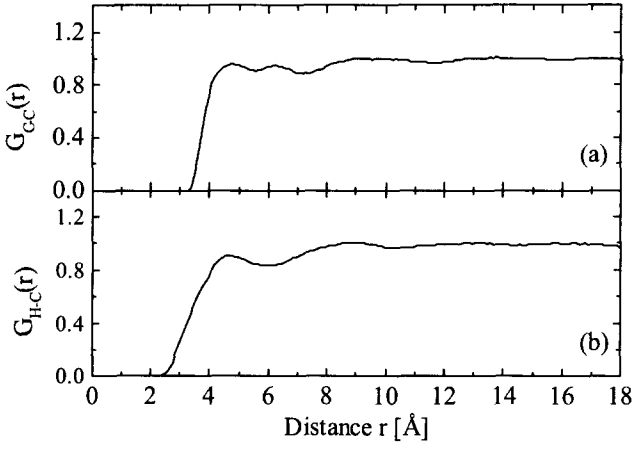


Figure 3: The calculated radial distribution function (RDF) between carbon atoms (a) and hydrogen atoms (b) of the β-carotene and carbon atoms of the acetone molecules, $G_{C-C}(r)$ and $G_{H-C}(r)$, respectively.

and its selected atom, all RDFs of β-carotene in solution, studied here, has the same shape with broad and low peaks (see figure 3). This just reflects the elongated geometry of the β-carotene and the wide spatial distribution of the carbon and hydrogen atoms. Therefore, these two RDF ($G_{C-Y}(r)$ and G_{H-Y})

can not help in the description of the distribution of solvation shells around the β-carotene.

The RDF between the center-of-mass of the solute and the solvent molecules, $G_{CM-CM}(r)$ is another natural possibility of describing the solvation shells around the β-carotene. In figure 4a, the RDF between the center-of-mass of β-carotene and acetone molecules is shown as an example of the liquid structure. This $G_{CM-CM}(r)$ presents a clear definition of four peaks that characterize the solvation shells around the center-of-mass of the β-carotene. The number of solvent molecules in each shell was obtained by integrating the peaks. In the case presented in figure 4a, 7 acetone molecules were found in the first shell (integrating until 6.35 $\AA$), 30 in the second shell (from 6.35 to 10.65 $\AA$), 46 in the third shell (from 10.65 to 13.85 $\AA$) and finally 108 in the fourth shell (from 13.85 to 18$\AA$). Figure 4b will be discussed soon below.

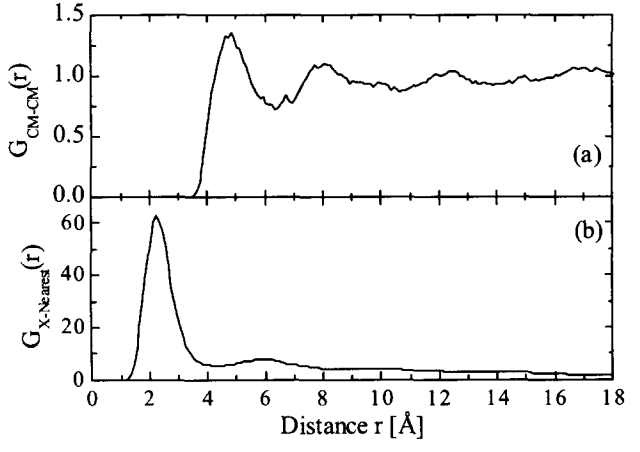


Figure 4: The calculated radial distribution function (RDF) between (a) the β-carotene center-of-mass and acetone center-of-mass, $G_{CM-CM}(r)$, and (b) all atoms of the β-carotene and its nearest atom of each acetone molecule, $G_{X-Nearest}(r)$.

Figure 5 illustrates typical configurations, generated in the simulation, of one β-carotene surrounded by the 7 and 37 acetone molecules corresponding to the first and second peaks, respectively, defined by $G_{CM-CM}(r)$. Although the center-of-mass RDF presented peaks that *a priori* could be considered as solvation shells, certainly the figure 5 shows that, in the case of β-carotene as the solute, these peaks can not be considered as solvation shells around the solute. As expected, the solvent molecules were distributed only in the central part of the β-carotene, close to the center-of-mass. Of course, even considering

the second or third solvation shells the $G_{CM-CM}(r)$ still gives a rather non-uniform distribution of solvent molecules around this elongated solute.

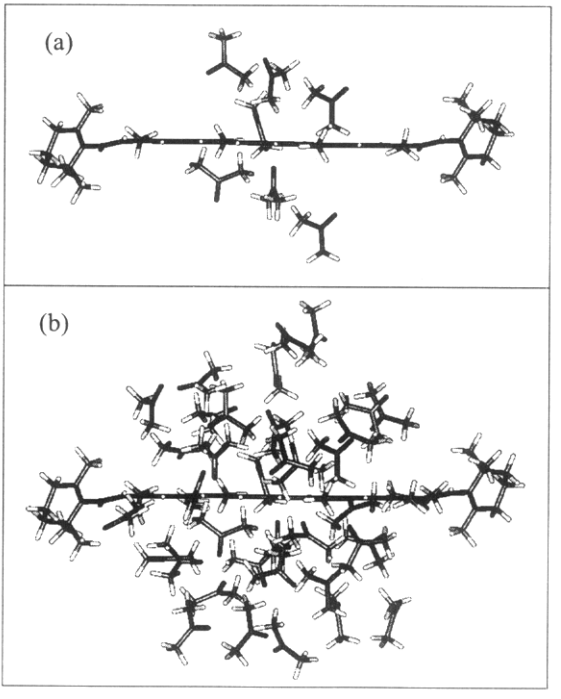


Figure 5: This illustration shows the β-carotene surrounded by (a) 7 acetone molecules and (b) 37 acetone molecules corresponding to the first and second peaks, respectively, defined by the center-of-mass RDF, $G_{CM-CM}(r)$.

To analyze the solvation shells of elongated molecules in solution, it is necessary to define a different kind of RDF that does not grow in a spherical form, but consider the shape of the solute. Then, we suggest here a nearest-neighbor RDF between *all atoms* of the elongated solute and its nearest atom of each one of the N_Y solvent molecules, $G_{X-Nearest}(r)$. The $G_{X-Nearest}(r)$ was calculated using the same definition of equation 2, but changing the assignment of types X and Y. Now, all atoms of the solute molecule are assigned as an unique type X and after testing all atoms of the solvent molecule, the nearest atom of each solvent molecule is assigned as an unique type Y. Thus, a list of nearest-neighbors was built taking into account not a fixed atom or the center-of-mass of the solute, but all atoms in the solute molecule. With this new list of neigh-

bors the shape of the elongate solute is taken into account in the distribution of solvent molecules. However for small solute, like, for instance, formaldehyde, this new list of neighbors generates a distribution of solvent molecules that is similar to that described by the center-of-mass distances. Thus, the nearest-neighbor distribution function generalizes the concept of solvation shells for a solute of any arbitrary shape. In figure 4b, the nearest-neighbor RDF of β-carotene in acetone is shown as an example of the liquid structure defined by $G_{X-Nearest}(r)$, that presents a clear definition of two solvation shells around the whole β-carotene. The number of solvent molecules in each solvation shell was obtained by integrating the peaks. In the case presented in figure 4b, 50 acetone molecules were found in the first solvation shell (integrating until 4.35 $\AA$) and 88 in the second shell (from 4.35 to 8.05 $\AA$). Figure 6 illustrates a typical configuration, generated in the simulation, of one β-carotene surrounded by 50 acetone molecules corresponding to the first solvation shell as defined by $G_{X-Nearest}(r)$. As expected, the solvent molecules are uniformly distributed

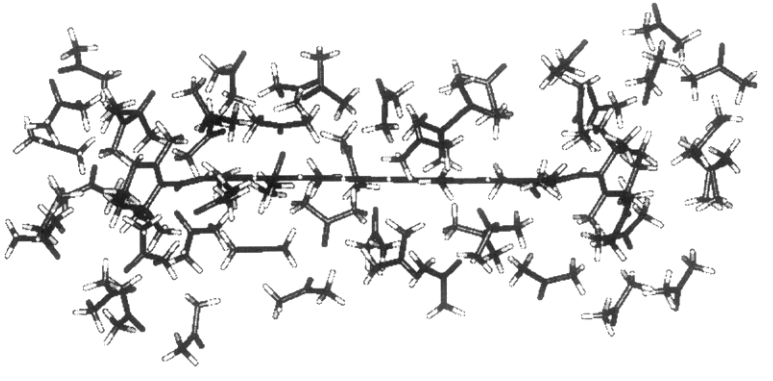


Figure 6: Typical supermolecule used in the QM calculations. This illustration shows the β-carotene surrounded by 50 first-neighbor acetone molecules corresponding to the first solvation shell as defined by the nearest-neighbor RDF, $G_{X-Nearest}(r)$.

around the β-carotene. For the other solvents, the minimum distance RDF presents the same shape as shown in figure 4b, a first well defined peak, a second less intense peak and a long structureless tail.

In table 3, a summary of the structural analysis is shown for all four solvents. Using the $G_{CM-CM}(r)$, the first neighborhood around the center-of-mass of the β-carotene is formed by either 5 isopentane or 7 acetone or 8 methanol or 9 acetonitrile. These solvent molecules are distributed approximately between 3.55 and 6.35 $\AA$ with maximum occurrency in the interval between 4.35 $\AA$ (methanol) and 5.05 $\AA$ (isopentane). However, the first solva-

Table 3
Structural information obtained from the first peak of the radial distribution functions. Distances are given in Å and N_s is the coordination number obtained from the integration of the first peak.

	$G_{CM-CM}(r)$				$G_{X-Nearest}(r)$			
Solvent	Start	Max.	End	N_s	Start	Max.	End	N_s
Isopentane	3.85	5.05	6.35	5	1.35	2.15	4.55	40
Acetone	3.55	4.65	6.35	7	1.35	2.25	4.35	50
Methanol	3.35	4.35	5.85	8	1.35	2.35	4.65	69
Acetonitrile	3.45	4.95	6.35	9	1.35	2.45	4.35	58

tion shell around the entire β-carotene using the nearest-neighbor distribution is formed by 40 isopentane or 50 acetone or 69 methanol or 58 acetonitrile. These solvent molecules have their nearest atom distributed approximately between 1.35 and 4.45 Å with a large maximum occurrence in $\sim$ 2.30 Å.

3 Quantum Mechanical Results

3.1 Solvatochromic Shifts

The quantum chemistry calculations are of the SCF type followed by configuration interaction over singly excited configuration state functions (CIS). This is the level of theory at which the INDO/S Hamiltonian was parametrized [60]. The low energy spectrum of β-carotene is dominated by excitations from the HOMO (π) molecular orbital to the LUMO (π^*) molecular orbital, where HOMO refers to the highest occupied molecular orbital and LUMO to the lowest unoccupied molecular orbital (illustrated in figure 7). As it can be seen the dipole moment is zero in both the ground and the excited state and the π orbitals are delocalized over the polyene chain. The gas phase absorption transition is not known experimentally. The value calculated here for the gas phase $\pi - \pi^*$ transition is 22230 cm^{-1}. In solvents of any polarity it is expected that this transition suffers a red shift. The magnitude of the shift, of course, depends on the solvent. To obtain these red shifts we have next performed the supermolecular calculations of β-carotene in acetone, acetonitrile, methanol and isopentane using the configurations generated by the MC simulation. As β-carotene is non-polar, it is expected that the solvation shift should not depend on solvent molecules that are situated much beyond the first shell. Table 4, thus gives the calculated $\pi - \pi^*$ transition energies of β-carotene in the four solvents considered here. As it can be seen, in all solvents this transition suffers a red shift compared with the gas phase value. These results are in

very good agreement with the experimental data. The relative positions are also correct for isopentane-methanol, for instance, but has a wrong sign (still within the statistical error) for the acetonitrile-acetone, if we use only the first shell of solvent molecules. Whereas the experimental result gives that the shift is larger for acetone over acetonitrile by 24 cm^{-1} the theoretical result as this stage gives the opposite by 12 cm^{-1}. Including half of the second shell changes the calculated value of acetone by only 3 cm^{-1}, showing that indeed this value is converged for the first shell. However, the other change for acetonitrile is slightly larger, by 36 cm^{-1}. As a result, including this half-shell corrects the

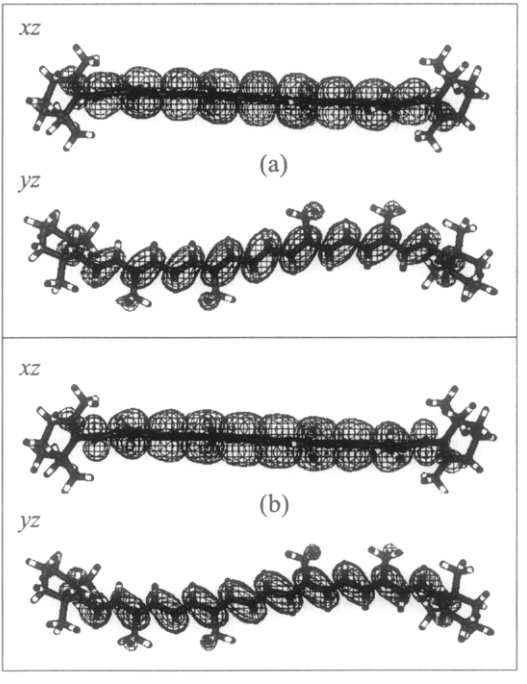


Figure 7: Illustration of the orbitals (a) HOMO and (b) LUMO involved in the first $\pi - \pi^*$ transition of β-carotene.

relative position of these two transitions. Note that the experimental shift of 24 cm^{-1} is now theoretically obtained as 21 cm^{-1}, in excellent agreement. The results obtained for acetone, for instance, is a result of 40 QM INDO/CIS calculations of one β-carotene surrounded by 77 acetone molecules. Each QM calculation is thus a 2064-valence-electron problem. The largest calculation is for β-carotene in isopentane, that involves 2104 valence electrons. The quali-

tative relative shifts of the $\pi - \pi^*$ transitions are well reproduced in agreement with the experimental results. Note that the solvents are of miscellaneous type, involving both protic and non-protic, polar and non-polar and also of low and high polarity. In this direction, it should be noted that these shifts do not follow the increase in dielectric constants or, even, in polarity, as it can be checked from the results given in tables 1 and 4. Thus it is not clear that a description based on macroscopic parameters can be obtained. Abe and co-workers[42], have found an approximate correlation but only after excluding protic solvents. Similarly, the correlations are different for non-polar and polar solvents. In this paper, we explicitly calculate these values using a methodology that combines statistical mechanics and quantum mechanics. The consideration of dispersive interaction is of great importance. Analyzing the calculated results given in table 4, one may conclude that our approach is very successful in describing the relative shifts of this very challenging system. Before concluding, it is rather appropriate to discuss the convergence of the calculated result with respect to the statistics; i.e. with respect to the number of structures used in the QM calculations.

Table 4
Summary of the calculated and experimental results for the first $\pi - \pi^*$ absorption transitions of the β-carotene in gas phase and in solution.

Solvent	1^{st} Shell		1^{st}+half Shell		Experiment[42]
	N	*Transition*	N	*Transition*	
Vacuum	1	22230			
Acetone	$1+50$	22071 ± 17	$1+77$	22074 ± 17	22046
Acetonitrile	$1+58$	22059 ± 19	$1+92$	22095 ± 11	22070
Methanol	$1+69$	22143 ± 6	$1+90$	22143 ± 7	22247
Isopentane	$1+40$	22181 ± 4	$1+59$	22182 ± 4	22364

3.2 Statistical Convergence Analysis

In several previous applications we have shown that the auto-correlation function of the energy can be used to obtain statistically converged values from a small number of uncorrelated structures. In the present applications, only 40 QM calculations have been performed and it thus seems quite appropriate to discuss the convergence problem. Figure 8 shows the distribution of the calculated individual transition energies for the case of β-carotene in acetone, including 50 acetone molecules (first shell). The calculated average value, as given before in table 4, is 22071 cm^{-1} and is shown as the horizontal line in

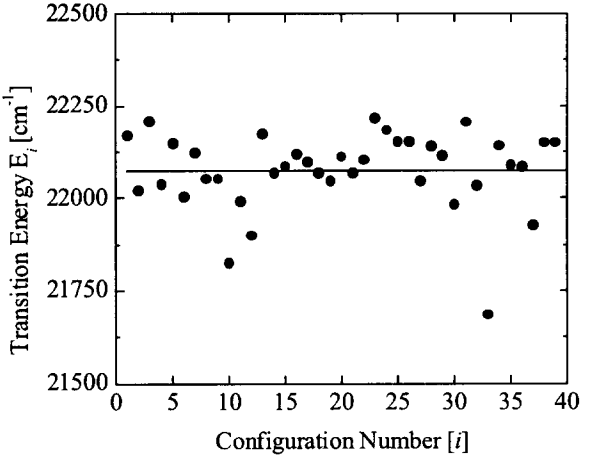


Figure 8: Distribution of the individual values of the $\pi - \pi^*$ transitions of the 40 MC configurations of β-carotene and 50 acetone molecules.

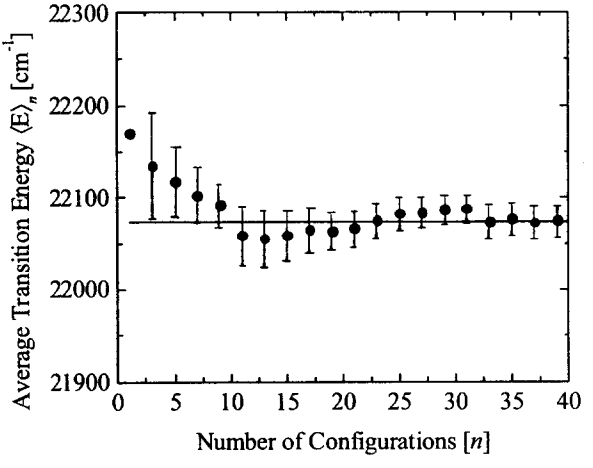


Figure 9: Convergence of the average value of the $\pi - \pi^*$ transition of β-carotene and 50 acetone molecules.

Figure 8. This distribution clearly shows that a single structure can not describe the liquid situation. Although, some structures can give a transition energy that is close to the average, this is rather fortuitous and a few other structures give transition energies that are far from the average. It is necessary to consider the average of several calculations. Figure 9 shows how the average value approaches the convergence with increased number of structures used. As stated before, and now clearly seen in figure 8, the calculated $\pi - \pi^*$ transition energy is a converged value, after using only 35 configurations. The results obtained for all the other solvents are similar to those shown for the acetone case. Converged values are obtained and a single structure can not represent the statistical nature of the liquid. It may be worth mentioning that these results demonstrate that using gas-phase optimized geometries in solute-solvent situations is rather artificial and this single structure clearly can not represent the liquid environment. Using the statistically uncorrelated structures obtained from an analysis of the auto-correlation function of the energy we obtain converged results after only a few QM calculations. The spread of the calculated results can be used to obtain the contribution of the liquid structure to the line broadening[25, 28, 38]. Truly uncorrelated configurations are obtained only with an infinite separation, because the auto-correlation follows an exponential decay[27, 28, 29]. In most of our applications we have used structures that are less than 10% correlated. It has been discussed before that using more structures is important for decreasing the statistical error but has no effect on the converged average value[32]. Clearly, the same analysis can be used for the calculation of other properties[32].

4 Summary and Conclusions

The solvatochromic shifts of the $\pi - \pi^*$ transition of all-*trans*-β-carotene in different solvents have been studied using a sequential Monte Carlo/quantum mechanics (S-MC/QM) methodology. In this procedure we first generate the structures of the liquid using Metropolis MC simulation and perform the QM calculations in selected structures generated by the simulation. These structures are selected after an analysis of the relative statistical correlation between successive configurations. This leads to a large decrease of the number of structures used in the QM calculations, without affecting the average converged value. In the present application it is shown that including only 40 QM calculations gives statistically converged results. To deal with the very elongated shape of the all-*trans*-β-carotene solute molecule the MC simulation has been first extended to a large rectangular box. The use of a spherical radial distribution function is criticized in this case and we developed a nearest-

neighbor distribution function between *all atoms* of the elongated solute and the nearest atom of each and every one of the solvent molecules. Although this has no effect on small and regular-shaped molecules it is of great importance in elongated solutes leading to a more appropriate distribution of neighbor molecules in solution. The nearest-neighbor distribution function, in fact, generalizes the concept of solvation shells for a solute of any arbitrary shape. Using only the first solvation shell the calculated results are found to be in very good agreement with the experimental results. However, to obtain the relative shifts in different solvents of varied properties, we found necessary to extend the number of solvent molecules. The relative shifts in isopentane, acetone, methanol and acetonitrile are calculated in excellent agreement with the experimental results. The different solvents are examples of systems of varied nature, differing in dielectric constants and covering a wide range of polarities, and including also polar and non-polar solvents.

As β-carotene itself is non-polar and the $\pi - \pi^*$ transition leads to a non-polar excited state, most of the solvatochromic shifts are consequence of the dispersive interaction. The solvation shift does not depend on solvent molecules that are situated much beyond the first solvation shell. In the present application, we find that inclusion of solvent molecules up to 6.0 $\AA$ is enough to give stable and accurate results, if the nearest-neighbor distribution function is used. This has also been found in the solvatochromic shifts of benzene in different solvents where the first solvation shell gives stable and accurate results[29]. This is, however, opposite to the case of formaldehyde (a polar molecule) in water (a protic solvent) where solvent molecules up to a distance of 10 $\AA$, were found to still affect the solvation shift[28]. The inclusion of dispersion interaction in the calculation of solvent effects has been recognized as one of the most important and difficult problems. It has been demonstrated[35] that although dispersion is a double excitation, calculation on a supermolecule that contains only single excitations includes dispersion interaction between the two subsystems when energy differences are taken between the ground state and low energy excited states in which single excitations dominate. Therefore the CIS calculations using supermolecular structures with explicit solute and solvent molecules seem to be an important step in this direction. Judging, from the qualitative and quantitative results of the solvatochromic shifts of β-carotene in different solvents, we are led to conclude that the most important contribution of dispersion is properly included.

5 Acknowledgments

It is our great privilege to dedicate this contribution to the memory of our teacher and friend Prof. Per-Olov Löwdin.This work is supported by CNPq, CAPES and FAPESP(Brazil).

References

[1] C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, Verlag Chemie, Weinheim, New York (1979).
C. Reichardt, *Chem. Rev.* **94,** 2319 (1994).

[2] C. J. Cramer and D. G. Truhlar, *Chem. Rev.* **99,** 2161 (1999).

[3] J. Tomasi and M. Persico, *Chem. Rev.* **94,** 2027 (1994).

[4] M. Orozco and F. J. Luque, *Chem. Rev.* **100**, 4187 (2000).

[5] L. Onsager, *J. Am. Chem. Soc.* **58,** 1486 (1936).

[6] J. G. Kirkwood, *J. Chem. Phys.* **2,** 351 (1934).
J. G. Kirkwood and F. H. Westheimer, *J. Chem. Phys.* **6,** 506 (1938).

[7] O. Tapia and O. Goscinski, *Molec. Phys.* **29,** 1653 (1975).

[8] J. L. Rivail and D. Rinaldi, *Chem. Phys.* **18,** 233 (1976).

[9] S. Miertus, E. Scrocco and J. Tomasi, *J. Chem. Phys.* **55,** 117 (1981).

[10] K. V. Mikkelsen, H Ågren and H. J. A. Jensen, *J. Chem. Phys.* **89,** 3086 (1988).

[11] E. Cances, B. Mennucci and J. Tomasi, *J. Chem. Phys.* **107,** 3032 (1997).

[12] A. Klamt and G. Schüürman, *J. Chem. Soc.Perkin Trans.* **2,** 799 (1993).
A. Klamt, V. Jonas, T. Burger and J. C. Lohrenz, *J. Phys. Chem. A* **102,** 5074 (1998).

[13] C. J. Cramer and D. G. Truhlar, *J. Am. Chem. Soc.* **113,** 8305 (1991).

[14] M. M. Karelson and M. C. Zerner, *J. Phys. Chem.* **96,** 6949 (1992).

[15] O. Christiansen and K. V. Mikkelsen, *J. Chem. Phys.* **110,** 1365 (1999).

[16] K. K. Baldridge and V. Jonas, *J. Chem. Phys.* **113,** 7511 (2000).
K. K. Baldridge, V. Jonas and A. D. Bain, *J. Chem. Phys.* **113,** 7519 (2000).

[17] J. Li, C. J. Cramer and D. G. Truhlar, *Int. J. Quantum Chem.* **77,** 264 (2000).

[18] A. Warshel and M. Levitt, *J. Mol. Biol.* **103,** 227 (1976)
A. Warshel, *J. Phys. Chem.* **83,** 1640 (1979).

[19] M.J. Field, P. A.Bash and M. Karplus, *J. Comput. Chem.* **11,** 700 (1990).

[20] J. Gao and X. Xia, *Science* **258,** 631 (1992).

[21] J. Gao, *J. Am. Chem. Soc.* **116,** 9324 (1994).

[22] J. Gao and K. Byun, *Theor. Chem. Acc.* **96,** 151 (1997).

[23] D. Bakowies and W. Thiel, *J. Phys. Chem.* **100,** 10580 (1996).

[24] M. A. Thompson, *J. Phys. Chem.* **100,** 14494 (1996).

[25] J. T. Blair, K. Krogh-Jespersen and R. M. Levy, *J. Am. Chem Soc.* **111,** 6948 (1989).

[26] J. Zeng, J. S. Craw, N. S. Hush and J. R. Reimers, *J. Chem. Phys.* **99,** 1482 (1993).

[27] K. Coutinho and S. Canuto, *Adv. Quantum Chem.* **28**, 90 (1997).

[28] S. Canuto and K. Coutinho, *Int. J. Quantum Chem.,* **77** 192 (2000).

[29] K. Coutinho, S. Canuto and M. C. Zerner, *J. Chem. Phys.* **112**, 9874 (2000).

[30] K. Coutinho, M. J. de Oliveira and S. Canuto, *Int. J. Quantum Chem.* **66**, 249 (1998).

[31] K. J. de Almeida, K. Coutinho, W. B. De Almeida, W. R. Rocha and S. Canuto, *Phys. Chem. Chem. Phys.* **3**, 1583 (2001).

[32] W. R. Rocha, K. Coutinho, W. B. De Almeida and S. Canuto, *Chem.Phys. Lett.* **335**, 127 (2001).

[33] P. Suppam and N. Ghoneim, *Solvatochromism*, Royal Soc. Chem., Cambridge (1997).

[34] N. Rösch and M. C. Zerner, *J. Phys. Chem.* **98**, 5817 (1994).

[35] S. Canuto, K. Coutinho and M. C. Zerner, *J. Chem. Phys.* **112**, 7293 (2000).

[36] W. Liptay, in *Modern Quantum Chemistry*, ed. O. Sinanoglu, Part II, p. 173, Academic Press, New York (1966).

[37] K. Coutinho, S. Canuto and M. C. Zerner, *Int. J. Quantum Chem.* **65**, 885 (1997).

[38] K. Coutinho and S. Canuto, *J. Chem. Phys.* **113**, 9132 (2000).

[39] K. Coutinho, N. Saavedra and S. Canuto, *J. Mol. Struct. (Theochem)* **466**, 69 (1999).

[40] K. J. de Almeida, W. R. Rocha, K. Coutinho and S. Canuto, *Chem.Phys. Lett.* **345**, 171 (2001).

[41] A. B. Myers and R. R. Birge, *J. Chem. Phys.* **73**, 5314 (1980).

[42] T. Abe, J.-L. Abboud, F. Belio, E. Bosch, J. I. Garcia, J. A. Mayoral, R. Notario, J. Ortega and M. Rosés, *J. Phys. Org. Chem* **11**, 193 (1998).

[43] C.-G. Zhan and D. M. Chipman, *J. Chem. Phys.* **109**, 10543 (1998), and refs. therein.

[44] P.A. Egelstaff, *An Introduction to the Liquid State*, Oxford Science Publ.,Oxford (1994).

[45] J. Applequist, *J. Phys. Chem.* **95**, 3539 (1991).

[46] M. C. Zerner, in *Problem Solving in Computational Molecular Science. Molecules in Different Environments*, Kluwer Acad. Publ., p. 249 (1997).

[47] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford, (1987).

[48] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

[49] D. R. Lide (ed.) *Handbook of Chemistry and Physics*, 73rd edition, 1992-1993, CRC-Press, Boca Raton, (1992).

[50] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996).

[51] C. Sterling, *Acta Cryst.* **17,** 1224 (1964).

[52] A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
A. D. Becke, *J. Chem. Phys.* **98** 5648 (1993).

[53] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).

[54] W. L. Jorgensen, J. M. Briggs and M. L. Contreras, *J. Phys. Chem.* **94**, 1683 (1990).

[55] W. L. Jorgensen and J. M. Briggs, *Mol. Phys.* **63**, 547 (1988).

[56] W. J. Jorgensen, J. D. Madura and C. J. Swenson, *J. Am. Chem. Soc.* **106**, 6638 (1984).

[57] W. L. Jorgensen, J. D. Madura and C. J. Swenson, *J. Phys. Chem.* **90**, 1276 (1986).

[58] K. Coutinho and S. Canuto, *DICE* (version 2.8): A general Monte Carlo program for liquid simulation, University of São Paulo, (2000).

[59] M. C. Zerner, *ZINDO*: A Semi-empirical Program Package, University of Florida, Gainesville, FL 32611.

[60] J. Ridley and M. C. Zerner, *Theor. Chim. Acta* **32,** 111 (1973).

Quantum chemistry of antimatter

Piotr Froelich

Department of Quantum Chemistry, Uppsala University
Uppsala 75120, Sweden

Abstract

A survey of the issues and incentives for research on antimatter is given. Some applications of atomic and molecular theory are discussed in the context of the current search for antimatter in outer space, the ongoing efforts to breed antihydrogen in the laboratory, and the forthcoming experiments with cold antihydrogen.

Contents

ADVANCES IN QUANTUM CHEMISTRY, VOLUME

0065-3276/02 $35.00

Till Per

Stående på ett avkylt korn av askan ser vi det långsamma förbleknande av solarna och försöker få en åskådlig bild av världarnas födelse.

Lemaître

1 Introduction

Antimatter, no longer a science-fiction concept nor an academic curiosity, will soon be manufactured in the laboratory. Research on antimatter has a high priority as evidenced by the large scale efforts to form, cool and trap antihydrogen (at CERN), or by the search for antimatter in the Universe (from the orbiting space station launched by NASA). These large scale experimental enterprises will profit from theoretical assistance in the area of atomic and molecular physics. Such assistance is expected to be important for the design, implementation and analysis of experiments on forming, cooling and trapping antihydrogen, as well as for astrophysical research.

The recent advances in producing, trapping and cooling antiprotons and positrons opened the possibility of antihydrogen formation in laboratory. This may allow the studies of antimatter and tests of fundamental physical principles such as charge - parity - time (CPT) invariance or the weak equivalence principle (WEP) for antiparticles. Such experiments are planned at the newly built CERN AD (Antiproton Decelerator) within ASACUSA, ATRAP and ATHENA projects, which have just started their operations.

ASACUSA (Atomic Spectroscopy and Collisions Using Slow Antiprotons) will carry out an atomic physics program on antiproton-containing species. This will feature antiproton - atom collisions and high precision spectroscopy on antiprotonic species. The ATHENA (AnTiHydrogEN Apparatus) experiment aims at producing antihydrogen atoms at low energies, capturing these atoms in a magnetic trap. The goal of ATRAP is to compare hydrogen and antihydrogen to the accuracy sufficient to test the possible CPT violations. One envisions very precise laser-spectroscopic measurements comparing the atomic and antiatomic transitions.

The advent of the Antiproton Decelerator at CERN (brought into operation in 2000) generates a need for theoretical calculations on small exotic systems involving positrons, antiprotons, and their collisional interactions including antihydrogen formation. Beyond the formation step, calculations are needed to predict the interaction of antihydrogen with matter, e.g. due to the imperfect vacuum in the trap, or the sympathetic cooling of antihydrogen in collisions with cold hydrogen. There will also be a need for calculations on the processes involved in the detection of antihydrogen. Finally, calculations will be needed to correctly interpret the results of actual experiments on antihydrogen, such as spectroscopic tests of CPT invariance or of WEP.

2 From antiparticles to antimatter

In 1928, while working on the theory of relativistic quantum mechanics, Dirac encountered unexpected and seemingly nonphysical solutions to the relativistically invariant equations describing relativistic electrons. The equations predicted the existence of electrons with negative energy (or negative mass). Not intimidated by the apparent conflict with physical reality, Dirac explained the non-observability of the negative energy states by proposing his famous hole theory [1]. He suggested that in the normal ground state of Nature (referred to as a vacuum state) all possible negative energy levels are already occupied. He then assumed that it is possible to raise a "negative energy electron" from the vacuum state to a positive energy state in which it behaves as a normal electron with positive mass. At the same time a "hole" is produced in the sea of negative energy electrons. Such a hole behaves as a particle with the mass of an electron, but with a *positive* charge. In such a way Dirac interpreted the absence of a negative particle with negative mass as the presence of a positive particle with positive mass: a positive electron.

The "hole theory" was perceived as a fictive mathematical construction and was initially rejected by prominent contemporary physicists such as Pauli and Bohr. The physical reality of "antiparticles" was not taken seriously even by Dirac himself. In 1931 he wrote about his anti-electron: "we should not expect to find it in Nature" [2]. Surprisingly, the first anti-electrons were discovered already in 1932 by Anderson, who studied cosmic rays in Caltech's magnet cloud chamber. Anderson noticed abnormally bending trajectories indicating the presence of light positively charged particles and, as related by Fowler [3], "could not resist the devastating conclusion" that they are caused by positive electrons! The first piece of antimatter, a positron, made its physical appearance.

The hunt for other antiparticles began. Soon it became obvious that the existence of a positron is not an esoteric exception, and that the general symmetries of Physics require the existence of the corresponding antiparticles to *all* particles. It was not until 1955 that the antiproton was discovered in high energy collisions generated in the Berkeley accelerator. The two basic ingredients of the simplest anti-atom were there. The concept of antimatter - a neutral conglomerate of antiparticles bound into antiatoms and antimolecules - was born. By the end of the fifties the presence of antiparticles was so firmly established that Alfvén [4] suggested the term *coino*-matter to distinguish ordinary matter from *anti*-matter.

In 1996, the first anti-atoms were synthesized at CERN and Fermilab [5, 6]. They were obtained in energetic antiproton-nuclei collisions whereby electron-positron pairs are created and the antiproton captures one of the positrons into a bound state. These antihydrogen atoms were however moving with relativistic velocities and were therefore not suitable for spectroscopic experiments. Accurate spectroscopic comparison of hydrogen and antihydrogen must involve the narrow transition lines that are associated with long lived states and long observation times. For such purpose *cold* antihydrogen is much preferred.

3 Matter-antimatter asymmetry in the Universe

As far as we can see into the Universe, we don't observe any primordial antimatter. Within the limits of our present observational horizon the Universe is seen to contain only matter and no antimatter. The presence of cosmic antimatter would lead to observable traces of annihilation; however the measurements of the extragalactic γ ray flux indicate an absence of annihilation radiation, and the microwave background spectrum lacks a corresponding distortion. These findings preclude the existence of a significant amount of antimatter within tens of Megaparsecs, which is the scale of super-clusters of galaxies.

Could that be so that the Universe was created with the preponderance of matter over antimatter? We have no support for such hypothesis. Einstein remarked: "If that's the way God made the world then I don't want to have anything to do with Him" [7]. Indeed, the contemporary Standard Model of Physics suggests that equal amounts of matter and antimatter were born during the Big-Bang. Where has the antimatter gone? What causes the apparent asymmetry between matter and antimatter? Obviously the antiparticles have been annihilated by particles - but apparently this process was not fully symmetric, since enough matter was left over for our Universe. We seem to be the result of an accident, caused by a a slight imperfection of Nature.

Within our present understanding based on the Standard Model, the matter-antimatter asymmetry originates from a slight deviation from the fundamental symmetry known as charge parity (CP) reversal. The parity reversal P reverses the direction of all vectors involving spatial coordinates (most notably position and momentum). The charge conjugation C reverses certain (additive) quantum numbers of particles (among them the electric charge) i.e. turns particles into antiparticles. Both operations are idempotent, i.e. give the same result when applied twice ($C^2 = P^2 = 1$). One might expect that even the combined CP symmetry should hold, i.e. that the parity reflected world of antiparticles should be the same as our world.

However, as noted by Lee and Yang in 1956 [8], such expectations reflect only our esthetical preferences or the intuition rooted in the laws of classical mechanics and electrodynamics. According to their prediction the weak interactions (responsible for the decay of particles) were not expected to obey the P symmetry. That prediction was confirmed experimentally by Wu in 1957 [9]. She utilized a simple property of the inversion operation - namely that it does not influence the so called axial vectors, such as the angular momentum or spin of particles. For the radioactive nuclei decaying through the emission of electrons and antineutrinos, parity conservation demands that the intensity of electrons emitted parallel to the nuclear spin should be the same as for electrons emitted antiparallel to the nuclear spin. The experiment studying the decay of ^{60}Co nuclei showed that this is *not* the case, i.e. that there is a preferred direction for the electron emission. Thus particles differing by spin orientation were shown to have different decay properties, and since particle decay (such as e.g. β-decay mentioned above) is governed by weak interactions, the latter must violate inversion symmetry.

Also the C symmetry is violated in weak interactions. Here we should distinguish between C invariance and invariance with respect to the operation of reversing only

the electric charge, Q. If C-invariance does not hold, why then do particles and antiparticles have exactly equal amounts of opposite charge? The answer to that question turned out to be given by a weaker (but more fundamental) condition - requiring the invariance of physical laws with respect to the *combined* operation of parity (P), charge (C) and time (T) reversal. This condition, known as the CPT theorem [10, 11, 12, 13] is believed to represent the minimal set of conditions for the existence of a relativistic field theory consistent with quantum mechanics and special relativity; this entails the general requirements of locality, Lorentz invariance and unitarity. The theorem assures that the laws of physics for the time-reversed, parity-reflected anti-world are the same as for our world (and for any other world obtained by the three operations C, P, T applied in arbitrary order).

The CPT theorem suffices to guarantee the equality of masses, lifetimes, spins, and exactly opposite charges and magnetic moments for particles and antiparticles. The following consequence is that the structure of bound species should be the same for both matter and antimatter; in particular the fine, hyperfine structure, and Lamb shift of antiatoms should be the same as that of atoms.

It should be clear that invariance with respect to the combined operation CPT (applied in any order) does not imply invariance with respect to the individual operations separately. Conversely, the theorem implies that if one of the symmetries is broken then some other must be broken, too. Indeed, as mentioned earlier the P and C symmetries were found *not* to hold separately. How about their combination? Are the physical laws for particles moving right the same as for antiparticles moving left?

In 1964 Cronin and Fitch [14, 15] showed experimentally by studying the decay of kaons that the combined operation CP is *not* an exact symmetry of Nature. This discovery is even more perplexing when viewed in the context of time evolution. If the (so far unchallenged) CPT theorem is to hold, then violations of joint CP symmetry imply violations of T symmetry (with T reversing the *motion* of particles). The laws of physics are therefore not the same when the time changes direction.

Even if allowed by the CPT theorem, the non-conservation of CP symmetry was hard to accept - not least because it was not consistent with the Standard Model. In 1972, Kobayashi and Maskawa predicted that CP violation would be consistent with the Standard Model provided that three generations of quarks exist (and only two were known at the time). The subsequent discoveries of the τ lepton by Pearls (1975) and of the top and bottom quarks at the Fermilab confirmed the existence of a third family, this resulted in the incorporation of the CP violation into the Standard Model.

It is believed that processes violating CP symmetry, therefore inducing imbalance in the number of particles and antiparticles, were occurring during the early and hot stage of the Universe. Some possible scenarios are suggested by the generalizations of the Standard Model known as Grand Unification Theories (GUT). They predict that at sufficiently high energy (GUT energy) the electromagnetic, strong and weak forces have the same strength, and electrons or positrons can be transformed into quarks. For instance it is possible that during the hot inflation time, due to the CP violations more antielectrons were transformed into quarks than electrons into antiquarks. The resulting excess of particles over antiparticles was probably not larger than 1 particle

in 10^{10}. After the Universe had cooled down to the stage when particles and antiparticles could no longer be created in collisions, the initial imbalance between matter and antimatter was permanently frozen. Annihilation, whereby particles and antiparticles are eliminated in equal amounts, could however still proceed and when all antiparticles were gone - our own Universe resulted.

The conditions required for a non-symmetric Universe were first put forward by Sakharov [16]: they include non-conservation of the baryon number, C and CP symmetry violation, and the existence of a period of thermal non-equilibrium during the evolution; however, the present limits on the proton lifetime (10^{33} years) are inconsistent with the first condition, and the small degree of CP symmetry violation displayed by kaons is not compatible with the second condition.

Even though the CP violation is allowed by the Standard Model, its predicted effect seems to be too small to explain the excess of matter (and the existence of our *coino*-matter Universe). This means that our way of understanding these CP violations and perheaps the Standard Model itself must be incomplete. Therefore a lot of effort goes into research on fundamental symmetries and setting bounds on their violations.

3.1 Cosmic antimatter

The matter - antimatter asymmetry in the Universe and its relation to the underlying symmetries of Nature is of such fundamental importance that a number of large-scale experiments have been undertaken to search for cosmic antimatter.

For instance, the content of positrons in cosmic rays has been previously studied with help of the High Energy Antimatter Telescope (HEAT), which was lifted above the densest part of the atmosphere on a balloon. The results indicate that the ratio of positrons to electrons in the cosmic wind is roughly 1:10. The Isotope Matter-Antimatter Experiment (IMAX) searched for the presence of antiprotons. The ratio of antiprotons to protons turned out to be about 1:10 000. However the presence of these energetic antiparticles observed in the upper atmosphere is of no consequence for the question about the symmetry of the Universe. These antiparticles have a secondary origin and are produced in interstellar space *via* collisions of the energetic particles propagated in the cosmic wind. The latter originate from events such as Super Nova explosions, and their energy is not sufficient to collisionally produce antiparticles heavier then an antiproton. (More exactly, the probability of obtaining heavier nuclei in such collisions drops by four orders of magnitude with each additional anti-nucleon [17].) Therefore the discovery of e.g. antihelium nuclei would imply its primordial cosmological origin, suggesting that antimatter might have been nucleosynthesised in the very early and hot stage of the Universe. Since heavier nuclei must be synthesized in nuclear reactions, the discovery of even single carbon antinuclei would imply the existence of antistars.

A major scientific experiment devoted to the search for antimatter originating from outside of our galaxy is presently being conducted by NASA. The experiment is a part of the project aiming at studies of matter, antimatter and dark matter in space. This project, coordinated by the Department of Energy (DOE), is a large collaboration of more than 30 universities, and will be conducted with the help of a state-of-the-art

particle detector placed in the orbit of Earth. This device, called Alpha Magnetic Spectrometer (AMS), will use the unique environment of space to study the properties and origin of cosmic particles and antiparticles and also may lead to the discovery of the dark matter that comprises up to 90% or more of the Universe.

The prototype of the detector has already been tested in space on the space shuttle. The AMS will be placed on the International Space Station (ISS) orbiting Earth in the fall of 2003. One of the most important experiments will be the search for antimatter of a cosmological origin.

4 Investigations of the basic symmetries in Nature

There are both theoretical and experimental reasons to search for CPT violations. The strong theoretical incentive is that, even though the CPT invariance is required to formulate a quantum field theory consistent with special relativity, it turns out to be difficult to construct a gravitational relativistic quantum field theory of the GUT type with the CPT symmetry maintained. In other words it is difficult to incorporate the CPT invariance in the GUT-type extensions of the Standard Model.

The most perceptible experimental reason is the evident asymmetry of the Universe. The CP violation alone, on the level allowed by the Standard model, is not sufficient to explain the excess of matter in the Universe [18]. Finding a suitable extension of the Standard Model is not easy: not only do the GUT theories tend to violate CPT invariance but they are also difficult to test. Specifically they can not be tested by means of the particle accelerator experiments, which so far have been very successful in extending our knowledge of Physics. This is because the GUT energy is some 10^{13} times larger than what we can now days produce in the most powerful accelerators. Consequently we must look for the *low-energy manifestations* of GUT and CPT invariance. Interestingly some possibilities for such tests lie within the realm of atomic and molecular physics.

Recent advances in the production and storage of positrons and antiprotons have made it possible to think about the synthesis of atomic antimatter in the laboratory. Parallely, contemporary advances in cooling and trapping atoms have led to an unprecedented accuracy of spectroscopic measurments. The important difference between the spectroscopy of atoms and antiatoms is that in the latter case, because of annihilation, the sample must be isolated from the surrounding environment.

It is interesting to observe that two important techniques for experimenting with antihydrogen, its cooling (needed for the elimination of the thermal motion which is a basic obstacle for accurate spectroscopy) and wall-free confinement (to eliminate annihilation) have been developed in conjunction with the search for Bose-Einstein condensations of atomic gases. The wall-free confinement of neutral antihydrogen atoms is possible in the magnetic traps of the same type that have been designed for obtaining the Bose-Einstein condensation in hydrogen [19].

An experiment aiming at the production of cold antihydrogen is presently being conducted at CERN. Hence it will be feasible to test CPT symmetry by the direct comparison of the spectroscopic properties of the simplest atom and its antiatom.

Hydrogen has played an important role in our understanding of the physics of matter. The *anti*hydrogen is expected to play a similar role in understanding of the antimatter. Since $\bar{\mathrm{H}}$ is an exact conjugate of H, the test of CPT invariance is limited only by the accuracy of the experiment. The accuracy of the spectroscopic measurements on H has already reached one part in 10^{14} and similar accuracy is expected for $\bar{\mathrm{H}}$. This makes antihydrogen a microscopic low energy laboratory for testing the fundamental symmetries of Nature.

What can be tested? As mentioned before, CPT invariance guarantees the equality of masses, charges and lifetimes of particles and antiparticles. This means that the experimental investigations of masses, charges, etc. of particle - antiparticle pairs are tests of CPT symmetry. Such experiments are not easy to do with the charged particles themselves (because of their interactions with stray fields). Comparison of neutral atom - antiatom pairs is much more convenient. In particular, the fine structure, hyperfine structure and Lamb shifts of atoms and antiatoms should be identical - and can be tested in laboratory.

The spectroscopic measurements, aiming at the comparison of the level structure of hydrogen and antihydrogen, are of special interest. This is because advances in the cooling and trapping of ordinary hydrogen have made possible very accurate measurements of the hydrogenic levels. For instance, the $2s-1s$ transition has been measured with the relative accuracy $(\Delta\nu_{1s-2s}/\nu_{1s-2s})$ of 3.4 parts in 10^{14} in the cold beams of hydrogen atoms [20].

The measurements of $2s-1s$ transitions in magnetically trapped hydrogen have achieved a relative accuracy of one part in 10^{12} [21] by means of two-photon spectroscopy which eliminates the first-order Doppler broadening. It is hoped that this technique will allow the measurement of the $1s-2s$ transition with the accuracy limited only by the shape of the transition line dictated by quantum electrodynamics, i.e. to a few parts in 10^{15}. Further, if the center of the $1s-2s$ line could be determined with the accuracy of a few parts in 10^3 of its width, the relative accuracy for this transition would increase to a few parts in 10^{18}.

But what precision is required? The answer to this question is two-fold. One scale of precision is set by the previous experimental tests of CPT invariance; the new tests must simply have better accuracy in order to be interesting. The so far best (albeit indirect) experimental test has been obtained from the comparison of the masses of neutral kaons K^0 and $\bar{K}^0$, which are known to be equal with the accuracy of 1 part in 10^{18} (i.e. $(m_{K^0}-m_{\bar{K}^0})/m_{K^0}\leq 10^{-18}$) [22].

Another scale of accuracy is set by theoretical considerations regarding the measure of the expected symmetry violations: the tests should be sensitive enough to detect the predicted effects.

In this respect the spectroscopic investigations of the $1s-2s$ transition in hydrogen and antihydrogen seem to be interesting. This is because the spectroscopic line-width corresponding to this transition is exceedingly narrow. The single photon transition from the $2s$ state is forbidden to the first order in electric dipole interaction, and the magnetic dipole single photon emission has a very long lifetime of 2 days. The decay of the meta-stable 2s state is therefore governed by a slow two-photon transition

which leads to a very long lifetime of $\tau = 1/8$ s and makes the corresponding linewidth very small, $\Gamma = \hbar/\tau = 2.6 \cdot 10^{-15}$ eV. If the transition is determined down to the quantum accuracy of its width, then the relative accuracy of the measurement is $\Gamma/\Delta_{1s-2s} = 2.58 \cdot 10^{-16}$. If in addition the center of the line could be determined with the precision of one part in 1 000, the measurement comparing hydrogen and antihydrogen spectra could be pushed to the accuracy of $2.58 \cdot 10^{-19}$. That exceeds the accuracy of any existing test of CPT invariance and makes the spectroscopic comparison of the $1s - 2s$ transition in hydrogen and antihydrogen of potential interest.

The gross energy difference between the energy levels is dictated by the Rydberg constant. The anti-Rydberg constant is sensitive to the ratio of the positron and antiproton masses and to the squares of the positron and antiproton charges. Although direct comparisons of the Rydberg and anti-Rydberg constants might sharpen the bounds on equality of masses and charges for the proton and antiproton, this seems not to be the best way of looking for the CPT violation in the conjugated atomic systems. Therefore, a more detailed analysis of atomic transitions and their sensitivity to CPT violating effects is needed.

4.1 CPT violation effects in atomic transitions

As mentioned before, the low energy manifestations of CPT violation (assumed to occur at GUT energies) might be suppressed. To establish what sort of measurements have the greatest chance to discover the CPT violations a help from the theory is useful. Of course, the theoretical predictions are by necessity limited by the validity of the extensions of the Standard Model applied in the analysis.

Extensions allowing CPT and Lorentz invariance violations [23] lead to atomic models that reflect the symmetry violations as shifts in the atomic energy levels. It has been argued that such effects can be discovered in the fine-structure of $1s - 2s$ transitions and also in the hyperfine structure of Zeeman transitions.

Concentrating on the $2s$ and $1s$ levels we notice that there are four transitions allowed by the selection rules for two-photon transitions, $\Delta F = 0$ and $\Delta m_F = 0$; they occur between states having the same lepton-hadron spin configuration. Perturbative analysis of the CPT violating terms in an extended Dirac equation leads to the energy shifts of the atomic levels. It turns out, however, that for free atoms these shifts are to leading order the same for $1s$ and $2s$ levels having the same spin configuration. In addition, the shifts are to leading order the same for hydrogen and antihydrogen. Therefore, to leading order, there are no CPT violating effects to be seen in the $1s - 2s$ spectra of hydrogen, antihydrogen, or in the experimental comparison of these spectra; i.e. $\delta\nu^{\mathrm{H}} = \delta\nu^{\bar{\mathrm{H}}} \simeq 0$.

Further analysis shows that the non-vanishing leading order effects of symmetry violation are expected in transitions between the hyperfine Zeeman states. Such effects can be observed while placing the atoms in the magnetic field - e.g. by confining the hydrogen/antihydrogen in the magnetic trap. Both $1s$ and $2s$ levels will then each split into four Zeeman components, designated a $(F = 0, M_F = 0)$, b $(F = 1, M_F = -1)$, c $(F = 1, M_F = 0)$, d $(F = 1, M_F = 1)$ in the order of increasing energy. Out of these

four states, only the low field seeking states of the c and d type can be confined in an Ioffe-Pritchard trap.

The CPT-related shift of the the un-mixed d type spin states ($|\uparrow\uparrow\rangle$) are the same to leading order, so that $\delta\nu^{\rm H}_{1d,2d} = \delta\nu^{\bar{\rm H}}_{1d,2d} \simeq 0$; however, the transitions between the c type spin-mixed states ($cos\theta_n|\uparrow\downarrow\rangle + sin\theta_n|\downarrow\uparrow\rangle$) are sensitive to CPT violating effects, resulting in a frequency shift $\delta\nu^{\rm H}_{2c}$ which is field dependent but of zeroth order with respect to α. This is possible thanks to the fact that the spin functions of the $1s-c$ and $2s-c$ states are not the same; the mixing of the spin components depends not only on the field strength but also on the main quantum number n. In addition, the shift for hydrogen $\delta\nu^{\rm H}_{1c.2c}$ is different from that for antihydrogen, $\delta\nu^{\bar{\rm H}}_{1c,2c}$, which leads to spectroscopic difference between hydrogen and antihydrogen: $\Delta\nu^{H-\bar{H}}_{1c,2c} = \nu^{\rm H}_{1c,2c} - \nu^{\bar{\rm H}}_{1c,2c}$.

Similarly, the hyperfine transitions involving the spin-mixed Zeeman states *within* the $1s$ level are expected to display frequency shifts due to CPT violation. In particular, the transition from the low field seeking $1s-c$ states to the $1s-d$ states could provide an unsupressed evidence of CPT violation that leads to differences between H and $\bar{\rm H}$: $\Delta\nu^{\rm H-\bar{H}}_{1d,1c} = \nu^{\rm H}_{1d,1c} - \nu^{\bar{\rm H}}_{1d,1c}$.

4.2 Gravitational effects in atomic transitions

Spectroscopic differences between hydrogen and antihydrogen do not necessarily imply the violation of CPT invariance. This might also be due to an anomalous gravitational redshift of the antiatom transition frequency (frequencies of the atomic clock change with the gravitational field). We then come to the question: how does antimatter interact with gravity? Do matter and antimatter behave in the same way in the gravitational field of Earth?

A short glance back into the past gives an interesting perspective on this question. Aristotle (350 BC) maintained that an iron ball falls faster than a feather, implying that the interaction of these two bodies with the Earth is different. Galileo (1604) has experimentally shown that all bodies fall with the same acceleration, a property known as the principle of universality of free fall. This principle can be derived from the Newton's gravitational law (1687) which, when combined with the Newton's equation of motion shows that the gravitational acceleration of all bodies is the same.

The attractive force between a given body B and the Earth E is proportional to their *gravitational masses* m^g_B and m^g_E. The resulting acceleration g_B is inversely proportional to the *inertial mass* of the body m^i_B. One gets

$$G\frac{m^g_E \cdot m^g_B}{R^2_{EB}} = m^i_B g_B \quad \Rightarrow \quad g_B = \frac{m^g_E G}{R^2_{EB}} \cdot \frac{m^g_B}{m^i_B} = const \cdot \frac{m^g_B}{m^i_B} \tag{1}$$

which shows that if $\frac{m^g_B}{m^i_B}$ is constant, then so is g_B, in which case all bodies, regardless of their mass and composition fall to Earth with the same acceleration. Could this then be true also for antibodies?

We notice that Newtonian mechanics adds an additional aspect to the Galilean principle of universal fall: the latter is true *provided* that the ratio m_g/m_i is constant

for all bodies. Hence, Newtonian mechanics relates the question about equal gravitational acceleration for all bodies to the question about the equivalence of inertial and gravitational masses.

Einstein has shown that the equivalence of inertial and gravitational mass, known as the Weak Equivalence principle, is a consequence of the more general Equivalence Principle in the general relativity theory [24]. (The latter tells that the reference frame of an observer in free gravitational fall is equivalent to the one at rest and free of gravitation.) Einstein's theory makes no distinction between the inertial and gravitational mass. This fact translates into Newtonian mechanics as the equivalence of inertial and gravitational masses. Indeed, in the general theory of relativity the quantity playing the role of the gravitational mass (or "gravitational charge") of a body is proportional to its energy-momentum tensor which in turn is proportional to the energy content mc^2 of the body; whereas the inertial mass is associated with the relativistic mass m itself. In the context of Newtonian mechanics (see eq. 1) it means that $m_g/m_i = const$ for all sorts of matter - but it still does not say anything about antimatter.

Assuming *a priori* that the gravitational acceleration of of matter g and antimatter $\bar{g}$ might be different ($\bar{g} = \eta g$, $\eta \neq 1$), one can derive the difference between the frequencies emitted by hydrogen and antihydrogen. Starting with the field-free atomic theory, the introduction of the gravitational field ϕ leads to the shift of atomic energy levels (even without WEP violation). Such a gravitational shift manifests itself at the level of the non-relativistic Schrödinger equation providing that the latter is appropriately modified. The effect of gravity can be effectively incorporated by multiplying the particle masses by a factor $(1 - 3\phi/c^2)$ and the simultaneous replacement of the dielectric constant for free space ϵ_0 by $\epsilon_0(1 - 2\phi/c^2)$ [25]. After such modifications, the energy difference between two levels of the hydrogen atom ΔE (which is proportional to $\frac{m_e}{\hbar}(\frac{Z^2e^2}{4\pi\epsilon_0})^2$ becomes in the presence of the gravitational field $\Delta E(\phi) = \Delta E(0)\frac{1-3\phi/c^2}{(1-2\phi/c^2)^2} \simeq (1 + \phi/c^2)$. The corresponding wavelength $\lambda(\phi)$ then becomes gravitationally *redshifted* relative to the field free wavelength $\lambda(0)$, $\lambda(\phi) = \lambda(0)\frac{(1-2\phi/c^2)^2}{1-3\phi/c^2} \simeq \lambda(0)(1 - \frac{\phi}{c^2})$.

Suppose that we compare the atomic energy levels of hydrogen with respect to their values at the surface of Earth, and we measure the gravitational potential with respect to its value at the surface, where we assume $\phi = 0$. If the energy difference between the two levels is $\hbar\omega(0)$ then at the top of a mountain (where the gravitational potential is ϕ) the energy difference between the same two levels is $\hbar\omega(\phi) = \hbar\omega(0)(1 + \phi/c^2)$. In the absence of CPT and WEP violations, the antihydrogen placed on the same mountain top would suffer the same energy shift, and emit the same frequencies as the hydrogen.

The *anomalous* redshift for the antihydrogen can be obtained by inserting into the Schrödinger equation an anomalous mass of the positron (caused by its different gravitational acceleration $\bar{g}$ due to WEP violation), $m_{\bar{e}} = m_e(1 - 3\eta\phi/c^2)$. This leads to the transition energy given by $\hbar\bar{\omega}(\phi) = \hbar\bar{\omega}(0)\frac{1-3\eta\phi/c^2}{(1-2\phi/c^2)^2} \simeq \hbar\bar{\omega}(0)\{1 - (3\eta - 4)\phi/c^2\}$; i.e., now the frequencies emitted by the antihydrogen are different from the hydrogenic frequencies.

If we were to observe the difference between hydrogenic and antihydrogenic transitions at the same given height (i.e. for a given ϕ), we still would not know if this difference could be attributed to the violation of WEP. For instance, it could have

been due to violation of CPT symmetry. Also, in the absence of the absolute scale for the gravitational field, we could not assume that hydrogen and antihydrogen atoms had the same spectra to start with, namely at the surface of Earth (where we have arbitrarily set $\phi = 0$). However, the WEP violation can be identified and separated from the CPT symmetry violation by considering the difference in *redshifts* of hydrogen ($\Delta\hbar\omega = \hbar\omega(\phi + \Delta\phi) - \hbar\omega(\phi)$) and antihydrogen ($\Delta\hbar\bar{\omega} = \hbar\bar{\omega}(\phi + \Delta\phi) - \hbar\bar{\omega}(\phi)$) arising when both atoms are taken through the same change $\Delta\phi$ in gravitational potential (rather than the difference between transition energies ($\hbar\bar{\omega}(\phi) - \hbar\omega(\phi)$) at the same ϕ). If we assume that other effects are independent of ϕ, they will cancel out from the redshift (even if they are present in the transition energy itself); therefore, in calculating $\Delta\hbar\bar{\omega}(0)$ we can be satisfied with the description including the gravitational effects alone. In such a case $\hbar\bar{\omega}(0) = \hbar\omega(0)$, and one obtains

$$\frac{\Delta\hbar\bar{\omega} - \Delta\hbar\omega}{\hbar\omega(0)} \simeq 3(1-\eta)\frac{\Delta\phi}{c^2}. \tag{2}$$

This differs from 0 when $\eta \neq 1$, i.e. when the gravitational accelerations of matter and antimatter are different.

5 Matter-antimatter interactions

In view of the current search for antimatter in outer space, the ongoing efforts to produce antihydrogen in laboratory, and the forthcoming experiments with cold antihydrogen we have undertaken the study of matter-antimatter interactions [26, 27, 28, 29, 30, 31, 32]. Such interactions are of interest both for CERN AD experiments on forming antihydrogen, and for the astrophysical search for the presence of antimatter in the Universe.

The theoretical investigations on the hypothetical existence of antimatter in outer space often refer to the so called B=0 (universally symmetric) Universe, in which the matter-antimatter domains are spatially separated. The interaction between such regions would lead to observable annihilation signals in the form of relic, redshifted annihilation photons contributing to the diffuse γ-ray spectrum in the cosmic radiation, or the distortion of the cosmic microwave background due to the scattering of energetic end-products (electrons/positrons) of the annihilation. The strength of such signals leads to bounds on the separation between the matter and antimatter regions, and the latter can be compared to estimations based on the observed uniformity of cosmic background radiation (CBR).

In view of the apparent absence of annihilation radiation the presumptive symmetric matter-antimatter Universe would have to be characterized by large voids between the matter and antimatter domains. In such a model the density of the Universe would have to vary on quite a large scale. On the other hand, the large scale "lumpiness" of the Universe is limited by the homogeneity of the CBR. The fluctuations of the CBR are related to the fluctuations of mass distribution through the square of the distance between the lumps of increased density [33]. This relation, given the fact that the observed inhomogeneity of the CBR is very small, limits the large scale lumpiness of

the Universe, regardless if it contains matter and antimatter, or only matter. These general considerations restrict the size and "emptiness" of the voids between the putative matter-antimatter domains, and set bounds on the anticipated signals due to the annihilation processes occuring at the overlapping matter and antimatter domains. Such bounds depend on the details of the cosmological models for the symmetric Universe; hence, experimental search for the debris of annihilation provides a test for the cosmological models.

In order to give a concrete estimate of the strength of the anticipated annihilation signals, the cosmological models must be quite explicit not only regarding the cosmological parameters, but also regarding the dynamics of annihilation. The latter aspect involves particle, plasma and atomic physics. Theoretical investigations of these questions use as input, among other parameters, the *microscopic cross sections* for hydrogen-antihydrogen collisions at the regional matter-antimatter boundaries [34]. Such cross sections are also needed to study the matter - antimatter interactions in general, and in particular to understand the prerequisites for the trapping, cooling and detection of antihydrogen at CERN.

5.1 Hydrogen-antihydrogen interactions

In recent work [26, 29] we have investigated the question of the stability of antimatter in contact with matter. The fundamental collisional interaction between hydrogen H and antihydrogen $\bar{\mathrm{H}}$ has been considered as the prototype reaction.

We have evaluated the cross sections for the most important collisional processes in the $\mathrm{H} - \bar{\mathrm{H}}$ scattering, notably the cross sections for elastic scattering (leading to antihydrogen cooling), and for various inelastic processes (leading to the loss of antihydrogen).

The cross sections have been obtained through a quantum mechanical calculation of the transition matrix elements in the distorted-wave approximation, for low and ultra-low collisional energies.

5.1.1 The rearrangement reaction

The rearrangement reaction results in the formation of protonium and positronium in the final channel, according to

$$\mathrm{H} + \bar{\mathrm{H}} \rightarrow p\bar{p} + e^+e^-. \tag{3}$$

This inevitably leads to the annihilation of anti-particles from the bound states of protonium ($\mathrm{Pn} \equiv p\bar{p}$) and positronium ($\mathrm{Ps} \equiv e^+e^-$). We found this reaction to be a very important mechanism for the loss of antihydrogen [26, 27, 29].

5.1.2 Proton-antiproton annihilation in flight

The $p - \bar{p}$ annihilation can also occur when H and $\bar{H}$ "fly by" each other, i.e. from the initial channel of the $\mathrm{H} - \bar{\mathrm{H}}$ collision, according to

$$\mathrm{H} + \bar{\mathrm{H}} \rightarrow e + e^+ + \text{decay products}. \tag{4}$$

This reaction constitutes an alternative loss of antihydrogen, and we found its rate to be of the same order of magnitude as the rate of loss through the rearrangement process (3) [26, 27].

5.1.3 Electron-positron annihilation in flight

When the two conjugated atoms approach each other, the leptons might in principle annihilate before the hadrons do. We have found that this is not the case. Even though the annihilation reaction constant for para-positronium is larger than that for protonium, the probability of $e^+ - e^-$ annihilation at any given interhadronic distance R is weighted by the hadronic probability density at that distance. Because of that, the $e^+ - e^-$ annihilation occurs mainly at $R \simeq 1$ whereas the hadrons annihilate basically at $R = 0$.

The cross section for the in flight annihilation of the $e - e^+$ pair according to

$$\mathrm{H} + \bar{\mathrm{H}} \rightarrow p + \bar{p} + \text{decay products} \tag{5}$$

has been found to be negligibly small and not important for the kinetics of a dilute gas at ultra low temperatures [31].

5.1.4 Radiative association

The radiative association, according to

$$\mathrm{H} + \bar{\mathrm{H}} \rightarrow \mathrm{H}\bar{\mathrm{H}} + h\nu \tag{6}$$

leads to the formation of a *novel meta-stable atom-antiatom molecule* [28]. The $H - \bar{H}$ molecule is formed during the slow collision of $\bar{H}$ and H atoms. The excess of the binding energy is radiated out *via* a single photon, electric dipole transition. Specifically, the formation process is a sum of transitions from the collisional continuum of $H + \bar{H}$ to the quasi-bound ro-vibrational states of the $H - \bar{H}$ molecule. The latter decays further to two different atoms: protonium (Pn) and positronium (Ps). This mode of decay competes with the intramolecular proton-antiproton annihilation. We have calculated the branching ratio for these two reactions and the lifetime of the $H - \bar{H}$ molecule [32].

The spectrum of photons emitted during formation ranges from infrared to hard X-rays. The soft photons are emitted during the transitions to the weakly bound states just below the $H + \bar{H}$ dissociation threshold, resembling rotational-vibrational states of an ordinary molecule. The X-rays come from transitions to the deeply bound states, where the antiproton and the proton orbit each other as in the ground state of the protonium.

For slow collisions the transitions occur from the continuum state of the $J = 0$ (spherical) symmetry to the bound state of the $J' = 1$ symmetry, and are induced by the electric dipole moment of the $H - \bar{H}$ system. This dipole moment is the expectation value of the dipole moment operator for all four particles with respect to the ground state leptonic wave function, $\mathbf{D} = \sum_i q_i \langle \mathbf{r}_i \rangle$, where $\mathbf{r}_i$ denote the positions of the particles, and q_i their charges.

It is interesting to note that a similar radiative association process is *not* possible for the two hydrogen atoms. On the symmetry grounds the dipole moment of the $H - H$ system (which is inversion symmetric) vanishes. In that case the nuclear dipole moment is identically 0 and the electronic dipole moments induced in the two approaching atoms have opposite orientations and cancel each other. For the $H - \bar{H}$ system (which lacks the inversion symmetry) the dipole moment (in the adiabatic and non-relativistic approximation) is finite. In that case the hadronic moment is $|e|R$ and the induced leptonic moments of H and $\bar{H}$ have the same orientations and add together to a non vanishing dipole moment (which tends to 0 in the limit of infinite separation R between the atoms).

In the association process discussed above the excess of binding energy is emitted as light, which provides a unique signature of the presence of antihydrogen in contact with hydrogen. If detectable, this mechanism might become a new diagnostic for probing the existence of antihydrogen in a sample of hydrogen gas, either in laboratory or in outer space.

5.1.5 Elastic scattering

The elastic scattering does not change the colliding fragments, but changes their collisional energies. This leads to equalization of the the collisional energies of the colliding fragments; a property that might be utilized for cooling the energetic atoms by contact with a colder atom cloud. We have computed the rate of elastic scattering and expressed its low energy limit in terms of a single characteristic quantity, the so called scattering length [26, 29]. By comparing the rates of elastic and inelastic scattering we were able to draw conclusions regarding the stability of hydrogen-antihydrogen mixtures and the collisional cooling of antihydrogen (see below).

5.1.6 Collisional cooling of antihydrogen

The *microscopic* cross sections for the processes mentioned above influence the *macroscopic* behaviour of hydrogen-antihydrogen mixtures under various conditions. This has implications on the prospects for trapping, cooling and detecting the antihydrogen. For instance, antihydrogen may have to interact with hydrogen because of the imperfect vacuum in the trap, or one may want to mix hydrogen with antihydrogen on purpose, in order to make possible an accurate comparison of the spectroscopic properties of the conjugated atoms under exactly the same conditions. An interesting question is whether it is possible to cool energetic antihydrogen *via* collisions within a sample of cold hydrogen. Such sympathetic cooling is based on equipartition of the collisional energy which is transferred from the hotter atoms to the surrounding bath. In the case of an hydrogen-antihydrogen mixture, the repetitive collisions not only cool the antihydrogen atoms but also lead to their depletion through annihilation. In particular, elastic scattering is responsible for cooling while the inelastic processes, particularly the rearrangement and the annihilation in flight, are responsible for losses of antihydrogen.

It is clear that cooling and loss processes compete and therefore the ratio of elastic to inelastic scattering becomes a crucial parameter dictating the survival times of anti-

hydrogen in contact with hydrogen. At very low energies, the rate for inelastic collisions behaves as $\lambda_{el} \sim \sqrt{T}$ and the rate for elastic collisions λ_{inel} is constant. We calculated the temperature at which these rates become equal [27, 29]. Above that temperature $\lambda_{el} > \lambda_{inel}$, i.e. cooling is faster then antihydrogen loss. Below this limiting temperature $\lambda_{inel} > \lambda_{el}$ and the annihilation resulting in the loss of antihydrogen dominates the cooling process. The limiting temperature is on the order of 50 mK. To validate these conclusions we are presently conducting an improved calculation incorporating the presence of the inelastic channels in the scattering process by means of the complex optical potential [30].

In order to see the time evolution of the temperature and density of the mixture of hydrogen and antihydrogen we have solved the rate equations describing the losses of kinetic energy and the density of antihydrogen. The solutions show how the size of σ_{el}, σ_{inel} and the ratio $\lambda_{inel}/\lambda_{el}$ determine the lowest temperature attainable in collisional cooling of antihydrogen, the cooling time, and the antihydrogen density loss during the cooling process. We have found that largely independently of the initial temperature, a sample of antihydrogen could be cooled to 0.05 - 0.1 K while loosing 90% of the atoms. We have also obtained estimates of the lifetime of antihydrogen stored under imperfect vacuum conditions, as well as the lifetime for a mixture of equal $H - \bar{H}$ concentrations suitable for spectroscopic comparisons of hydrogen and antihydrogen.

6 Epilogue

I feel privileged to be given this opportunity to pay tribute to Per-Olov Löwdin. His premature demise is a great loss for international science, but most of all for those of us who knew Per-Olov personally. Per-Olov is dearly missed not only by his family but also by the greater scientific community. He acted as a god-parent for his close collaborators, most certainly for me and my family. He was a born leader, both at work and in leisure activities. He liked to work hard and play hard. I will always remember our joint hikes in the Swedish and Norwegian mountains, the soccer games, excursions to St. Augustine, the ocean beaches and springs of Florida, and the welcoming atmosphere of Per's and Karin's home. Per-Olov was always in the forefront of our work and recreation. He liked to achieve and share with others. For me he was not only a teacher of science but also a teacher of life, a person whom I will always admire.

Per-Olov was not only an outstanding scholar who made lasting contributions to science, but also an exceptionally warm person, full of natural kindness and goodwill. We cherish his scientific achievements as well as his memory.

References

[1] P.A.M. Dirac. *Proc. R. Soc. (London) A*, 126:360, 1930.

[2] P.A.M. Dirac. *Proc. Roy. Soc. A*, 133:61, 1931.

[3] ed. R.J. Weiss. *The discovery of Antimatter.* World Scientific, 1999.

[4] H. Alfvén and O. Klein. *Arkiv f. Fysik*, 23:187, 1962.

[5] G. Baur *et al. Phys. Lett. B*, 368:251, 1996.

[6] D.G. Christian *et al. Bull. Am. Phys. Soc.*, 42:951, 1997.

[7] S. Dimopoulos and L. Susskind. *Phys. Rev. D*, 18:4500, 1978.

[8] T.D. Lee and C.N. Yang. *Phys. Rev.*, 104:254, 1956.

[9] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, and R.P. Hudson. *Phys. Rev.*, 105:1413, 1957.

[10] J. Schwinger. *Phys. Rev. A.*, 91:713, 1953.

[11] G. Lüders. *Kong. Danske Vidensk. Selsk. Mat.- Fys. Medd.*, 28:nr. 5, 1954.

[12] W. Pauli. *Niels Bohr and the Development of Physics, ed. W. Pauli.* Pergamon Press, New York, 1955.

[13] J.S. Bell. *Proc. Roy. Soc. A*, 231:479, 1955.

[14] J.P. Christenson, J.W. Cronin, V.L. Fitch, and R. Turlay. *Phys. Rev. Lett.*, 13:138, 1964.

[15] A. Abashian, R.J. Abrams, D.W. Carpenter, G.P. Fisher, and B.M.K. Nefkens. *Phys. Rev. Lett.*, 13:243, 1964.

[16] A.D. Sakharov. *JETP Letters*, 5:24, 1967.

[17] F. Iazzi. *Nucl. Phys. A*, 655:371c, 1999.

[18] A.G. Cohen, D.B. Kaplan, and A.E. Nelson. *Annu. Rev. Nucl. Part. Sci.*, 43:27, 1993.

[19] D.G. Fried *et al. Phys. Rev. Lett.*, 81:3811, 1998.

[20] T. Udem *et al. Phys. Rev. Lett.*, 79:2646, 1997.

[21] C.L. Cesar *et al. Phys. Rev. Lett.*, 77:255, 1996.

[22] R. Carosi *et al. Phys. Lett. B*, 237:303, 1990.

[23] R. Bluhm, V.A. Kostelecky, and N. Russell. *Phys. Rev. Lett.*, 82:2254, 1999.

[24] A. Einstein. *Jahrb. Rad. Elektr.*, 4:411, 1907.

[25] W.E. Thirring. *Ann. Phys.*, 16:96, 1961.

[26] P. Froelich, S. Jonsell, A. Saenz, B. Zygelman, and A. Dalgarno. *Hydrogen - Antihydrogen Collisions. Phys. Rev. Lett.*, 84:4577, 2000.

[27] S. Jonsell, A. Saenz, and P. Froelich. *Low Energy Hydrogen - Antihydrogen Collisions. Nuclear Physics A*, 663:959c, 2000.

[28] B. Zygelman, P. Froelich, A. Saenz, S. Jonsell, and A. Dalgarno. *Radiative Association of Atomic Hydrogen with Antihydrogen at Sub-Kelvin Temperatures. Phys. Rev. A*, 63:052722, 2001.

[29] S. Jonsell, A. Saenz, P. Froelich, A. Dalgarno, and B. Zygelman. *Stability of Hydrogen - Antihydrogen Mixtures at Low Energies. Phys. Rev. A*, 64:052712, 2001.

[30] B. Zygelman, P. Froelich, A. Saenz, S. Jonsell, and A. Dalgarno. *Optical Potential for Hydrogen - Antihydrogen Collisions at Cold Temperatures - preprint.*

[31] S. Eriksson, A. Saenz, S. Jonsell, and P. Froelich. *Leptonic Annihilation during Hydrogen - Antihydrogen Collisions - preprint.*

[32] S. Eriksson, A. Saenz, S. Jonsell, and P. Froelich. *The lifetime of the metastable hydrogen - antihydrogen molecule - preprint.*

[33] R.K. Sachs and A.M. Wolfe. *The Astrophysical Journal*, 147:73, 1967.

[34] A.G. Cohen, A. de Rujula, and S.L. Glashow. *Astrophys. J.*, 495:539, 1998.

Quantum-Chemical Studies of Metal Oxides for Photoelectrochemical Applications

P. Persson[a], R. Bergström[a], L. Ojamäe[b], and S. Lunell[a]

[a] *Department of Quantum Chemistry, Uppsala University, Box 518, SE-751 20 Uppsala, Sweden*

[b] *Department of Chemistry, IFM, Linköping University, SE-581 83 Linköping, Sweden*

Abstract

A review of recent research, as well as new results, are presented on transition metal oxide clusters, surfaces, and crystals. Quantum-chemical calculations of clusters of first row transition metal oxides have been made to evaluate the accuracy of *ab initio* and density functional calculations. Adsorbates on metal oxide surfaces have been studied with both *ab initio* and semi-empirical methods, and results are presented for the bonding and electronic interactions of large organic adsorbates, e.g. aromatic molecules, on TiO_2 and ZnO. Defects and intercalation, notably of H, Li, and Na in TiO_2 have been investigated theoretically. Comparisons with experiments are made throughout to validate the calculations. Finally, the role of quantum-chemical calculations in the study of metal oxide based photoelectrochemical devices, such as dye-sensitized solar cells and electrochromic displays, is discussed.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 41

0065-3276/02 $35.00

Contents

1 Introduction

Binary compounds formed between metals and group 6 or group 7 elements usually occur in the form of ionic crystals rather than as isolated molecules. The most typical example is, of course, given by the alkali halides, studied by Löwdin in his classic treatise from 1948 [1]. Another important class of ionic crystals, with somewhat different properties, are the metal oxides, which play a central role in many contexts in chemistry and physics. To mention only one example, their catalytic properties have long been recognized and subject to extensive study, and have given rise to numerous applications of enormous practical importance.

We have for some time been interested in the use of first row transition metal oxides, primarily TiO_2 and to a certain extent ZnO, in electrochemical applications. This includes, above all, their use as the backbone of dye-sensitized photoelectric cells, so called Grätzel cells. Such cells using nanocrystalline-nanoporous TiO_2, photosensitized with Ru-bipyridyl complexes, as electrodes have shown strikingly high photovoltaic conversion efficiencies [2, 3]. Investigations have also revealed efficient intercalation of Li into nanostructured TiO_2 which can be used for applications in the field of batteries (energy densities up to five times those of conventional lead batteries have been reached, [4]) and electrochromic devices with high coloration efficiencies and fast switching times [5]. Another approach for electrochromism is to attach so called viologens to the TiO_2 surface, different types of viologens giving different colours [6], which can be used in multicoloured displays.

The underlying motivation of the work presented in this paper is to provide a theoretical understanding of basic physical and chemical properties and processes of relevance in photoelectrochemical devices based on nanostructured transition metal oxides. In this context, fundamental problems concerning the binding of adsorbed molecules to complex surfaces, electron transfer between adsorbate and solid, effects of intercalated ions and defects on electronic and geometric structure, etc., must be addressed, as well as methodological aspects, such as efficiency and reliability of different computational schemes, cluster models versus periodic ones, etc..

We will in the presentation below review some of our recent work in this area. Those results which have already been published elsewhere will only be briefly summarized, while previously unpublished material will be presented in greater detail. In particular, the material in Section 2, and parts of 3.2, and 4.1, 4.2, and 4.3 is essentially new. Some outlooks for the future, connected especially with the application of ultrafast laser techniques, are given in section 5.

2 Free Clusters

Small molecular systems containing transition metal atoms are interesting both in themselves and as model systems (see e.g. the recent reviews by Harrison [7] and Noguera [8]). They are also useful as (quite difficult) "trial systems" for theoretical methods.

In an early investigation TiO_2 (rutile) clusters were studied with the semiempirical intermediate neglect of differential overlap (INDO) model [9] which can be applied to very large clusters. Some early studies of titanium-oxygen clusters were also performed using the *ab initio* Hartree Fock method to investigate the structure and stability of these systems [10, 11]. These studies gave information about possible isomers of small titanium-oxygen clusters, and have later been refined and extended by more advanced methods (DFT) by other groups (see e.g. [12, 13, 14]).

In order to get more detailed information about, e.g., bond strengths and equilibrium geometries in transition metal systems it is necessary to include electron correlation. This can be done either by traditional *ab initio* quantum chemistry models, e.g., CI-methods and coupled cluster methods, or by density functional theory (DFT) based methods. Correlated *ab initio* methods are often computationally very demanding, especially in cases where multi-reference based treatments are needed. Also, the computational cost of these methods increases dramatically with the size of the system. This implies that they can only be applied to rather small systems.

We have earlier investigated the performance of a large number of density functional theory (DFT) methods when applied to small titanium/oxygen compounds [15]. The DFT methods were in general found to give good results even with a DZP basis set but, for e.g. the titanium dimer Ti_2 a larger basis set seemed to be necessary. Since we found the DFT methods promising and since they have the potential for being applied to larger systems than those accessible to advanced *ab initio* methods, we have extended these studies to investigate what kind of functionals are best suited for transition metal systems.

In this section we present some results from a larger, ongoing, study of DFT methods applied to transition metal systems. Here we consider only diatomic metal oxides.

3d transition metal monoxides are much studied. Some of them are important in astrophysics due to their presence in stellar atmospheres. They are also interesting model systems for investigating metal-oxygen bonds, of importance, e.g., for many metal surfaces. The spectroscopy of the 3d transition metal oxides have been reviewed by Merer [16].

The bonding in these systems is rather complicated and consists of a mix-

ture of ionic and covalent contributions. For a detailed discussion about the bonding the reader is referred to refs [17, 18, 16]. The early oxides ScO – VO have the largest bond strengths in the series (>6 eV) and they have significant triple bond character. The following oxides, CrO – NiO, have double bond character (bond strengths between 3.8 and 4.8eV), and CuO has basically a single bond (binding energy 2.9eV) [17].

There are many computational investigations of transition metal oxides, see, e.g., the recent review by Harrison [7]. Some studies have included the whole sequence of 3d metal oxides. In one of these studies, Bauschlicher and Maitre [17], employed different high-level *ab initio* methods. It was found that ScO – MnO and CuO were well described by single-reference based techniques and that the CCSD(T) method gave spectroscopic constants (r_e, ω_e and D_0) in good agreement with experiments. For FeO – NiO multi-reference based techniques (CASSCF/ICACPF) were necessary to get good results.

Piechota and Suffczynski [19] studied the systems using DFT with the local-spin-density (LSD) approximation with non-local gradient corrections to the exchange-correlation energy added in a perturbative way; the gradient terms were evaluated from a density generated by LSD calculations. This gave results in fairly good agreement with experiment for bond lengths, vibrational frequencies, ionization potentials and dipole moments but not for binding energies, which were drastically overestimated.

A more sophisticated DFT study by Bridgeman and Rothery has also been published [18]. They used a functional consisting of a LSD part with non-local corrections from the Becke 1988 functional for exchange and the Perdew 1986 functional for correlation. The results were slightly better than those of Piechota and Suffczynski but the dissociation energies were still systematically overestimated.

2.1 Methods

In this study of the 3d metal oxides (ScO – ZnO) the popular hybrid functional B3LYP was used. The Gaussian 94 [20] and Gaussian 98 [21] programs were used. In these, the B3LYP exchange-correlation functional is of the form [22]:

$$E_{B3LYP} = 0.8E_x^{Slater} + 0.2E_x^{HF} + 0.72E_x^{Becke} + 0.81E_c^{LYP} + 0.19E_c^{VWN}$$

Basis sets [23]: For the Sc-Cu metal atoms the Ahlrichs VTZ basis sets [24] were used with extra polarization and diffuse functions (two p-functions, one d-function) from the Wachters basis set [25] and (one f-function) from ref. [26]. For Zn the Ahlrichs VTZ basis set was used with addition of two p-functions, one d-function and one f-function [27]. For the oxygen atom Ahlrichs TZV basis set [28] was used with two polarization d-functions and one diffuse s and

Table 1: Atomic excitation energies and ionization potentials, IP (eV)

	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn
B3LYP $E(3d^n4s^2 \rightarrow 3d^{n+1}4s^1)$	0.94	0.23	-0.21	[a]	1.41	[b]	-0.10	-0.41	-1.80	
Exp[c]	1.44	0.81	0.24	-1.00	2.14	0.88	0.42	-0.03	-1.49	
"Exp"[c] (non-rel)	1.31	0.67	0.08	-1.21	1.94	0.61	0.12	-0.39	-1.92	
B3LYP IP	6.58	6.84	6.74[d]	7.03	7.53	[b]	7.71[d]	7.91	8.02	9.43
Exp IP[e]	6.56	6.82	6.74	6.77	7.43	7.90	7.86	7.64	7.73	9.39
"Exp" IP (non-rel)[f]	6.54	6.78	6.49	6.68	7.38	7.84	7.41	7.49	7.58	9.23

[a] The excited state Cr atom ($3d^44s^2$) was not found in the calculations. With multiplicity 5 the resulting Cr-state is highly spin-contaminated and very far from the correct valence electron occupation numbers. [b] For the Fe atom the lowest energy state was found to be an impure state with a 4s-occupation of 1.81 and 3d-occupation of 6.19. The energy difference between this state and the $3d^74s^1$ was 0.28eV. The ionization potential calculated for ionization from the neutral state to the Fe^+ ($3d^64s^1$) is 8.00 eV. [c] Experimental values taken from ref. [7]. [d] Calculated using the B3LYP ground state for the neutral atom. [e] As quoted in [35]. [f] Adjusted experimental value using computed relativistic and non-relativistic HF ionization energies in ref. [36].

p function from the aug-cc-pVTZ basis set [29]. The resulting basis sets have the contraction schemes [62111111/3311111/3111/3] for the metal atoms and [621111/4111/11] for oxygen.

2.2 Results

Bonding in transition metal compounds may involve different electronic states of the metal atoms. It is important to know if a given theoretical method can describe the energy differences between different states of the atoms and, in cases where ionized states are involved, also the ionization potential of the atoms. If the atomic states are poorly described it may sometimes be advantageous to compensate for these errors when computing atomization energies. E.g., in the TiO molecule the Ti $4s^1 3d^3$ state is responsible for the bonding but the $4s^2 3d^2$ state is the ground state of the atom. If the method employed does not reproduce the energy difference between these states it may be advantageous to compute the atomization energy of the system as the difference in energy between the molecule and the excited state Ti atom and then use the experimental splitting of the Ti states rather than the calculated one [15].

2.2.1 Atomic excitation energies and ionization potentials

The calculated excitation energies and ionization potentials for the 3d transition metal atoms are given in Table 1 together with, on the one hand, measured experimental values, on the other hand, experimental values with estimated relativistic contributions to the energies removed. B3LYP clearly has problems with the atomic $3d^n 4s^2 \rightarrow 3d^{n+1} 4s^1$ excitation energies. In all cases B3LYP gives too low energies for the $3d^{n+1} 4s^1$ states compared to the $3d^n 4s^2$. For V and Co, B3LYP even predicts the wrong ground states. Part of the failure of B3LYP can be related to relativistic contributions to the excitation energies but even if these are removed the deviations of B3LYP from the non-relativistic "experimental" values are fairly large.

It should be noted that in some cases it was difficult to find reasonably "pure" states for the atoms with B3LYP. For Fe the B3LYP ground state was found to have a 4s-occupation of 1.81 and a 3d-occupation of 6.19. A state with 4s-occupation 2 could not be found.

The B3LYP ionization potentials were in better agreement with experimental values, but for Cr, Ni and Cu B3LYP overestimates IP by ca 0.3 eV. The calculated electron affinity (EA) of oxygen was 1.67 eV, which is 0.21 eV higher than the experimental value.

Table 2: Bond lengths, r_e (Å)

	ScO	TiO	VO	CrO	MnO	FeO	CoO	CoO	NiO	CuO	ZnO	ZnO
	$^2\Sigma^+$	$^3\Delta$	$^4\Sigma^-$	$^5\Pi$	$^6\Sigma^+$	$^5\Delta$	$^4\Delta$	$^4\Sigma^-$	$^3\Sigma^-$	$^2\Pi$	$^1\Sigma^+$	$^3\Pi$
B3LYP[a]	1.663	1.614	1.582	1.613	1.636	1.614	1.629	1.596	1.632	1.761	1.714[b]	1.886
LSD + BP[c]	1.68	1.63	1.60	1.63	1.64	1.62	1.68	1.62	1.638	1.73	1.71	
ab initio[d]	1.680	1.628	1.602	1.634	1.646	1.609	1.621		1.626	1.771	1.719	1.857
Exp[e]	1.668	1.620	1.589	1.615	1.646	1.616	1.629		1.627	1.724		
Relativ. contrib.[f]	-0.0005	-0.002	0	-0.007	-0.011					-0.034		

[a] This study. [b] Using spin restricted B3LYP. Unrestricted B3LYP (UB3LYP) gives a longer bond length: $r_e = 1.741$ Å. [c] Pure DFT results (VWN local spin density approximation + non-local corrections (BP) from refs. [37, 18]. [d] CCSD(T) [ScO-MnO,CuO] and CASSCF/ICACPF [FeO-NiO] results Ref. [17]. CCSD(T) results for ZnO from ref. [31]. [e] Taken from ref. [7]. [f] Estimate taken from ref. [17].

Table 3: Harmonic vibrational frequencies, $\omega_e(cm^{-1})$

	ScO	TiO	VO	CrO	MnO	FeO	CoO	NiO	CuO	ZnO	ZnO
	$^2\Sigma^+$	$^3\Delta$	$^4\Sigma^-$	$^5\Pi$	$^6\Sigma^+$	$^5\Delta$	$^4\Delta$	$^3\Sigma^-$	$^2\Pi$	$^1\Sigma^+$	$^3\Pi$
B3LYP[a]	1003	1047	1045	884	874	900	885	858	614	746	501
ab initio[b]	971	1014	1028	888	794	885	909	850	572	727	567
Exp[c]	965[d]	1009	1011	898	840	880	853	838	640	805±40[e]	

[a] This study. [b] CCSD(T) [ScO-MnO,CuO] and CASSCF/ICACPF [FeO-NiO] results Ref. [17]. CCSD(T) results for ZnO from ref. [31]. [c] Taken from ref. [7] unless otherwise noted. [d] $\Delta G_{\frac{1}{2}}$-value. Taken from ref. [16]. [e] Taken from ref. [38].

Table 4: Dissociation energies, D_0 (eV) for the neutral metal oxides

	ScO	TiO	VO	CrO	MnO	FeO	CoO	NiO	CuO	ZnO	ZnO
	$^2\Sigma^+$	$^3\Delta$	$^4\Sigma^-$	$^5\Pi$	$^6\Sigma^+$	$^5\Delta$	$^4\Delta$	$^3\Sigma^-$	$^2\Pi$	$^1\Sigma^+$	$^3\Pi$
B3LYP D_0	6.91	6.90	6.34[a]	4.36	4.01	4.29[b]	3.84[c]	3.60	2.74	1.20[d]	1.32
B3LYP BSSE	0.015	0.015	0.017	0.031	0.017		0.019	0.022	0.054	0.027	0.021
B3LYP D_0 (BSSE-corr)	6.89	6.88	6.32[a]	4.33	4.00	(4.27)[b]	3.82[c]	3.57	2.68	1.17	1.29
LSD + BP[e]	7.45	8.05	7.50	5.01	5.23	4.95	4.97		3.11	1.58	
ab initio[f]	6.90	6.84	6.26	4.30	3.56	3.65	3.64	3.75	2.66	1.63	1.38
Relativ. contrib. to D_0[g]	-0.04	-0.05	+0.03	+0.22	-0.03				+0.08		
Exp[h]	7.01 ± 0.12	6.87 ± 0.07	6.44 ± 0.20	4.77 ± 0.09[i]	3.83 ± 0.08	4.17 ± 0.08	3.94 ± 0.14	3.87 ± 0.03[j]	2.94 ± 0.12[k]	1.61 ± 0.04[l]	

[a] Dissociation energy for VO calculated as dissociation to B3LYP ground state V ($3d^4 4s^1$). If calculated towards the experimental ground state ($3d^3 4s^2$) D_0 (BSSE-corrected)=6.53 eV. [b] Dissociation energy for FeO computed for dissociation to the B3LYP stable Fe atom, which is an "impure" state ($3d^{6.19} 4s^{1.81}$). If we instead dissociate to the excited state Fe atom ($3d^7 4s^1$) and subtract the experimental excitaion energy for the atom we get D_0=3.81 eV. The BSSE for FeO was *estimated* to be ca 0.02 eV. [c] Dissociation energy for CoO calculated as dissociation to B3LYP ground state Co ($3d^8 4s^1$). If calculated towards the experimental ground state ($3d^7 4s^2$) D_0=3.94 eV. [d]Calculated using spin restricted B3LYP for ZnO. Unrestricted B3LYP gives a slightly lower energy for ZnO, leading to an increase of D_0 by ca 0.04 eV. [e] Pure DFT results (VWN local spin density approximation + non-local corrections (BP) from refs. [37, 18]. [f] CCSD(T) [ScO-MnO,CuO] and CASSCF/ICACPF [FeO-NiO] results from Ref. [17]. CCSD(T) results for ZnO from ref. [31]. [g] Relativistic contribution to the binding energy estimated in the study by Bauschlicher and Maitre [17] using first order perturbation theory. [h] Taken from ref. [7] unless otherwise noted. [i] Kang and Beauchamp [39] as quoted by Espelid and Borve in [40]. [j] Watson et. al [41]. [k] Rodgers, Walker, and Armentrout [42]. [l] Clemmer, Dalleska, and Armentrout [43].

2.2.2 Bond lengths

The calculated equilibrium bond lengths are given in Table 2. For comparison the results of the studies of Bridgeman [18] and Bauschlicher and Maitre [17] are also given, as well as experimental bond lengths.

The B3LYP calculations give bond lengths in very good agreement with the experimental values, in some cases even better than the *ab initio* calculations by Bauschlicher and Maitre. The B3LYP bond lengths are within 0.01 Å of the experiments for all the 3d transition metal oxides except CuO. For CuO the B3LYP bond length is 1.761 Å, an overestimation by 0.037Å. The large deviation for CuO is probably mainly due to relativistic "effects". Bauschlicher and Maitre estimated the effect of relativity on the bond length of CuO to be a shortening of the bond length by ca 0.034Å.

2.2.3 Vibrational frequencies

In Table 3 the computed and experimental harmonic frequencies are given. In general the B3LYP frequencies are within 4% of the experimental values. In all cases, except CrO and CuO, B3LYP overestimates the frequencies. The B3LYP results are of similar accuracy as the *ab initio* results of Bauschlicher and Maitre. The average absolute deviations from the experimental values are 3.2% in both studies. A noteworthy difference is seen for MnO, where B3LYP overestimates the vibration frequency by 34 cm^{-1} (4%) and the *ab initio* CCSD(T) method underestimates it by 46 cm^{-1} (5%).

2.2.4 Dissociation energies

The dissociation energies, D_0, of the neutral metal oxide molecules were calculated as the energy difference between the B3LYP ground state of the molecule and the B3LYP ground state of the atoms. The computed zero-point vibrational energies were included in the calculations of D_0. Basis set superposition errors (BSSE) were estimated using the counterpoise method [30]. The BSSEs were relatively small for most of the oxides. The largest BSSE was computed for CuO: 0.054 eV.

In Table 4 the calculated D_0-values are presented together with experimental values and results from other calculations. B3LYP gives results that are very similar to the results of the *ab initio* study by Bauschlicher and Maitre [17], with the exception of MnO and FeO. In general the B3LYP results are very good, but D_0 for CrO and NiO are significantly underestimated. At least for CrO this could be related to a fairly large relativistic contribution to the binding energy. Bauschlicher and Maitre calculated a relativistic contribution of 0.22eV for CrO.

The B3LYP dissociation energies are generally in much better agreement with experiment than the pure DFT results of Bridgeman and Rothery [18]. The exception is ZnO, where B3LYP fails, with a considerable underestimation of the bond strength, while the pure DFT method gives a dissociation energy in excellent agreement with experiment. Bauschlicher and Partridge [31] have studied ZnO and they found the ground state to be $^1\Sigma^+$. B3LYP, however, places the $^3\Pi$ state lower in energy. Our results for ZnO are practically identical to the B3LYP results of Bauschlicher and Partridge. The incorrect ordering of the states by B3LYP could be related to the fact that the $^1\Sigma^+$ state is not as well described by a single configuration as the $^3\Pi$ state. Using spin unrestricted B3LYP (UB3LYP) lowers the energy of the singlet ZnO, but only by ca 0.04eV.

2.2.5 Dipole moments

The calculation of accurate dipole moments of transition metal oxides is very difficult (see, e.g., ref [17]). Our calculated dipole moments are given in Table 5. In most of the cases where experimental values are available B3LYP overestimates the dipole moments. The exception is ScO, which is underestimated by 0.55D. The largest deviation is for TiO (+0.7D). The LSD calculations of Piechota and Suffczynski [19] give dipole moments in better agreement with experiment, but for the two worst cases (ScO and TiO) the values are only marginally better than the B3LYP results.

2.2.6 Ionization potentials

The ionization potentials (IP) of the metal oxides were computed as the energy difference between the (calculated) ground state of the neutral system and the energy of the lowest lying ionized state. Zero-point energy corrections were made, although the differences in the ZPEs between the neutral and ionized molecules were in all cases very small (<0.015eV). The IPs are given in Table 6. The B3LYP IPs are in reasonable agreement with experimental values but at least for TiO, VO and CrO the IPs are overestimated by more than 0.2eV. For ZnO, B3LYP underestimates IP by 0.18eV, about the same deviation as that for D_0 of ZnO.

2.2.7 MO^+ dissociation energies

Several groups have investigated the transition metal oxide cations. B3LYP was used in at least two studies. Sodupe, Branchadell, Rosi and Bauschlicher used B3LYP and CCSD(T) methods [32]. Nakao, Hirao and Taketsugu [33]

used several different methods, including *ab initio* multireference-based methods and three different DFT methods. Our results using B3LYP are, naturally, quite similar to the already published values but since the basis sets differ it may be interesting to compare the results.

We present the dissociation energies of the oxide cations in Table 7. The basis sets used in this study are slightly larger than those used by Sodupe et al. The largest differences in dissociation energies are found for the early MO^+ ions. For ScO^+, TiO^+ and CrO^+ we get ca 0.1eV larger dissociation energies than Sodupe et al. This may be an indication that even the B3LYP method requires fairly large basis sets for transition metal systems, if accurate energies are wanted. The early transition metals seem to be extra demanding.

Nakao et al. used effective core potentials to describe the Ne core of the transition metal atoms and treated the $3s^2 3p^6 3d^m 4s^n$ electrons by a (8s7p6d1f)/[6s5p3d1f] basis set and the oxygen atom by a fairly large basis set: (11s6p3d2f)/[5s4p3d2f]. Comparing our results to those of Nakao et al. the differences are not very systematic. The largest difference in bond strength is seen for TiO^+, where Nakao et al. get 0.32 eV lower dissociation energy than our all-electron calculation. The differences are larger than 0.1eV also for ScO^+, MnO^+ and FeO^+. For CrO^+ Nakao et al. report a $^4\Sigma^-$ ground state, while our results, as well as those of Sodupe et al., give a $^4\Pi$ ground state.

2.3 Free cluster conclusions

In general, the B3LYP method gives results in reasonably good agreement with experimental bond lengths, binding energies, vibrational frequencies and ionization potentials for the 3d transition metal monoxides. The method is clearly competitive with traditional *ab initio* methods. However, the method is not perfect, which is not surprising considering the fact that the parameters used in the functional were optimized using a set of main- group molecules [34] and no transition metal compounds. More work is under way investigating the performance of B3LYP and other DFT functionals for a broader set of transition metal compounds.

Table 5: Dipole moments, μ(D)

	ScO	TiO	VO	CrO	MnO	FeO	CoO	CoO	NiO	CuO
	$^2\Sigma^+$	$^3\Delta$	$^4\Sigma^-$	$^5\Pi$	$^6\Sigma^+$	$^5\Delta$	$^4\Delta$	$^4\Sigma^-$	$^3\Sigma^-$	$^2\Pi$
B3LYP[a]	4.00	3.71	3.70	4.20	5.11	5.25	4.66	5.24	4.73	4.71
LSD[b]	4.06	3.69	3.43	3.92	4.83	4.59		4.25	3.81	4.30
ab initio[c]	3.91	3.52	3.60	3.89	4.99	4.17	3.46		3.91	5.11
Exp[d]	4.55	2.96	3.355	3.88		4.7				4.45
		$\pm$ 0.05	$\pm$0.014	$\pm$ 0.13		$\pm$ 0.2				$\pm$ 0.3

[a] This study. [b] Pure DFT results using the local spin density approximation from ref. [19]. [c] CCSD(T) [ScO-MnO,CuO] and CASSCF/ICACPF [FeO-NiO] results from ref. [17]. [d] Taken from ref. [7].

Table 6: Ionization potentials, IP (eV)

	ScO	TiO	VO	CrO	MnO	FeO	CoO	NiO	CuO	ZnO
	$^2\Sigma^+$	$^3\Delta$	$^4\Sigma^-$	$^5\Pi$	$^6\Sigma^+$	$^5\Delta$	$^4\Delta$	$^3\Sigma^-$	$^2\Pi$	($^3\Pi$)
B3LYP	6.56	7.04	7.48	8.22	8.83[b]	8.84	8.75	9.24	9.39	9.16
Exp[a]	6.6	6.8198	7.25	7.85	8.65	8.9	8.9	9.5	9.32	9.34
	± 0.3[c]	± 0.0007[d]	± 0.01	± 0.02	± 0.2[e]	± 0.16[e]	± 0.2[e]	± 0.2[e]	± 0.17[f]	± 0.02[g]

[a] Taken from ref. [7] unless otherwise noted. [b] The MnO^+ ion has rather high spin contamination $< S^2 >= 6.75$. [c] as quoted in [44]. [d] From ref [45]. [e] From ref [46]. [f] From ref. [42]. [g] From ref. [43].

Table 7: Oxide-cations dissociation energies, D_0 (eV)

	ScO^+	TiO^+	VO^+	CrO^+	CrO^+	MnO^+	FeO^+	CoO^+	NiO^+	CuO^+	ZnO^+
	$^1\Sigma^+$	$^2\Delta$	$^3\Sigma^-$	$^4\Pi$	$^4\Sigma^-$	$^5\Pi$	$^6\Sigma^+$	$^5\Delta$	$^4\Sigma^-$	$^3\Sigma^-$	$^2\Pi$
B3LYP D_0[a]	6.91	6.70	5.59	3.16	3.08	2.74[e]	3.46	2.82	2.28	1.38	1.57
B3LYP-Sodupe[b]	6.81	6.59	5.52	3.05	3.00	2.66	3.42	2.87	2.31	1.35	
B3LYP-ECP[c]	6.74	6.38	5.56		3.14	2.61	3.25	2.92	2.35	1.48	1.54
Exp[d]	7.14 ± 0.06	6.88 ± 0.07[f]	5.85 ± 0.16	3.72 ± 0.12[g]		2.95 ± 0.13[g]	3.47 ± 0.06	3.25 ± 0.05	2.74 ± 0.07[g]	1.35 ± 0.12[h]	1.67 ± 0.05[i]

[a] This study. Basis sets "Ahlrichs" (see text) metal: (14s11p6d3f)/[8s7p4d1f], oxygen: (12s7p2d)/[6s4p2d]. [b] Taken from [32]. Basis set "Wachters/Dunning" metal: (14s11p6d)/[8s6p4d], oxygen: (10s6p1d)/[5s3p1d]. [c] Taken from Nakao et al. [33]. Metals: Ne core described by effective core potential, "valence" basis set (8s7p6d1f)/[6s5p3d1f], oxygen: (aug-cc-pVTZ): (11s6p3d2f)/[5s4p3d2f] . [d] From ref [47] unless stated otherwise. [e] The MnO^+ ion has rather high spin contamination $< S^2 >= 6.75$. [f] From ref. [44]. [g] From ref. [35]. [h] from ref. [42]. [i] From ref. [43].

3 Surface-Adsorbate Interactions

Detailed knowledge about the interactions of molecular adsorbates with metal oxide surfaces is today available only for a limited number of systems. Most experimental and theoretical investigations that aim to determine the adsorption geometry or the interfacial electronic interactions have focussed on small adsorbates including atoms or small molecules like hydrogen, oxygen, or water [48, 49, 50]. Such investigations are important to understand e.g. heterogeneous catalysis and corrosion processes.

There are, however, several fields of current research in which a corresponding level of understanding would be of interest also for large molecular adsorbates. For example, adsorbate-substrate interactions are relevant in the general areas of biocompatibility [51] and chemical sensors [52]. The requirement of dye-sensitization of metal oxide semiconductors also makes this an important aspect of many molecular photovoltaic devices. In fact, a good interfacial contact between dye and substrate, characterized by long-term stability and intimate electric contact, is vital for the efficiency of e.g. the dye-sensitized solar cells which have been at the center of our attention for the last five years.

In order to realistically model complete dye-surface interfaces, however, the need to handle large system sizes involving hundreds of atoms, is an inevitable problem which has until now largely prevented the use of quantum-chemical calculations. For aromatic adsorbates, such as pyridine [53], specific aspects of the interaction with metal oxide surfaces can be studied in detail, but few studies have been published which give a complete account of the bonding interactions on a molecular level. For even larger adsorbates, such as the most common ruthenium dyes, the situation is even worse. Either compromises have to be made in terms of methodological sophistication, e.g. by resorting to semi-empirical methods, or if *ab initio* and density functional calculations are required, very time-consuming calculations must be performed.

The structural studies discussed here deal specifically with the interactions of molecules anchored with carboxylic acids to transition metal oxide surfaces. Bi-isonicotinic acid interacting with metal oxide surfaces combines interest in local carboxylic acid binding, and aromatic interaction with the surfaces. Electronic interactions have been investigated quantum-chemically for several aromatic molecules, including benzoic acid, bi-isonicotinic acid and catechol, all strongly anchored to TiO_2 substrates. Together, these systems represent a significant step towards studies of dye-sensitized metal oxide surfaces in photoelectrochemical devices.

3.1 Methods

In addition to the cluster approach discussed in section 2, it can sometimes be advantageous to use periodic calculations to model bulk and surface phenomena. Some of the methods which we have found useful are described briefly below.

3.1.1 Periodic *ab initio* Hartree-Fock calculations

The structures of the surfaces, the surface adsorption and the alkali-doped crystal and the atom diffusion path (cf. Section 4) were investigated by different quantum-chemical methods. We used foremost *ab initio* methodologies. The main computational tool utilized was the program CRYSTAL [54]. This program makes it possible to treat molecules and in particular crystalline solids and surfaces at an *ab initio* level of theory: for surfaces and solids the periodic boundary conditions are applied in 2 or 3 dimensions [55]. The familiar Gaussian basis sets can be used for systems ranging from crystals to isolated molecules, which enables systematic comparative studies of chemical properties in different forms of matter. In our studies, split-valence basis sets were used [56].

The calculations were performed employing either pure *ab initio* Hartree-Fock (HF) methods, or hybrid HF-DFT functionals, in particular B3LYP [22]. The hybrid functionals have several advantages. One is that they are commonly applied with great success in computational studies of molecules and clusters, thus making it possible to benefit from the gathered experience from molecular studies. Another is their recently noted ability to accurately model band gaps in semiconductor compounds [57].

3.1.2 Plane-wave density functional calculations

Also pure density-functional methods combined with plane-wave basis sets and ultrasoft pseudopotentials [58] were used in our studies of extended systems [59]. The computational efficiency of these methods enables larger systems and to some extent dynamical processes to be studied. Generalized-gradient approximation (GGA) or spin-polarized GGA DFT functionals [60, 61] were employed in the electronic structure calculations.

Some calculations involving large organic adsorbates on metal oxide surfaces were made with the CPMD program [62].

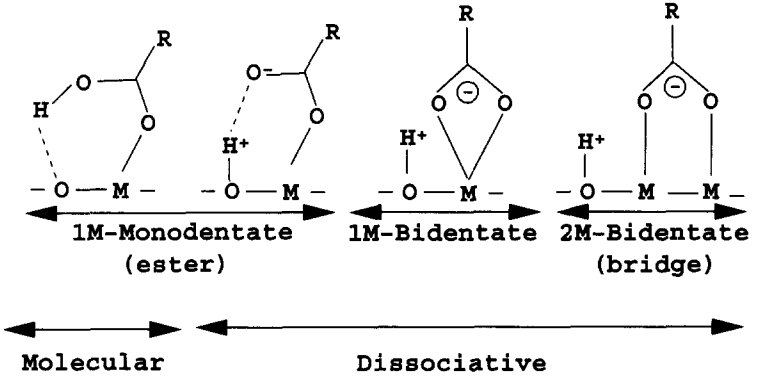


Figure 1: Common binding modes of carboxylic acids on metal oxide surfaces.

3.1.3 Periodic semiempirical calculations

Semiempirical methodologies make it possible to study even larger systems compared to the DFT methods, albeit with some sacrifice of chemical accuracy. In the present semiempirical calculations a version of the intermediate neglect of differential overlap (INDO) method, modified for crystal calculations [63, 64, 65], was used. This program was initially parameterized for solid state applications involving ionic crystals, and we have used it extensively in the past to model intercalation phenomena [65, 66, 67]. We have also developed a set of parameters for the most common atoms in organic molecules, including carbon and nitrogen. This has made it possible to study large aromatic adsorbates on TiO_2 surfaces [68].

3.2 Adsorption geometry

Direct bonds between substrate and adsorbate are loosely divided into weak physical adsorption (physisorption), and stronger chemical bonding (chemisorption). We are here focusing on chemisorption cases, where strong substrate-adsorbate interactions make it reasonable to first consider the direct interactions between the adsorbate and the substrate. In physisorption, this interaction is likely to be competing with the interactions between neighbouring adsorbates which may be of similar strength.

There is a wide range of possibilities for reactions between organic adsorbates and metal oxide surfaces, and to date there are few general rules to rely on for predictions of what binding will prevail in new systems or under altered

	TiO_2 rutile (110)	TiO_2 anatase (101)	ZnO($10\bar{1}0$)
M – M	2.90	3.78	3.28
M – O	3.44	1.97, 3.87	1.87, 3.31
Binding	bridge[a]	monodentate/bridge[b]	bridge[c]

[a] From Bates et al. [71]. [b] From Vittadini et al. [73]. [c] Periodic HF [59].

Table 8: Bond lengths (in Å) between chemically active surface metal (M) and oxygen (O) atoms in different metal oxides. Large differences in M-M and M-O distances between the different surfaces may be an important factor for the relative stability of different adsorption modes which typically involve different surface atoms in the surface-adsorbate bonding. It can be noted that two binding structures were proposed for TiO_2 anatase (101), depending on environmental influences.

conditions. Some of the common binding modes for a carboxylic anchor group are presented in Fig. 1. It is clear, however, that the presence of both metal and oxygen atoms in a surface will make the surface reactivity different compared both to pure metal surfaces, and covalent solids such as graphite and silicon.

A particular feature of metal oxide surfaces is that, depending on the oxide stoichiometry and the specific surface, surface cell dimensions may be quite large. In TiO_2 rutile (110), for example, it is 6.49 Å between neighbouring 5-fold surface Ti atoms in the ($1\bar{1}0$) direction. For a small adsorbate which adsorbs preferentially on metal atoms, this may largely exclude direct adsorbate-adsorbate interactions in certain directions. A comparison of some atomic distances with proposed binding modes for formic acid on some examples of metal oxide surfaces are presented in Table 8.

A first step towards an understanding of the structure of dye-surface interfaces is to understand the interaction of the functional groups used to anchor the dyes to the substrates. In particular, carboxyl and phosphonate groups have been used to anchor dyes to TiO_2 and other nanostructured metal oxide materials. Both experimental and theoretical studies of organic adsorbates on metal oxide surfaces have appeared in recent years, and several of these have dealt with the adsorption of organic molecules through a carboxyl group, see for example [69, 70, 71, 72, 73]. In addition to their use as anchoring group in dye-sensitized solar cells, carboxylic acids are believed to be involved in heterogeneous catalysis [74]. This is a double-edged situation, where carboxylic acids are investigated both because of their reactivity, and because of their stability.

3.2.1 Molecular and dissociative adsorption

Molecular adsorption is the conceptually simplest form of adsorption, with an essentially unchanged adsorbate binding to a surface. However, many stable adsorption geometries are formed only after bonds within the adsorbate have been broken. This is for example the case with the bridge-adsorbed carboxylates on both TiO_2 and ZnO. There may be significant barriers to such dissociative adsorption, and as in any chemical reaction, this could prevent the structure to form on kinetic grounds, even if it would be the most stable thermodynamically. From computational evidence, this does not appear to be a limiting factor for formic acid adsorption on ZnO $(10\bar{1}0)$ [56, 59]. The ease of proton transfer to the surface seems reasonable in view of the acidic nature of the adsorbate. Periodic *ab initio* calculations have now been done for formic acid adsorption on several of the surfaces of interest. In Table 8 key distances are listed together with suggested binding modes for ZnO$(10\bar{1}0)$, as well as TiO_2 rutile (110) and anatase (101).

3.2.2 Adsorbate induced surface relaxation

When the strength of the adsorbate-surface interactions is significant compared to the binding within the solid, the adsorbate is likely to induce changes in the surface structure [75]. In case of formic acid adsorption on the ZnO $(10\bar{1}0)$ surface (Fig. 2), this effect was found to be substantial, and the adsorption energy of the favoured bridge binding was stabilized by ca. 20 kcal/mol [59]. It appears that the adsorption process partly restores the full co-ordination of the surface atoms, which as a consequence relax outwards. Adsorption energies calculated for adsorbate induced relaxed surfaces are likely to be intermediate compared to calculations using a completely unrelaxed surface, and using a relaxed free surface. If these two extreme cases are known to give substantially different adsorption energies, the inclusion of adsorbate induced surface relaxation may be crucial to obtain reliable adsorption energies. It is also noteworthy that the change in adsorption energy due to this effect depends on the adsorbate, and the binding mode. The calculated change in adsorption energy for the monodentate binding of formic acid to ZnO was found to be only about half that of the bridge binding.

3.2.3 Adsorbate-adsorbate interactions

At high coverages, adsorbate molecules interact not only with the surface, but also with neighbouring adsorbate molecules. In the most obvious case this involves the formation of bonds between neighbouring adsorbates. It has, for

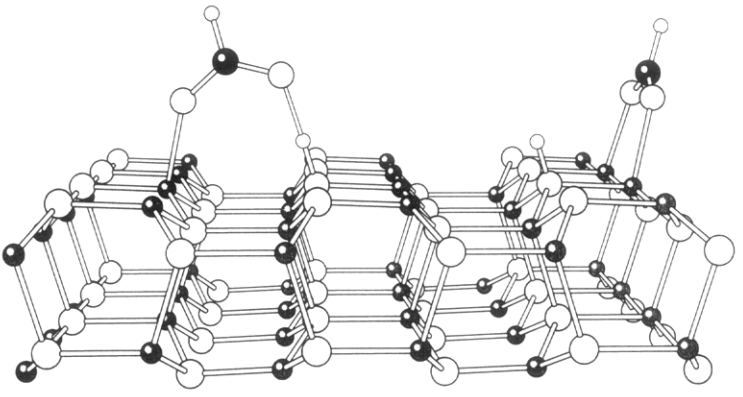


Figure 2: Adsorption of formic acid on ZnO($10\bar{1}0$) in the monodentate B (left), and bridge (right) modes.

example, been suggested that water forms a two-dimensional hydrogen bonded network on Mica surfaces at low temperatures [76].

However, also when the local adsorption geometries are not strongly affected, the adsorbate-adsorbate interactions may influence the stability of the adsorbate layer. This is the case with formic acid on ZnO($10\bar{1}0$) [56, 59], where the dissociated adsorbates are not directly bonding to neighbour adsorbates, but where large dipole interactions between the adsorbates substantially change the adsorption energy, depending on the coverage and overlayer structure. The electrostatic interactions seem to have a significant influence on the adsorption energies for coverages in the range of 0.1 – 1 monolayer.

The adsorbate dipole interactions are strongly direction dependent [59], and can act both stabilizing and de-stabilizing. An illustration is given in Fig. 3 for monodentate adsorption on ZnO ($10\bar{1}0$). This case is unusually clear as the adsorbate dipole points along the ($000\bar{1}$) principal surface direction. Differences of 20 kcal/mol for different packings contribute substantially to the calculated adsorption energies which are ca. 40 kcal/mol. With such directional adsorbate-adsorbate electrostatic interactions, it seems likely that the overlayer structures will be affected. This may be most important at intermediate coverages, when adsorbates are relatively unhindered to move on the surface, while still close enough to interact with other adsorbates in the vicinity. Given favourable experimental conditions relative to e.g. mobility barriers, one might hope to discover adsorbate strands along directions of favourable dipole-dipole interactions, and not just randomly shaped adsorbate islands.

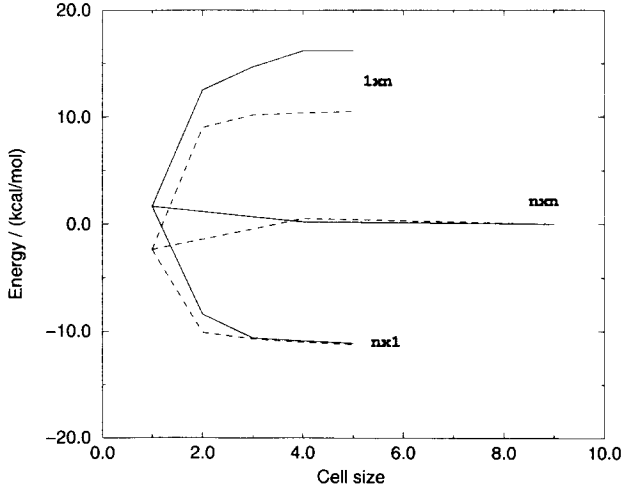


Figure 3: Adsorption energy of monodentate-B adsorbed formic acid on ZnO($10\bar{1}0$) as a function of coverage. Calculated values (solid lines) are compared to a simple electrostatic model (dashed line) based on the atomic charges. 1xn coverages refer to surface cells extended in the ($000\bar{1}$) direction, and nx1 to extensions in the ($11\bar{2}0$) direction. nxn cells have been extended in both the ($000\bar{1}$) and the ($11\bar{2}0$) directions.

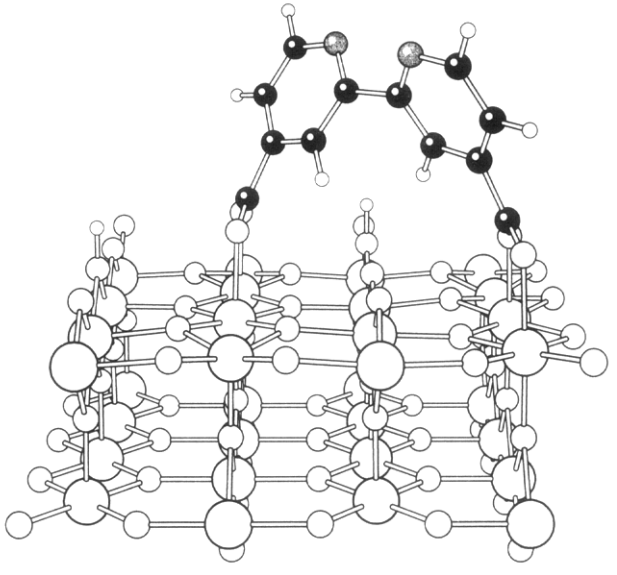


Figure 4: Structure of bi-isonicotinic acid adsorbed on rutile TiO_2 (110). The geometry was optimized with periodic density functional calculations.

3.2.4 Multiple adsorption sites

Large adsorbates, such as bi-isonicotinic acid, may bind to a surface at several sites which are sufficiently far apart not to interact strongly in a direct way. This kind of system is by necessity large and complex, and few detailed studies have been reported on such systems. Various structural aspects of bi-isonicotinic acid adsorption on rutile and anatase TiO_2 surfaces have been presented in several recent studies [68, 77, 78]. Bi-isonicotinic acid adsorption on TiO_2 surfaces is not only taken as a problem of direct interest to the photoelectrochemical applications, but also serves as a model system for surface science investigations of phenomena connected to the adsorption of large organic adsorbates on metal oxide surfaces.

Both carboxylic acid groups of the adsorbate want to bind to the surface in a fashion similar to that of formic acid, while the rest of the molecule is only weakly interacting with the surface. The constituent functional groups of the adsorbate are, however, rather rigid, and the flexibility of the bi-isonicotinic acid is mainly restricted to dihedral twists between the functional groups. The

property	semi-empirical[1]	DFT[2]	experimental[3]
O equivalence	Yes	Yes	Yes
Azimuthal Orientation	47	45	43
Vertical Orientation	17	27	25
R(Ti-O)	2.0	2.12	-
ring-carboxyl twist	25	36	-

[1] From refs. [68, 77]. [2] Car-Parrinello MD calculations [80]. [3] From [77].

Table 9: Comparison of structural parameters between calculations and experiment for bi-isonicotinic acid adsorbed on TiO_2 rutile (110).

limited flexibility of the adsorbate is found to severely limit the number of reasonable adsorption possibilities for the investigated TiO_2 surfaces. While this may be a cause for concern when designing effective anchoring ligands, it does facilitate the elucidation of the adsorption geometry significantly. It would have been difficult to study this large a system theoretically if the problem had not been naturally limited to a few strongly bonded adsorption possibilities. According to the comparison between adsorbed bi-isonicotinic acid and that of two formic acids using first principles calculations [80], the proposed double-bridge binding on TiO_2 shown in Fig. 4 only induces a moderate weakening of the surface binding by about 20% due to additional adsorbate strain.

Regardless of the twists between the functional groups, there remains a common orientational axis of the two pyridine rings in bi-isonicotinic acid. This makes it a good test case for a determination of the molecular adsorption geometry, which can be compared to experimental determinations of adsorbate orientation, e.g. using NEXAFS spectroscopy. Such experiments indicate a "diagonal upright" adsorption, with a horizontal orientation very close to 45° relative to the (001) direction of the (110) surface of rutile TiO_2. As shown in figure 5 (left) this would from simple geometrical arguments appear to be consistent with a "2 shift" adsorption, i.e. a binding conformation where one of the anchoring carboxylic groups is shifted two lattice units along the (001) direction relative to the other carboxylic anchor.

This is, however, an illustration of the importance of taking adsorbate deformations into account, and the usefulness of theoretical modelling for doing so. A computationally optimized "1 shifted" adsorption geometry is shown in figure 5 (right), and this has an adsorbate orientation very close to the desired 45°, and not the simplistically expected 60°. Instead, the deformed "2 shift" structure does not have a 45° orientation, but rather one of 35°.

Furthermore, the deformation required to optimize the binding makes each

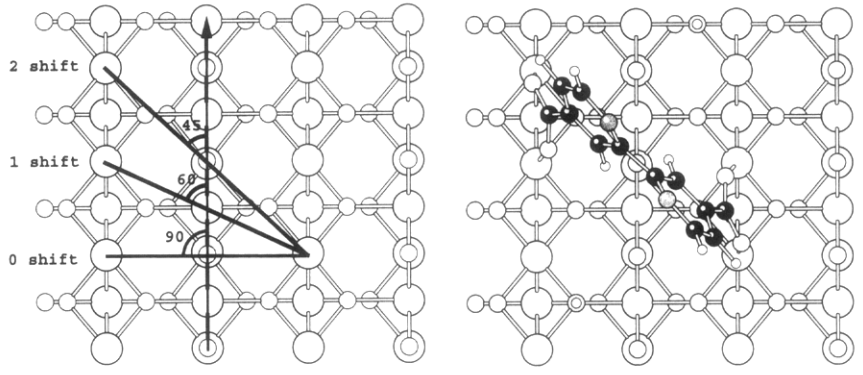


Figure 5: Bird's eye view of the orientation of bi-isonicotinic acid adsorbed on rutile TiO_2 (110). Simple substrate-anchoring based models (left). Calculated deformed 1-shifted geometry (right).

of the pyridine rings tilt relative to the surface normal, also in good agreement with the experimental findings. In the geometry shown in Figs. 4 and 5 the rings are twisted relative to each other, while the NEXAFS experiments cannot distinguish between an overall tilted geometry, and one involving a twist between the two rings. The combination of experimental and theoretical evidence directly is thus seen to be highly useful to elucidate bonding interactions [77, 79, 78]. Until recently, however, studies of as complicated systems like the one presented here could not have been carried out experimentally nor theoretically with a sufficient level of confidence. In Table 9 a comparison is presented for experimental and theoretical structural data on the adsorption. The agreement between the different methods is generally very good.

3.2.5 Sensitizer-substrate linking in solar cells

Having considered adsorption of organic compounds under controlled experimental conditions, it is desirable to make a connection to the binding in photoelectrochemical devices. The most interesting system in this respect is the binding of ruthenium dyes, such as N3, to nanostructured anatase TiO_2. The many uncertainties that exist about local oxide surface morphology, defect concentration, solvent effects etc. make this a far more complex problem to address in detail. It is even probable that there is a range of adsorption structures present in the cells. Some results which shed light also on these systems have appeared in the last years. The anatase (101) surface is espe-

cially important in this respect, as it is the most prominent surface in ordinary preparations of nanocrystalline TiO_2 [3].

The nature of the dye-surface linking has been a constant concern since the beginning of the development of the dye-sensitized solar cells. In 1977, Fujihira et al. [87] concluded that a rhodamine B sensitizer was ester-linked to the TiO_2 substrate. This assignment seems to have referred primarily to anchoring of the sensitizer via the carboxyl group, rather than through an amide-link. Gradually ester-linking came to refer specifically to a 1M-monodentate binding of the carboxyl group to the substrate, see e.g. [88], an assignment supported by some vibrational data [89, 90]. Such linking, in which only one of the carboxyl oxygens binds to the surface has, however, not usually been listed among the stronger binding possibilities according to the above described detailed studies on smaller carboxylic acid adsorbates on relevant metal oxide surfaces.

In the last few years IR studies [91, 92], as well as the combined XPS and computational study referred to above[79], have instead suggested that dyes bind to the surface via a couple of carboxyl groups in which both oxygens are linked to the surface. Often it is difficult experimentally to distinguish between bridge and bidentate binding, and neither IR, nor XPS gives information about the overall binding, only about the local anchoring by the carboxyl groups. From the calculations, the bridge structure seems to represent the stronger binding possibility compared to the bidentate binding, in spite of the considerable steric constraints. Questions regarding the relative positions of the anchoring groups on the surface can be also be considered in detail by means of calculations [81].

In addition to the interactions of individual dyes with the substrate, the ordering of dyes on the surface has been explored, starting from crystallographic data on the N3 structure and electron microscopic evidence of the ordering [93]. Mainly binding possibilities involving one carboxyl group from each of the two bi-isonicotinic acid ligands were considered. While the proposed structures certainly seem reasonable, our studies on the doubly anchored individual ligand show that such arrangements are possible, and also need to be explored. One recent development which contributes to make such single-anchoring-ligand binding interesting is the use of more highly differentiated dyes, in which one of the bi-isonicotinic acid ligands is replaced by a ligand with other desirable properties, such as for example hydrophobic tails replacing the carboxylic acids. This leaves just one bi-isonicotinic anchoring ligand per dye molecule, which would preferentially bind to the surface through both its carboxyl groups.

While some crucial molecular aspects of the surface-bonding of dyes have been elucidated in the last few years, several questions remain to be answered, not the least regarding cells made with alternative components. In any case,

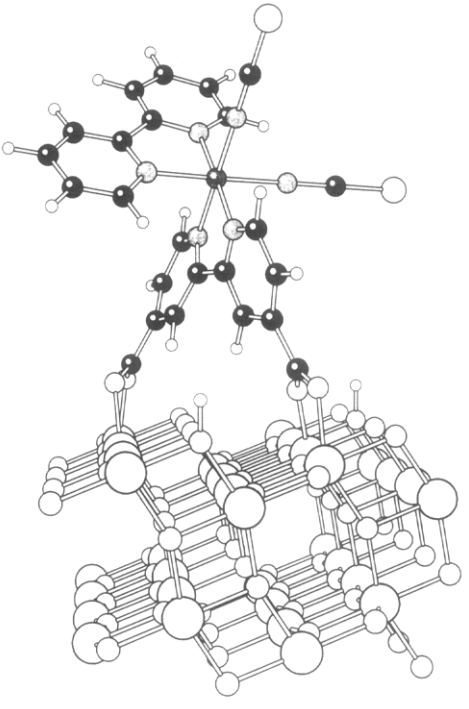


Figure 6: Molecular model of adsorption of $Ru(bpy)(bpdc)(ncs)_2$ on anatase TiO_2 (101). With only one bi-isonicotinic acid surface-anchoring group, functional groups can be added to the other bipyridine ligand to tailor the dye properties.

a collection of methods with which these questions can be systematically addressed now exist.

3.3 Interfacial electronic interactions

Electron transfer processes are at the heart of electrochemistry, and often the focus is on events at electrode surfaces. While the theory for electron-transfer in solution [94], and at metal surfaces [95] is rather extensive, a comprehensive theory for electron transfer at metal oxide-organic interfaces [96, 97] is still under development. This section is devoted to a discussion of some of the key elements of the surface electron transfer in dye-sensitized solar cells, illustrated by results from recent calculations.

It is often assumed that, to a first approximation, the electronic interactions between the surface and the adsorbate can be understood in terms of weak interactions between the electronic structures of the two components, which are assumed not to change significantly on adsorption compared to their separate properties. The first sections of this chapter therefore deal with relevant aspects of the electronic structure of the adsorbates and the metal oxides by themselves. In the final section, interfacial interactions are discussed.

3.3.1 Sensitizer electronic properties

Organic molecules are interesting for two reasons in the present context: firstly they provide a simplified starting point for detailed investigations compared to organometallic dyes, and secondly some organic species are interesting by themselves as redox, or photoactive species. Adsorbates belonging to both categories have been considered in the research presented here. Bi-isonicotinic acid is, for example, a good model for many popular ruthenium dyes, as it is the surface-binding ligand of the most efficient such dyes. Catechol, on the other hand, is known to shift the absorption threshold by itself, which has made it the subject of several experimental studies.

Triarylamine is a purely organic molecule which is interesting as a chromophore in e.g. display technology. The molecule can be switched between a reduced colourless, and an oxidized blue state. The sensitization to nanostructured TiO_2 electrodes provides the substantial surface area required for a strong coloration. It is, however, believed that the electron transfer involved in the redox reaction takes place mainly within the sensitizer layer, and does not involve the substrate. Instead, there is an eventual electrical contact between the back-contact and the sensitizer layer [98]. For a quantum chemical modelling of the system, the inclusion of the substrate is in this case not likely to be essential. For a molecule of this size, it is possible to apply standard quantum

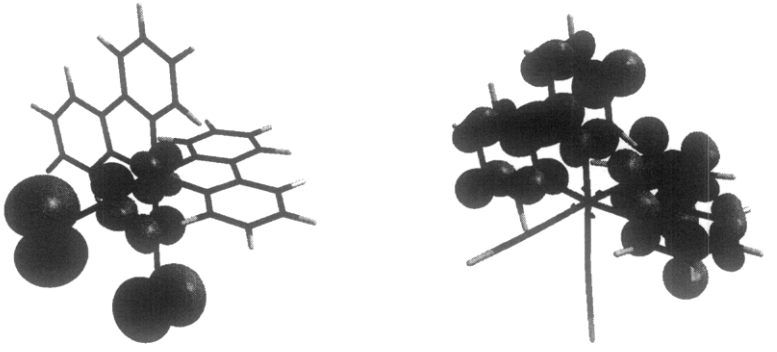


Figure 7: Calculated HOMO (left) and LUMO (right) molecular orbitals of $Ru(bpy)_2(NCS)_2$.

chemical techniques such as B3LYP DFT calculations using standard basis sets [99]. Both geometry optimizations and electronic structure calculations are interesting. The steric interactions between the rings promote a propeller structure, with the calculations suggesting that the difference in geometry is small between the reduced and the oxidized state. According to the calculations, the HOMO orbital of the reduced species, and the SOMO orbital of the oxidized species are both delocalized π orbitals. The calculated difference in total density, however, shows that oxidation essentially involves the removal of an electron from the central nitrogen.

The most efficient sensitizers to date for photoelectrochemical applications are organometallic ruthenium dyes. These compounds are coloured in solution, typically due to optical absorption through electronic excitations from the metal to the ligands, so called Metal to Ligand Charge Transfer (MLCT) excitations, see e.g. Fig. 7. The excitations in e.g. N3 are reasonably well understood, see e.g. [82], and the main light-absorbing excitations are from Ru(4d) levels to π orbitals on the pyridine ligands. Ruthenium is in its ground state in oxidation state II, i.e. with a formal charge of +2. When the dye is excited, this increases to oxidation state III, with a corresponding increase in the charge to +3.

3.3.2 Electronic structure of metal oxides

The electrodes used in the dye-sensitized solar cell are wide band-gap metal oxides. There is a filled valence band, which is separated from an empty conduction band, by a band-gap of several eV. Both for the standard TiO_2

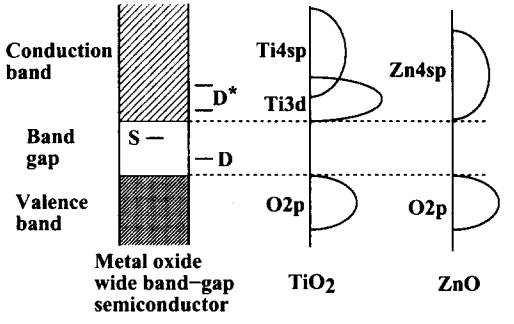


Figure 8: Metal oxide band structure. Comparison between schematic (left), TiO_2 (middle), and ZnO (right) DOS close to the band gap. S are surface states in the band gap. D and D* are the dye ground and excited states, respectively. Note that low excited dye states interact primarily with 3d levels in TiO_2, but with 4sp levels in ZnO.

electrode, and the alternative ZnO electrode [100], the bulk band gap is about 3 eV. Also, the band-gap is situated at almost the same energy for these two materials. An important difference between the two materials is, however, that Ti is an early transition metal, while Zn is a post-transition metal. This means that Ti has a nominally empty 3d-shell in TiO_2, while in ZnO the 3d-shell is full. The valence band in TiO_2 is O2p, and the conduction band Ti3d, overlapping at higher energies with the Ti4sp band (cf. Fig. 8). In ZnO, the valence band is also O2p, but the conduction band is Zn4sp. The Zn3d band is instead completely occupied, and lies below the valence band edge [83]. The interactions of low-lying excited states of the adsorbates with the lower part of the conduction band are therefore fundamentally different in TiO_2 and ZnO.

For very small metal-oxide clusters, the band-gap increases. This becomes a concern for clusters which are just a few nanometer in diameter. This effect is most readily understood in the quantum chemical model of a cluster, where the bands develop due to interactions between more and more atoms. For clusters with, say, up to 1000 atoms, the addition of new atoms to the cluster broadens the bands, and thus narrows the band-gaps. For large clusters this effect levels off, and for clusters with a diameter of more than a few nanometers band-widths and band-gaps stabilize at their bulk values.

3.3.3 Dye-surface interactions

When sensitizers are brought in physical contact with metal oxide semiconductors, electronic interactions become important. This has been studied using quantum-chemical calculations for some aromatic organic adsorbates: catechol [84], benzoic acid [84], and bi-isonicotinic acid [78].

As already mentioned, the electronic interactions involved at the metal oxide-adsorbate interface have not been studied nearly as extensively as, for example, metal surfaces. Some notable experimental progress has, however, taken place in the last few years, see e.g. [101, 102], and some relevant theoretical models have recently been proposed [103, 104, 105, 106, 107, 108]. However, little is known about the perhaps single most important factor determining the interaction: the electronic coupling strength between the excited adsorbate levels and the metal oxide conduction band.

Also, not much is known about the weakening of the coupling strength due to the anchoring group inserted in-between the aromatic dye ligands, and the metal oxide surface. It has been suggested that this significantly weakens the interaction for typical aromatic adsorbates, in particular as the lowest unoccupied carboxylate orbitals are found at high energies compared to the first unoccupied aromatic π levels [103, 110, 108]. It is possible to measure the injection time for electron-transfer processes in dye-sensitized solar cells using pump-probe spectroscopy [111, 112]. With increased time resolution, the upper bound for electron-transfer to the metal oxide conduction band has been pushed down to 20-50 fs [113, 114, 115, 116]. There is still some uncertainty as to the exact value of this fast injection, and some experiments also indicate the presence of more slowly decaying components in the picosecond regime [117].

Using resonant effects in core-level spectroscopic investigations of model chromophore adsorbates, such as bi-isonicotinic acid, on metal-oxide surfaces under UHV condition, even faster injection times have been tentatively proposed [85]. The injection time is observed to be comparable to the core-hole decay time of ca. 5 fs. It is also possible to resolve different injection times for different adsorbate electronic excited states with this technique. While the core-excitations themselves provide a perturbation to the system, and it cannot be ruled out that this influences the detailed interactions, the studies provide some of the first local molecular, state-specific injection time analysis with good temporal resolution in the low femtosecond regime. The results provide information about which factors determine the injection time on a molecular level.

Electronic effects involving a solid are conceptually and computationally more difficult than purely molecular systems, as the electronic bands provide a quasi-infinite number of electronic levels. It is therefore a natural first step to

try to build models from a molecular starting-point. These can then, hopefully, be generalized to take the electronic bands into account. In this respect, the highly successful Marcus theory for electron transfer in solution often serves as a first attempt for electron transfer models [94]. It is based on considerations concerning motion along a reaction co-ordinate together with a weak electronic coupling at a reactant-to-product crossing. It has successfully explained trends for electron transfer between solvated metal ions, and it is tempting to extend the concepts to more complicated cases, such as those encountered at electrode surfaces. With measured electron injection times in the low femtosecond region, and with strong anchoring of dyes to the metal oxide surfaces, this may not be an adequate starting point for a quantitative modelling. But even if the electron transfer initially is too fast for nuclear relaxations to be substantial, the change in the charge distribution due to the electron transfer will lead to a relaxation of the molecular structure. At solid-gas phase interfaces, this involves the specific relaxation of the adsorbate which has lost an electron to the surface. In wet cells, there will also be important solvent rearrangements. Computational modelling of these dynamic effects presents a major challenge at present, but progress can hopefully be made in the next few years.

It is also interesting to consider charge-transfer models developed primarily for metal surfaces. There are clear parallels to the metal oxide case in that there is an interaction between discrete molecular orbitals on one side, and electronic bands on the other side of the interface. The Newns-Anderson model [118] qualitatively accounts for the interactions between adsorbed atoms and metal surfaces. The model is based on resonance of adatom levels with a substrate band. In particular, the model considers an *energy shift* in the adatom level, as well as a *broadening* of that level. The width of the level is taken as a measure of the interaction strength with the substrate bands [118]. Also femtosecond electron dynamics have been studied at electrode interfaces, see e.g. [119]. It needs to be established, however, to what extent metal surface models are valid also for organic adsorbates on metal oxides in view of the differences between the metal an the metal oxide band structures. The significance of the band gap, as well as of surface states in it, must in any case be considered [102].

Ultimately an understanding of electron transfer processes in dye-sensitized solar cells must be expressed in terms of a model which takes the specific nature of metal oxide surfaces into account [97]. Moreover, the nanostructured devices often involve oxide nanoparticles which approach the limit where quantum-size effects become important. It would be a great step forward if this could be incorporated into an electron-transfer model.

It is important to establish the validity and accuracy of theoretical models that aims to quantitatively predict physical or chemical properties. The situation for metal oxide electrodes today is probably similar to that for metal

electrodes 20 years ago, when Bockris and Kahn concluded that "The situation is difficult in respect to calculations of adsorption, because of the lack of comparisons between experimental values and work in which the answers have *not* been obtained by substituting arbitrary parameters. The extension to electrochemical situations must begin with interpretations at a quantum mechanical level of simple phenomenon" [120].

3.3.4 Quantum-chemical models

Theoretical models of electron transfer based on a quantum mechanical description of the system are clearly desirable, in that they go beyond the phenomenological models, and stand on a firm theoretical basis. One can categorize the quantum methods according to whether they rely on a static or a time-dependent description of the transfer. For systems which are small enough to be treated dynamically, this is clearly attractive as one can directly follow the motion of the involved particles, be they electrons, atoms, or molecules. Where such an approach is not possible, much valuable information can instead often be obtained from time-independent quantum chemical calculations. In particular, information about the electronic coupling strength between a donor and an acceptor state can give a good indication about the ease and speed with which electron transfer can be expected to occur.

A second concern for quantum mechanical models of electron-transfer is the level at which the model is constructed. There is a wide-range of possibilities, ranging from Hückel and tight binding models which can be used for qualitative reasoning, to sophisticated *ab initio* methods. Likewise, time-dependent studies can be made with classical molecular dynamics (MD) simulations, or time-dependent quantum mechanical calculations.

Very recently, some models which explicitly deal with with surface electron transfer at semi-conductor or metal oxide electrode surfaces have been published [102, 103, 104, 106, 107, 108]. These models have the merit of being relatively general, and can thus explain experimentally observed trends. The success of these methods largely depends on the successful choice of parameters to describe the system. Some important parameters can be specified easily, such as the width of the TiO_2 conduction band. Other factors may be more difficult to predict *a priori*, including the electronic coupling strength between an excited sensitizer state and the TiO_2 conduction band. Either the parameters can be estimated in order to provide guesses of injection times, or the reverse process of fitting parameters to measured injection times, in order to estimate the parameters, can be tried.

We have explored the possibilities to make explicit quantum chemical calculations using methods which are known to provide near-quantitative spec-

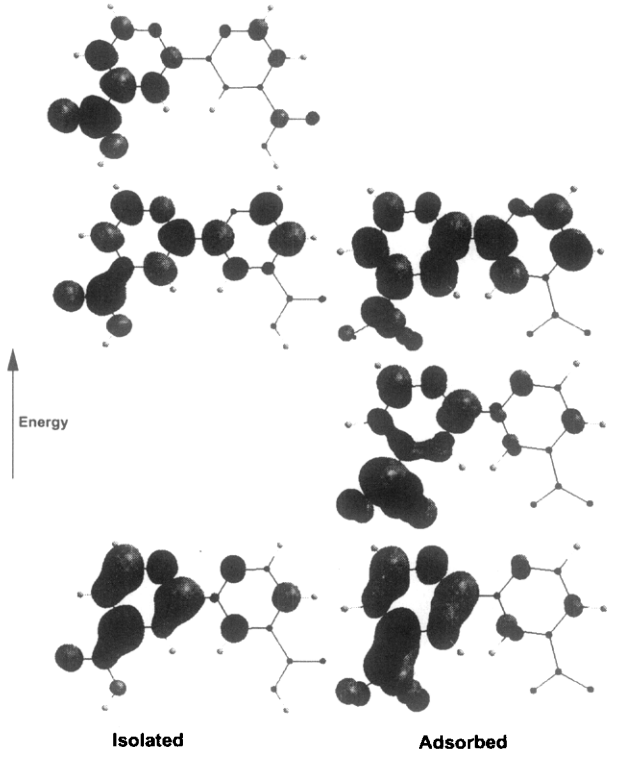


Figure 9: Isolated and adsorbed bi-isonicotinic acid bound unoccupied orbitals important in the N1s XAS excitations. The calculations show that the adsorbate unoccupied orbitals both change appearance, and shift significantly in energy on adsorption. This is in good agreement with experimental results [78].

troscopic accuracy. For example, the ZINDO method has been found to be capable of achieving high accuracy for spectroscopic phenomena of complex molecular systems at low computational cost [109]. In particular, it is important to clarify the role of the metal oxide band structure in these interfacial interactions. This means that calculations of essentially molecular systems, which are routinely used to calculate e.g. absorption properties of the dyes, here are of somewhat limited value. It is also sufficiently inexpensive to allow calculations of sensitized nanocrystals. It will be a challenge in the near future to make corresponding calculations with *ab initio* and density functional calculations.

Adsorbate level shifts

Bi-isonicotinic acid adsorption on the TiO_2 rutile (110) surface was presented above as a model for structural interactions between a large organic molecule and and a transition metal oxide surface. With such unusually detailed structural information at hand, it is realistic to proceed to questions concerning the electronic coupling. Evidence from XAS spectroscopy and INDO calculations were combined to show that the binding to the surface induces important changes to the electronic structure of the adsorbate [78]. This is shown graphically in Fig. 9. In particular, the carboxyl group interacts with surface Ti atoms, so that the positions of the unoccupied carboxyl orbitals change relative to the bipyridine π orbitals. Initially well above the lowest unoccupied bipyridine π orbitals, the first unoccupied carboxyl orbital of π character is brought down, and intermixed with the bipyridine π orbitals. The result is the appearance a new level in the XAS spectrum, almost 2 eV away from any observable peak for the free molecule. Only the lowest unoccupied bipyridine π orbital remains safely below the carboxyl component according to the presented calculations. This implies a strengthened coupling to the substrate, compared to what would be expected from a simple matching of unperturbed adsorbate orbitals with the substrate conduction band. With regards to the coupling in the dye-sensitized solar cells, this means that the coupling may be stronger than anticipated, with a corresponding decrease in the injection time. An exception is the lowest excited state, which is coupled more weakly to the substrate. This picture is consistent with the preliminary resonant core-spectroscopic evidence [85] which suggests that the electron transfer into the substrate from the second excited state and up proceeds at only a few femtoseconds, while the injection time of the first excited state is much longer.

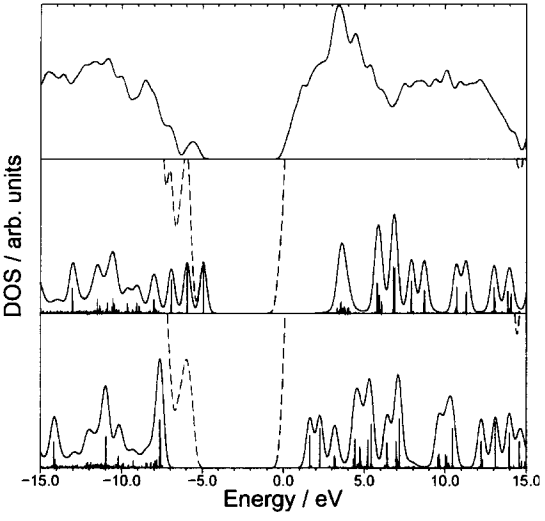


Figure 10: TiO_2 DOS (top), catechol PDOS (middle), and benzoic acid PDOS (bottom). The TiO_2 DOS is for a clean $Ti_{38}O_{76}$ cluster, while the PDOS plots were obtained from the same cluster sensitized by the respective molecules.

Electronic coupling strength

The surface electron transfer processes in which optical excitations lead to a transfer of an electron from a dye to a metal oxide electrode has traditionally been view as a Marcus type process in which the transfer proceeds via a well-defined intermediate excited state of the dye. The establishment of injection times of just a few femtoseconds for tightly adsorbed dyes requires this picture to be challenged. In the limit of strong coupling between dye and surface, one can no longer talk about a well-defined excited molecular state existing prior to the injection. Indeed, in some systems the excitation seems to promote the electron directly from a molecular state to a level in the substrate conduction band [122, 123]. Such processes can evidently proceed even in the absence of an excited molecular state that could be reached energetically. The strongly coupled processes are therefore best thought of in terms of Mulliken charge-transfer excitations [124, 125].

The possibility to calculate excited states of sensitized TiO_2 nanoclusters, to model photoelectrochemical properties, has been explored theoretically [84]. Catechol and benzoic acid were used as test cases, because they are examples

of the characteristic "anchored-aromatic" structures in these devices. Furthermore, both molecules have been investigated experimentally [122], and catechol has shown an unusual low-energy absorption threshold [122, 126, 127], while the electronic structure of benzoic acid is similar to that of the carboxylated pyridines in many Ruthenium dyes. There is a promising agreement between the calculations and the experiments for the absorption thresholds. Having confirmed the validity of the calculations, the results can be analyzed in greater detail to provide insight into the dye-surface coupling.

For the excited states of the catechol and benzoic acid sensitizers, the Partial Densities Of States (PDOS) provide graphic illustrations of varying coupling strengths between the adsorbate orbitals and the substrate bands, seen in Fig. 10. The coupling strength goes from weak (~meV) to strong (~eV), depending on which excited state is considered. Important factors influencing the strength appear to be the shape of the excited state, as well as the position of the excited state relative to the conduction band. The strongest energy broadening, of about 0.5 eV, is seen for the first excited catechol state, which in the calculations coincides with the conduction band maximum DOS. The catechol ring π orbitals are also close to the surface, as there is only an oxygen between the aromatic ring, and the surface. The weakest coupling is seen for the two lowest benzoic acid excited states which display essentially no spreading of the adsorbate PDOS. These orbitals are close to the bottom of the conduction band, and have relatively little intermixing with carboxyl π orbitals which lie significantly higher in energy. This is similar to the situation for bi-isonicotinic acid. It should be noted that there is a cluster size limited spectral resolution, which for the present cluster means that only coupling strengths down to ca. 25 meV can be determined with confidence.

An alternative method to calculate adsorbate-substrate electronic interactions accurately it to perform periodic surface calculations. By calculating band-structures, the inherent limitation of the cluster calculations with respect to the coupling strength is overcome, if a sufficiently thick slab can be used in the calculations. Such calculations are presently becoming manageable for aromatic adsorbates also with *ab initio* Hartree-Fock and density functional based methods. We have, for example, recently investigated the interactions of isonicotinic acid with ZnO and TiO_2 surfaces using periodic calculations [121] with the B3LYP hybrid functional which is well established in molecular calculations, and which also has been shown recently to describe semiconductor electronic structures well. Figures 11 and 12 show the molecular and electronic structures of isonicotinic acid on a TiO_2 rutile (110) surface, respectively. The contributions from the adsorbate are investigated by a projection of the DOS. The adsorbate levels in the band-gap are found to interact weakly with with the substrate, as they are narrow and sharp. The adsorbate levels overlapping

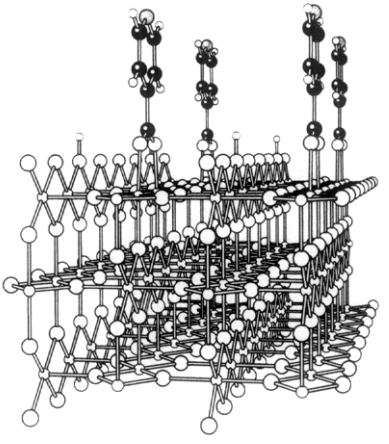


Figure 11: Periodic model for isonicotinic acid adsorbed on TiO_2 rutile (110).

the valence and conduction bands, on the contrary, are significantly broadened, which is a sign of strong interfacial interactions.

Both the cluster and the periodic calculations indicate a similarity to the Newns-Anderson model for metal adsorbates, in that both energy shifts, and broadenings need to be included in models of electron transfer, as shown schematically in Fig. 13. It will be a challenge in the near future to incorporate the increasingly accurate calculations of the crucial electronic coupling-strength parameter in existing dynamical models of the surface electron transfer processes.

Charge-transfer excitations

In addition to the general information of the electronic interaction for the various adsorbate levels from the PDOS of benzoic acid and catechol on TiO_2 nanoparticles, an experimentally observed absorption-edge lowering with catechol needs to be explained. This is believed to be an intramolecular ligand-to-titanium charge-transfer transition within the surface complex [127]. The actual role of the substrate conduction band has, however, not been settled by the experiments. The feature is well reproduced by INDO/S-CI cluster calculations using a one nanometer sized TiO_2 anatase cluster (cf. Fig. 14). It has been suggested that the low-energy absorption threshold for the catechol sensitized TiO_2 can be explained by a different electron transfer mechanism

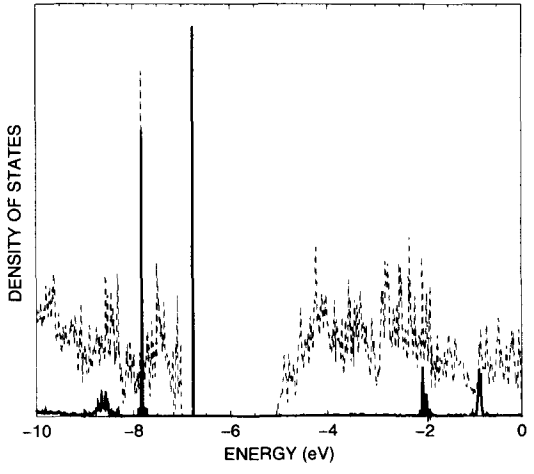


Figure 12: Calculated electronic structure for isonicotinic acid adsorbed on TiO_2 rutile (110) in the upper valence band, band-gap, and lower conduction band regions. The thin dashed line shows the total DOS, and the thick solid line the projected DOS for the adsorbate.

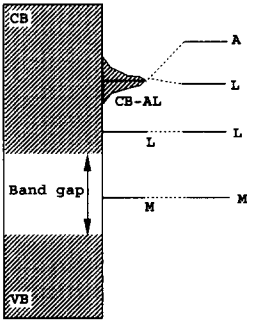


Figure 13: Interaction between a wide band-gap metal oxide such as TiO_2, and an anchored dye molecule such as N3. The dye orbitals are labeled according to dominant components: A for Anchor, L for Ligand, and M for metal. The metal oxide valence band (VB), and conduction band (CB) are shown. Sensitizer orbitals are shifted in energy relative to their normal positions, especially those involving the anchor group. Sensitizer orbitals are also broadened if they interact significantly with the substrate conduction band.

than is usually assumed for the sensitized metal oxides. No involvement of excited catechol states are seen in the calculations. This suggests that this is a direct charge-transfer photoinjection rather than the usual two-step excitation-injection [84]. The two electron transfer schemes are contrasted in Fig. 15. A direct injection should be extremely fast, and occur on the same time-scale as ordinary photoexcitations, i.e. in just a couple of femtoseconds. Direct injection is, however, not necessarily advantageous for efficient photoelectrochemical devices, as one must also consider the back-transfer times. These may well be too short for a functioning device in the case of catechol, as there are considerable interactions between the lower part of the conduction band, and the adsorbate ground state.

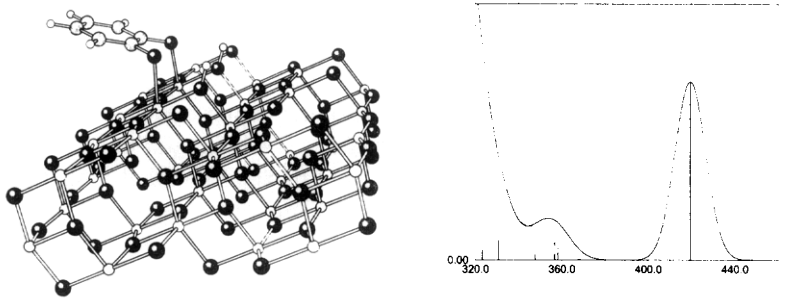


Figure 14: Structure of the cluster model of catechol adsorbed on TiO_2 (anatase) particle (left). Calculated absorption threshold of catechol sensitized TiO_2 as a function of wavelength in nm (right). The solid line refers to a catechol sensitized TiO_2 cluster, while the dashed line refers to a bare TiO_2 cluster.

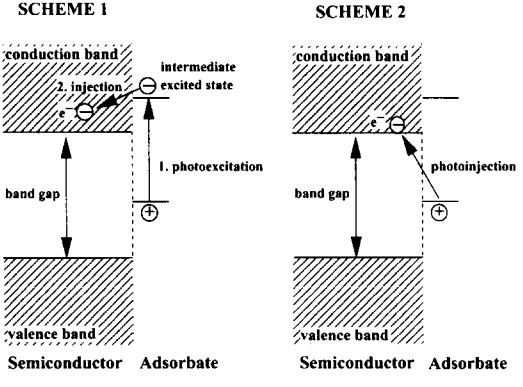


Figure 15: Electron transfer schemes for dye-sensitized metal oxides. Scheme 1 is the standard excitation-injection scheme, while Scheme 2 is a direct photoinjection, consistent with the calculated absorption threshold of catechol sensitized TiO_2

4 Defects and Intercalation in TiO_2

Intercalation of cations into a framework of titanium dioxide is a process of wide interest. This is due to the electrochromic properties associated with the process (a clear blue coloration results from the intercalation) and to the system's charge storage capabilities (facilitated by the reversibility of the process) and thus the potential application in rocking-chair batteries. We have studied alkali-metal intercalation and ion diffusion in the TiO_2 anatase and spinel crystals by theoretical methods ranging from condensed-phase *ab initio* to semiempirical computations [65, 66]. Structure relaxation, electron-density distribution, electron transfer, diffusion paths and activation energies of the ion intercalation process were modeled.

4.1 Equilibrium structures

In an early investigation [66] Li and Na intercalation in bulk TiO_2 anatase was studied by periodic *ab initio* Hartree–Fock (using CRYSTAL95 [128]) and semiempirical INDO methods. The composition of the Li_xTiO_2 and Na_xTiO_2 crystals varied from x=0.0625 to 0.5 in the INDO calculations whereas the higher concentration was used in the *ab initio* calculations (experimentally it has been found that anatase can accommodate Li ratios at least up to about $x = 0.7$ [129]). Later, $LiTiO_2$ was studied by Mackrodt [130] also using the Hartree–Fock method and the CRYSTAL95 code.

In our previous *ab initio* computations it was assumed that the anatase structure maintained its tetragonal structure on intercalation of the alkali metal atoms. (This was necessitated by the at that time considered heavy computational requirements and the lack of efficient automated procedures for geometry optimizations.) Experimentally, it is known that the cell symmetry decreases from tetragonal for TiO_2 to orthorhombic when $Li_{0.5}TiO_2$ is formed [129, 133].

New calculations using a more efficient CRYSTAL code [54] and optimization scheme [135] are presented here. In these, also the spin-polarized B3LYP hybrid functional was used and the orthorhombic structure of $Li_{0.5}TiO_2$ was studied. The same basis set was used as in [66]. Recently, the orthorhombic $Li_{0.5}TiO_2$ structure was studied by plane-wave generalized gradient-approximation (GGA) DFT calculations by Koudriachova et al. [136]. GGA calculations were performed also in the present study for comparison.

The geometry-optimized structure parameters (the only free parameter in the space group) of pure anatase calculated by different methods can be seen in Table 10. The computations reproduce the experimental structure quite well, although HF slightly overestimates the length of the c axis.

Table 10: The geometry-optimized cell dimensions and the oxygen z-coordinate of pure anatase from experiment and HF, B3LYP and GGA calculations.

	a/Å	c/Å	z(O)
Expt [131]	3.784	9.515	0.2081
INDO [64]	3.69	10.07	0.208
HF [132]	3.763	9.851	0.2025
B3LYP[a]	3.833	9.526	0.2082
GGA[b]	3.808	9.575	0.2075
B3LYP[c] [137]	3.792	9.824	0.2033

[a] This work, using the split-valence basis set from refs. [132, 66].

[b] Plane-wave GGA density-functional calculations [58] and similar computational conditions as used in the study by Koudriachova et al. [136], in the present work a kinetic energy cutoff of 380 eV and a k-point spacing of 0.06 $Å^{-1}$.

[c] From ref. [137], where a modification of the 6-31G basis set was used.

[d] For studies of pure anatase employing local density approximation (LDA) DFT calculations, see e.g. ref. [138].

The space group of the experimental structure of the $Li_{0.5}TiO_2$ crystal is *Imma*. In this structure, 2 Li atoms are disordered over 4 equivalent octahedral interstices. In our calculations we place two Li in two orthorhombic cavities at maximum distance, thereby reducing the symmetry further to *Imm2*. The calculated structural parameters for the $Li_{0.5}TiO_2$ crystal can be seen in Table 11. The structure of $Li_{0.5}TiO_2$ obtained from our B3LYP calculations is displayed in Fig. 16. The HF calculations overestimate the c-axis length, whereas both GGA and B3LYP partly remedy the situation. All methods manage to reproduce the shortening of the c-axis observed when Li ions are intercalated.

In the orthorhombic $Li_{0.5}TiO_2$ crystal, the Li ions are situated in octahedral sites but within the interstices the coordination of Li by O is more accurately described as 5-fold rather than 6-fold: in the z-direction the Li-O separations are 2.17 and 2.80 Å. These distances are fairly well reproduced by the B3LYP calculations, 1.99 and 2.98 Å, although the asymmetry is further accentuated. The corresponding numbers deduced from the GGA structure in Table 11 are 2.00 and the even longer 3.18 Å. There are two sets of Ti-Ti distances in the structure, the shorter of which is 2.89 Å, resulting in zig-zag Ti-Ti chains along the y-direction. The corresponding B3LYP and GGA results are 2.91 and 2.92 Å, respective.

Table 11: The geometry-optimized cell dimensions and coordinates of $Li_{0.5}TiO_2$ in the space group *Imm2* from experiment and HF, B3LYP and GGA calculations. The Ti_I, O_I and O_{II} y-coordinates are 0.5 by symmetry, the remaining coordinates are 0.

	Expt[a]	HF	B3LYP	GGA[b]
a/Å	3.8082	3.825	3.827	3.791
b/Å	4.0768	4.072	4.114	4.111
c/Å	9.0526	9.578	9.212	9.243
$z(Ti_I)$	0.8871	0.8789	0.8853	0.8831
$z(Ti_{II})$	0.1129	0.1146	0.1093	0.1072
$z(O_I)$	0.1030	0.0804	0.0915	0.0954
$z(O_{II})$	0.6521	0.6659	0.6552	0.6525
$z(O_{III})$	0.8970	0.8964	0.8900	0.8987
$z(O_{IV})$	0.3479	0.3444	0.3507	0.3383
z(Li)	0.6570	0.6709	0.6742	0.6818

[a] Reference [129]. The experimental coordinates in the *Imma* space group were transformed to the corresponding coordinates in the *Imm2* space group under the assumption of Li ordering.

[b] Plane-wave GGA density-functional calculations [58] and similar computational conditions as used in the study of this phase by Koudriachova et al. [136], in the present work a kinetic energy cutoff of 380 eV and a k-point spacing of 0.10 $Å^{-1}$.

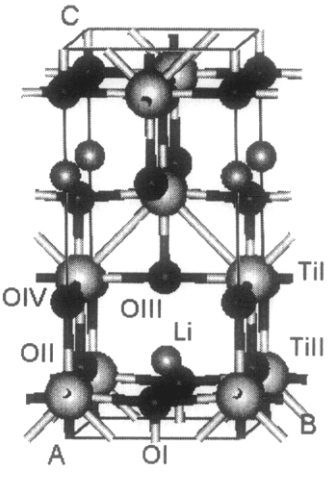


Figure 16: The structure of the $Li_{0.5}TiO_2$ crystal from B3LYP geometry optimization.

4.2 Electronic properties

The valence electron of the Li atom will not be situated near the Li nucleus when the atom is intercalated. Instead, Li is ionized and the electron is transfered to an empty d-orbital on the Ti ions. From the B3LYP computations, the Mulliken atomic charge of Li is +0.69, which is somewhat lower than the HF result +0.79 from ref. [66]. In that study, also Na intercalation was studied, and the Na atoms were ionized to an even larger extent (+0.99), which can be expected. The B3LYP charges for the two Ti ions are about +1.89, and those of the O ions between -0.82 and -0.94. The spin density (or more precisely, the difference between Mulliken α and β charge populations) is evenly distributed between the Ti ions (spin 0.71 on the Ti ion that is nearest neighbour to Li in the yz-plane, and 0.87 on the Ti ions in the approximate xy-plane), whereas the spin density on Li is zero. This is somewhat at variance with the result in ref. [66] where the unpaired spin was localized at the Ti ion closest to Li in the xy-plane. (Such a spin localization occurs also in the new HF-optimized structure: the Mulliken charges on the two Ti ions are +2.56 and +2.32, and the spins are 0.00 and 1.01.)

The B3LYP electronic density of states spectra for $Li_{0.5}TiO_2$ can be seen in Fig. 17. For pure TiO_2, the top of the valence band essentially consists of

O2p orbitals, whereas the conduction band is mainly Ti3d. In the intercalated compound, part of the Ti3d band has been filled, and constitutes a donor state below the conduction band. In these calculations the crystal was required to be a semiconductor (or actually an insulator in the absence of thermal excitations since the calculations are carried out at 0 Kelvin), in agreement with the experimental observation by Cava et al. [129]. Even so, the gap between the occupied d-orbital state and the valence band is vanishingly small, so that small variations in the composition of the compound might drastically change its properties. The electronic structure obtained from GGA computations corresponded to a conducting material.

4.3 Ion diffusion

In ref. [66] the activation energies for the diffusion process of Li and Na in $Li_{0.5}TiO_2$ were calculated using *ab initio* and semiempirical methods. The activation energies of both Li and Na diffusion were in excellent agreement with the results from chronocoloumetric experiments presented in that paper. The diffusion of the alkali metal ions was assumed to occur along a path connecting the octahedral vacant hole sites. The energy of the transition state structure in between two octahedral holes was obtained from INDO, HF and B3LYP computations. In these, the structure was geometry-optimized, where the intercalated ion was confined to the transition state position and all other free parameters were allowed to vary (when the alkali metal ion is displaced along the reaction path the symmetry is reduced and the resulting space group is $Pmn2_1$). The diffusion path is shown schematically in Fig. 18.

The activation energies for the diffusion path, i.e. the energy difference between the transition state and the equilibrium state per diffusing atom, for various compositions of Me_xTiO_2 is displayed in Table 12.

The agreement between both INDO and *ab initio* HF results with the experimental results for Li and Na is good. For Na there is a discrepancy between the different experimental results. Our computational results clearly favor the chronocoloumetric measurements [66] rather than earlier impedance spectroscopy results [134]. The B3LYP computation appears to underestimate the diffusion barrier by a substantial amount. In ref. [136] the diffusion barrier in Li_xTiO_2 was calculated using the GGA method to be 0.5 eV, in fair agreement with our HF results [66]. (As mentioned above, it might be that the physical situations represented by the different calculations are quite different, since in ref. [136] the crystal was a conductor whereas in ref. [66] an insulator.)

In ref. [66] it was found that at equilibrium the unpaired electron spin was found on the Ti_{II}, whereas at the transition state the spin had transferred to the Ti_I ion. There thus appears to be a concerted motion of electrons and ions

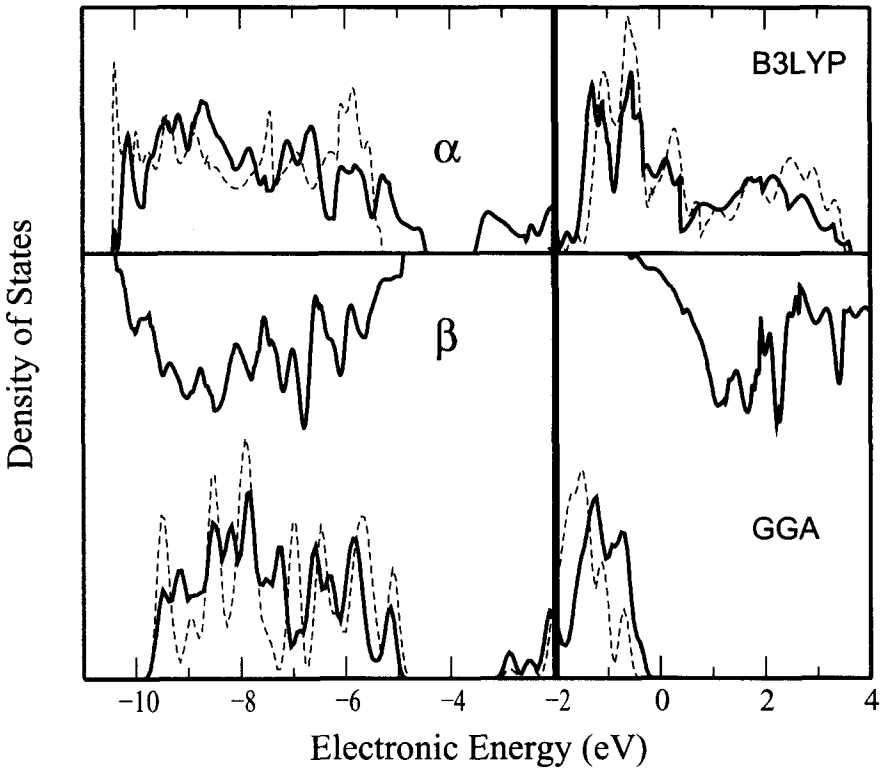


Figure 17: The electronic density-of-states spectrum for $Li_{0.5}TiO_2$ (solid line) and pure anatase (dashed line) from B3LYP (top) and GGA (bottom) calculations. In the B3LYP spectrum, the DOS is divided into the contributions from alpha (positive) and beta (negative) states. The vertical line denotes the Fermi level of $Li_{0.5}TiO_2$. The $Li_{0.5}TiO_2$ spectra were shifted by −9.6 eV in order to make the B3LYP and GGA Fermi levels coincide, and the anatase spectra were shifted by +0.3 (B3LYP) and −8.9 (GGA) eV to make the left onsets of the $Li_{0.5}TiO_2$ and anatase bands coincide.

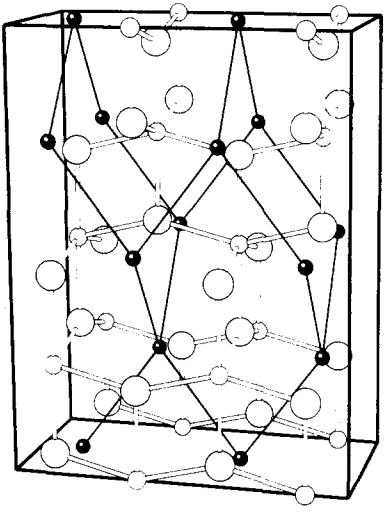


Figure 18: Schematic representation of the alkali metal diffusion paths. Octahedral interstitial sites are represented by black ions. The transition states are situated at the midpoints connecting the interstitial sites.

Table 12: The alkali-metal ion diffusion activation energies (eV) in TiO_2 anatase.

	theory			experiment	
	INDO[a]	HF	B3LYP	ref. [134][b]	ref. [66]
Li_xTiO_2				0.51,0.66	0.6±0.1
x=0.0625	0.51				
x=0.25	0.54				
x=0.5	0.56	0.60[a],0.57	0.31		
Na_xTiO_2				1.60,1.78	0.52±0.1
x=0.0625	0.54				
x=0.25	0.55				
x=0.5	0.56	0.55[a]			

[a] From ref. [66] [b] The two experimental values from ref. [134] correspond to $x < 0.5$ and $x > 0.5$, respective.

during the diffusion, even though the electron is not situated on the diffusing ion but in neighbouring Ti d-orbitals.

5 Conclusions

Quantum chemistry has so far had little impact on the field of photoelectrochemistry. This is largely due to the molecular complexity of the experimental systems, which has prevented reliable computational methods to be used on realistic model systems, although some theoretical approaches to various aspects of the performance of nanostructured metal oxide photoelectrochemical systems have appeared in the last 10 years, see e.g. [139, 140, 141]. Here we have focussed on quantum-chemical cluster and surface calculations of a number of relevant problems including adsorbates and intercalation. These calculations illustrate the emerging possibilities of using quantum chemical calculations to model complicated dye-sensitized photoelectrochemical systems.

Extensive tests have been carried out to establish the reliability of quantum-chemical schemes for metal oxide investigations. This includes schemes at a variety of levels of sophistication suitable for calculations of very large systems. In particular density functional methods offer good possibilities to treat sufficiently large systems to be applicable to many central problems in the field of photoelectrochemistry with reasonable accuracy and at very competitive computational costs. Semiempirical methods still offer a last possibility to perform reasonably accurate calculations on nanostructured systems containing several hundred atoms where first principles methods still cannot be applied routinely.

Periodic calculations of intercalation, in particular in TiO_2 have successfully provided support and interpretations of several experiments designed to investigate fundamental properties of potential displays and batteries.

Structural and energetic effects of molecular adsorption on metal oxide surfaces have been studied. Also adsorption effects on the electronic structure of anchored aromatic adsorbates have been investigated. The strength of the electronic interactions is seen to vary significantly (from meV to eV), depending on the shape and position of the involved adsorbate level. Both shifts in energy positions, and broadenings of energy levels have been seen for aromatic adsorbates on TiO_2.

The implications of the obtained structural and electronic information on the binding and the surface electron transfer models in dye-sensitized solar cells have been discussed. Calculated strong binding, and strong electronic surface-adsorbate interactions, is consistent with experimentally observed ultrafast photoinjection processes in stable dye-sensitized electrochemical devices. It will be important, however, to combine results from explicit calculations of

coupling strengths with general models of surface electron transfer, in an effort to obtain direct comparisons with experimental measurements of electron injection times. Comparisons of calculations with ultrafast laser spectroscopy investigations, where theory and experiment are applied to identical systems, should in the near future provide detailed information of many charge-transfer processes central to the photoelectrochemical devices.

It should now be possible to study interactions between many other interesting surfaces and adsorbates. In particular, it will be important to try to model the surface interactions of organometallic dyes like N3. There are, however, many factors which are yet to be accounted for, including solvent and thermal effects, as well as defects and surface states. With increasing computer power, improved theoretical models and better computational schemes, quantum chemistry has good prospects of becoming a valuable tool in the field of photoelectrochemistry.

Acknowledgments

It is a privilege to dedicate this paper to the memory of Per-Olov Löwdin, one of the pioneers of quantum chemistry, and an outstanding teacher whose energy and enthusiasm, and passion for rigor and clarity, and his pedagogical skill, formed a whole generation of quantum chemists, not only in Uppsala but all over the world.

This work was sponsored by the Swedish Research Council (VR) and the European Community Joule III program. Many of the calculations were made possible by grants for computer time at the Swedish National Supercomputer Center (NSC). SL and PP gratefully acknowledge a grant from the Göran Gustafsson Foundation.

References

[1] P.-O. Löwdin, *A Theoretical Investigation into Some Properties of Ionic Crystals*, Uppsala Dissertation (1948)

[2] A. Hagfeldt, and M. Grätzel, Chem. Rev. 95, 49 (1995)

[3] A. Hagfeldt and M. Grätzel, Acc. Chem. Res., 33, 269 (2000)

[4] S. Huang, L. Kavan, I. Exnar and M. Grätzel, J. Electrochem. Soc. 142, 142 (1995)

[5] A. Hagfeldt, N. Vlachopoulos and M. Grätzel, J. Electrochem. Soc. 141, L82 (1994)

[6] A. Hagfeldt, L. Walder and M. Grätzel, Proc. Soc. Photo-Opt. Instrum. Engn. 2531, 297 (1995)

[7] J. F. Harrison, Chem. Rev. 100, 679-716 (2000)

[8] C. Noguera, Surf. Rev. Lett. 8, 121 (2001)

[9] A. Hagfeldt, H. Siegbahn, S.-E. Lindquist, and S. Lunell, Int. J. Quantum Chem. 44, 477 (1992)

[10] A. Hagfeldt, R. Bergström, H. O. G. Siegbahn, S. Lunell, J. Phys. Chem. 97, 12725 (1993)

[11] A. Hagfeldt, S. Lunell, H. O. G. Siegbahn, Int. J. Quantum Chem. 49, 97 (1994)

[12] K. S. Jeong, Ch. Chang, E. Sedlmayr, and D. Sülzle, J. Phys. B: At. Mol. Opt. Phys. 33, 3417 (2000)

[13] T. Albaret, F. Finocchi, and C. Noguera, Appl. Surf. Sci. 144-145, 672 (1999)

[14] T. Albaret, F. Finocchi, and C. Noguera, J. Chem. Phys. 113, 2238 (2000)

[15] R. Bergström, S. Lunell, and L. A. Eriksson, Int. J. Quantum Chem. 59, 427 (1996)

[16] A.J. Merer, Ann. Rev. Phys. Chem. 40, 407 (1989)

[17] C.W. Bauschlicher Jr. and P. Maitre, Theor. Chim. Acta 90, 189 (1995)

[18] A.J. Bridgeman and J. Rothery, J. Chem. Soc., Dalton Trans. 211 (2000)

[19] J. Piechota and M. Suffczynski, Z. Phys. Chem. 200, 39 (1997)

[20] Gaussian 94, Revision D.4, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.

[21] Gaussian 98, Revision A.6, A.7, A.9, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1998.

[22] Æ. Frisch, and M.J. Frisch, Gaussian 98 User's Reference, Gaussian Inc. Pittsburgh, PA, 1998 and references therein.

[23] Some of the basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institue for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller or Karen Schuchardt for further information.

[24] A. Schäfer, H. Horn and R. Ahlrichs, J. Chem. Phys. 97, 2571 (1992)

[25] A. J. H. Wachters, J. Chem. Phys. 52, 1033 (1970). A. J. H. Wachters, IBM Tech. Rept. RJ584 (1969)

[26] C. W. Bauschlicher, Jr. S. R. Langhoff and L. A. Barnes, J. Chem. Phys. 91, 2399 (1989)

[27] Zn p-functions: exponents 0.256 and 0.0748; d-function: exponent 0.18; f-function: contracted (3f)/[1f] exponents 9.420, 3.378, 1.447, contraction coefficients the same as for the other metal atoms

[28] A. Schäfer, C. Huber and R. Ahlrichs, J. Chem. Phys. 100, 5829 (1994)

[29] T.H. Dunning, Jr. J. Chem. Phys. 90, 1007 (1989). R.A. Kendall, T.H. Dunning, Jr. and R.J. Harrison, J. Chem. Phys. 96, 6769 (1992)

[30] S.F. Boys and F. Bernardi, Mol. Phys. 19, 553 (1970)

[31] C.W. Bauschlicher, Jr and H. Partridge, J. Chem. Phys. 109, 8430 (1998)

[32] M. Sodupe, V. Branchadell, M. Rosi, and C.W. Bauschlicher, Jr., J. Phys. Chem. A 101, 7854 (1997)

[33] Y. Nakao, K. Hirao, and T. Taktsugu, J. Chem. Phys. 114, 7935 (2001)

[34] A. D. Becke, J. Chem. Phys. 98, 5648 (1993)

[35] E. R. Fisher, J.L. Elkind, D.E. Clemmer, R. Georgiadis, S.K. Loh, N. Aristov, L.S. Sunderlin, and P.B. Armentrout, J. Chem. Phys. 93, 2676 (1990)

[36] R.L. Martin, and P.J. Hay, J. Chem. Phys. 75, 4539 (1981)

[37] A.J. Bridgeman, J. Chem. Soc., Dalton Trans., 4555 (1996)

[38] C.A. Fancher, H.L. de Clercq, O.C. Thomas, D.W. Robinson, and K.H. Bowen, J. Chem. Phys. 109, 8426 (1998)

[39] H. Kang and J.L. Beauchamp, J. Am. Chem. Soc. 108, 5663 (1986)

[40] Ø. Espelid and K.J. Borve, J. Phys. Chem. A 101, 9449 (1997)

[41] L.R. Watson, T.L. Thiem, R.A. Dressler, R.H. Salter, and E.J. Murad, J. Phys. Chem. 97, 5577 (1993)

[42] M.T. Rodgers, B. Walker, and P.B. Armentrout, Int. J. Mass Spectrom. 182/183, 99 (1999)

[43] D. E. Clemmer, N.F. Dalleska, and P.B. Armentrout, J. Chem. Phys. 95, 7263 (1991)

[44] D. E. Clemmer, J. L. Elkind, N. Aristov, and P. B. Armentrout, J. Chem. Phys. 95, 3387 (1991)

[45] H.-P. Loock, B. Simard, S. Wallin, and C. Linton, J. Chem. Phys. 109, 8980 (1998)

[46] P.B. Armentrout, L.F. Halle, and J.L. Beauchamp, J. Chem. Phys. 76, 2449 (1982)

[47] P.B. Armentrout and B.L. Kickel, p. 1-45 Organometallic Ion Chemistry; Freiser, S. B., Ed.; Kluwer Academic Publishers: Dordrecht (1996)

[48] H.-J. Freund, Faraday Discussion 114, 1 (2000)

[49] B. G. Daniels, R. Lindsay, and G. Thornton, Surf. Rev. Lett. 8, 95 (2001)

[50] R. Lindsay and G. Thornton, J. Phys.: Condens. Matter 13, 11207 (2001)

[51] F.H. Jones, Surf. Sci. Reports, 42, 75 (2001)

[52] G. E. Brown, Jr., V. E. Heinrich, W. H. Casey, D. L. Clark, C. Eggleston, A. Felmy, D. W. Goodman, M. Grätzel, G. Maciel, M. I. McCarthy, K. H. Nealson, D. A. Sverjensky, M. F. Toney, J. M. Zachara, Chem. Rev. 99, 77 (1999)

[53] St. Hövel, C. Kolczewski, M. Wühn, J. Albers, K. Weiss, V. Staemmler, Ch. Wöll, J. Chem. Phys. 112, 3909 (2000)

[54] V. R. Saunders, R. Dovesi, C. Roetti, M. Causà, N. M. Harrison, R. Orlando, and C. M. Zicovich-Wilson, *Crystal98 User's Manual*, University of Torino, Torino, 1998.

[55] C. Pisani, R. Dovesi and C. Roetti, *Ab Initio Treatment of Crystalline Solids*, Lecture Notes in Chemistry 48 (Springer, Berlin, 1988)

[56] P. Persson and L. Ojamäe, Chem. Phys. Lett. 321, 302 (2000)

[57] J. Muscat, A. Wander, N. M. Harrison, Chem. Phys. Lett. 342, 397 (2000)

[58] M. C. Payne, M. P. Teter, D. C. Allen, T. A. Araias and J. D. Joannopolous, Rev. Mod. Phys. 64, 1045 (1992)

[59] P. Persson, L. Ojamäe, and S. Lunell, Int. J. Quantum. Chem. 00, 0000 (2002)

[60] J. P. Perdew and Y. Yang, Phys. Rev. B 45, 13244 (1992)

[61] J. P. Perdew, Phys. Rev. B 34, 7406 (1986)

[62] CPMD version 3.3, written by J. Hutter *et al.*, Max-Planck-Institut für Festkörperforschung, Stuttgart, IBM Research laboratory Zürich (1995-2000).

[63] E. V. Stefanovich, E. K. Shidlovskaya, A. LSchluger, M. A. Zakharov, Phys. Status Solidi B 160, 529 (1990)

[64] A. Stashans, S. Lunell, and R. W. Grimes, J. Phys. Chem. Solids, 57, 1293 (1996)

[65] A. Stashans, S. Lunell, R. Bergström, A. Hagfeldt, S.-E. Lindquist, Phys Rev. B 53, 159 (1996)

[66] S. Lunell, A. Stashans, L. Ojamäe, H. Lindström, and A. Hagfeldt, J. Am. Chem. Soc. 119, 7374 (1997)

[67] A. Stashans and S. Lunell, Int. J. Quantum Chem. 63, 729 (1997)

[68] P. Persson, A. Stashans, R. Bergström, and S. Lunell, Intern. J. Quantum Chem. 70, 1055 (1998)

[69] R. Davies, J. F. Walsh, C. A. Muryn, G. Thornton, V. R. Dhanak, and K. C. Prince, Surf. Sci. 298, L196 (1993)

[70] Q. Guo, I. Cocks, and E. M. Williams, J. Chem. Phys. 106, 2924 (1997)

[71] S. P. Bates, G. Kresse, and M. J. Gillan, Surf. Sci. 409, 336 (1998)

[72] H. Nakatsuji, M. Yoshimoto, Y. Umemura, S. Takagi, and M. Hada, J. Phys. Chem. 100, 53 (1998)

[73] A. Vittadini, A. Selloni, F. P. Rotzinger, and M. Grätzel, J. Phys. Chem. B 104, 1300 (2000)

[74] I. M. Campbell, *Catalysis at Surfaces*, Chapman and Hall Ltd, London New York (1988)

[75] P. Zapol, J. B. Jaffe, and A. C. Hess, Surf. Sci. 422, 1 (1999)

[76] M. Odelius, M. Bernasconi, and M. Parrinello, Phys. Rev. Lett. 78, 2855 (1997)

[77] L. Patthey, H. Rensmo, P. Persson, K. Westermark, L. Vayssieres, A. Stashans, Å. Petersson, P. A. Brühwiler, H. Siegbahn, S. Lunell, and N. Mårtensson, J. Chem. Phys. 110, 5913 (1999)

[78] P. Persson, S. Lunell, P. A. Brühwiler, J. Schnadt, S. Södergren, J. N. O'Shea, O. Karis, H. Siegbahn, N. Mårtensson, M. Bässler, and L. Patthey, J. Chem. Phys. 112, 3945 (2000)

[79] H. Rensmo, K. Westermark, S. Södergren, O. Kohle, P. Persson, S. Lunell, and H. Siegbahn, J. Chem. Phys. 111, 2744 (1999)

[80] M. Odelius, P. Persson, and S. Lunell, to be published.

[81] P. Persson and S. Lunell, Sol. Energy Mater. Sol. Cells 63, 139 (2000)

[82] H. Rensmo, S. Lunell, and H. Siegbahn, J. Photochem. Photobio. A 114, 117 (1998)

[83] V. E. Heinrich, and P. A. Cox, *The Surface Science of Metal Oxides*, Cambridge University Press, Cambridge (1994)

[84] P. Persson, R. Bergström, and S. Lunell, J. Phys. Chem. B 104, 10348 (2000)

[85] L. Patthey, J. Schnadt, S. Södergren, O. Karis, J. N. O'Shea, P. A. Brühwiler, H. Siegbahn, N. Mårtensson, M. Bässler, P. Persson, and S. Lunell, *MAX-LAB Activity Report 1998*, 84 (1999)

[86] Y. Iwasawa, H. Onishi, K.-i. Fukui, S. Suzuki, and T. Sasaki, Faraday Discussion 114, 259 (2000)

[87] M. Fujihira, N. Ohishi, and T. Osa, Nature, 268, 226 (1977)

[88] M. Dare-Edwards, J. B. Goodenough, A. Hamnett, K. R. Seddon, and R. D. Wright, Faraday Discuss. Chem. Soc, 70, 285 (1981)

[89] T. Meyer, G. J. Meyer, B. W. Pfenning, J. R. Schoonover, C. J. Timpson, J. F. Wall, C. Kobusch, X. Chen, B. M. Peek, C. G. Wall, W. Ou, B. W. Erickson, and C. A. Bignozzi, Inorg. Chem. 33, 3952 (1994)

[90] K. Murakoshi, G. Kano, Y. Wada, S. Yanagida, H. Miyazaki, M. Matsumoto, and S. Murasawa, J. Electroanal. Chem. 361, 27 (1995)

[91] N. W. Duffy, K. D. Dobson, K. C. Gordon, B. H. Robson, and A. J. McQuillan, Chem. Phys. Lett. 266, 451 (1997)

[92] K. S. Finnie, J. R. Bartlett, and J. L. Woolfrey, Langmuir 14, 2744 (1998)

[93] V. Shklover, Yu. E. Ovchinnikov, L. S. Braginsky, S. M. Zakeeruddin, and M. Grätzel, Chem. Mater. 10, 2533 (1998)

[94] R. A. Marcus, and N. Sutin, Biochimica et Biophysica Acta 811, 265 (1985)

[95] E. Ilisca and K. Makoshi (eds.), *Electronic Processes at Solid Surfaces*, World Scientific Publishing Co. Pte. Ltd. (1996)

[96] M. Grätzel, *Heterogeneous photochemical electron transfer*, CRC Press, Inc. (1989)

[97] R. J. D. Miller, G. L. McLendon, A. J. Nozik, W. Schmickler, and F. Willig, *Surface Electron Transfer Processes*, VCH Publishers, Inc. (1995)

[98] P. Bonhôte, E. Gogniat, S. Tingry, C. Barbé, N. Vlachopoulos, F. Lenzmann, P. Comte, and M. Grätzel, J. Phys. Chem. B, 102, 1498 (1998)

[99] K. Westermark, S. Tingry, P. Persson, H. Rensmo, S. Lunell, A. Hagfeldt, and H. Siegbahn, J. Phys. Chem. B 105, 7182 (2001)

[100] H. Rensmo, K. Keis, H. Lindström, S. Södergren, A. Solbrand, A. Hagfeldt, S.-E. Lindquist, L. N. Wang, and M. Muhammed, J. Phys. Chem. B 101, 2598 (1997)

[101] T. A. Heimer, S. T. D'Arcangelis, F. Farzad, J. M. Stipkala, and G. J. Meyer, Inorg. Chem. 35, 5319 (1996)

[102] F. Willig, C. Zimmermann, S. Ramakrishna, and W. Storck, Electrochim. Acta 45, 4565 (2000)

[103] S. Ramakrishna, and F. Willig, J. Phys. Chem. B, 104, 68 (2000)

[104] Y. Q. Gao, Y. Georgievskii, and R. A. Marcus, J. Chem. Phys. 112, 3358 (2000)

[105] Å. Petersson, M. Ratner, and H. O. Karlsson, J. Phys. Chem. B 104, 8498 (2000)

[106] S. Ramakrishna, F. Willig, and V. May, Phys. Rev. B 62, R16330 (2000)

[107] S. Ramakrishna, and F. Willig, J. Chem. Phys. 115, 2743 (2001)

[108] S. Ramakrishna, F. Willig, and V. May, Chem. Phys. Lett. 351, 242 (2002)

[109] M. C. Zerner and co-workers, ZINDO: A Semiempirical Program Package, University of Florida, Gainesville, FL 32611, USA.

[110] B. Burfeindt, S. Ramakrishna, C. Zimmermann, B. Meißner, Th. Hannappel, W. Storck, and F. Willig, In *Ultrafast Phenomena XI*; Springer-Verlag: Berlin, p. 636 (1998)

[111] Y. Tachibana, J. E. Moser, M. Grätzel, D. R. Klug, and J. R. Durrant, J. Phys. Chem. 100, 20056 (1996)

[112] T. Hannappel, B. Burfeindt, W. Storch, and F. Willig, J. Phys. Chem. B 101, 6799 (1997)

[113] B. Burfeindt, C. Zimmermann, S. Ramakrishna, T. Hannappel, B. Meißner, W. Storch, and F. Willig, Z. Phys. Chem. 212, 67 (1999)

[114] R. J. Ellingson, J. B. Asbury, S. Ferrere, H. N. Ghosh, J. R. Sprague, T. Lian, and A. J. Nozik, Z. Phys. Chem. 212, 77 (1999)

[115] J.-E. Moser, M. Wolf, F. Lenzmann, and M. Grätzel, Z. Phys. Chem. 212, 85 (1999)

[116] G. Benko, J. Kallioinen, J. E. I. Korppi-Tommola, A. P. Yartsev, and V. Sundström, J. Am. Chem. Soc. 124, 489 (2002)

[117] J. B. Asbury, E. Hao, Y. Wang, H. N. Ghosh, and T. Lian, J. Phys. Chem. B 105, 4545 (2001)

[118] J. P. Muscat and D. M. Newns, Prog. Surf. Sci. 9, 1 (1978)

[119] N.-H. Ge, C. M. Wong, and C. B. Harris, Acc. Chem. Res. 33, 111 (2000)

[120] J. O'M. Bockris and S. U. M. Kahn, *Quantum Electrochemistry*, Plenum Press, New York and London (1979)

[121] P. Persson, S. Lunell, and L. Ojamäe, to be published

[122] J. Moser, S. Punchihewa, P. P. Infelta, and M. Grätzel, Langmuir 7, 3012 (1991)

[123] I. Martini, J. H. Hodak, and G. V. Hartland, J. Phys. Chem. B 102, 9508 (1998)

[124] R. S. Mulliken and W. B. Person, *Molecular Complexes - A Lecture and Reprint Volume*, Wiley-Interscience (1969)

[125] J. N. Murrell, *Theory of the Electronic Spectra of Organic Molecules*, Methuen & CO LTD (1963)

[126] R. Rodriguez, M. A. Blesa, and A. E. Regazzoni, J. Colloid and Interface Sci. 177, 122 (1996)

[127] Y. Liu, J. I. Dadap, D. Zimdars, and K. B. Eisenthal, J. Phys. Chem. B, 103, 2480 (1999)

[128] R. Dovesi, V. R. Saunders, C. Roetti, M. Causà, N. M. Harrison, R. Orlando, and E. Aprà, *Crystal95 User's Manual*, University of Torino, Torino, 1996.

[129] R. J. Cava, D. W. Murphy, S. Zahurak, A. Santoro, and R. S. Roth, J. Solid State Chem. 53, 64 (1984)

[130] W. C. Mackrodt, J. Solid State Chem. 142, 428 (1999)

[131] S. C. Abrahams, J. L. Bernstein, Acta Cryst. B 25, 1233 (1969)

[132] A. Fahmi, C. Minot, B. Silvi and M. Causà, Phys. Rev. B 47, 11717 (1993)

[133] D. W. Murphy, R. J. Cava, S. M. Zahurak, and A. Santoro, Solid State Ionics 9 & 10, 413 (1983)

[134] M. Strömme Mattson, M. Vezelei, G. A. Niklasson, C. G. Granqvist, A. Stashans, and S. Lunell. In *Electrochromic Materials and Their Applications III*, Eds. K. C. Ho, C. B. Greenberg, and D. M. MacArthur. Electrochem. Soc. Proc. Series 96-24, 229 (1997)

[135] C. M. Zicovich-Wilson, private communication.

[136] M. V. Koudriachova, N. M. Harrison, S. W. de Leeuw, Phys. Rev. Lett. 86, 1275 (2001)

[137] A. Beltrán, J. R. Sambrano, M. Calatayud, F. R. Sensato, and J. Andrés, Surf. Sci. 490, 116 (2001)

[138] M. Mikami, S. Nakamura, O. Kitao, H. Arakawa, and X. Gonze, Jpn. J. Appl. Phys. 39, L847 (2000)

[139] S. Södergren, A. Hagfeldt, J. Olsson, and S.-E. Lindquist, J. Phys. Chem. 98, 5552 (1994)

[140] J. Nelson, Phys. Rev. B 59, 15374 (1999)

[141] J. Ferber, and J. Luther, J. Phys. Chem. B 105, 4895 (2001)

Subject Index