



ADVANCES IN QUANTUM CHEMISTRY

Volume 25

John R. Sabin &
Michael C. Zerner

EDITORIAL BOARD

David P. Craig (Canberra, Australia)
Raymond Daudel (Paris, France)
Ernst R. Davidson (Bloomington, Indiana)
Inga Fischer-Hjalmars (Stockholm, Sweden)
Kenichi Fukui (Kyoto, Japan)
George G. Hall (Kyoto, Japan)
Frederick A. Matsen (Austin, Texas)
Roy McWeeney (Pisa, Italy)
Joseph Paldus (Waterloo, Canada)
Ruben Pauncz (Haifa, Israel)
Siegrid Peyerimhoff (Bonn, Germany)
John A. Pople (Pittsburgh, Pennsylvania)
Alberte Pullman (Paris, France)
Bernard Pullman (Paris, France)
Klaus Ruedenberg (Ames, Iowa)
Henry F. Schaefer III (Athens, Georgia)
Au-Chin Tang (Kirin, Changchun, China)
Rudolf Zahradnik (Prague, Czechoslovakia)

ADVISORY EDITORIAL BOARD

David M. Bishop (Ottawa, Canada)
Jean-Louis Calais (Uppsala, Sweden)
Giuseppe del Re (Naples, Italy)
Fritz Grein (Fredericton, Canada)
Mu Shik Jhon (Seoul, Korea)
Mel Levy (New Orleans, Louisiana)
Jan Linderberg (Aarhus, Denmark)
William H. Miller (Berkeley, California)
Keiji Morokuma (Okazaki, Japan)
Jens Oddershede (Odense, Denmark)
Pekka Pyykkö (Helsinki, Finland)
Leo Radom (Canberra, Australia)
Mark Ratner (Evanston, Illinois)
Dennis R. Salahub (Montreal, Canada)
Isaiah Shavitt (Columbus, Ohio)
Per Siegbahn (Stockholm, Sweden)
Harel Weinstein (New York, New York)
Robert E. Wyatt (Austin, Texas)
Tokio Yamabe (Kyoto, Japan)

ADVANCES IN QUANTUM CHEMISTRY

EDITOR-IN-CHIEF

PER-OLOV LÖWDIN

PROFESSOR EMERITUS

QUANTUM CHEMISTRY GROUP
UPPSALA UNIVERSITY
UPPSALA, SWEDEN

AND

QUANTUM THEORY PROJECT
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA

EDITORS

JOHN R. SABIN AND MICHAEL C. ZERNER

QUANTUM THEORY PROJECT
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA

VOLUME 25



ACADEMIC PRESS

A Division of Harcourt Brace & Company

San Diego

New York

Boston


London

Sydney

Tokyo

Toronto

Academic Press Rapid Manuscript Reproduction

This book is printed on acid-free paper. 

Copyright © 1994 by ACADEMIC PRESS, INC.

All Rights Reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Academic Press, Inc.

525 B Street, Suite 1900, San Diego, California 92101-4495

United Kingdom Edition published by

ACADEMIC PRESS LIMITED

24-28 Oval Road, London NW1 7DX

International Standard Serial Number: 0065-3276

International Standard Book Number: 0-12-034825-X

PRINTED IN THE UNITED STATES OF AMERICA

94 95 96 97 98 99 QW 9 8 7 6 5 4 3 2 1

Contributors

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

Emili Besalú (255), Institut de Química Computacional, Universitat de Girona, 17071 Girona, Spain

David M. Bishop (3), Department of Chemistry, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5

Blanca Calabuig (255), Institut de Química Computacional, Universitat de Girona, 17071 Girona, Spain

Ramon Carbó (255), Institut de Química Computacional, Universitat de Girona, 17071 Girona, Spain

Wolfgang Förner (207), Friedrich-Alexander University, Erlangen-Nürnberg, D-91058 Erlangen, Germany

Peter M. W. Gill (143), Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

P. P. Schmidt (49), Chemistry and Materials Divisions, Office of Naval Research, Arlington, Virginia 22217

Leonel Vera (255), Institut de Química Computacional, Universitat de Girona, 17071 Girona, Spain

Preface

In investigating the highly different phenomena in nature, scientists have always tried to find some fundamental principles that can explain the variety from a basic unity. Today they have shown not only that all the various kinds of matter are built up from a rather limited number of atoms, but also that these atoms are composed of a few basic elements of building blocks. It seems possible to understand the innermost structure of matter and its behavior in terms of a few elementary particles: electrons, protons, neutrons, photons, etc., and their interactions. Since these particles obey not the laws of classical physics but the rules of modern quantum theory of wave mechanics established in 1925, there has developed a new field of “quantum science” which deals with the explanation of nature on this ground.

Quantum chemistry deals particularly with the electronic structure of atoms, molecules, and crystalline matter and describes it in terms of electronic wave patterns. It uses physical and chemical insight, sophisticated mathematics, and high-speed computers to solve the wave equations and achieve its results. Its goals are great, and today the new field can boast of both its conceptual framework and its numerical accomplishments. It provides a unification of the natural sciences that was previously inconceivable, and the modern development of cellular biology shows that the life sciences are now, in turn, using the same basis. “Quantum Biology” is a new field which describes the life processes and the functioning of the cell on a molecular and submolecular level.

Quantum chemistry is hence a rapidly developing field which falls between the historically established areas of mathematics, physics, chemistry, and biology. As a result there is a wide diversity of backgrounds among those interested in quantum chemistry. Since the results of the research are reported in periodicals of many different types, it has become increasingly difficult for both the expert and the nonexpert to follow the rapid development in this new borderline area.

The purpose of this serial publication is to present a survey of the current development of quantum chemistry as it is seen by a number of internationally leading research workers in various countries. The authors have been invited to give their personal points of view of the subject freely and without severe space limitations. No attempts have been made to avoid overlap—on the contrary, it seems desirable to have certain important research areas reviewed from different points of view.

The response from the authors and the referees has been so encouraging that a series of new volumes is being prepared. However, in order to control production

costs and speed publication time, a new format involving camera-ready manuscripts has been in use since Volume 20. A special announcement about the new format was enclosed in that volume (page xiii).

In the volumes to come, special attention will be devoted to the following subjects: the quantum theory of closed states, particularly the electronic structure of atoms, molecules, and crystals; the quantum theory of scattering states, dealing also with the theory of chemical reactions; the quantum theory of time-dependent phenomena, including the problem of electron transfer and radiation theory; molecular dynamics; statistical mechanics and general quantum statistics; condensed matter theory in general; quantum biochemistry and quantum pharmacology; the theory of numerical analysis and computational techniques.

As to the content of Volume 25, the Editors thank the authors for their contributions, which give an interesting picture of part of the current state of the art of the quantum theory of matter: From nonlinear-optical calculations, over a study of ion motion in molecular channels, a treatment of molecular integrals over Gaussian basis functions, and an investigation of soliton dynamics in *trans*-polyacetylene, to applications of quantum molecular similarity measures.

It is our hope that the collection of surveys of various parts of quantum chemistry and its advances presented here will prove to be valuable and stimulating, not only to the active research workers but also to the scientists in neighboring fields of physics, chemistry, and biology who are turning to the elementary particles and their behavior to explain the details and innermost structure of their experimental phenomena.

PER-OLOV LÖWDIN

ASPECTS OF NON-LINEAR-OPTICAL CALCULATIONS

DAVID M. BISHOP

Department of Chemistry
University of Ottawa
Ottawa, Canada

Table of Contents

1. INTRODUCTION
2. DEFINITIONS AND UNITS
3. ELECTRIC FIELDS
 - 3.1. Electronic Hyperpolarizabilities
 - 3.1.1. Calibration-quality calculations on small systems
 - 3.1.2. *Ab initio* calculations on medium sized systems
 - 3.2. Vibrational Hyperpolarizabilities
 - 3.3. Dispersion Formulas
4. MAGNETIC FIELDS
5. FUTURE DIRECTIONS

1. INTRODUCTION

I start, not with an apology, but with an explanation. This book is entitled "Advances in ..." not "Reviews of ...". The reader therefore should not expect to find a compendium of recently published papers of no-matter-what quality but rather a perspective, necessarily with bias, of a given sub-discipline. My intention is to indicate, in a personal way† (and I shall break the rules by using the first person singular), where we are at, as quantum chemists, in the so-called 'hot' field of non-linear optics. My reader of choice is a post-doctoral fellow who is wondering what to work on next. For others, and my friends, my words may simply form an idiosyncratic essay.

Let me start by saying why there is a *practical* interest in non-linear optics. There are two aspects to this. The first is the development of optical signal-processing devices such as the extension of the frequency range of a laser. If high energy light from a semi-conductor infrared laser could be converted to the visible, then the density of information recorded (optical data storage) could be greatly increased. An example of this would be second harmonic generation (SHG) and the corresponding theoretical property, the first hyperpolarizability $\beta(-2\omega; \omega, \omega)$. The second practical aspect is in telecommunication transmission as in electro-optical modulators and switches, where information is coded on a low energy carrier by placing a dc electric field across the material. The theoretical basis is the induced change in the refractive index (the Pockels effect) and the relevant theoretical property is $\beta(-\omega; \omega, 0)$. Developments in this area could lead to all-optical computers.

These practical aspects are governed by the interaction of electromagnetic radiation with matter; at the molecular level this interaction polarizes the charge distribution and alters the propagated field. The linear response is described by the polarizability (α) and the non-linear response (the subject of this article) is described by the hyperpolarizabilities (β, γ , etc.). The word hyperpolarizability was first used by Coulson, Macoll and Sutton [1] some 40 years ago.

Currently-used non-linear-optical materials are single inorganic crystals and they have obvious drawbacks such as high manufacturing costs. The drive is

† Even a cursory reading of the first volume of the *Journal of Chemical Physics* will show how much we have lost over the last sixty years in the conveyance of science in a personal manner.

therefore to develop cheap organic materials (e.g. long chain polymers) which will exhibit appreciable non-linear optical effects. Quantum chemistry can help in this task by leading to an understanding of the processes involved, defining the molecular characteristics which are required, and pre-screening any proposed candidates for synthesis.

It is apparent that non-linear-optical processes rely on a dynamic or frequency-dependent property. I will therefore, in general, restrict this article to calculations made at this level and, with one or two exceptions, I will consider only *ab initio* theory. Much work has been done on the static hyperpolarizabilities (as well as the static and dynamic polarizability α) but in order to make a *direct* connection with experiment, I have chosen to exclude this work. Also, again with one or two exceptions, I will only deal with molecules and exclude atoms.

One of the hurdles in this field is the plethora of definitions and abbreviations: in the next section I will attempt to tackle this problem. There then follows a review of calculations of non-linear-optical properties on small systems (He, H₂, D₂), where quantum chemistry has had a considerable success and to the degree that the results can be used to calibrate experimental equipment. The next section deals with the increasing number of papers on *ab initio* calculations of frequency-dependent first and second hyperpolarizabilities. This is followed by a sketch of the effect that electric fields have on the nuclear, as opposed to the electronic, motions in a molecule and which leads, in turn, to the vibrational hyperpolarizabilities (a detailed review of this subject has already been published [2]). Section 3.3. is a brief look at the dispersion formulas which aid in the comparison of hyperpolarizabilities obtained from different processes.

The poor sister of electric-field-induced non-linear optics is that of magnetic-field-induced non-linear optics, of which the Cotton-Mouton effect is the most important example; attention is drawn to this in Section 4. Finally, some thoughts on the future of quantum chemistry in this subject are given in Section 5. It is an area which is a tough test of computational quantum chemistry but recent achievements show it to be one well worth pursuing.

Previous complementary review articles are: Buckingham [3], Buckingham and Orr [4], Bogaard and Orr [5], and Dykstra, Liu and Malik [6]; and as well there is the first issue of Volume 43 of the *International Journal of*

Quantum Chemistry (1992) which is entirely devoted to molecular non-linear optics. The foundations of time-dependent coupled-perturbed Hartree Fock theory, referred to in Section 3.1.2, are beautifully described in Sections 12.5-12.7 of Roy McWeeny's book [7].

2. DEFINITIONS AND UNITS

Since non-linear optics has been developed by both physicists and chemists, and by both theoreticians and experimentalists, there are, as might be expected, differences in names, abbreviations and definitions of the pertinent properties, as well as the use of different units. In this section I hope to both clarify and simplify the situation. Let us start with the usual expansion [3] of the change in energy of a species when it is subjected to a static uniform electric field (F) and a static uniform magnetic field (B); strictly speaking, B is the magnetic flux density. We can write:

$$\begin{aligned} \Delta E = & -\mu_{\alpha}F_{\alpha} - \frac{1}{2}\alpha_{\alpha\beta}F_{\alpha}F_{\beta} - \frac{1}{6}\beta_{\alpha\beta\gamma}F_{\alpha}F_{\beta}F_{\gamma} - \frac{1}{24}\gamma_{\alpha\beta\gamma\delta}F_{\alpha}F_{\beta}F_{\gamma}F_{\delta} + \dots \\ & - m_{\alpha}B_{\alpha} - \frac{1}{2}\chi_{\alpha\beta}B_{\alpha}B_{\beta} - \frac{1}{6}Y_{\alpha\beta\gamma}B_{\alpha}B_{\beta}B_{\gamma} - \frac{1}{24}\zeta_{\alpha\beta\gamma\delta}B_{\alpha}B_{\beta}B_{\gamma}B_{\delta} + \dots \\ & - G_{\alpha\beta}F_{\alpha}B_{\beta} - \frac{1}{2}\xi_{\alpha\beta\gamma}F_{\alpha}B_{\beta}B_{\gamma} - \frac{1}{2}X_{\alpha\beta\gamma}F_{\alpha}F_{\beta}B_{\gamma} - \frac{1}{4}\eta_{\alpha\beta\gamma\delta}F_{\alpha}F_{\beta}B_{\gamma}B_{\delta} + \dots \end{aligned} \quad (1)$$

Here, μ_{α} and m_{α} are the electric and magnetic dipole moment functions, $\alpha_{\alpha\beta}$, $\beta_{\alpha\beta\gamma}$, $\gamma_{\alpha\beta\gamma\delta}$ are the polarizability and first and second hyperpolarizabilities, $\chi_{\alpha\beta}$ is the magnetizability, and, of the other terms, only the hypermagnetizability $\eta_{\alpha\beta\gamma\delta}$ will be of interest (it relates to the Cotton-Mouton effect). The Greek subscripts α , β , ... denote vector or tensor quantities and can be equal to the Cartesian coordinates x, y, or z. Einstein summation over these subscripts is implied both here and elsewhere. Differentiation of this expression with respect to F (or B) leads to an expression for the dipole moment (or polarization) of the species in the presence of the perturbing fields and it is clear that β , γ , η , etc. will govern the non-linear terms in the induced electric (or magnetic) dipole moment - hence, *non-linear* optics.

Table 1. Names and abbreviations.				
	Property	Process	Abbreviation	Alternative
1.	$\beta(0; \omega, -\omega)$	Optical Rectification	OR	
2.	$\beta(-2\omega; \omega, \omega)$	Second Harmonic Generation	SHG	
3.	$\gamma(-\omega; \omega, \omega, -\omega)$	Intensity Dependent Refractive Index	IDRI	Degenerate Four Wave Mixing (DFWM)
4.	$\gamma(-\omega_1; \omega_1, \omega_2, -\omega_2)$	ac-Kerr effect	ac-K	Optical Kerr Effect (OKE)
5.	$\gamma(-3\omega; \omega, \omega, \omega)$	Third Harmonic Generation	THG	
6.	$\beta(-\omega; \omega, 0)$	dc-Pockels effect	dc-P	Electro-optical Pockels Effect (EOPE)
7.	$\gamma(-\omega; \omega, 0, 0)$	dc-Kerr effect	dc-K	Electro-optical Kerr Effect (EOKE)
8.	$\gamma(0; \omega, -\omega, 0)$	dc-Optical Rectification	dc-OR	Electric Field Induced Optical Rectification (EFIOR)
9.	$\gamma(-2\omega; \omega, \omega, 0)$	dc-Second Harmonic Generation	dc-SHG	Electric Field Induced Second Harmonic Generation (EFISH)
10.	$\eta(-\omega; \omega, 0, 0)$	Cotton-Mouton Effect	CME	

special cases of process 4.

The Kerr and Cotton-Mouton experiments stand apart from the others in the sense that it is the electric-field-induced *anisotropy* in the refractive index which is measured (i.e. the difference in two quantities) rather than a single quantity such as an intensity. This means that, e.g., for the Cotton-Mouton effect, the experimentally-useful quantity is

$$\Delta\eta = (3\eta_{\alpha\beta\alpha\beta} - \eta_{\alpha\alpha\beta\beta})/15 \quad (3)$$

A similar formula for the Kerr experiment will be given later.

Experimentally-determined properties naturally relate to laboratory axes and the connection to the above properties, tacitly assumed to be defined in relation to molecular axes, is made by taking classical orientational averages (in quantum-mechanical terms we would be taking into account rotation) and this has been discussed in detail in Ref. [2]. For example, for $\langle\gamma\rangle_{ZZZZ}$ (commonly referred to as $\bar{\gamma}$), if Z is a space fixed axis and k_α is the cosine of the angle between Z and the molecular axis α , we have

$$\begin{aligned} \langle\gamma\rangle_{ZZZZ} &= (4\pi)^{-1} \int_0^{2\pi} \int_0^\pi \gamma_{\alpha\beta\gamma\delta} k_\alpha k_\beta k_\gamma k_\delta \sin\theta \, d\theta \, d\phi \\ &= (\gamma_{\xi\xi\eta\eta} + \gamma_{\xi\eta\xi\eta} + \gamma_{\xi\eta\eta\xi})/15 \end{aligned} \quad (4)$$

A similar prescription for β leads to zero, but this is not the relevant experimental quantity. The useful one is obtained from the perturbation-theoretic expression for β in an applied field (F):

$$\bar{\gamma} = (\bar{\beta}_{ZZZ}^F / F)_{F \rightarrow 0} = \langle\gamma\rangle_{ZZZZ} + \mu\bar{\beta} / 3kT \quad (5)$$

and this leads to

$$\bar{\beta} = (\beta_{z\xi\xi} + \beta_{\xi z\xi} + \beta_{\xi\xi z})/5 \quad (6)$$

where z is the molecule-fixed electric dipole axis. If one is to be pedantic $\bar{\beta}$ is not a *mean* or *average* hyperpolarizability. Notice that with Ward and

co-workers' definitions, equations (4) and (6) will be multiplied by $\frac{1}{5}$ and $\frac{1}{2}$, respectively, and equation (5) will be replaced by

$$\bar{\gamma} = \langle \gamma \rangle_{ZZZZ} + \mu \bar{\beta} / 9kT \quad (7)$$

For the dc-Pockels and dc-Kerr effects, we have

$$\beta_{\parallel} = \langle \beta \rangle_{ZZZ} = [\beta_{\xi\xi z}(-\omega; \omega, 0) + 2\beta_{\xi z\xi}(-\omega; \omega, 0)]/5 \quad (8)$$

$$\beta_{\perp} = \langle \beta \rangle_{XXZ} = [2\beta_{\xi\xi z}(-\omega; \omega, 0) - \beta_{\xi z\xi}(-\omega; \omega, 0)]/5 \quad (9)$$

$$\gamma_{\parallel} = \langle \gamma \rangle_{ZZZZ} = [\gamma_{\xi\xi\eta\eta}(-\omega; \omega, 0, 0) + 2\gamma_{\xi\eta\xi\eta}(-\omega; \omega, 0, 0)]/15 \quad (10)$$

$$\gamma_{\perp} = \langle \gamma \rangle_{XXZZ} = [2\gamma_{\xi\xi\eta\eta}(-\omega; \omega, 0, 0) - \gamma_{\xi\eta\xi\eta}(-\omega; \omega, 0, 0)]/15 \quad (11)$$

$$\beta^K = (3/2) (\langle \beta \rangle_{ZZZ} - \langle \beta \rangle_{XXZ}) \quad (12)$$

and

$$\gamma^K = (3/2) (\langle \gamma \rangle_{ZZZZ} - \langle \gamma \rangle_{XXZZ}) \quad (13)$$

The latter two terms (β^K and γ^K) are the ones which are related to the measurable quantities in the dc-Pockels and dc-Kerr experiments. In practice both β^K and γ^K are obtained from the temperature-dependent analysis of the Kerr experiment, hence the notation β^K rather than $\beta^{\text{dc-P}}$.

It is evident that molecular symmetry will drastically reduce the number of independent, non-zero components of the β , γ , and η tensors: tables in Refs. [3] and [5] say how many there are; in general, the 'mean' or average values will involve only a few components. Depending on the process, there will also be equalities arising from interchanging subscripts α , β , ... along with their concomitant frequencies.

A frequently-used approximation is to invoke Kleinman symmetry [9]; this holds rigorously true for static fields, is reasonably good for low frequencies and is totally unacceptable near a resonance. The Kleinman relations are:

$$\beta_{\xi\xi z} = \beta_{\xi z\xi} = \beta_{z\xi\xi} \quad (\xi \neq z) \quad (14)$$

and

$$\gamma_{\xi\xi\eta\eta} = \gamma_{\eta\eta\xi\xi} = \gamma_{\xi\eta\eta\xi} = \gamma_{\eta\xi\xi\eta} \quad (\xi \neq \eta) \quad (15)$$

and in Eq. (14) z is the molecule-fixed electric dipole axis. These relations are usually introduced in order to cut down on computational costs.

As for units, I think it is fair to say that for the theoretician the only acceptable ones are atomic units and for the experimentalist the only acceptable ones are the *Système Internationale*. However, the cgs-esu system dies a slow death and for the sake of making comparisons, but by no means condoning the use of the cgs-esu system, I give in Table 2 the relations between all commonly used units in this subject. The fundamental constants used were taken from Cohen and Taylor [10].

Table 2. Units and equivalences.			
Property	Atomic unit	SI equivalent	cgs-esu equivalent
β	$e^3 a_0^3 E_h^{-2}$	3.20636×10^{-53} $C^3 m^3 J^{-2}$	8.63922×10^{-33} $Fr^3 cm^3 erg^{-2}$
γ	$e^4 a_0^4 E_h^{-3}$	6.23538×10^{-65} $C^4 m^4 J^{-3}$	5.03670×10^{-40} $Fr^4 cm^4 erg^{-3}$
η	$e^4 a_0^4 m_e^{-1} E_h^{-2}$	2.98425×10^{-52} $C^2 m^2 J^{-1} T^{-2}$	2.68211×10^{-44} $Fr^2 cm^2 erg^{-1} G^{-2}$

3. ELECTRIC FIELDS

3.1. Electronic Hyperpolarizabilities

Since rotation is handled, for the most part, classically through orientational averaging (a quantum mechanical treatment is discussed in [2]), we are left, in these processes, with the effects of electromagnetic radiation on the vibrational and electronic motions. The former concern, relatively new, is discussed in Section 3.2. and the latter we consider in this section in two parts. The first concerns those computations which are of sufficient quality that we can

justifiably use them to calibrate the experimental measurements. In fact, this is exactly what is done when it comes to the dc-SHG experiments, where ratios of intensities of two species are measured (absolute intensities being notoriously difficult to come by) and these, in turn, are related directly to the ratios of the hyperpolarizabilities. One absolute (theoretical) γ value is then needed for conversion of these ratios to absolute values. Our helium results, reproduced here, are the usually-chosen ‘blank’.

In the second part of this section (3.1.2.), with one exception, I will limit my survey to other calculations which have used *ab initio* techniques to determine *frequency-dependent* hyperpolarizabilities. It is unfortunate that, again with one exception, none of these calculations takes account of nuclear vibrations, not even to the extent of zero-point vibrational averaging (i.e. a fixed nuclear geometry is assumed). Any close agreement with experiment, which doesn’t happen often, must therefore be considered coincidental. To redress (somewhat) the balance of this section, I will also report on an excellent paper dealing with a set of non-*ab-initio* calculations.

3.1.1. Calibration-quality calculations on small systems

To obtain hyperpolarizabilities of calibration quality, a number of standards must be met. The wavefunctions used must be of the highest quality and include electronic correlation. The frequency dependence of the property must be taken into account from the start and not be simply treated as an *ad hoc* ‘add-on’ quantity. Zero-point vibrational averaging coupled with consideration of the Maxwell-Boltzmann distribution of populations amongst the rotational states must also be included. The effects of the electric fields (static and dynamic) on nuclear motion must likewise be brought into play (the results given in this section include these effects, but exactly how will be left until Section 3.2.). All this is obviously a tall order and can (and has) only been achieved for the simplest of species: He, H₂, and D₂. Comparison with dilute gas-phase dc-SHG experiments on H₂ and D₂ (with the helium theoretical values as the standard) shows the challenge to have been met.

Though the published calculations deal with several different non-linear optical processes, it will be simpler, and the points can still be made, to restrict ourselves to dc-SHG and dc-Kerr processes. The helium results were obtained by

Bishop and Pipin [11] and the H_2 (D_2) ones by Bishop, Pipin, and Cybulski [12]. The starting off point in all cases is the set of perturbation-theoretic sum-over-states formulas given by Orr and Ward [8] (I include for completeness those for α and β , though only that for γ is germane at this point):

$$\alpha_{\alpha\beta}(-\omega_\sigma; \omega_1) = \hbar^{-1} \sum_P \sum'_m \langle 0 | \hat{\mu}_\alpha | m \rangle \langle m | \hat{\mu}_\beta | 0 \rangle (\omega_m - \omega_\sigma)^{-1} \quad (16)$$

$$\beta_{\alpha\beta\gamma}(-\omega_\sigma; \omega_1, \omega_2) = \hbar^{-2} \sum_P \sum'_{m,n} \langle 0 | \hat{\mu}_\alpha | m \rangle \langle m | \bar{\mu}_\beta | n \rangle \langle n | \hat{\mu}_\gamma | 0 \rangle \times (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_2)^{-1} \quad (17)$$

$$\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma; \omega_1, \omega_2, \omega_3) = \hbar^{-3} \sum_P \left[\sum'_{m,n,k} \langle 0 | \hat{\mu}_\alpha | m \rangle \langle m | \bar{\mu}_\beta | n \rangle \langle n | \bar{\mu}_\gamma | k \rangle \langle k | \hat{\mu}_\delta | 0 \rangle \times (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_2 - \omega_3)^{-1} (\omega_k - \omega_3)^{-1} - \sum'_{m,n} \langle 0 | \hat{\mu}_\alpha | m \rangle \langle m | \hat{\mu}_\beta | 0 \rangle \langle 0 | \hat{\mu}_\gamma | n \rangle \langle n | \hat{\mu}_\delta | 0 \rangle \times (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_3)^{-1} (\omega_n + \omega_2)^{-1} \right] \quad (18)$$

In these equations: $\omega_\sigma = \sum \omega_i$, $\bar{\mu}_\alpha$ is the α component of the fluctuation electric dipole moment operator, $\bar{\mu}_\alpha = \hat{\mu}_\alpha - \langle 0 | \hat{\mu}_\alpha | 0 \rangle$, which in fact, for the cases at hand, is simply $\hat{\mu}_\alpha$. \sum_P in Eq. (16) infers a sum over the two permutations of the pairs $(-\omega_\sigma, \hat{\mu}_\alpha)$ and $(\omega_1, \hat{\mu}_\beta)$ and in Eq. (17) to a sum over the six permutations of the pairs $(-\omega_\sigma, \hat{\mu}_\alpha)$, $(\omega_1, \hat{\mu}_\beta)$, and $(\omega_2, \hat{\mu}_\gamma)$ and in Eq. (18) to a sum over the 24 permutations of the pairs $(-\omega_\sigma, \hat{\mu}_\alpha)$, $(\omega_1, \hat{\mu}_\beta)$, $(\omega_2, \hat{\mu}_\gamma)$, and $(\omega_3, \hat{\mu}_\delta)$. The primes on the sums indicate the exclusion of the ground state (0), and ω_m is the circular, electronic transition frequency between the m^{th} excited state, with wavefunction $|m\rangle$, and the ground state.

It should be remarked that these expressions for α , β , and γ are all invariant to frequency phase change. This is easy to see for α , β , and the first part of γ - simply interchange the frequency/operator pairs so that the operators in the numerator are in reverse order and then change the indices m, n, k . For the second part of γ , it requires writing out the 24 terms and making a number of

algebraic manipulations.

To apply these equations, we need the wavefunctions $|m\rangle$ in order to get the dipole moment transition elements and the frequencies ω_m . In order that we need not worry about the continuum in the summations, we use what is called a pseudo-spectral series, where only the ground state need be near-exact. This is done by diagonalizing the Hamiltonian matrix formed from a large number of basis functions (which implicitly include the interelectronic coordinate and thus electron correlation). We do this for each symmetry state that is involved. All the ensuing eigenvalues and eigenvectors are then used in the sum-over-states expressions. For helium we require S, P, and D states and for H_2 (or D_2) Σ , Π , and Δ states.

The helium basis functions we have used were based on an ansatz of Thakkar and Smith [13,14] and which we have extended by additional types of functions in the D states: our generalized formula for the wavefunctions is:

$$\Psi^L(r_1, r_2, r_{12}, \theta_1, \theta_2, \phi_1, \phi_2) = \sum_{l_1, l_2} \sum_{k=1}^{N^{l_1 l_2}} C_k^{l_1 l_2} (1 + P_{12}) r_1^{l_1} r_2^{l_2} \mathcal{Y}_{L l_1 l_2}^M(\hat{r}_1, \hat{r}_2) \times \exp(-\alpha_k^{l_1 l_2} r_1 - \beta_k^{l_1 l_2} r_2 - \gamma_k^{l_1 l_2} r_{12}) \quad (19)$$

where, in vector-coupling notation:

$$\mathcal{Y}_{L l_1 l_2}^M(\hat{r}_1, \hat{r}_2) = \sum_{m_1, m_2} (l_1 m_1 l_2 m_2 | l_1 l_2 L M_L) Y_{l_1}^{m_1}(\hat{r}_1) Y_{l_2}^{m_2}(\hat{r}_2) \quad (20)$$

and r_1 , θ_1 , and ϕ_1 are the coordinates of electron 1; r_{12} is the interelectronic separation and Y_l^m are spherical harmonics. The non-linear parameters were chosen, in a pseudo-random way, according to:

$$\alpha_k^{l_1 l_2} = (A_2^{l_1 l_2} - A_1^{l_1 l_2}) \llcorner \frac{1}{2} k(k+1) \sqrt{2} \gg + A_1^{l_1 l_2} \quad (21)$$

$$\beta_k^{l_1 l_2} = (B_2^{l_1 l_2} - B_1^{l_1 l_2}) \llcorner \frac{1}{2} k(k+1) \sqrt{3} \gg + B_1^{l_1 l_2} \quad (22)$$

$$\gamma_k^{l_1 l_2} = (G_2^{l_1 l_2} - G_1^{l_1 l_2}) \llcorner \frac{1}{2} k(k+1) \sqrt{5} \gg + G_1^{l_1 l_2} \quad (23)$$

where $\langle x \rangle$ denotes the fractional part of x . The parameter ranges, A , B , and G , were found by optimizing those with $(l_1, l_2) = (0, 0)$ so that the energy is minimized, $(l_1, l_2) = (0, 1)$ so that $\alpha(0; 0)$ is maximized, and $(l_1, l_2) = (0, 2)$ and $(1, 1)$ so that $\gamma(0; 0, 0, 0)$ is maximized. Convergence of α with respect to the number of basis functions ($N = N^{00} = N^{01}$) is seen in Table 3 and there is excellent agreement with experiment [15]. Table 4 shows a comparison with previous

Table 3. Helium results (atomic units).		
N	E	$\alpha_{zz}(0)$
100	-2.903 724 3748	1.383 192
120	-2.903 724 3758	1.383 192
140	-2.903 724 3758	1.383 192
Experiment [15]		1.383 20 \pm 0.00007

Table 4. Helium results (atomic units).	
Author	$\gamma_{zzzz}(0)$
Buckingham and Hibbard [16]	43.10
Grasso <i>et al.</i> [17]	42.81
Sitz and Yaris [18]	42.6
Klingbeil [19]	42.8
Jaszuński and Roos [20]	43.7
Boyle <i>et al.</i> (expt.) [21]	51.6 \pm 8.6
Bishop and Pipin [11]	43.104

static values of $\gamma_{zzzz}(0; 0, 0, 0)$ [16-21]. Of key interest is the difference between our result and that of Sitz and Yaris [18], since for many years their *dynamic* value was used to calibrate dc-SHG experiments. Values of all the components of γ for a range of frequencies as well as for certain standard laser frequencies were determined for several processes. A general dispersion formula was then constructed; this will, more conveniently, be discussed in Section 3.3. There are two comparisons with experiment which we can make: the ratio of $\gamma_{\parallel}/\gamma_{\perp} =$

$\gamma_{zzzz}/\gamma_{xxzz}$ for dc-SHG at $\lambda = 4880 \text{ \AA}$ is known to be 2.949 a.u. [22], the calculated value is 2.950 a.u. A dc-Kerr experiment [23] at 5145 \AA gives γ^K , see Eq. (13), to be $47.2 \pm 3 \text{ a.u.}$, the calculated value is 44.73 a.u.

For H_2 and D_2 the procedure is much the same but there is more experimental information with which we can compare. Explicitly electron-correlated functions of the James-Coolidge type were chosen as basis functions. They were defined in terms of elliptical coordinates as:

$$\psi_i^\sigma(1,2) = \exp[-\alpha(\xi_1 + \xi_2)] \xi_1^{m_i} \xi_2^{n_i} \eta_1^{k_i} \eta_2^{l_i} \rho^{q_i} \quad (24)$$

$$\psi_i^{\pi(x)}(1,2) = \psi_i^\sigma(1,2) (\xi_1^2 - 1)^{\frac{1}{2}} (1 - \eta_1^2)^{\frac{1}{2}} \cos \phi_1 \quad (25)$$

$$\psi_i^{\pi(y)}(1,2) = \psi_i^\sigma(1,2) (\xi_1^2 - 1)^{\frac{1}{2}} (1 - \eta_1^2)^{\frac{1}{2}} \sin \phi_1 \quad (26)$$

$$\psi_i^{\delta(x^2-y^2)}(1,2) = \psi_i^\sigma(1,2) (\xi_1^2 - 1) (1 - \eta_1^2) \cos 2\phi_1 \quad (27)$$

$$\psi_i^{\delta'(x^2-y^2)}(1,2) = \psi_i^\sigma(1,2) (\xi_1^2 - 1)^{\frac{1}{2}} (1 - \eta_1^2)^{\frac{1}{2}} (\xi_2^2 - 1)^{\frac{1}{2}} (1 - \eta_2^2)^{\frac{1}{2}} \cos(\phi_1 + \phi_2) \quad (28)$$

$$\psi_i^{\delta(xy)}(1,2) = \psi_i^\sigma(1,2) (\xi_1^2 - 1) (1 - \eta_1^2) \sin 2\phi_1 \quad (29)$$

$$\psi_i^{\delta'(xy)}(1,2) = \psi_i^\sigma(1,2) (\xi_1^2 - 1)^{\frac{1}{2}} (1 - \eta_1^2)^{\frac{1}{2}} (\xi_2^2 - 1)^{\frac{1}{2}} (1 - \eta_2^2)^{\frac{1}{2}} \sin(\phi_1 + \phi_2) \quad (30)$$

where $\rho = 2r_{12}/R$ and R is the internuclear separation. The state wavefunctions were then taken as the following linear combinations:

$$\Psi[\Sigma_g^+] = \sum_{i=1}^{249} a_i (1 + P_{12}) \psi_i^\sigma(1,2) \quad k_i + l_i \text{ even} \quad (31)$$

$$\Psi[\Sigma_u^+] = \sum_{i=1}^{113} b_i (1 + P_{12}) \psi_i^\sigma(1,2) \quad k_i + l_i \text{ odd} \quad (32)$$

$$\left. \begin{aligned} \Psi[\Pi_u^{(x)}] &= \sum_{i=1}^{190} c_i (1 + P_{12}) \psi_i^{\pi(x)}(1,2) \\ \Psi[\Pi_u^{(y)}] &= \sum_{i=1}^{190} c_i (1 + P_{12}) \psi_i^{\pi(y)}(1,2) \end{aligned} \right\} k_i + l_i \text{ even} \quad (33)$$

$$\left. \begin{aligned} \Psi[\Pi_g^{(x)}] &= \sum_{i=1}^{214} d_i (1+P_{12}) \psi_i^{\pi(x)}(1,2) \\ \Psi[\Pi_g^{(y)}] &= \sum_{i=1}^{214} d_i (1+P_{12}) \psi_i^{\pi(y)}(1,2) \end{aligned} \right\} k_i + l_i \text{ odd} \quad (34)$$

$$\begin{aligned} \Psi[\Delta_g^{(x^2-y^2)}] &= \sum_{i=1}^{190} e_i (1+P_{12}) \psi_i^{\delta(x^2-y^2)}(1,2) \\ &+ \sum_{i=1}^{36} e_i' (1+P_{12}) \psi_i^{\delta'(x^2-y^2)}(1,2) \quad k_i + l_i \text{ even} \end{aligned} \quad (35)$$

$$\Psi[\Delta_g^{(xy)}] = \sum_{i=1}^{190} e_i (1+P_{12}) \psi_i^{\delta(xy)}(1,2) + \sum_{i=1}^{36} e_i' (1+P_{12}) \psi_i^{\delta'(xy)}(1,2) \quad (36)$$

where P_{12} is the electron permutation operator. The non-linear parameter α was found by optimization of the ground state energy for each chosen internuclear separation (R) and the powers m , n , k , l , and q were chosen in a rational way such as to achieve convergence for each symmetry state.

With each value of R , frequency, process, and component, Eq. (18) was evaluated and the results used to form the following orientational averages:

$$\gamma_{\parallel}^{\text{dc-K}}(R) = (8\gamma_{xxxx} + 8\gamma_{xzzx} + 2\gamma_{xxzz} + 2\gamma_{zzxx} + 3\gamma_{zzzz}) / 15 \quad (37)$$

$$\gamma_{\perp}^{\text{dc-K}}(R) = (\gamma_{xxxx} + 5\gamma_{xxyy} + 4\gamma_{xxzz} + 4\gamma_{zzxx} - 4\gamma_{xzzx} + \gamma_{zzzz}) / 15 \quad (38)$$

$$\begin{aligned} \gamma_{\parallel}^{\text{dc-SHG}}(R) &= (8\gamma_{xxxx} + 2\gamma_{xxzz} + 2\gamma_{zzxx} + 4\gamma_{xzzx} \\ &+ 4\gamma_{zxzx} + 3\gamma_{zzzz}) / 15 \end{aligned} \quad (39)$$

$$\begin{aligned} \gamma_{\perp}^{\text{dc-SHG}}(R) &= (\gamma_{xxxx} + 5\gamma_{xxyy} + 4\gamma_{xxzz} + 4\gamma_{zzxx} \\ &- 2\gamma_{xzzx} - 2\gamma_{zxzx} + \gamma_{zzzz}) / 15 \end{aligned} \quad (40)$$

These values were, in turn, averaged over the manifold of rovibrational wavefunctions belonging to the vibrational ground state. Finally, a mean value was found on the assumption that there is a Maxwell-Boltzmann distribution of molecules in the rotational levels (for this the temperature 295K was assumed).

Pure vibrational contributions (to be discussed in Section 3.2.) were also added to these numbers. D₂ as well as H₂ was considered.

We are now in a position to compare *directly* with experiment: it has always been my philosophy to calculate as nearly as possible what is actually measured, in this case the ratios of hyperpolarizabilities for dc-SHG and differences for dc-Kerr. Tables 5 and 6 show the high degree of agreement between theory and experiment [22,24-26]. In Table 6 the 'fitted' experimental values are those obtained by smoothing out the experimental errors by essentially

Table 5. H ₂ and D ₂ results for $R = \gamma_{\parallel}^{\text{dc-SHG}}(\text{total})/\gamma_{\perp}^{\text{dc-SHG}}(\text{total})$.				
$\lambda(\text{\AA})$	R(H ₂)		R(D ₂)	
	Theory	Expt.	Theory	Expt.
5900	2.896	2.890	2.907	2.909
5145	2.873	2.867	2.881	2.881
4880	2.860	2.858	2.868	2.872

Table 6. H ₂ and D ₂ results for ratios of total $\gamma_{\parallel}^{\text{dc-SHG}}$ terms.					
$\lambda(\text{\AA})$	H ₂ / D ₂		H ₂ / He		
	Theory	Expt.	Theory	Expt.	Fitted Expt.
13190	0.988	0.987	15.10	15.10	15.06
10640	1.003	0.998	15.81	15.77	15.80
6943	1.024	1.023	17.48	17.54	17.50
6328	1.026	1.022	17.99	17.88	18.03
5900	1.028	1.029	18.44	18.62	18.51
5145	1.032	1.032	19.58	19.81 *	19.72
4880	1.033	1.030	20.14	20.20	20.31
* A new experiment at this frequency gives 19.57 (D.P. Shelton, private communication, 1992).					

putting a curve through the experimental points. The experimental data *can* be extrapolated to zero frequency (as can the equivalent theoretical data) and comparison made with the *directly* computed static quantity $\bar{\gamma}(0; 0, 0, 0)$; this will be discussed in Section 3.3., when we deal with dispersion formulas.

The dc-Kerr experimental data [23,27,28] is not such a happy story; there are obvious differences between the experiments and no agreement with the theory (see Table 7). Furthermore, a general dispersion formula shows that the

Table 7. H ₂ results for dc-Kerr: $\gamma^K = \frac{3}{2} [\gamma_{\parallel}(\text{total}) - \gamma_{\perp}(\text{total})]$, in atomic units.		
$\lambda(\text{\AA})$	theory	expt.
6328	726	550 \pm 60 [27]
		666 \pm 19 [28]
5145	759	820 \pm 49 [23]

dc-SHG and Kerr experimental data are incompatible. By such a dispersion formula we have also predicted the ac-Kerr hyperpolarizabilities for H₂ and D₂, without making a direct calculation, and, as will be seen in Section 3.3., the agreement between theory and this experiment is much better. This suggests that the dc-Kerr experiments should be re-assessed.

[A recent, but as yet unpublished, dc-Kerr experiment made by D.P. Shelton at $\lambda = 6328 \text{ \AA}$ finds γ^K to be 729 ± 4 a.u. and this measurement is in very good agreement (see Table 7) with our value.]

To conclude this section, I hope that it is apparent that quantum chemistry can and *does* play an important rôle in monitoring and assessing experimental non-linear-optical data.

3.1.2. *Ab initio* calculations on medium sized systems

Though the sum-over-states formalism of the previous section can be used in semi-empirical treatments, e.g. Pierce [29], it is not appropriate in *ab initio* calculations for anything but the smallest of chemical systems. This is because of the need for a highly extensive and highly electron-correlated pseudo-spectral series. *Faute de mieux*, one must turn to less accurate *ab initio* treatments. For

static hyperpolarizabilities the way is clear - use standard electron-correlated methods (many body perturbation theory, etc.) to calculate the energy of the molecule in the presence of a uniform static electric field and then determine β , γ , etc. through Eq. (1). This is the finite field method. Equivalently, analytical derivative techniques can be applied in place of Eq. (1). But the question is how to extend such treatments so as to incorporate frequency dependence (dispersion) and obtain properties of *direct* relevance to non-linear optics. The initial approach has been to obtain the frequency-dependent properties at a lower level of theory (no electron correlation), e.g., time-dependent coupled-perturbed Hartree-Fock theory (TDCPHF or, simply, TDHF), and then use this information to adjust an accurate (electron-correlated) static value. Whether this correction should be done by scaling or simply additivity has, at this date, not been really resolved. Only very recently have both dispersion and electron correlation been treated simultaneously and then only at the level of second-order Møller-Plesset perturbation theory (MP2) [30].

It was in 1986 that Sekino and Bartlett [31] first used TDCPHF theory to evaluate frequency-dependent β 's and γ 's. There then followed a period of silence, but since 1990 there has been feverish activity, with four couples playing the major rôles: Sekino and Bartlett [32,33], Karna and Dupuis [34,35], Rice and Handy [30,36], and Parkinson and Oddershede [37].

As is often the case in any new development in quantum chemistry, the fundamental formalism has been around for some time, in this case TDCPHF based on Frenkel's variational principle and density matrices [7]. However, the detailed theory will not be our immediate concern, but rather how it has been put to use.

As I have said, Sekino and Bartlett [31] were the first to show how to proceed to calculate frequency-dependent hyperpolarizabilities within the TDCPHF approximation. They developed an infinite-order recursive procedure, using density matrices, and, by solving the equations iteratively at each order, could, in principle, calculate any non-linear optical property. Their first application was to H_2 , FH (the work on FH was analysed in detail in another paper [38]), CH_4 and the fluoromethanes. The processes SHG, OR, dc-SHG, dc-OR, IDRI and THG were considered but not all hyperpolarizability components were computed (the assumption of Kleinman symmetry was made).

Only one frequency ($\lambda = 6943 \text{ \AA}$) was used. The results for β were smaller than the experimental ones and for γ were quite far from experiment. This deficiency was blamed (as is usual in these cases) on basis-set-incompleteness and lack of electron correlation. However, for the fluoromethanes, they found that the ratios for $\bar{\gamma}(\text{dc-SHG}) / \bar{\gamma}(\text{THG})$ were in excellent agreement with experiment. It was this which doubtless inspired them to believe that dispersion effects, at least, could be adequately treated at the SCF level. If this were true, of course, the whole problem would be greatly simplified.

Following new experimental results [39], their calculation on FH was later expanded and given detailed scrutiny [38]. Values of the static β and γ , obtained from a finite field application to highly-correlated field-perturbed energies, were adjusted by the TDCPHF dispersion corrections (in this case by simple addition rather than by scaling). The effects of vibration (see Section 3.2.) were also considered. Nonetheless, their best results for $\bar{\beta}(\text{SHG})$ and $\bar{\gamma}(\text{dc-SHG})$ were still 81% and 69%, respectively, of the experimental ones. This non-meeting of experimental and theoretical minds remains to this day. One possible explanation is the presence of $(\text{FH})_2$ in the experiment, but this is mere conjecture.

Sekino and Bartlett's 1991 and 1992 papers [32,33], which include some methodological and computational improvements (though still in the 1991 paper maintaining Kleinman symmetry), report results for β and γ for CO, and γ for N_2 and *trans*-butadiene. Correlation corrections at the static level were estimated by using a second-order many-body-perturbation theory and a coupled-cluster singles and doubles relaxed density method. Correlated, frequency-dependent results were obtained by scaling the static correlated ones rather than by assuming additivity (no explanation for this change of strategy was given). For N_2 , the dc-SHG results were in good agreement with experiment. This was not the case for *trans*-butadiene, where the theoretical result (at $\lambda = 6943 \text{ \AA}$) was 40% larger than the experimental value. Both correlation and dispersion effects were large for this molecule and these circumstances have the potential to undermine the "additivity/scaling" procedure. The SHG and dc-SHG results for CO agree well with experiment (it would have been helpful if they had stated for which frequency their results pertain).

The two papers written in 1990 and 1991 by Karna and Dupuis [34,35],

the latter containing the methodology and the former application of it to the molecular series CHX_3 ($\text{X} = \text{F}, \text{Cl}, \text{Br}, \text{I}$), are based on the same principles as those used by Sekino and Bartlett [31]. They are, however, much more complete and are also models of clarity. The 1991 paper, using TDCPHF theory, gives explicit expressions for all non-linear optical processes up to third order and makes use of the $(2n+1)$ rule of perturbation theory by which knowledge of the wavefunction through order n , allows the energy to be known through order $2n+1$. Considering their applications paper [34] (β and γ for several non-linear optical processes for CHX_3), there is fortuitously good agreement with experiment in the case of $\beta(\text{SHG})$ for CHF_3 , but elsewhere fortune does not smile. There are two obvious problems: (a) in many cases the experimental data relate to the condensed phase, and (b) there is no account taken of electron correlation. Their use of Kleinman symmetry is not likely to be critical. Karna and co-workers [40-43] have continued this work with applications to (a) *p*-nitroaniline (the results are considerably smaller than experiment); (b) hexapentaene, hexadiyne, and divinylacetylene (a comparative study for the semi-empirical INDO method); (c) benzene [$\gamma(\text{dc-SHG})$ was within 5-14% of the measured value but there were greater differences for THG and IDRI]; (d) ethylene, *trans*-butadiene, *trans*-hexatriene, *trans*-octatetraene (this paper questions the use of Kleinman symmetry even at low frequencies and the results compare favourably with the dc-SHG experiments).

Almost simultaneously with the work of Karna and Dupuis, Rice *et al.* [36] gave a prescription for TDCPHF calculations of the first hyperpolarizability, the differences are only in the details; it too improves the Sekino and Bartlett approach by introducing the $(2n+1)$ theorem. To avoid solving second-order perturbation equations, two methods were described: the first involves the determination of appropriate response vectors which arise in the first-order perturbation equations at frequency ω (the method of Handy and Schaeffer [44]), and the second involves solving first-order equations at zero frequency, ω and 2ω . The effects of electron correlation were assumed to be *additive* and presumed to be adequately accounted for by finite-differencing the second-order MP2 energy corrections (these were obtained from analytic second derivatives). Applications were made for SHG and dc-P (or OR via an interchange of components) for HCHO and CH_3F . For CH_3F , their values were 71% and 81% of the

experimental ones for dc-P and SHG, respectively. For HCHO no experimental data were available; one interesting point, however, was noted: Kleinman symmetry was lost to the extent of 10% for SHG at $\hbar\omega = 0.05 E_h$.

Finally, and most recently, the polarization propagator technique has been used by Parkinson and Oddershede [37]. By evaluating the quadratic response function (QRF) in the random phase approximation (RPA), this is equivalent to the TDCPHF technique. Dalgaard [45] had previously given the formalism and expressions for the QRF but without any implementation. Computationally, Parkinson and Oddershede used the reduced-linear-equation technique whereby the RPA eigenvalue problem is avoided by solving systems of linear equations and extracting directly the first polarizabilities. The authors claim that this method had certain advantages over the usual TDCPHF method. Both dipole-length and mixed-velocity representations were used as a way of investigating basis set errors. They calculated $\beta(\text{SHG})$ for FH and H_2O but had to admit that, though dispersion was well represented, overall the results were poor. This approach, however, can be expanded by evaluating the QRF with coupled cluster reference functions and this would allow electron correlation and dispersion to be simultaneously accounted for.

This last point leads us to the latest paper by Rice and Handy [30] where electron correlation (at the MP2 level) and dispersion are treated together. The basis of the method is the introduction of a quasi-energy (this gets around the problem that we cannot write a true energy expression when an oscillating field is present but only an expression for the polarization) and it follows along the lines of their earlier exposition [46]. The quasi-energy ansatz allows the armoury of analytic differentiation to come into full force. The theory is developed for the first hyperpolarizability and the prescription leads to two terms, the first of which can be found via a finite-field static calculation and the second of which involves the frequency explicitly. Detailed expressions (extremely complex and a type-setter's nightmare) are given for $\beta(\text{SHG})$. $\beta(\text{dc-P})$ or $\beta(\text{OR})$ may be straight-forwardly obtained from static finite-field calculations on the similarly-obtained $\alpha(-\omega; \omega)$. Extension to higher-order hyperpolarizabilities has yet to be explored.

Comparison of results for HCHO with the earlier treatment (where dispersion, at the SCF level, was treated as an *additive* correction to the MP2

static value) showed that $\beta(\text{dc-P})$ was underestimated by 3% and $\beta(\text{SHG})$ by 12% at $\hbar\omega = 0.1 E_h$. For NH_3 these percentages rise to 12% and 33%, respectively. However, had dispersion (at the SCF level) been used as a *scaling* correction to the MP2 static results, there would, in fact, be perfect agreement with the MP2-frequency-dependent results for NH_3 and good agreement for HCHO (even better at lower frequencies). With more cases investigated, it will be possible to be *sure* that this latter technique is viable and even appropriate for cases where electron correlation and dispersion are both large (the latter depending, of course, on the proximity of the optical frequency to an electronic resonance). For a true test we will also need frequency-dependent values at a higher order of correlation than MP2 theory with which to make the comparison.

In summary, it seems to me that the *ab initio* techniques are far from having been fully exploited. There is no doubt that these high-order properties will not only be sensitive to basis-set selection and electron correlation, as we would expect, but also to the coupling of electron correlation with dispersion. And we are still treating molecules as if their nuclear geometry was fixed! On the other hand, it is also true that the experimental data can be perfected and certainly more gas-phase work would be useful.

In order that I do not give the false impression that I think that the *ab initio* route is the only way to go, I would like to give one example of a semi-empirical study (there are many). I choose, because of its thoroughness and elegance, the article by Pierce [29] on the non-linear optical properties of benzene and seven linear polyenes. The scope of this work and the systems treated puts it outside of current *ab initio* techniques. What Pierce does is to separate γ into σ and π components and to treat the former by bond-additivity rules. γ_π is found from the sum-over-states formula through the use of semi-empirical INDO all-valence-molecular-orbitals coupled with single- and double-excitation configuration interactions of singlet π -electron configurations. A number of revealing conclusions result: (a) a single set of INDO parameters can service both γ and the one- and two-electron absorption spectra; (b) γ is strongly influenced by chain length (something known from other studies); (c) the degree of dispersion is also sensitive to chain length; (d) γ can be defined by z-polarized virtual electronic transitions between just three states. His overall agreement with experiment is roughly 20%, which is not that much different from

the *ab initio* studies which I have just surveyed. Nonetheless, the *ab initio* studies *are* required if we want to understand (and we should) the approximations made in hyperpolarizability determinations such as those by Pierce.

Shortly after the above was written, a paper by Jaszuński *et al.* [47] on FH appeared. They used the polarization propagator technique (a Danish tradition!) and calculated the QRF using a multi-configurational-self-consistent-field reference state (the previously unimplemented formalism had been developed by Olsen and Jørgensen [48]). They also took into account zero-point-vibrational averaging. For β (SHG) at $\lambda = 6943 \text{ \AA}$, and using CAS 4220 functions (more than 125000 determinants), their final value was (in the 'Ward' convention) -7.8 a.u. or 71% of the experimental value; they considered that Sekino and Bartlett's [38] estimate to be an exaggeration of "how close to experiment theory can get". However, their own value is, in fact, within the limits that Sekino and Bartlett proposed.

3.2. Vibrational Hyperpolarizabilities

There are two ways in which molecular vibrations affect non-linear optical properties. The first, which is well understood, is zero-point-vibrational averaging of the calculated electronic properties. This need not delay us long. The second comes about from the effect that the electromagnetic radiation has on the *vibrational* motions themselves and this leads to the vibrational polarizabilities and hyperpolarizabilities which are the exact counterparts of the electronic ones which stem from the effect that the radiation has on the *electronic* motions. This phenomenon is now receiving long overdue attention and will be the main subject of this section. A more extensive review is available elsewhere [2].

For diatomic molecules, if the potential curve is known, numerical vibrational wavefunctions can be determined by the standard Numerov-Cooley technique [49] and for which computer programmes are available [50]. The vibrational averaging for any state is then simply the numerical integration of $\langle v|P|v \rangle$, where $|v \rangle$ is the vibrational wavefunction for the state of interest (usually the ground state). This, obviously, requires knowledge of the property P for a number of internuclear separations (R). Alternatively, simple perturbation

theory leads to the approximate formula:

$$\langle v|P|v \rangle = P_e + (v + \frac{1}{2}) (B_e/\omega_e) (d^2P/d\xi^2 - 3adP/d\xi) \quad (41)$$

where v is the vibrational quantum number, $\xi = (R - R_e)/R_e$ and B_e , ω_e , and a are the rotational constant, harmonic vibrational frequency, and anharmonic constant, respectively; the subscript e denotes an equilibrium value. This equation is a transparent demonstration of the fact that the vibrational averaging correction will depend on both the degree of mechanical anharmonicity (through a) and the degree of electrical anharmonicity (through $dP/d\xi$ and $d^2P/d\xi^2$). For polyatomics, perturbation theory is the *only* practical method and the relevant expressions are more complex; the reader is referred to Refs. [51-53]. Outside of H_2 , very seldom have the zero-point corrections to β and γ been made; Ref. [47] is an exception. This is a serious lacuna since the known strong dependence of these properties on geometry can mean that the corrections can be 10% or more.

Now let us consider the second effect. As long ago as 1924 [54] it was recognized that there was a vibrational contribution to the polarizability α (rather unfortunately called, at the time, an 'atomic polarizability'). This contribution can be directly related to infrared intensities and, in this fashion, semi-empirical values have been compiled [55]. For β and γ , if the fields are all static and the molecule is a diatomic, the vibrational contribution can be obtained by the finite field technique [56]. That is to say, potential curves are found for a sequence of electric field strengths. The vibrational equation is then solved numerically and repeatedly to produce *total* energies which are field dependent. Numerical differencing of these energy values gives, by analogy with Eq. (1), the total (electronic plus vibrational) polarizabilities and hyperpolarizabilities. Unfortunately, this technique is not applicable to polyatomic molecules or for frequency-dependent hyperpolarizabilities. We must therefore fall back on the direct use of perturbation theory and this has been exploited by Elliott and Ward [57], Shelton [58], Bishop [59], and Kirtman and Bishop [60-62].

Our starting point is the set of expressions in Eqs. (16-18), which are now generalized by replacing the electronic wavefunctions by vibronic wavefunctions (the exclusion of rotation is discussed in [2]). It will be sufficient, to get the idea across, to consider only β since the results for γ follow the same logic. Eq. (17) is

now written as:

$$\begin{aligned} \beta_{\alpha\beta\gamma}(-\omega_{\sigma}; \omega_1, \omega_2) \\ = \hbar^{-2} \sum_P \sum'_{m,M} \sum'_{n,N} \langle 0,0 | \hat{\mu}_{\alpha} | M,m \rangle \langle m,M | \hat{\mu}_{\beta} | N,n \rangle \langle n,N | \hat{\mu}_{\gamma} | 0,0 \rangle \\ \times (\omega_{mM} - \omega_{\sigma})^{-1} (\omega_{nN} - \omega_2)^{-1} \end{aligned} \quad (42)$$

where the vibronic wavefunctions are taken as the products $\psi_M \phi_m = |M,m\rangle$, where M and m refer, respectively, to the electronic and vibrational states. We will take the vibrational contribution to β to be those terms in Eq. (42) in which at least one of the indices M,N is zero, i.e. the electronic ground state. This allows us to write the vibrational hyperpolarizability β^v as:

$$\beta^v(-\omega_{\sigma}; \omega_1, \omega_2) = [\mu\alpha] + [\mu^3] \quad (43)$$

where $[\mu\alpha]$ is the contribution from $M=0, N \neq 0$ and $M \neq 0, N=0$ and $[\mu^3]$ from $M=0, N=0$ (the square bracket notation will make more sense in a moment). These components can be simplified by using certain approximations which allow for closure over the vibrational states. For example, we can assume that, when the electronic ground state is not involved, $\omega_{mM} \gg \omega$ (an optical frequency), and also that the vibrational energy differences are small compared to the electronic energy differences, hence $\omega_{mM} \equiv \omega_{0M} \equiv \omega_M$ ($M \neq 0$). Then, for example,

$$[\mu\alpha] = \frac{1}{2} \hbar^{-1} \sum_P \sum'_m (\mu_{\alpha})_{0m} (\alpha_{\beta\gamma})_{m0} [(\omega_m + \omega_{\sigma})^{-1} + (\omega_m - \omega_{\sigma})^{-1}] \quad (44)$$

where $(\mu_{\alpha})_{0m}$ etc. are vibrational transition matrix elements, between the lowest and the m^{th} vibrational states, of the dipole moment function (in the electronic ground state) etc. I should remark that, by using closure, $\alpha_{\beta\gamma}$ is to some extent frequency dependent. In the static limit these formulas are identical to those that could be obtained by applying the perturbation

$$E' = -\mu_{\alpha} F_{\alpha} - \frac{1}{2} \alpha_{\alpha\beta} F_{\alpha} F_{\beta} + \dots \quad (45)$$

to the Hamiltonian in the Schrödinger vibrational equation, using perturbation

theory and collecting terms in F^2 , F^3 , etc.

For diatomic molecules, expressions such as (44) can be evaluated very accurately and this has been done for H_2 , where the vibrational hyperpolarizability γ^v (β^v is zero) is:

$$\begin{aligned}\gamma_{\alpha\beta\gamma\delta}^v(-\omega_\sigma; \omega_1, \omega_2, \omega_3) &= [\alpha^2] + [\mu\beta] + [\mu^2\alpha] + [\mu^4] \\ &= [\alpha^2] \\ &= \frac{1}{4} \hbar^{-1} \sum_p \sum'_m (\alpha_{\alpha\beta})_{0m} (\alpha_{\gamma\delta})_{m0} (\omega_m - \omega_2 - \omega_3)^{-1}\end{aligned}\quad (46)$$

We have used the fact that for a homonuclear diatomic $(\mu)_{0m}$ is zero. The $(\alpha_{\alpha\beta})_{0m}$ can be evaluated numerically from values of the polarizability function at different R values and the numerical vibrational wavefunctions. Table 8 is an

Table 8. Importance of the pure vibrational component for dc-SHG for H_2 (in atomic units).			
$\hbar\omega/E_h$	$\bar{\gamma}_{ }(\text{elec})$	$\gamma_{ }(\text{vib})$	$\gamma_{ }(\text{total})$
0.00	682.7	183.7	866.4
0.01	685.3	-323.5	361.9
0.02	693.4	-1009.9	-316.5
0.03	707.1	-88.2	618.9
0.04	726.9	-43.0	683.9
0.05	753.6	-24.3	729.3

example of what we have found for $\gamma(\text{dc-SHG})$ [11] and it is fairly typical: at zero frequency (the static case) the vibrational component is very significant and increases as the optical frequency approaches a vibrational resonance, for a laser frequency of, say, $\hbar\omega = 0.05 E_h$, the effect is diminished but still important. It is obvious that dynamic vibrational hyperpolarizabilities will always be less important than the static ones since the nuclei just do not have the ability to keep up with a rapidly oscillating electric field. Incidentally, for H_2^+ and for a static field, $\gamma^e = 193$ a.u. and $\gamma^v = 2197$ a.u. [63]!

For polyatomics, in order to find the electric-property-function

Table 9. Non-zero terms in the perturbation expansion of the various components that contribute to the vibrational polarizabilities.

(n,m) =	(0,0)	(1,0)	(2,0)	(0,1)	(1,1)
$[\mu^2]$	✓		✓		✓
$[\mu\alpha]$	✓		✓		✓
$[\mu^3]$		✓		✓	
$[\alpha^2]$	✓		✓		✓
$[\mu\beta]$	✓		✓		✓
$[\mu^2\alpha]$		✓		✓	
$[\mu^4]$			✓		✓

transition-matrix elements, it is necessary to make further approximations and what now follows is based on a collaboration with Bernard Kirtman [60-62]. What we do is to use an initial harmonic-oscillator-approximation and then extend it by calculating the corrections due to mechanical anharmonicity in the vibrational wavefunctions and to electrical anharmonicity in the normal-coordinate expansion of the electrical properties. The former requires writing the vibrational potential as:

$$V = V^o + \frac{1}{2} \sum_a \lambda_a Q_a^2 + \frac{1}{6} \sum_{a,b,c} F_{abc} Q_a Q_b Q_c + \dots \quad (47)$$

where F_{abc} are the cubic anharmonic force constants and Q_a , Q_b , etc. are the normal coordinates. The latter requires, for example, writing:

$$\mu_\alpha = \mu_\alpha^o + \sum_a (\alpha/a) Q_a + \frac{1}{2} \sum_{a,b} (\alpha/ab) Q_a Q_b + \dots \quad (48)$$

where $(\alpha/a) = (\partial\mu_\alpha/\partial Q_a)$, $(\alpha/ab) = (\partial^2\mu_\alpha/\partial Q_a\partial Q_b)$, etc. There are similar equations with the dipole moment function $\mu_\alpha \equiv \alpha$ replaced by the polarizability function $\alpha_{\alpha\beta} \equiv \alpha\beta$ or the hyperpolarizability function $\beta_{\alpha\beta\gamma} \equiv \alpha\beta\gamma$. The components $[\mu^2]$ etc. can now be evaluated to various orders $[\mu^2]^{n,m}$ etc., where n and m are the orders in electrical and mechanical anharmonicity, respectively. So far we have only chosen $(n,m) = (0,0)$, $(1,0)$, $(2,0)$, $(0,1)$, and $(1,1)$ and in Table 9 a specific list of the non-zero $[\dots]^{n,m}$ is given. The most difficult and

tedious part comes next: that is to figure out all the possible intermediate vibrational states (when the anharmonicities are present) and to evaluate the integrals and, through algebraic manipulation, to simplify and combine the ensuing terms (in the worst case there are over two hundred of these). The final formulas are tantalizingly compact, making one suspect that there must be a simpler way of getting them. They are:

$$[\mu^2]^{0,0} = \frac{1}{2} \sum P_{-\sigma,1} \sum_a (\alpha/a) (\beta/a) \lambda_a^{\pm\sigma} \quad (49)$$

$$[\mu^2]^{2,0} = \frac{\hbar}{4} \sum P_{-\sigma,1} \sum_{a,b} (\alpha/ab) (\beta/ab) \omega_a^{-1} \lambda_{ab}^{\pm\sigma} \quad (50)$$

$$[\mu^2]^{1,1} = -\frac{\hbar}{4} \sum P_{-\sigma,1} \sum_{a,b,c} [2F_{abc}(\alpha/ab) (\beta/c) \omega_a^{-1} \lambda_{ab}^{\pm\sigma} \lambda_c^{\pm\sigma} + F_{aab}(\alpha/bc) (\beta/c) \omega_a^{-1} \omega_b^{-2} \lambda_c^{\pm\sigma}] \quad (51)$$

$$[\mu^3]^{1,0} = \frac{1}{2} \sum P_{-\sigma,1,2} \sum_{a,b} (\alpha/a) (\beta/ab) (\gamma/b) \lambda_a^{\pm\sigma} \lambda_b^{\pm 2} \quad (52)$$

$$[\mu^2\alpha]^{1,0} = \frac{1}{4} \sum P_{-\sigma,1,2,3} \sum_{a,b} [(\alpha/a) (\beta\gamma/ab) (\delta/b) \lambda_a^{\pm\sigma} \lambda_b^{\pm 3} + 2 (\alpha/a) (\beta/ab) (\gamma\delta/b) \lambda_a^{\pm\sigma} \lambda_b^{\pm 23}] \quad (53)$$

$$[\mu^3]^{0,1} = -\frac{1}{6} \sum P_{-\sigma,1,2} \sum_{a,b,c} F_{abc} (\alpha/a) (\beta/b) (\gamma/c) \lambda_a^{\pm\sigma} \lambda_b^{\pm 1} \lambda_c^{\pm 2} \quad (54)$$

$$[\mu^2\alpha]^{0,1} = -\frac{1}{4} \sum P_{-\sigma,1,2,3} \sum_{a,b,c} F_{abc} (\alpha/a) (\beta/b) (\gamma\delta/c) \lambda_a^{\pm\sigma} \lambda_b^{\pm 1} \lambda_c^{\pm 23} \quad (55)$$

$$[\mu^4]^{2,0} = \frac{1}{2} \sum P_{-\sigma,1,2,3} \sum_{a,b,c} (\alpha/a) (\beta/ab) (\gamma/bc) (\delta/c) \lambda_a^{\pm\sigma} \lambda_b^{\pm 23} \lambda_c^{\pm 3} \quad (56)$$

$$[\mu^4]^{1,1} = -\frac{1}{2} \sum P_{-\sigma,1,2,3} \sum_{a,b,c,d} F_{abc} (\alpha/a) (\beta/b) (\gamma/cd) (\delta/d) \lambda_a^{\pm\sigma} \lambda_b^{\pm 1} \lambda_c^{\pm 23} \lambda_d^{\pm 3} \quad (57)$$

where $\lambda_{xy}^{\pm ij} \dots$ is defined by

$$\lambda_{xy}^{\pm ij} \dots = [(\omega_x + \omega_y + \dots) + (\omega_i + \omega_j + \dots)]^{-1} \times [(\omega_x + \omega_y + \dots) - (\omega_i + \omega_j + \dots)]^{-1} \quad (58)$$

and $\omega_x, \omega_y, \dots$ are harmonic vibrational frequencies and $\omega_i, \omega_j, \dots$ are the optical frequencies. It should be noted that the sums over the normal modes (a, b, c, d) do not exclude a=b, etc. The remaining components which are required can be found by using Table 10.

Table 10. Conversion of formulas for $[\mu^2]$ to those for $[\mu\alpha]$, $[\alpha^2]$, and $[\mu\beta]$.				
Term	Starting formula	Multiply by	Permutation sum change	Change of parameters and frequencies
$[\mu\alpha]^{0,0}$	Eq.(49)	1	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2}$	$\beta \rightarrow \beta\gamma$
$[\mu\alpha]^{2,0}$	Eq.(50)	1	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2}$	$\beta \rightarrow \beta\gamma$
$[\mu\alpha]^{1,1}$	Eq.(51)	$\frac{1}{2}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2}$	Each term is replaced by two, in the first $\beta \rightarrow \beta\gamma$ and in the second $\alpha \rightarrow \beta\gamma$ and $\beta \rightarrow \alpha$
$[\alpha^2]^{0,0}$	Eq.(49)	$\frac{1}{4}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2,3}$	$\alpha \rightarrow \alpha\beta, \beta \rightarrow \gamma\delta, \sigma \rightarrow 23$
$[\alpha^2]^{2,0}$	Eq.(50)	$\frac{1}{4}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2,3}$	$\alpha \rightarrow \alpha\beta, \beta \rightarrow \gamma\delta, \sigma \rightarrow 23$
$[\alpha^2]^{1,1}$	Eq.(51)	$\frac{1}{4}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2,3}$	$\alpha \rightarrow \alpha\beta, \beta \rightarrow \gamma\delta, \sigma \rightarrow 23$
$[\mu\beta]^{0,0}$	Eq.(49)	$\frac{1}{3}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2,3}$	$\beta \rightarrow \beta\gamma\delta$
$[\mu\beta]^{2,0}$	Eq.(50)	$\frac{1}{3}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2,3}$	$\beta \rightarrow \beta\gamma\delta$
$[\mu\beta]^{1,1}$	Eq.(51)	$\frac{1}{6}$	$P_{-\sigma,1} \rightarrow P_{-\sigma,1,2,3}$	Each term is replaced by two, in the first $\beta \rightarrow \beta\gamma\delta$ and in the second $\alpha \rightarrow \beta\gamma\delta$ and $\beta \rightarrow \alpha$

So far, we have not exploited these equations very much - just some preliminary work on FH and CO₂ [61]. For FH, where the static SCF electronic $\bar{\gamma}^e$ is 298 a.u., $\gamma^v(\text{dc-K})$ is 9 a.u. and $\gamma(\text{dc-SHG})$ is -5 a.u. at $\hbar\omega = 0.07 E_h$. For CO₂, where the 'experimental' static value of $\bar{\gamma}^e$ is 1350 a.u., $\gamma^v(\text{dc-K})$ is -102 a.u. and $\gamma^v(\text{dc-SHG})$ is -87 a.u. at $\hbar\omega = 0.07 E_h$. We hope shortly to have further results for CO₂ as well as for NH₃ and H₂O.

It should be remarked that, if one wishes to go to higher orders of perturbation theory in the above development, then consistency will demand that anharmonic effects (zero at first order) be taken into account in the energy denominators.

A final comment: analytical-derivative programs are now common fare and a part of most 'packages' (e.g. CADPAC), so that evaluating (α/a) , (α/ab) ,

$(\alpha\beta/a)$, $(\alpha\beta/ab)$ and, with a little difficulty, $(\alpha\beta\gamma/a)$ and $(\alpha\beta\gamma/ab)$ should be possible. The hardest quantities to come by are the anharmonic force constants F_{abc} . It is the size of the derivatives and the anharmonic force constants which will determine the overall size of the vibrational contribution to the non-linear optical properties and therefore knowledge of their magnitudes will alert one as to the necessity, or not, of doing extended vibrational calculations.

3.3. Dispersion Formulas

A dispersion formula is a mathematical function which expresses a physical property's dependence on the optical frequencies used to measure it. The prototype is the Sellmeier dispersion formula for the polarizability α and which can be written as:

$$\alpha(-\omega; \omega) = \alpha(0; 0)(1 - C\omega^2)^{-1} \quad (59)$$

where C is a constant. There are several reasons why dispersion formulas are useful: (a) they are a compact way of synthesizing both theoretical and experimental data, (b) they can allow for interpolation and extrapolation to frequencies where data is not available (this is particularly useful, for instance, in assessing experimental dynamic hyperpolarizabilities if only a static theoretical value is known, as is often the case), (c) they can allow for a comparison of data (theoretical or experimental) on the same species but for different processes.

For α , β , and γ , changing the sign of all of the optical frequencies is equivalent to taking their conjugate complex, and, since these quantities are real, this implies that α , β , and γ are invariant to this sign change and that they are all therefore expressible in expansions in the squares of the frequencies. In fact, over the years, many different and *ad hoc* dispersion formulas (as a function of ω^2) have been introduced for β and γ . But the most acceptable form is the one which expresses the dispersion as an expansion in ω_L^2 where

$$\omega_L^2 = \omega_\sigma^2 + \sum_i \omega_i^2 \quad (60)$$

This idea was first proposed by Shelton [64] in 1985 and later shown, at least at

low frequency, to be theoretically justified [65-69]. It has the advantage of being quite general, that is process-independent.

The justification [65-68] was made by expanding Eqs. (17) and (18) in powers of the squares of the frequencies and collecting terms. It was found that for the experimentally measurable quantities ($\beta_{\parallel} = \beta_{ZZZ}$, $\gamma_{\parallel} = \gamma_{ZZZZ}$, $\gamma_{\perp} = \gamma_{XXZZ}$), the following expansions existed:

$$\beta_{\parallel}(-\omega_{\sigma}; \omega_1, \omega_2) = \beta_{\parallel}(0; 0, 0) (1 + A\omega_L^2 + \dots) \quad (61)$$

where A is the same for all processes (there is no comparable formula for β_{\perp});

$$\gamma_{\parallel}(-\omega_{\sigma}; \omega_1, \omega_2, \omega_3) = \gamma_{\parallel}(0; 0, 0, 0) (1 + A\omega_L^2 + A'\omega_L^4 + \dots) \quad (62)$$

where A, again, is the same for all processes and A' is the same for dc-SHG and dc-K;

$$\gamma_{\perp}(-\omega_{\sigma}; \omega_1, \omega_2, \omega_3) = \gamma_{\perp}(0; 0, 0, 0) (1 + B\omega_L^2 + B'\omega_L^4 + \dots) \quad (63)$$

where $B = p + qa$, and $a = (\omega_{\sigma}\omega_3 - \omega_1\omega_2)/\omega_L^2$, p and q being frequency independent, so that, for example, the values of B for dc-K, IDRI (DFWM), dc-SHG, and THG are in the ratio $1 : (1 + k/2) : (1 + k/3) : (1 + k/6)$ where k is a constant for any given species;

$$\gamma_{\parallel}(-\omega_{\sigma}; \omega_1, \omega_2, \omega_3)/\gamma_{\perp}(-\omega_{\sigma}; \omega_1, \omega_2, \omega_3) = 3(1 + C\omega_L^2 + \dots) \quad (64)$$

where $C = r(1 - 6a)$ and r is frequency independent, C is a measure of deviation from Kleinman symmetry and is naturally zero for THG.

Rice *et al.* [36] have used Eq. (61) to analyse their calculations of β_{\parallel} for HCHO and CH₃F. Bishop and Pipin [11] have used Eqs. (62) and (63) to express their low-frequency helium data ($\hbar\omega \leq 0.02 E_h$). They found the dispersion coefficients which are given in Table 11. It is apparent that A is indeed the same for all processes, A' is nearly the same for dc-K and dc-SHG and that B follows quite precisely the ratios given above if $k = 0.966$.

For H₂ [12], Eq. (62) has been used as a means to investigate the

Table 11. Helium coefficients for Eqs. (62) and (63) using data for $\omega = 0, 0.005, 0.01, 0.015$, and 0.02 a.u. ^a .				
	dc-Kerr	IDRI	dc-SHG	THG
A	2.189	2.189	2.189	2.189
A'	3.494	3.681	3.507	3.464
B	1.885	2.797	2.493	(2.189) ^b
B'	2.714	5.835	4.489	(3.464) ^b
^a $\gamma_{\parallel}^0 = 43.104$ a.u., $\gamma_{\perp}^0 = 14.368$ a.u.				
^b Necessarily, $A=B$ and $A' = B'$ for THG.				

extrapolated experimental static value of γ_{\parallel} (dc-SHG). In Table 12 four different fits of the experimental data and consequently four different values of $\gamma_{\parallel}(0; 0, 0, 0)$ are shown. They range from 685.2 to 686.6 a.u. and can be

Table 12. Values of the coefficients, using atomic units, in Eq. (62) for dc-SHG for H_2 .			
Least squares fit to data	$\bar{\gamma}_0$	A	A'
Theory, 7 points ^a (unweighted)	684.4 ^b	6.034	44.27
Expt., 7 points ^a (unweighted)	685.7	5.953	48.84
Expt., 7 points ^a (weighted ^c)	686.6	5.746	53.27
Expt., 21 points ^d (unweighted)	685.2	6.026	47.57
Expt., 21 points ^d (weighted ^e)	686.4	5.780	52.30
^a $\lambda = 13190, 10640, 6943, 6328, 5900, 5145$ and 4880 Å.			
^b Our actual theoretical value (unfitted) is $\gamma_0 = 682.7 e^4 a_0^4 E_h^{-3}$.			
^c Values at $\lambda = 13190$ and 10640 Å weighted by a factor of 4.			
^d All the frequencies used in Ref. [25].			
^e This was the fit published in Ref. [25].			

compared with the directly computed static value of 682.7 a.u. However, and this is the point, if we *fit* the *theoretical* dynamic values at the same frequencies as chosen in one of the experimental fits, we get 684.4 a.u. (compared with experiment's 685.7 ± 3 a.u.). This is clearly not only in better agreement but well within the experimental error. The moral is to calculate exactly what is measured and, if there is any further analysis, to carry that analysis out on the theoretical data in like manner.

Another way in which Eqs. (62) and (63) were used for H_2 in [12], was to predict from the dispersion constants for the dc-K process, those for the ac-K process. This led to the estimate that, when $\omega_1 = 14797 \text{ cm}^{-1}$ and $\omega_2 = 17751 \text{ cm}^{-1}$, $\gamma^K(-\omega_\sigma; \omega_1, \omega_2, -\omega_2) = 796.7$ and 771.9 a.u. for H_2 and D_2 , respectively. The comparable experimental values are in good agreement: 789 ± 38 and 774 ± 39 a.u. [69]. Lastly, for H_2 , it was through Eq. (62) that it became clear that the dc-SHG and dc-K results were incompatible.

In Section 3.1.2. reference was made to the scaling of accurate static hyperpolarizabilities by a ratio of the SCF dynamic/static values. Judging from Eqs. (61) and (62) this is the correct thing to do (rather than simple additivity) if the frequency is small and the constant A is independent of electron correlation.

In 1989, experimental dc-SHG data for Ne [70] suggested that at low frequencies $\gamma_{||}$ decreased with increase in ω^2 (or ω_L^2) - so-called anomalous dispersion - that is to say A in Eq. (62) must be negative. Later, it turned out that this was an experimental artefact [71]; *ab initio* calculations [72,73] had already questioned the observation. However, before that happened, I was able to use Eqs. (18) and (62) and, in a very simple way [74], show that indeed A could not be negative (or, at least, not to the extent suggested). It is only the second part of Eq. (18) which can be negative and, in terms of the standard sum rules S_{-k-1} and an expansion in ω_L^2 , I found this term contributes to A the amount

$$-\frac{1}{2} (3S_{-4}S_{-3} + 7S_{-2}S_{-5})/\gamma_{||}(0; 0, 0, 0)$$

or, using high-quality semi-empirical values of S_{-k-1} , -0.38 a.u.; the original experimentally determined total value of A was over twenty times as large. Hence, there was a clear indication that something was experimentally wrong, as is now known to be the case.

Incidentally, while we are with neon, it has been fascinating to watch, over the past few years, the 'ups and downs' of the theoretical and experimentally-derived values of $\gamma_{||}(0; 0, 0, 0)$ [25,70-72,75-80] - Table 13 shows what I mean.

Table 13. Static $\gamma_{ }$ values for neon.		
Year	Experiment	Theory
1988	84 ± 9 [75]	
1989	116 ± 2 [70]	114 [76] 119 ± 4 [77]
1990	119.2 [25]	111 [78]
1991		99 ± 6 [72]
1992	108 ± 2 [71]	110 ± 3 [79] 111 ± 4 [80]

4. MAGNETIC FIELDS

Non-linear optical effects can also be induced by magnetic fields; the Cotton-Mouton Effect (CME) [81,82], which is the magnetic analogue of the dc-Kerr effect (i.e. field-induced refractive index anisotropy), is probably the most famous one. Magnetic field effects, in general, are much weaker than electrical ones and consequently very powerful (and uniform) magnetic fields must be introduced for the effects to be observed. This, doubtless, accounts for the paucity of *accurate* gas-phase measurements. Nonetheless, the requirement of a strong magnetic field is not an obstacle for a theoretician and it is surprising that theoretical chemistry has not contributed more to this subject. Recently, there have been quantum-chemical investigations of the CME for H_2 (and D_2) at the SCF [83] and the electron-correlated [84] levels and for the rare gases [83,85-87] and that is it. The work in [83-87] will be the subject of this section.

The experimental quantity of interest in the CME is the hypermagnetizability anisotropy ($\Delta\eta$) and this has already been defined in Eq. (3). The static hypermagnetizability tensor itself is defined by Eq. (1) and its dynamic

counterpart is $\eta_{\alpha\beta\gamma\delta}(\omega)$, which, unlike its dc-Kerr effect cousin, is composed of two parts: one paramagnetic (p) and one diamagnetic (d). We have:

$$\eta_{\alpha\beta\gamma\delta}(\omega) = \eta^p(-\omega; \omega, 0, 0) + \eta^d(-\omega; \omega, 0) \quad (65)$$

These two terms, as we will see, in a certain way match $\gamma(\text{dc-K})$ and $\beta(\text{dc-P})$, respectively.

In their work on H_2 and D_2 (as well as He and Li^+), Fowler and Buckingham [83] ignored the frequency dependence and calculated the paramagnetic and diamagnetic components of the η tensor by computing the paramagnetic and diamagnetic magnetizabilities (χ in Eq. (1)) at the CPHF level in the presence of a finite, uniform, static, electric field. They then employed numerical differentiation to get η^p and η^d . Consequently electron correlation was also ignored. They assessed (approximately) the corrections for zero point vibrational averaging and they also estimated the *vibrational* hyper-magnetizability anisotropy through the expression:

$$\Delta\eta^v = (4/15) \sum_n \langle 0 | \Delta\chi | n \rangle \langle n | \Delta\alpha | 0 \rangle / \omega_n \quad (66)$$

(though only the first vibrational excited state was actually used). This formula is the counterpart to Eq. (46). Their final results for two different basis sets are given in Table 14.

Table 14. Cotton-Mouton effect for H_2 and D_2 , static values (in atomic units).		
	$\Delta\eta$	
	H_2	D_2
Ref. [83] Basis A	11.6	11.1
Basis B	12.2	11.7
Ref. [84]	10.806	10.571

However, for H_2 (and other two-electron systems) it is possible to carry out a much more sophisticated treatment along the lines of Section 3.1.1. and to

include both frequency dependence and electron correlation. This was done by myself, Janusz Pipin and Mark Cybulski [84]. We took the Orr and Ward expressions in Eqs. (17) and (18) for β and γ and adapted them to the case at hand, namely η^d and η^p , and introduced the following equations:

$$\begin{aligned}\eta^d(\omega) &= \eta_{\alpha\beta\gamma\delta}^d(-\omega; \omega, 0) = \eta_{\alpha\beta\gamma\delta}^d(-\omega_\sigma; \omega_1, \omega_2) \\ &= \hbar^{-2} \sum_p \sum_m \sum_n (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_2)^{-1} \langle 0 | \hat{\mu}_\alpha^e | m \rangle \langle m | \hat{\mu}_\beta^e | n \rangle \langle n | \hat{\chi}_{\gamma\delta}^d | 0 \rangle\end{aligned}\quad (67)$$

$$\begin{aligned}\eta_{\alpha\beta\gamma\delta}^p(\omega) &= \eta_{\alpha\beta\gamma\delta}^p(-\omega; \omega, 0, 0) = \eta_{\alpha\beta\gamma\delta}^p(-\omega_\sigma; \omega_1, \omega_2, \omega_3) \\ &= \hbar^{-3} \sum_p \left[\sum_m \sum_n \sum_p' (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_2 - \omega_3)^{-1} (\omega_p - \omega_3)^{-1} \right. \\ &\quad \times \langle 0 | \hat{\mu}_\alpha^e | m \rangle \langle m | \hat{\mu}_\beta^e | n \rangle \langle n | \hat{\mu}_\gamma^m | p \rangle \langle p | \hat{\mu}_\delta^m | 0 \rangle \\ &\quad - \sum_m \sum_n (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_3)^{-1} (\omega_n + \omega_2)^{-1} \\ &\quad \left. \times \langle 0 | \hat{\mu}_\alpha^e | m \rangle \langle m | \hat{\mu}_\beta^e | 0 \rangle \langle 0 | \hat{\mu}_\gamma^m | n \rangle \langle n | \hat{\mu}_\delta^m | 0 \rangle \right]\end{aligned}\quad (68)$$

where, now, in the relevant places the electric dipole moment operator ($\hat{\mu}_\eta^e$) has been replaced by the magnetic dipole moment operator ($\hat{\mu}_\eta^m$) or the diamagnetic magnetizability operator ($\hat{\chi}_{\eta\phi}^d$). We used our usual explicit-electron-correlated wavefunctions and made calculations for a range of frequencies (as well as for the standard laser frequencies) and internuclear distances. Our final numbers were obtained after rovibrational averaging and summing over a Maxwell-Boltzmann population distribution in the manifold of rotational states in the ground vibrational state (this distinguishes the H_2 and D_2 results). We also included the vibrational component $\Delta\eta^v$. Our static values are compared in Table 14 with the SCF values. It is hard to be sure, since the SCF values (either basis A or B) may not be at the Hartree-Fock limit, but it looks as if electron correlation is not very significant for this study. Changes brought about by the other factors (dispersion, vibrational averaging, $\Delta\eta^v$) are summarized in Table 15. When it comes to comparing the results with experiment, Refs. [88] and [89], with the latter as corrected in [90], the situation is not healthy; see Table 16. The Hüttner *et al.* results are certainly too large and though those of Scuri *et al.* are

Table 15. Contributions to $\Delta\eta$ (H_2).

correlation	small
vib ⁿ averaging	~10%
pure vibration	~2%
dispersion (static \rightarrow 5145 Å)	~4%

Table 16. Cotton-Mouton effect for H_2 and D_2 (in atomic units).

			$\Delta\eta$	
			H_2	D_2
	$\lambda / \text{\AA}$			
Hüttner <i>et al.</i> [88]	expt.	6328	15.7 ± 2	15.3 ± 3
Bishop <i>et al.</i> [84]	theor.	6328	10.547	10.374
Scuri <i>et al.</i> [89]	expt.	5145	9 ± 2	2 ± 3
Bishop <i>et al.</i> [84]	theor.	5145	10.386	10.175

good for H_2 , they are quite unreasonable for D_2 .

For atoms, the sum-over-states formulas can be simplified considerably and we can write:

$$\Delta\eta^p(\omega) = -\frac{1}{4} (e^2/\hbar m_e^2) \sum_m' |<0|\hat{\mu}^e|m>|^2 [(\omega_m + \omega)^{-3} + (\omega_m - \omega)^{-3}] \quad (69)$$

and

$$\Delta\eta^d(\omega) = -\frac{1}{4} (|e|/m_e) B(\omega) \quad (70)$$

where $B(\omega)$ is the dynamic dipole-dipole-quadrupole polarizability given by:

$$\begin{aligned} B(\omega) &= B(-\omega; \omega, 0) = B(-\omega_\sigma; \omega_1, \omega_2) \\ &= \hbar^{-2} \sum_p \sum_m' \sum_n' <0|\hat{\mu}_\alpha^e|m><m|\hat{\mu}_\beta^e|n><n|\hat{\Theta}_{\gamma\delta}|0> (\omega_m - \omega_\sigma)^{-1} (\omega_n - \omega_2)^{-1} \end{aligned} \quad (71)$$

Applying these equations to helium, and using our earlier-mentioned electron-correlated wavefunctions, $\Delta\eta(\omega)$ has been found for a number of frequencies

[86]. The static value is 1.06061, compared with the SCF value [83] of 0.96 a.u. At $\hbar\omega = 0.07 E_h$ ($\sim 6328 \text{ \AA}$), we have 1.0561 a.u. against the experimental value [91] of $1.1 (\pm 1.1)$ a.u. At $\hbar\omega = 0.09 E_h$ ($\sim 5145 \text{ \AA}$) we have 1.0530 a.u. compared with the experimental value [92] of 0.80 ± 0.16 a.u. and it looks as if the error bars in this last experiment have been set too low.

In the static limit, Eqs. (69) and (71), for an atom, combine to give:

$$\Delta\eta(0) = -\frac{1}{4} (|e|/m_e) B(0) - \frac{1}{2} (e^4 a_0^4 / m_e E_h^2) S_{-4} \quad (72)$$

where S_{-4} is the usual sum rule. Values of S_{-4} and $B(0)$ for He, Ne, Ar, Kr, and Xe are in the literature and with them we have found [86] the $\Delta\eta$ values in Table 17. Compared with the experimental ones of Hüttner [91], the agreement is fairly good; on the other hand the experimental bounds are very large.

Table 17. Comparison of theoretical and experimental values of $\Delta\eta$, in units of $10^{-44} \text{ Fr}^2 \text{ cm}^2 \text{ erg}^{-1} \text{ G}^{-2}$, of the rare gases.		
	Theory ^a	Experiment ^b
He	2.84	3 (3)
Ne	8.27	11 (8)
Ar	72.7	70 (4)
Kr	154	121 (8)
Xe	371	286 (14)
^a Static values.		
^b $\lambda = 6328 \text{ \AA}$, uncertainties in parentheses, [91].		

For neon there has been a good dynamic calculation carried out by Jaszuński *et al.* [87] along the lines of their hyperpolarizability calculation for FH [47] which was mentioned at the end of Section 3.1.2.: that is using a QRF with MCSCF reference functions. At $\lambda = 5145 \text{ \AA}$ they estimate $\Delta\eta$ to be 2.670 a.u., which can be compared with experimental values of 1.25 ± 0.07 a.u. at this frequency [93] and 4.10 ± 3.00 a.u. at $\lambda = 6328 \text{ \AA}$ [91] and with our static estimate of 3.08 a.u. [86] (their static value is 2.687 a.u.). It is clear that both experiments must be re-evaluated (the second one because of the large error

bars). In fact, it is my opinion that all CME experiments, which are very expensive, should be very critically examined in the light of existing theoretical results.

5. FUTURE DIRECTIONS

I believe that this account has shown the enormous interplay between quantum-chemical computation and experiment in the science of non-linear optics over the past few years. It is heartening to realize that calibrational-quality results have led to the un-masking of quite obscure technical problems in some of the electric-field-induced experiments. This has still to happen for the magnetic-field-induced experiments, but I am sure that it will.

As far as standard *ab initio* calculations go, it appears to me that in the future there will be improvements to the computational techniques, though my guess is that at the moment they are ahead of the gas-phase experimental techniques. An example of such a computational improvement [94] is the ‘frequency-dependent moment method’ in which hyperpolarizabilities are evaluated from configuration-interaction matrices without diagonalizing them. However, it should be remembered that calculations of high-order properties, such as the ones we have been discussing here, are a very tough test of computational quantum chemistry.

It will be a pity if economic pressures and industrial competition force the adoption of approximate theoretical strategies before they have been properly assessed by comparison with rigorous *ab initio* results. These pressures, at present, appear to be being resisted and the best semi-empirical work [29,95] keeps its eye on the underlying fundamental theory. Probably the near-future will contain many more such semi-empirical calculations and, thereby, a body of information which can be used to make predictions (one is reminded of the rôle, years ago, of Hückel theory). In the long run, the ‘useful’ species will likely be so large as to be beyond *ab initio* calculations. As well, the near-future will almost certainly see improved measurements.

Besides all this, the theoretician faces, as always, the problem of condensed phases. The way into this subject could come by considering cluster

calculations [96] and the calculation of the hyperpolarizability of long-range interacting species, where, so far, only the polarizability has been investigated, see, e.g. Ref. [97].

Other interesting related phenomena, worthy of a quantum chemist's attention, are the study of the hyperpolarizabilities of molecules trapped in a cage, e.g. a zeolite [98] or on a surface [99-102], and the whole relevance of these properties to surface property calculations, where, at the local level, there are enormous electric fields. There is also the area of very high-order harmonic generation, such as can occur in very intense laser beams and where conventional perturbation theory breaks down [103]. Finally, there is magnetic non-linear optics and here the surface, computationally, has, so to speak, only been scratched.

Acknowledgments: This work has been funded by the Natural Sciences and Engineering Research Council of Canada. I thank Mark Cybulski, Michel Dupuis, Bernard Kirtman, Jens Oddershede, Janusz Pipin, Julia Rice, David Shelton, and Ajit Thakkar who have all helped me at one time or another, and to Per-Olov Löwdin, who asked me to write this article and whose science and personality I have always admired. Finally, I acknowledge all those who have contributed to this subject, but whose work, by intent or negligence, I have ignored.

References

1. C.A. Coulson, A. Macoll, and L.E. Sutton, *Trans. Faraday Soc.* **48**, 106 (1952).
2. D.M. Bishop, *Rev. Mod. Phys.* **62**, 343 (1990).
3. A.D. Buckingham, *Adv. Chem. Phys.* **12**, 107 (1967).
4. A.D. Buckingham and B.J. Orr, *Quart. Rev.* **21**, 195 (1967).
5. M.P. Bogaard and B.J. Orr, in *International Review of Science, Physical Chemistry, Molecular Structure, and Properties*, edited by A.D. Buckingham (Butterworths, London, 1975), Ser. 2, Vol. 2, p. 149.

6. C.E. Dykstra, S.-Y. Liu, and D.J. Malik, *Adv. Chem. Phys.* **75**, 37 (1989).
7. R. McWeeny, *Methods of Molecular Quantum Mechanics* (2nd Ed., Academic Press, San Diego, 1989).
8. B.J. Orr and J.F. Ward, *Mol. Phys.* **20**, 513 (1971).
9. D.A. Kleinman, *Phys. Rev.* **126**, 1977 (1962).
10. E.R. Cohen and B.N. Taylor, *Rev. Mod. Phys.* **59**, 1221 (1987).
11. D.M. Bishop and J. Pipin, *J. Chem. Phys.* **91**, 3549 (1989).
12. D.M. Bishop, J. Pipin, and S.M. Cybulski, *Phys. Rev. A* **43**, 4845 (1991).
13. A.J. Thakkar and V.H. Smith, Jr., *Phys. Rev. A* **15**, 1 (1977).
14. A.J. Thakkar, *J. Chem. Phys.* **75**, 4496 (1981).
15. D. Guban and G.W. Michel, *Mol. Phys.* **39**, 783 (1980).
16. A.D. Buckingham and P.G. Hibbard, *Symp. Farad. Soc.* **2**, 41 (1968).
17. M.N. Grasso, K.T. Chung, and R.P. Hurst, *Phys. Rev.* **167**, 1 (1968).
18. P. Sitz and R. Yaris, *J. Chem. Phys.* **49**, 3546 (1968).
19. R. Klingbeil, *Phys. Rev. A* **7**, 48 (1973).
20. M. Jaszuński and B.O. Roos, *Mol. Phys.* **52**, 1209 (1984).
21. L.L. Boyle, A.D. Buckingham, R.L. Disch, and D.A. Dunmur, *J. Chem. Phys.* **45**, 1318 (1966).
22. V. Mizrahi and D.P. Shelton, *Phys. Rev. A* **31**, 3145 (1985).
23. S. Carusotto, E. Iacopini, E. Polaco, F. Scuri, G. Stefanini, and E. Zavattini, *Il. Nuovo Cim. D* **5**, 328 (1985).
24. D.P. Shelton and Z. Lu, *Phys. Rev. A* **37**, 2231 (1988).
25. D.P. Shelton, *Phys. Rev. A* **42**, 2578 (1990).
26. V. Mizrahi and D.P. Shelton, *Phys. Rev. A* **32**, 3454 (1985).
27. A.D. Buckingham and B.J. Orr, *Proc. Roy. Soc. (London) A* **305**, 259 (1968).
28. R. Tammer and W. Hüttner, *Chem. Phys.* **146**, 155 (1990).
29. B.M. Pierce, *J. Chem. Phys.* **91**, 791 (1989).
30. J.E. Rice and N.C. Handy, *Int. J. Quant. Chem.* **43**, 91 (1992).
31. H. Sekino and R.J. Bartlett, *J. Chem. Phys.* **85**, 976 (1986).
32. H. Sekino and R.J. Bartlett, *J. Chem. Phys.* **94**, 3665 (1991).
33. H. Sekino and R.J. Bartlett, *Int. J. Quant. Chem.* **43**, 119 (1992).
34. S.P. Karna and M. Dupuis, *Chem. Phys. Lett.* **171**, 201 (1990).
35. S.P. Karna and M. Dupuis, *J. Comp. Chem.* **12**, 487 (1991).

36. J.E. Rice, R.D. Amos, S.M. Colwell, N.C. Handy, and J. Sanz, *J. Chem. Phys.* **93**, 8828 (1990).
37. W.A. Parkinson and J. Oddershede, *J. Chem. Phys.* **94**, 7251 (1991).
38. H. Sekino and R.J. Bartlett, *J. Chem. Phys.* **84**, 2726 (1986).
39. J.W. Dudley II and J.F. Ward, *J. Chem. Phys.* **82**, 4673 (1985).
40. S.P. Karna, P.N. Prasad, and M. Dupuis, *J. Chem. Phys.* **94**, 1171 (1991).
41. S.P. Karna, Z. Laskowski, G.B. Talapatra, and P.N. Prasad, *J. Phys. Chem.* **95**, 6508 (1991).
42. S.P. Karna, G.B. Talapatra, and P.N. Prasad, *J. Chem. Phys.* **95**, 5873 (1991).
43. S.P. Karna, G.B. Talapatra, W.M.K.P. Wijekoon, and P.N. Prasad, *Phys. Rev. A* **45**, 2763 (1992).
44. N.C. Handy and H.F. Schaefer, *J. Chem. Phys.* **81**, 5031 (1984).
45. E. Dalgaard, *Phys. Rev. A* **26**, 42 (1982).
46. J.E. Rice and N.C. Handy, *J. Chem. Phys.* **94**, 4959 (1991).
47. M. Jaszuński, P. Jørgensen, and H.J.A. Jensen, *Chem. Phys. Lett.* **191**, 293 (1992).
48. J. Olsen and P. Jørgensen, *J. Chem. Phys.* **82**, 3235 (1985).
49. J.W. Cooley, *Math. Comp.* **15**, 363 (1963).
50. R.J. LeRoy, *University of Waterloo Chemical Physics Res. Rept.* CP23OR (1986).
51. A.D. Buckingham and W. Urland, *Chem. Rev.* **75**, 113 (1975).
52. C.W. Kern and R.L. Matcha, *J. Chem. Phys.* **49**, 2081 (1968).
53. W.T. Raynes, P. Lazzeretti, and R. Zanasi, *Mol. Phys.* **64**, 1061 (1988).
54. L. Ebert, *Z. Phys. Chem.* **113**, 1 (1924).
55. D.M. Bishop and L.M. Cheung, *J. Phys. Chem. Ref. Data* **11**, 119 (1982).
56. L. Adamowicz and R.J. Bartlett, *J. Chem. Phys.* **84**, 4988 (1986); *ibid* **86**, 7250 (1987).
57. D.S. Elliott and J.F. Ward, *Mol. Phys.* **51**, 45 (1984).
58. D.P. Shelton, *Mol. Phys.* **60**, 65 (1987).
59. D.M. Bishop, *J. Chem. Phys.* **86**, 5613 (1987).
60. B. Kirtman and D.M. Bishop, *Chem. Phys. Lett.* **175**, 601 (1990).
61. D.M. Bishop and B. Kirtman, *J. Chem. Phys.* **95**, 2646 (1991).
62. D.M. Bishop and B. Kirtman, *J. Chem. Phys.* (in press).

63. D.M. Bishop and S.A. Solunac, *Phys. Rev. Lett.* **55**, 1986 and 2627 (1985).
64. D.P. Shelton, *J. Chem. Phys.* **84**, 404 (1985).
65. D.M. Bishop, *Phys. Rev. Lett.* **61**, 322 (1988).
66. D.M. Bishop, *Chem. Phys. Lett.* **153**, 441 (1988).
67. D.M. Bishop, *J. Chem. Phys.* **90**, 3192 (1989).
68. D.M. Bishop, *J. Chem. Phys.* **95**, 5489 (1991).
69. G.J. Rosasco and W.S. Hurst, *J. Opt. Soc. Am. B* **3**, 1251 (1986).
70. D.P. Shelton, *Phys. Rev. Lett.* **62**, 2660 (1989).
71. D.P. Shelton, *Chem. Phys. Lett.* (in press).
72. H.J.A. Jensen, P. Jørgensen, H. Hettema, and J. Olsen, *Chem. Phys. Lett.* **187**, 387 (1991).
73. J.E. Rice, *J. Chem. Phys.* **96**, 7580 (1992).
74. D.M. Bishop, *Phys. Rev. Lett.* **65**, 1688 (1990).
75. D.P. Shelton and Z. Lu, *Phys. Rev. A* **37**, 3813 (1988).
76. G. Maroulis and A.J. Thakkar, *Chem. Phys. Lett.* **156**, 87 (1989).
77. P.R. Taylor, T.J. Lee, J.E. Rice, and J. Almlöf, *Chem. Phys. Lett.* **163**, 359 (1989).
78. D.P. Chong and S.R. Langhoff, *J. Chem. Phys.* **93**, 570 (1990).
79. J.E. Rice, G.E. Scuseria, T.J. Lee, P.R. Taylor, and J. Almlöf, *Chem. Phys. Lett.* **191**, 23 (1992),
80. P.R. Taylor, T.J. Lee, J.E. Rice, and J. Almlöf, *Chem. Phys. Lett.* **189**, 197 (1992).
81. A. Cotton and H. Mouton, *Comp. Rend.* **141**, 317, 349 (1907).
82. A. Cotton, *Rapp. Cons. Phys. Solvay*, p. 418 (1932).
83. P.W. Fowler and A.D. Buckingham, *Mol. Phys.* **67**, 681 (1989).
84. D.M. Bishop, S.M. Cybulski, and J. Pipin, *J. Chem. Phys.* **94**, 6686 (1991).
85. M.J. Jamieson, *Chem. Phys. Lett.* **154**, 521 (1989); *ibid* **176**, 467 (1991).
86. D.M. Bishop and J. Pipin, *Chem. Phys. Lett.* **186**, 195 (1991).
87. M. Jaszuński *et al.*, *Chem. Phys. Lett.* **191**, 599 (1992).
88. W. Hüttner, H. Träuble, H.U. Wieland, and H. Müller, *Chem. Phys. Lett.* **140**, 421 (1987).
89. F. Scuri, G. Stefanini, E. Zavattini, S. Carusotto, E. Iacopini, and E. Polacco, *J. Chem. Phys.* **85**, 1789 (1986).

90. A.D. Buckingham and J.H. Williams, *J. Chem. Phys.* **86**, 5883 (1987).
91. W. Hüttner, private communication.
92. R. Cameron *et al.*, *Phys Lett. A* **157**, 125 (1991).
93. R. Cameron *et al.*, *J. Opt. Soc. Am. B* **8**, 520 (1991).
94. T. Inoue and S. Iwata, *Chem. Phys. Lett.* **167**, 566 (1990).
95. J.O. Morley and D. Pugh, *J. Chem. Soc. Faraday Trans.* **87**, 3021 (1991).
96. G.H. Wagnière and J.B. Hutter, *J. Opt. Soc. Am. B* **6**, 693 (1989).
97. D.M. Bishop and J. Pipin, *J. Chem. Phys.* (in press).
98. A.M. Goulay, F. Clairet, and E. Cohen de Lara, *Mol. Phys.* **73**, 845 (1991).
99. F. Depasse and J.M. Vigoureux, *Chem. Phys. Lett.* **160**, 311 (1989); *ibid* **177**, 167 (1991).
100. P.R. Antoniewicz, *Phys. Rev. Lett.* **87**, 432 (1974).
101. L. Galatry and C. Girard, *Chem. Phys. Lett.* **101**, 242 (1983).
102. R.G. Forbes, *Surf. Sci.* **108**, 311 (1981); Y.R. Shen, *Ann. Rev. Phys. Chem.* **40**, 327 (1989).
103. J.L. Krause, K.J. Shafer and K.C. Kulander, *Chem. Phys. Lett.* **178**, 573 (1991).

The Mechanics of Ionic Motion in Molecular Channels

By P. P. Schmidt

Chemistry and Materials Divisions
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217-5000

The Mechanics of Ionic Motion in Molecular Channels

By P. P. Schmidt

Chemistry and Materials Divisions
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217-5000

1. Introduction

2. The Channel Model

2.1 Location of wall sources

2.2 Model potential energy functions

2.3 Axial and transverse properties of the model potential

3. Classical mechanical harmonic oscillator analyses

3.1 Choice of parameters: far IR spectra of solvated ions

3.2 Mechanical stability in channel configurations

3.3 Classical vibrations of ions at minima in the channels

4. Free and nearly free ions

4.1 Free ions

4.2 Nearly free ions: general properties

4.3 Nearly free ions: direct use of the model potential

4.4 Mobility in the nearly free ion limit

5. Tight binding analysis

5.1 One-dimensional harmonic oscillator analysis

5.2 Improvement in accuracy: anharmonic potentials

5.3 Analysis in three-dimensional Gaussian basis: use of the full potential

5.4 Matrix elements of the screened potential

5.5 Matrix elements of the exponential three-body potential

5.6 Mobile ion-channel phonon coupling

5.7 Ionic mobility in the small polaron-like limit

6. Estimates for two or more ions per channel

7. A model of ion-gating into biological conduction channels

7.1 The membrane model and mechanism of ionic flux gating

7.2 Tunneling probes of membrane surfaces and channels

8. Summary

9. Appendices

A: Gaussian screened Coulomb potential

B: Cartesian force constants for ions in channels

C: Matrix elements for nearly free ions

D: Matrix elements in pseudo-Bloch functions

- E: Matrix elements of the truncated harmonic potential
F: Matrix elements in Gaussian primitive basis functions

10. References

1. INTRODUCTION

Facile ionic conduction characterizes many systems. The migration of small ions through molecular channels is a regular part of the model of impulse conduction in active membranes such as nerves and muscles [1-4]. Ionic migration in many non-biological systems may involve conduction along tight channels as well [5]. The recent discovery [6,7] of the tubular fullerenes raises the possibility of highly regular ion conduction channels that might possibly be fabricated into extremely small ion-selective probes. With reference to biologically interesting ionic channel compounds, a number of calculations have been carried out over the years; there is a considerable literature on the structure of ionic conduction channels examined from the vantage points of both quantum and classical mechanics [1]. Until now, however, there has been relatively little consideration of the nature of the states available to a mobile ion within a channel and the consequence of these states on the time-dependent mobility of the ion. It is clear, even at the outset, that conduction bands for an ionic species as large as sodium, for example, will be extremely narrow—vanishingly small in comparison to the band widths associated with electrons in metals and semiconductors. Nevertheless, a band picture needs to be formulated if one is to address properly certain essential processes that can take place involving ion conduction channels [see, in particular, ref. 8]. For example, gated ionic conduction across a membrane into the channel controls excitable phenomena such as the nerve impulse and muscle contraction. An adequate model of this process can only be constructed if one begins with an appropriate representation of the states available to the migratory ion when it is trapped first outside the channel and finally when it has crossed the barrier and moves somewhere inside the channel. The channel itself can be regarded as long (effectively infinite for mathematical tractability) or short. Moreover, it is entirely possible to imagine (and perhaps one day to find) strong ion-channel wall source interactions that *dress* the ion in channel phonons. Such a situation can lead easily to a large effective mass associated with self-trapped states.

This paper begins to address the issue of the conduction of small ions in channels by examining a model of the system as a regular double helix of polar sources that interact with a small spherical ion, for example sodium, located on the helical axis. Both classical and quantum mechanical analyses are carried out to determine energy states and possible vibrations of the ions that may be used to probe these channel conductors. Some issues associated with ionic transport across barriers,

and along the channel will be considered here. Much of our initial effort addresses the formal treatment of various processes. I investigate spectroscopic properties with the use of simple model representations. More complete approaches are indicated. The theory developed requires the assumption of complete, analytical potential energy functions for the interaction of the ion in the channel with the atoms in the molecules that make up the wall of the channel. An essential feature of this investigation is the use of discrete ion-molecule potentials to describe the interaction between the ion and the wall of the conduction channel. Tunneling or delocalization of the ion is handled by using a wavefunction for the state of the ion in the whole channel as a linear combination of local functions expanded at each potential energy minimum. The treatment is analogous to those used for molecular electronic problems. Delocalization and the associated band structure is evident. Although we avoid the use of *ad hoc* models, such as the well-known Kronig-Penney box model [9], at this point we nevertheless employ relatively simple model potentials, such as the intersecting harmonic well function and Morse potential to explore the tight binding limit. Directions past this representation to consider a fully anharmonic treatment are also discussed. In principle, these potential energy functions are available as fitted forms from more extensive, and perhaps complete, quantum chemical calculations of the kind already mentioned [1]. In addition, novel spectroscopic experiments may provide data with which to refine the model potential energy functions.

The first goal of this work is to develop a sound theoretical foundation for the description of ion transport along a channel. Once this description is established, it is possible to consider refinements: interactions with channel wall vibrations and ion transfer across interfaces that control the flow of ions from solution, for example, into the channel. In this paper, I examine a model for ion transport in screened, but otherwise electrically neutral channels. Band states may exist for ions in such systems. There is evidence [10] that ion conduction channels do not need to have incorporated water to solvate mobile ions effectively; aromatic π -electrons are sufficiently polarizable to interact strongly with a simple cation to create an association that is as effective as water solvation. Thus, the models constructed assume only that the sources (molecules) that make up the channel walls interact directly with the ions through various forms of ion-fixed charge (*i.e.*, polarization) forces. Whether these band states are narrow or broad, for some conditions (perhaps very low temperatures) there should be evidence of their existence. I therefore present calculations of possible spectral features, particularly the far infrared-active vibrational frequencies for an ion within a single channel. Measured spectra, if obtained and assigned, should yield considerable information about the structure of the channels.

The height of the barrier to ionic migration along the axis between well minima in an *actual* channel may be any value from zero to one that effectively traps the ion in a local well. Ionic conduction may therefore occur by means of acceleration in the near continuum of levels in the case of a highly delocalized system or by tunneling (or hopping)

in systems with finite periodic barriers [11]. Ionic mobility in a sequence of relatively deep local traps may resemble the conduction of an electron in the small polaron model. The consequences of tunneling, or its absence, on spectroscopic properties in particular will be explored in this paper. In 1968 Fraser and Frey [12] reported observing microwave radiation emitted from active biological systems. The possibility exists that this radiation might be connected with ion transfer between local sites or ion acceleration in channels. In biological systems, for example, it is likely that wells exist along the channel axis, but their presence may only be visible spectroscopically when the system is examined at low temperature. The calculations carried out [unpublished, using methods reported in this paper] indicate that the well depth is sensitive to the effective channel radius. That is, if some of the wall sources are closer to the ion than others, it is possible to generate a fairly deep well without any other changes in the parameters of the potential energy function. Although it may not be possible to carry out spectral studies of biologically active channels *in vivo*, representative model systems may be synthesized and collected *in vitro* so that spectroscopic measurements could be interpreted with this model, or modifications of it, to give insight into the structure that assists the conduction. On the other hand, in a frozen system, even a biological sample, it may be possible to carry out Scanning Tunneling Microscopic (STM) measurements that simultaneously locate channel openings and activate the ion transfer. At ambient temperatures, ionic motion in biologically important systems is likely to conform to classical mechanical limits. Nevertheless, examination at low temperatures, at which point quantum mechanical effects may be detected, should tell much about the structure of the conduction channels in any real system.

The following sections contain both classical and quantum mechanical calculations to predict expected spectral signatures for the vibrations of an ion in the channel. Although the calculations use model potential energy functions, it is possible to assign or refine parameters using additional molecular electronic calculations. The properties of ions in regular sequences of wells must reflect to some degree the influence of tunneling or delocalization on the spectra and on the conduction properties. For this reason, it is essential to carry out analyses, based in quantum mechanics, that can adequately account for the expected behavior. Notable initial findings are that far infrared active vibrations normal to the conduction axis should be observable and that for ionic channels with multiple occupancy the axial vibrations should be density dependent. In addition to these issues, consideration is given to ionic mobility, especially in the tight-binding, self-trapped limit. The point of view adopted in this work is applied to a discussion of a possible, reasonable model for ion-gating across membrane surfaces. Finally, the possibility of using proton and small ion tunneling as a microscopy is also explored.

2. THE CHANNEL MODEL

We begin the consideration of the mechanics of ionic motion in channels by constructing a representation of the channel that the ion will occupy. Although real channels are complex as a result of the aggregation of different molecular species, when the predominant chemical type is organic with polar oxygen and nitrogen groupings, the effective potential felt by an ion is mostly uniform. As a result, it is much simpler to model the channel as a regular molecular framework. The subsequent theory evolves using the translational symmetry built into the model. And, in spite of the fact that natural systems are usually bewilderingly complex, the regularity of the model generally reproduces the major features of the physics. This is what we seek.

2.1 LOCATION OF WALL SOURCES

An ion channel is modeled for simplicity as a rigid, regular double helix of sources. Although there is no particular requirement to represent the channel as a double helix, a single helix will reflect minima for the ion-source interaction that are also helical. The virtue of the double helix, therefore, is an axial symmetry that makes the subsequent analysis reasonably simple. It is in fact possible to carry out a similar analysis for the general single helix system using an extension of the Bloch theorem that has been explored by Mintmire et al. [13] for polymer systems.

The locations of sources on the wall of the double helix are given by the equation [14]

$$\mathbf{S}_u \mathbf{R} = \begin{pmatrix} \cos\vartheta & -\sin\vartheta & 0 \\ \sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \zeta \end{pmatrix}; \quad (1)$$

\mathbf{S}_u is the operator of the helical group that generates the set of coordinates for the wall sources located on the *upper* helical strand of the double helix that defines the channel; the subscript *u* indicates *upper*. Here, ζ is the displacement along the helical axis and ϑ is the helical twist angle in a right-handed rotation down the axis. The second helical strand can be generated from the first simply by reflecting the first source in the x-z plane:

$$\mathbf{S}_l \mathbf{R} = \begin{pmatrix} \cos\vartheta & \sin\vartheta & 0 \\ \sin\vartheta & -\cos\vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \zeta \end{pmatrix}. \quad (2)$$

Here, the subscript l stands for *lower* corresponding to the second chain whose source is, from the viewer's perspective, on the bottom of the first unit cell. Successive sources are located according to the application of the operators S_u and S_l m times to the vector \mathbf{R}_0 , which is the position of the first wall source:

$$\mathbf{R}_{m(u,l)} = \left[\mathbf{S}_{u,l} \right]^m \mathbf{R}_0. \quad (3)$$

In constructing the helices, the assumption is made that a minimum distance between source atoms in the channel wall can be tolerated. Because of the nature of the repeat units, it is easy to establish a critical angle to ensure appropriate separation. One finds, therefore, that if d is a minimum, or contact, distance, then

$$\vartheta = \arccos \left[1 - \frac{d^2 - \zeta^2}{2r} \right], \quad (4)$$

and here r is the radial distance from the axis to the wall source.

The double helix that defines the channel is constructed as a sequence of neutral, but polar and polarizable sources distributed on the z -axis. The same parameters for the interaction of the helical source atoms with the ion are used for each atom. Vibrational structure of the wall-forming helix is not considered at this time; the wall of the conduction channel is considered to be rigid. In general, the channel radii used were between 3.2 and 3.6 Å, values that are close to the solvation radius for a sodium in water. Replica distances used for the sources along the helical axis were of the order of 1.2 Å to 2.0 Å. In order to have a distance of about 3 Å between neighboring atoms on the helix wall, and to make sure that the double strands were uniformly separated from one another, twist angles were determined with eq (4) and generally were found to have values of around 50°.

2.2 MODEL POTENTIAL ENERGY FUNCTIONS

Part of the potential energy function is a simple pair potential:

$$V(r) = D \exp[-\alpha(r-r_0)] \left[\exp[-\alpha(r-r_0)] - 2 \right] - \frac{Ze^2}{r} \left(\text{erf}(br) - 1 \right). \quad (5)$$

In this equation, \mathbf{r} is the vector between the ion located on the channel axis and an individual source in the channel wall. The total potential for the interaction of an ion with the whole channel is the sum of interactions of the form of eq (5) or an equivalent potential. The first term is the Morse function and the second is a Gaussian screened Coulomb

potential. In eq (5) D is the "bond energy" of the Morse term, α the associated exponent, r_0 is an equilibrium distance that enters here as an adjustable parameter, and $\text{erf}(x)$ is the error function. Ze is the charge on the ion that interacts with the core and an equal amount of screening electronic charge; the character of the screening is based on the assumption of a s -type Gaussian electronic charge distribution on the wall atom (the derivation of the specific form of the potential is discussed in Appendix A). It is also possible to use the Yukawa, or Debye-Hückel, potential with Gaussian basis functions to construct matrix elements by employing the transformation [15]

$$\frac{1}{r} \exp(-ar) = \int_0^\infty ds (\pi s)^{-1/2} \exp(-a^2 r^2/4s). \quad (6)$$

The resulting integrals are standard; the integrals are discussed in later sections and in the appendices. The virtue of the Gaussian-based potential is the fact that it can be manipulated easily and directly in terms of sums and integrals of Gaussian functions [16]. Thus, the evaluation of quantum mechanical matrix elements in a Cartesian Gaussian basis is straightforward.

In addition, a simple three-body potential energy function has been considered that has the formula

$$V_3 = \mathfrak{A} \exp\left[-g(r_G - R_{G0}) - h(r_H - R_{H0})\right] \quad (7)$$

where r_G and r_H are the distances of a single ion from two source atoms at different points in the wall of the channel. The effect of this potential on the axial vibrations of a channel ion, however, is found to be relatively small in the classical harmonic limit calculations. Further consideration is given to this form of the potential later in the chapter.

Finally, permanent dipoles were embedded in the source atoms to provide an attractive ion-dipole interaction. The orientations of the dipoles were along the vectors normal to the helical axis. No account was taken at this time, however, of induced dipole, or polarization, interaction.

General features of the potential energy function, for various choices of the parameters, are examined here. The complete function is not particularly sophisticated; Morse potentials behave in a representative, but not wholly accurate, manner [17]. The Coulomb part is chosen for ease in computation; the Gaussian screening is probably more severe than is the case for real screening. As noted, induced dipolar polarization is not considered, although it can be included; the Gaussian screening, although not a complete substitute for induced dipolar polarization, in all likelihood represents the effect to some degree.

2.3 AXIAL AND TRANSVERSE PROPERTIES OF THE MODEL POTENTIAL

The first issue to consider is the nature of the potential for the degrees of freedom of the ion along the line directly joining two wall sources, in particular, reflected sources. This type of system was investigated a number of years ago by Löwdin [18,19] in connection with the problem of the hydrogen bond in the DNA double helix. The analysis used here shows similar properties for the potential. Linear triatomic systems that involve two Morse potentials and a repulsive interaction have also been investigated by Hall [20]; it is instructive to review some of his conclusions.

The form of Hall's potential is

$$V_H = A \exp(-ax) \left(\exp(-ax) - 2 \right) + B \exp(-by) \left(\exp(-by) - 2 \right) + 2H \exp(-ax) \exp(-by) \quad (8)$$

with

$$x = r_a - r_{a0} \quad y = r_b - r_{b0} \quad (9)$$

This potential has a genuine saddle point somewhere along the line joining the three atoms when the following conditions, listed by Hall, are satisfied: $H \neq 0$ and $H > A$ and $H > B$. If the end atoms of the triatomic system are free to move with respect to the central species, a stable triatomic species forms when $H < A$ and $H < B$. When the Morse potential terms are combined with other terms, as done here, it is no longer easy to establish a simple relationship for the saddle point. With fixed end atoms, a barrier can arise for the the sum of two Morse potentials as used here (with $H = 0$) or alone without additional contributions. It is the case, however, that when $H = 0$ the height of that barrier cannot exceed the dissociation energy D . It seems appropriate to avoid double minima with respect to the transverse axes within the channel for the types of systems discussed here; although, when considering DNA it is of course appropriate to consider sets of parameters that yield double minima.

The (here unwanted) transverse double wells arise when r_0 for the Morse potential is considerably less than the actual distance to the wall sources. This situation is shown in Figure 1. Case (a) is the Morse potential alone. The effect of a diffuse charge, $b = 0.5 \text{ \AA}^{-1}$, is shown in case (b); the diffuseness of electronic charge in effect exposes the underlying positive core charge that then interacts repulsively with the cation in the channel. The next curve in Figure 1, (c), consists of the Morse term, a compact, and effectively screened Coulomb interaction ($b =$

2.0\AA^{-1}), and the three-body Hall interaction. The general effect of this type of three-body term is to shift the potential upwards without substantially altering other characteristics of the potential. For the purposes of this paper, this term is therefore neglected for studies of axial motions. Finally, the inclusion of a dipolar interaction (d) lowers and slightly alters the shape of the curve.

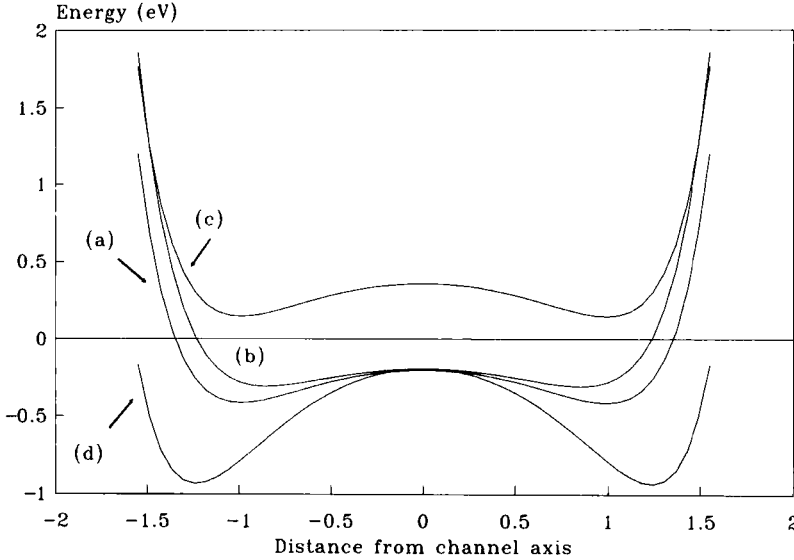


FIGURE 1: Potential energy of the ion moving along the y -axis perpendicular to the channel axis. The parameters in the potential are the following: for both (a) and (b); $d_{\min} = 3\text{\AA}$, $l = 1.5\text{\AA}$, $\rho = 3.3\text{\AA}$, $D = 0.4\text{ eV}$, $a_{\text{Morse}} = 2.0\text{\AA}^{-1}$, $R_0 = 2.3\text{\AA}$, $Z = 1.0$, $a_{\text{Gauss}} = 2.0\text{\AA}^{-1}$, $V_{3b} = 0.4\text{ eV}$, $a_{3b} = 2.0\text{\AA}^{-1}$. For (a), $\mu = 0.0$; (b) $\mu = 1.0$ Debye.

Figure 2 illustrates the behavior of the channel ion for motion *along* the channel through the midpoint of a line joining a pair of surface sources. The parameters have been chosen to illustrate the effect of just the Morse potential for the case that the spacing of the wall sources with respect to the channel axis yields unstable motion at the center.

As more channel wall sources are added to the system, the character of the potential changes slightly. Figure 3 shows the case that the channel consists of 22 wall sources. For the ionic motion along the channel axis in the center cell, one sees now that motion in between two

wall sources is unstable. Positions of equilibrium arise at the half

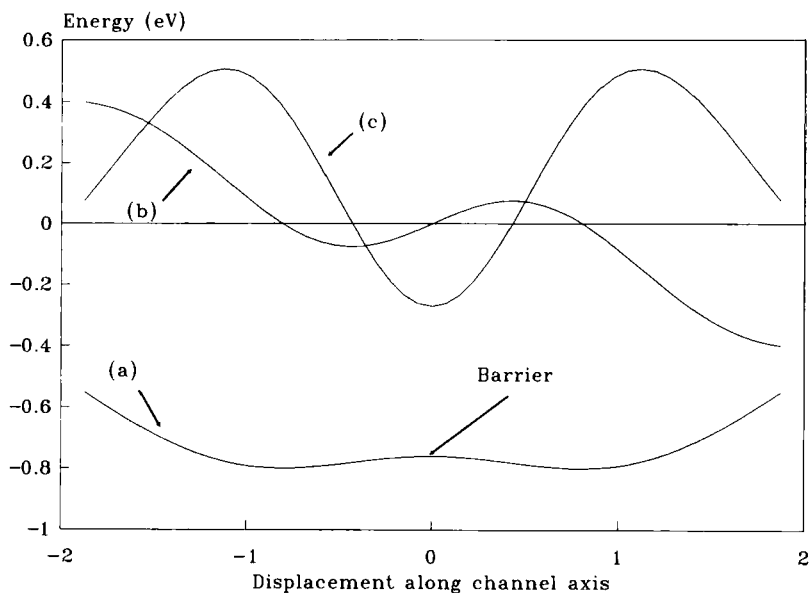


FIGURE 2: The behavior of an ion along the axis under the influence of only a Morse potential. Curve (a) is the potential, (b) the first derivative, and (c) the second derivative.

spacings between the locations of the pairs of wall sources. The depth of the emergent energy wells is not great; and, as the analysis of the next sections shows, the associated frequencies for axial motions are small.

Finally, in Figure 4 when the system consists of 202 wall sources, axial motion in the central cell appears to be representative of all the unit cells. As with the case of 22 sources, positions of stable equilibrium are found at the half spacings between the source pairs that make up the helical wall. These studies indicate that a system of 202 wall sources is sufficient to calculate properties of the system. The contributions to the potential for sources greater than about 12 unit cell lengths from the center appear to be vanishingly small. This can also be seen by comparing the barrier of Figure 3 with Figure 4.

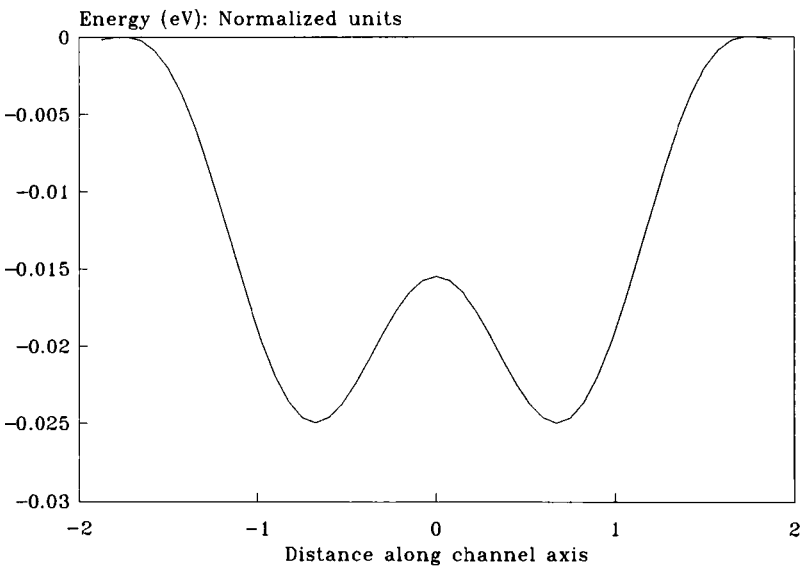


FIGURE 3. The ion potential along the axis when the channel is only 11 repeat units long (22 wall sources).

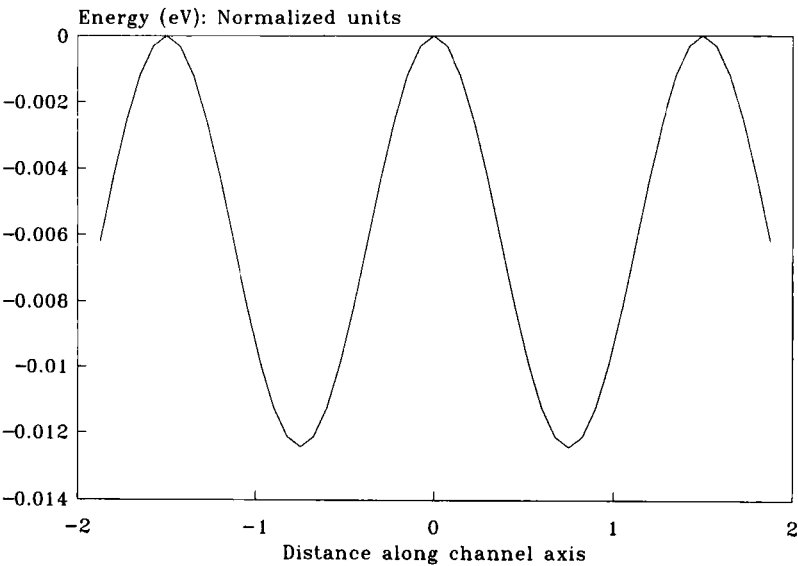


FIGURE 4. The potential for the motion of the ion along the channel axis when there are 101 repeat units (202 wall sources).

3. CLASSICAL MECHANICAL HARMONIC OSCILLATOR ANALYSES

Although the regular spacing of the helix source atoms can generate a sequence of wells into and out of which the ion can tunnel, it is instructive to carry out a classical mechanical analysis of the vibrations one can expect for an ion arbitrarily held in one particular well. Moreover, it is possible to determine the axial force and the various harmonic force constants as the ion moves along the axis through the saddle points. These types of analyses are helpful in quickly seeing the character of the surface. The formulae needed to compute Cartesian force constants for the ion at any position of equilibrium are given in Appendix B.

At this point, the methods developed are not matched to any particular system. The objective, as stated, is to examine a general class of phenomenon—ionic conduction in molecular channels. Thus, a relatively flexible, simple model is preferred over one that attempts too soon (and inadequately) to represent all the many molecular species and interactions that are indeed present.

3.1 CHOICE OF PARAMETERS: EXAMPLE, FAR IR SPECTRA OF SOLVATED IONS

A necessary issue has to do with the reasonable values one can use for the parameters in the potential energy function. Whether a simple spherical ion is found in an aqueous or some other polar environment, it is now known that its spectroscopic signature in the far infrared is similar in a broad class of systems [see, for example, the discussion in refs. 21-26]. Far infrared frequencies associated with vibrations of a solvated spherical ion are probably similar to the vibrations normal to the axis of an ion found within a channel. The analysis of representative spectra fixes values of the parameters in the potential energy function. This is followed by an examination of the expected classical vibrations for the ion at one point on the helical axis. Delocalization and band-like behavior is taken up after that.

The far infrared spectra of the Group Ia alkali metal ions were investigated extensively over twenty years ago [21-25]. At that time, the experimental evidence pointed to the fact that the far infrared bands arise predominantly from the vibrations of the ion relative to its position of equilibrium in the solvation cavity. Subsequent analysis, based on methods I developed together with several colleagues, indicated that for regular solvation [26], *viz.*, tetrahedral or octahedral, the classical harmonic vibrations are sensitive to the repulsive potentials that operate between the ion and the solvent. Thus, for example, with the use of a simple Born exponential model of repulsion [27], it is possible to assign values of the parameters. As a result, using

straightforward and simple potential energy functions, it is possible to account reasonably accurately for the behavior of an ion in a complicated environment.

If one assumes tetrahedral or octahedral solvation about an ion, such as sodium, then using eq (B1) and (B2) one sees that

$$k_{xx} = k_{yy} = k_{zz} = k_r, \quad (10)$$

which is a radial force constant. The specific functional form of k_r is [26]

$$k_r = \frac{N_C}{3} I_{20}(R) \quad (11)$$

in which N_C is the coordination number of the ion in the particular solvent environment. For the combination of the Morse and Gaussian screened Coulomb potentials, one has

$$k_r = \frac{N_C}{3} \left\{ 2a^2 D e^{-a(R-R_0)} \left[2(1 - 1/aR) e^{-a(R-R_0)} - 1 + \frac{2}{aR} \right] + 4Ze^2 b^{3/2} e^{-bR^2} \right\}. \quad (12)$$

Note, classical dipole terms do not contribute to the harmonic frequencies because of symmetry. In the case that $R = R_0$, this simplifies considerably

$$k_r = \frac{N_C}{3} \left\{ 2a^2 D + 4Ze^2 b^{3/2} e^{-bR_0^2} \right\} \quad (13)$$

Although it is the case that the classical ion-dipole interactions do not explicitly contribute to the vibrational frequencies of an ion inside a regular cubic cage [26], it is nevertheless still true that dipole interactions play a role. The difficulty one may encounter is the fact that a reasonable frequency can be found for a configuration that may be unreasonable with respect to the dipolar interaction. To account for this possibility, the calculations were carried out for the ion-solvent aggregate relaxed to a radial equilibrium configuration. The following table indicates selected, representative values of the parameters to give frequencies in the 190 cm^{-1} range with optimized ion-solvent radii.

TABLE 1

Representative values of the parameters in the potential energy to simulate frequencies for sodium in solution

ρ	D	α	R_0	b	ω
3.059	0.6	1.00	3.2	1.00	198.3
3.059	0.6	1.00	3.2	1.75	198.4
3.311	0.8	1.00	3.3	0.25	190.7
3.171	0.6	1.00	3.3	1.00	195.9
3.171	0.6	1.00	3.3	1.75	195.9
3.400	0.8	1.00	3.4	0.25	193.5
3.282	0.6	1.00	3.4	1.00	193.7
3.282	0.6	1.00	3.4	1.75	193.7
3.491	0.8	1.00	3.5	0.25	195.8
3.391	0.6	1.00	3.5	1.00	191.7
3.391	0.6	1.00	3.5	1.75	191.7
3.584	0.8	1.00	3.6	0.25	197.7
3.223	0.3	1.75	3.2	0.25	196.3

ρ is the solvation radius.

As the table indicates, only in the case of very diffuse electronic screening charge is there any pronounced effect on the vibrational frequencies. For values of the Gaussian exponent of $b \geq 1$, one sees that pairs of values of the remaining parameters arise. This, of course, is to be expected. Large values of the Gaussian exponent indicate an electronic density that is a maximum close to the core; as a result, the positive core charge is effectively screened completely by the tightly bound electronic charge.

3.2 MECHANICAL STABILITY IN CHANNEL CONFIGURATIONS

In addition to the values of the parameters that one can estimate to get reasonable vibrational frequencies, there is the question of the mechanical stability of an ion in any configuration of channel or solvent molecules [28]. Whether or not an ion will occupy a particular position of equilibrium depends on the configuration of molecules around the ion as well as on the potential energy function. The consideration of ionic mobility in tubular channel systems frequently requires one to find positions of stable equilibrium, corresponding to well minima, in order to locate a reasonable coordinate origin for the basis set used in a quantal calculation. In addition, in instances in which local self-

trapping of an ion takes place, it is necessary to be able to locate not only positions of equilibrium for the initial and final states for the elementary diffusive act, but the location in coordinate and energy space of the transition state is also required. The approach taken in the next few paragraphs is suited to systems with axial symmetry, but also of course applies in general form to all problems.

In certain instances, regular polygons can have positions of equilibrium in the plane and at the center of the structure. As the cross-sectional helical structure of the channel can have the character of an annulus, it is worth while examining the nature of the intrinsic stability in 'rings' for the bearing it has on the positions of equilibrium in the channel. Mechanical stability of an ion located on the line joining two wall sources is the limiting regular polygon for the model system we examine here. The following discussion explores the general question of the mechanical stability of an ion at the center of a regular planar array of sources. The same potential energy function is used. A Taylor series expressed in terms of spherical coordinates can be written as [29]

$$G(\mathbf{r}+\mathbf{R}) = \sum_{n=1}^{\infty} \frac{r^n}{n!} \sum_{lm} \frac{A_{nl}}{2l+1} Y_{lm}^*(\hat{\mathbf{r}}) Y_{lm}(\hat{\mathbf{R}}) I_{nl}(R) \quad (14)$$

where the Y_{lm} are the spherical harmonic functions [30] and $I_{nl}(R)$ is given by [28,29]

$$I_{nl}(R) = R^l \left(\frac{d}{R dR} \right)^l \frac{1}{R} \left(\frac{d}{dR} \right)^{n-l} R F(R) \quad (15)$$

and $F(R)$ is the purely radial part of the potential. The coefficients A_{nl} are defined by [31]

$$A_{nl} = \frac{n!(2l+1)}{(n-l)!(n+l+1)!!} \quad \text{for } n \geq l \text{ and } n-l = \text{even} \quad (16)$$

$$= 0 \quad \text{for } n < l \text{ and } n-l = \text{odd.} \quad (17)$$

For regular planar distributions of matter about a central point, it is possible to expand the Taylor series in the following manner [28,29]

$$G(\mathbf{r}+\mathbf{R}) = N_s \sum_{n=1}^{\infty} \frac{r^n}{n!} \sum_l A_{nl} P_l(\cos \theta_R) P_l(\cos \theta_r) I_{nl}(R) \quad (18)$$

in which the $P_l(x)$ are the Legendre polynomials. The angle θ_r is measured between the vector \mathbf{r} , which defines the instantaneous location of the ion, and the z -axis; θ_R measures the angle between the vector \mathbf{R} , drawn from the origin to the wall source, and the z -axis. The number of source molecules at a distance R from the center of the distribution is N_s . This form of expansion applies both to the discrete and the continuous distributions. In the continuous case, N_s is replaced by density of charge per unit angle of circumference.

The first order quantity determines positions of equilibrium or saddle points along the axis through the center:

$$\begin{aligned} G_1 &= r A_{11} P_1(\cos \theta_R) P_1(\cos \theta_r) I_{11}(R) \\ &= z \mathfrak{F}. \end{aligned} \quad (19)$$

From this expression, one finds that the solution to the equation

$$\begin{aligned} \mathfrak{F} &= \frac{Z}{R} \frac{dV}{dR} \\ &= 0 \end{aligned} \quad (20)$$

locates the position of equilibrium. At the center of the ring, $Z = 0$ and $\cos \theta_R = 0$; thus $\theta_R = \pi/2$. The derivative of the potential may also be zero at the center of the ring, but its vanishing is alone not necessary to establish equilibrium. The stability of equilibrium at the center of the ring is found from the second order term G_2 . For $\theta_R = \pi/2$, $P_2(0) = -1/2$. Thus,

$$G_2 = \frac{r^2}{2} \frac{1}{3} \left\{ I_{20}(R) - P_2(\cos \theta_r) I_{22}(R) \right\}. \quad (21)$$

For axial motion along the Z -axis, $\theta_r = 0$ and

$$\begin{aligned} G_{2z} &= \frac{r^2}{2} \frac{1}{3} \left\{ I_{20}(R) - I_{22}(R) \right\} \\ &= \frac{z^2}{2} \frac{dV}{R dR}. \end{aligned} \quad (22)$$

On the other hand, for radial motion normal to the axis, $\cos \theta_r = 0$, and

$$G_{2xy} = \frac{\rho^2}{4} \left\{ \frac{d^2 V}{dR^2} + \frac{1}{R} \frac{dV}{dR} \right\}. \quad (23)$$

Stability of motion requires $G_2 > 0$. In some instances, it is possible to determine stability from an inspection of the functional form of V alone. This is true of a single Morse potential [28]. When combined with other functional forms, however, such as the Gaussian screened Coulomb interaction, it is no longer trivial to extract stability conditions.

With reference to the stability of the ion at the center of the line joining two sources from the two chains, we have already seen that for the parameters chosen, axial motion is stable. As the number of sources in the walls of the channel increase, the position of mechanical stability shifts to a position midway between wall sources along the z -axis. This fact can be accounted for with the formulae generated above [32].

3.3 CLASSICAL VIBRATIONS OF IONS AT MINIMA IN THE CHANNELS

As was noted in the last section, for all reasonable values of the parameters in the potential energy function, the regular sequence of potential energy wells evolve to show minima *between* the locations of the wall sources. This is true even for the case in which, for two or a small number of wall sources, the original position of equilibrium lies in the middle of the configuration. It is clear, therefore, that if in reality the ion occupies a position of equilibrium in the plane of a group of wall sources, strong angular many-body terms must be working to maintain this configuration. The situation seen in these examples arises most likely from the nearly exclusive use of pair potentials. The many-body potential in the form tested is not sufficiently bond-specific to force the location of the ion to a mid-plane position of equilibrium with respect to the wall sources.

For the parameters investigated earlier to obtain representative frequencies for sodium (for example) in a polar solvent, we obtain the following frequencies in the full channel system.

For some choices of the parameters, it is possible to obtain very small values for the axial vibrations of the ion along the channel. As we will see, band states for these systems correspond to the nearly free ion limit. On the other hand, it is possible to find other axial frequencies reaching nearly 20 cm^{-1} . In these cases, we have treated (*post*) the bands in the tight binding limit. With distortions of the channel walls in the vicinity of the ion, a case not specifically illustrated here, it is possible to see the formation of deep traps of several decades in axial frequency. In these cases, the ion transfer along the chain would occur *via* an activated site-hopping mechanism. The formalism necessary to treat such a case is considered later in Section 5.7.

The frequencies were calculated for the zeroth cell. The x -polarized mode is lower than the y -polarized mode because it corresponds to a vibration perpendicular to the line between opposed sources that make up the helical wall. In a channel composed of (at

least) three helical strands, the transverse vibrations perpendicular to the channel axis would be more nearly degenerate.

TABLE 2.

Representative frequencies for the vibrations of ions in channels

$R_0(\text{\AA})$	$\rho(\text{\AA})$	$D(eV)$	$\alpha(\text{\AA}^{-1})$	Θ_{twist}	$b(\text{\AA}^{-2})$	$\omega_{\text{cl}}(\text{cm}^{-1})$	ω_x	ω_y	ω_z
3.2	3.06	0.6	1.00	50.255	1.00	198.12	117.62	233.96	10.33
3.2	3.06	0.6	1.00	50.260	1.00	198.35	117.68	234.04	10.34
3.3	3.17	0.6	1.00	48.365	1.00	195.77	113.14	233.54	8.92
3.3	3.17	0.6	1.00	48.368	1.00	195.89	113.17	233.58	8.92
3.4	3.28	0.6	1.00	46.635	1.00	193.63	109.03	233.20	7.72
3.4	3.28	0.6	1.00	46.636	1.00	193.69	109.04	233.22	7.72
3.5	3.39	0.6	1.00	45.045	1.00	191.70	105.23	232.97	6.69
3.5	3.39	0.6	1.00	45.046	1.00	191.73	105.23	232.98	6.70
3.4	3.40	0.3	1.75	45.000	1.75	195.17	52.56	215.83	15.92

In all the calculations, the step distance is $\Delta = 1.5 \text{ \AA}$. Remaining quantities defined in the equations. ω_{cl} is the classical harmonic frequency associated with regular solvation, see Table 2.

4. FREE AND NEARLY FREE IONS

At this point, we begin to consider the nature of band states in channel conductors. As we shall see, the bands can be very narrow. This is due in part to the fact that ionic masses are large compared to electrons for which we understand band states in solids [33-35] and the energies associated with their vibrational states are small. Whether or not a band picture has any practical bearing on the mechanical properties of ions in channels depends on the phenomenon studied. The band width generally is too small to have any real effect on vibrational spectral band broadening in channels at ambient temperatures. On the other hand, for ion tunneling across a channel gate, as would be the case for one of the channel conductors in an axonal membrane, the band states may figure importantly in the interpretation of the tunnel transfer rate constant. As indicated at the beginning of this article, however, the goal here is to attempt to formulate a band picture as an adequate representation of the system as one sees it. Where the classical limit is appropriate, the correspondence limit can be used. Subsequent incorporation of vibrational states of the channel wall, not specifically done in this article, may lead to self-trapping. A band representation still applies. The system, however, exhibits ion transfer in such a limit as an activated process. This particular general aspect

of ionic conductivity in the tight binding limit is discussed in Section 5.

4.1 FREE IONS

For some parameters, the potential energy along the helical axis is essentially flat. This will probably be particularly true of the tubular fullerenes [36]. An infinite, flat channel is a one-dimensional ionic conductor for which the band of states has energies characterized simply by [33]

$$\frac{\hbar^2 k^2}{2m} . \quad (24)$$

It is of course well known from solid state physics that degenerate energies arise in periodic systems that result in band gaps [37,38]. Thus, this overly simple representation is inadequate as a general picture of the band states of the ionic system.

4.2 NEARLY FREE IONS: GENERAL PROPERTIES

When the ion (in analogy to the electronic conduction problem [37,38]) interacts with a periodic sequence of energy minima, the problem becomes one of the states of a *nearly free ion* in the channel. There are certainly bands of vibrational states for the transverse motions of the ion normal to the channel axis, and discussion of those vibrations is found in the next section.

The usual analysis assumes standing waves of the form [37]

$$\psi = a e^{ikz} + b e^{i(k-g)z} \quad (25)$$

where g is the reciprocal lattice vector with values

$$g = 2\pi m/\zeta \quad m = 0, 1, 2, 3, \dots \quad (26)$$

and ζ is the length of the unit cell. The Hamiltonian operator is

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dz^2} + V(z) \quad (27)$$

where $V(z)$ is the effective periodic potential along the axis, viz., the potential measured on the axis.

A standard variational calculation yields energies [37]

$$E(k) = \epsilon(k) \pm | \langle k | V | k-g \rangle | \quad (28)$$

with

$$\epsilon(k) = \hbar^2 k^2 / 2m. \quad (29)$$

The matrix element, $\langle k | V | k-g \rangle$, of the potential is given by

$$\langle k | V | k-g \rangle = \frac{1}{\zeta} \int_0^\zeta dz e^{2\pi i m z / \zeta} V(z). \quad (30)$$

The average value of the potential, $\langle V \rangle = \langle k | V | k \rangle$, is a constant that is absorbed into the value of E .

With this type of analysis, the bandwidth is

$$\Delta E = E(\pi) - E(0). \quad (31)$$

For an ion with the mass of sodium, this is a very small quantity; it is roughly 42,297 (*viz.*, the mass of sodium versus the mass of the electron) times smaller than an equivalent band for an electron. Moreover, as indicated later, the bands associated with the transverse motions normal to the helical axis are also narrow.

TABLE 3
Parameters for the nearly free ion
(Calculated by a Fast Fourier Transform)

Variable	Value
d_{\min}	3.00 Å
ζ	1.50 Å
ρ	3.20 Å
D	0.01 eV
a	2.00 Å ⁻¹
R_0	3.30 Å
Z	1.00 esu
b	0.25 Å ⁻²
μ	0.05 Debye
a	44.67 Å ⁻²
Gap	9.768×10^{-4} eV

Note that the $k=\pi$ interband transition corresponds to a microwave frequency of 237 GHz, or a wavelength of 8 cm⁻¹.

An estimate of the size of the matrix element $\langle k | V | k-g \rangle$ takes advantage of the fact that eq (30) is simply the Fourier transform of $V(z)$. The axial potential is a composite of interactions of the ion with all the source atoms in the wall of the channel. Here, however, one

notes that if the values of the potential are known at a sequence of points within the first unit cell, it is possible accurately to estimate the value of the matrix element by means of a fast Fourier transform [39]. Table 3 lists the band gap found using matrix elements evaluated *via* the fast Fourier transform.

Even though the parameters of the potential energy function may not be accurate, they yield a potential that satisfies the nearly free ion limit. The Morse component, for example, is probably too small. In most molecular systems, the non-bonding interactions that arise from van der Waals attractive forces should exceed the value used. Second, the value of the Gaussian scale in the potential is far too small, indicating an extremely diffuse charge that inadequately screens the valence core charge of 1 esu . Finally, it is probably the case that the dipole moment is unrealistically small. The character of the potential is largely one of a weak dipolar attraction and a core repulsion. Nevertheless, the fast Fourier transform yields a value of the band gap that is close to the value found below *via* a more detailed analysis.

4.3 NEARLY FREE IONS: DIRECT USE OF THE MODEL POTENTIAL

An ion in a channel moves both in longitudinal and transverse directions under the influence of interactions with sources in the channel wall; this much is evident from the classical limit analysis above. Because the effective potential involves a sum of all the wall interactions, that function is not easily separable into Cartesian components. Instead, matrix elements of the ion, referred to a coordinate origin for the channel system, must be evaluated for each contribution and the result summed. The Hamiltonian operator for the system is given by

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \quad (32)$$

in which the potential is the sum of the Morse, Coulomb and possible dipole terms, for simplicity. For an adequate representation of the low-lying transverse vibrations in the nearly free ion case it is sufficient to consider s - and p_x - and p_y -type contributions in the plane normal to the axis. The degrees of freedom along the axis still dictate a plane wave.

The basis functions for the nearly free ion are the following

$$\phi_s(\mathbf{r}) = \sqrt{2a/\pi N\zeta} \exp(-ap^2 + ikz) \quad (33)$$

$$\phi_{x(y)}(\mathbf{r}) = \sqrt{8a^2/\pi N\zeta} x(\text{or } y) \exp(-ap^2 + ikz) \quad (34)$$

normalized to $N\zeta$ where again ζ is the unit cell-length and N is total number of cells. The coordinate ρ is

$$\rho = \sqrt{x^2 + y^2}. \quad (35)$$

The state function is a linear combination of the three basis functions:

$$\Psi = \sum \left[a_i \phi_i(k) + b_i \phi_i(k-g) \right]. \quad (36)$$

A variational treatment yields three block diagonal 2×2 secular determinants, although this result is not initially immediately obvious. The structure of the potential energy matrix elements is discussed in detail in Appendix C. The diagonal matrix elements are readily found to be

$$H_{ss} = \frac{\hbar^2}{2m} (2a + k^2) + \langle V \rangle \quad (37)$$

$$H_{qq} = \frac{\hbar^2}{2m} (4a + k^2) + \langle V \rangle \quad \text{with } q = x, y \quad (38)$$

where $\langle V \rangle$ is the average value of the potential; equated here to zero. The primed matrix elements, *e.g.*, $H_{s's'}$, have k replaced by $k - g$. The off-diagonal matrix elements are

$$H_{qq'} = \langle q(k) | V | q' (k-g) \rangle \quad q = s, x, y. \quad (39)$$

As with electrons in metals [37], here as well one sees that there are band gaps for values of $k = g/2 = \pi m/\zeta$. For the s -band, for example, there are forbidden zones, or gaps, of $2 | \langle s(\pi m/\zeta) | V | s(-\pi m/\zeta) \rangle |$.

In order to see the emergence of the 2×2 block diagonal structure of the problem, it is instructive to specialize these general formulae to the problem. In the first place, it is relatively easy to see from the structure of the matrix elements of the Coulomb and ion-dipole interaction operators that s - $p_{x(y)}$ mixed matrix elements vanish when the individual contributions are summed over all the possible interactions of the chain. In particular, for the $\langle s | V | x \rangle$ type of matrix element, one has

$$\langle s | V | x' \rangle = \langle X \rangle \mathfrak{F}_{-1}(R, g) \quad (40)$$

and the average value of X , $\langle X \rangle$, vanishes for the whole chain. This is true for the Morse potential contributions as well as for the

electrostatic terms. It is not true, however, that terms of the form $\langle q|V|q' \rangle$ vanish.

TABLE 4
Parameters for the nearly free ion
(Evaluated "exactly")

Variable	Value
d_{\min}	3.00 Å
ζ	1.05 Å
ρ	3.20 Å
D	0.01 eV
α	2.00 Å ⁻¹
R_0	3.30 Å
Z	1.00 esu
b	0.25 Å ⁻²
μ	0.05 Debye
a	44.67 Å ⁻²
Gap	1.155×10^{-4} eV

Here the $k = \pi$ interband transition frequency corresponds to 28 GHz.

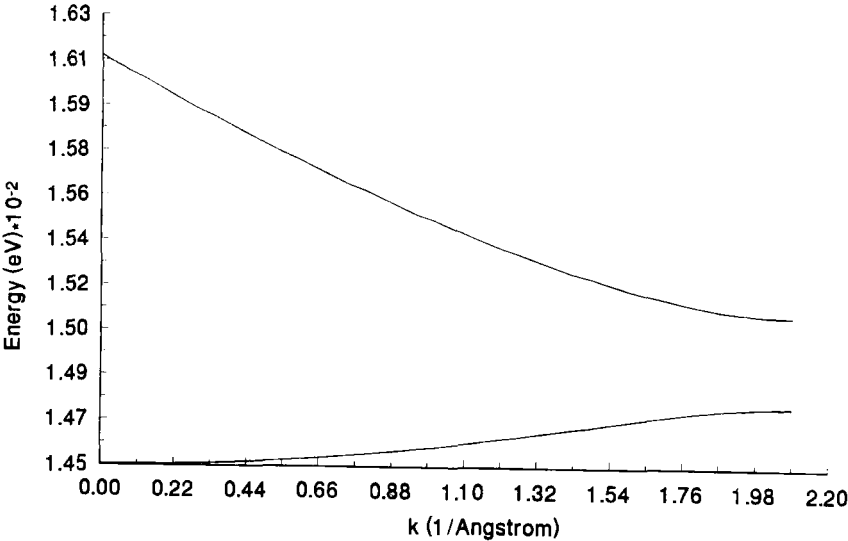


FIGURE 5. The *s*-band in the nearly free ion limit

The application of this method of detailed analysis with the same system parameters examined above (using the fast Fourier transform [39]) yields the following value for the band gap shown in Table 4.

The character of the band states is illustrated in Figs (5) and (6). As noted above, the parameters used were unrealistic, but chosen to satisfy a limit that allows a proper set of bands and gaps to arise. By choosing $D = 0.1 \text{ eV}$, $b = 2.0 \text{ \AA}^{-2}$ and $\mu = 0 \text{ Debye}$, one gets the bands shown in Figures 5 and 6; for non-zero μ , the contribution of ion-dipole interactions is found to be small. As can be seen, the band gap is greater than the width of the lower band. For values of the parameters only slightly larger, the lower band shows inverse behavior; *i.e.*, the band energy decreases from $k = 0$ to $k = g/2$. At this point, it is appropriate to consider the tight binding limit. Even though this limit arises for very small well depths, for the nature of the system, and the fact that these are zero Kelvin calculations, the tight binding approximation applies because free or nearly free ions are no longer possible.

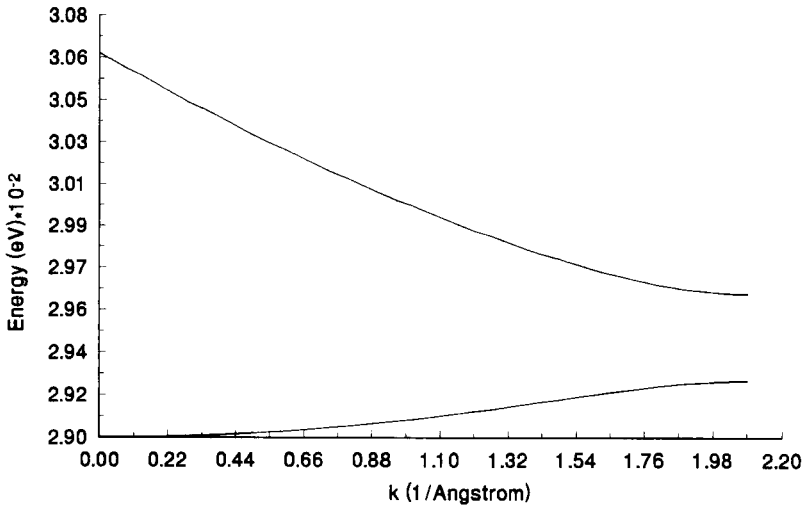


FIGURE 6. The p_x -band in the nearly free ion limit

4.4 MOBILITY IN THE NEARLY FREE ION LIMIT

The formalism that has been used for the free and nearly free ion limits is essentially the same as that used for conduction electrons in

metals. As a result, for all practical purposes, matters of ionic conductivity in this limit are also the same as for a conduction electron. Ionic conduction depends on the lifetime τ of the ion in any particular band state. The calculation of this quantity is well-known in all text books of solid state physics [see, for example, refs. 33-35,37], and need not be duplicated here. It is worth noting, however, the fact that the lifetime τ will depend on the scattering of the ion by vibrations of the atoms in the wall of the channel. It is certainly the case that for vibrations of the channel ion transverse to the channel axis, the interaction with the wall source vibrations should be large. Whether or not this interaction will lead to self-trapping of the ion and an associated low conductivity needs to be explored.

It is necessary to note, in concluding this section, that the states of the ion discussed here are *single* ion states. The system consists of one ion in an otherwise empty, infinitely long channel. As a consequence, the band states arise only because of delocalization due to overlap of the ion at separate locations. These states are not phonon modes that arise from the cooperative motion of a chain of sources vibrating in unison. Because it is likely that the channels ought to be multiply occupied, phonon modes should arise from the cooperative vibrations of the interacting ions in the chain. If there are empty lattice sites between ions, then tunneling also contributes. The issue of channel ion phonon modes is discussed briefly in Section 5.6.

5. TIGHT BINDING ANALYSIS

When, due to the nature of the potential energy functions used and the spacing of the wall molecules, there arises a regular sequence of reasonably deep wells along the axis, it is necessary to consider the use of the tight-binding approximation. Band states arise only because of the translational symmetry on the z -axis. As a result, rather simple expressions can be derived that owe much to similar quantities that have been derived for the cyclic conjugated polyenes [40]. The treatment developed here follows reasonably closely analogous approaches to molecular electronic and solid state structure calculations.

5.1 ONE-DIMENSIONAL HARMONIC OSCILLATOR ANALYSIS

Some time ago Schmidt and Korzeniewski [11] presented a simple one-dimensional tight binding limit analysis of the problem of band states for ionic conduction. Although the potential energy function used was harmonic, delocalization was considered in the same manner as was used years ago in the Wall-Glockler [41] analysis of the inversion doubling of ammonia. Because of the simplicity of the calculation, and its ease of application, I summarize it here. The treatment is limited, at this

stage, to one dimension and therefore does not address the issue of vibrations normal to the transfer axis.

The system is one-dimensional with closed, cyclic boundary conditions. The Hamiltonian operator is also one-dimensional, as in the previous section, but the potential now is represented as a sequence of intersecting harmonic wells [11]:

$$V = \frac{1}{2} k \sum_{m=-\infty}^{\infty} (z \pm 2(m+1)\zeta)^2 \quad (41)$$

for which $m = 0, 1, 2, \dots$, is the well index, k is a uniform force constant, and ζ is the lattice spacing. Values for the force constant are conveniently derived with the use of the formulae given in Section 3 for the classical vibrations of the channel system.

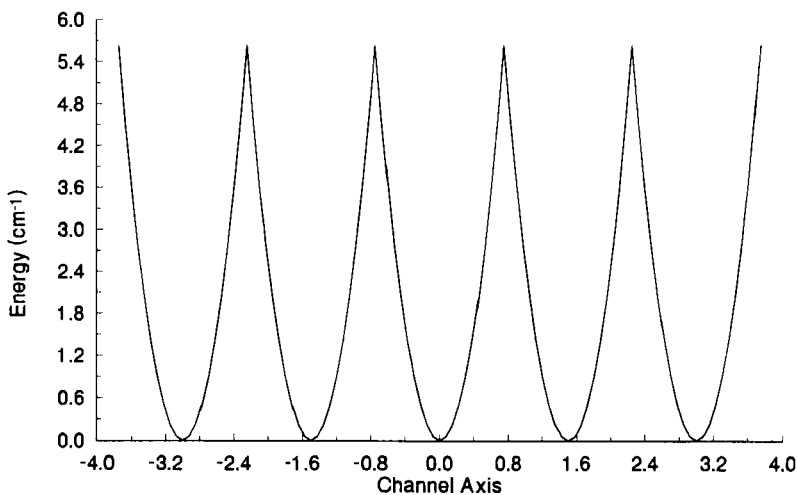


FIGURE 7. The representation of the channel as a sequence of intersecting harmonic wells.

A state function Ψ is expanded in terms of Bloch functions for the ion in its ground and first excited state vibrational states. (The particular form of the functions and the resulting individual matrix elements are discussed shortly.) Thus

$$\Psi = A \sum_n e^{in\theta} \phi_{0n} + B \sum_n e^{in\theta} \phi_{1n} \quad (42)$$

in which θ is $2\pi/N$ where N is the number of cells in the chain. Given an operator O of the Hamiltonian, including the overlap operator $\mathbf{1}$, matrix elements between the ground and first excited state basis functions satisfy

$$O_{n,-m}^{1\ 0} = - O_{m,-n}^{1\ 0} \quad (43)$$

and one obtains, in the usual manner, Hermitian matrices. The eigenvalue equation has complex conjugate off-diagonal elements; as a consequence, the energies for the two bands are given by

$$E = \frac{1}{2(s_{00}s_{11}-s_{01}^2)} \left\{ h_{00}s_{00} + h_{11}s_{11} + 2h_{01}s_{01} \right. \\ \left. \pm \left[(h_{11}s_{00} + h_{00}s_{11} - 2h_{01}s_{01})^2 + 4(s_{00}s_{11}-s_{01}^2)(h_{01}^2 + h_{00}h_{11}) \right]^{1/2} \right\} \quad (44)$$

with the overlap matrix elements for the first zone given by

$$s_{qq} = 1 + 2 \sum_m \cos(m\theta) S_m^{qq} \quad (45)$$

in which q, p are indices that label the type of basis function, ground and excited: q, p (0,1).

$$s_{qp} = 2 \sum_m \sin(m\theta) S_m^{qp} \quad (q \neq p). \quad (46)$$

The matrix elements of the Hamiltonian operator are given in turn by

$$h_{qq} = \alpha_q + 2 \sum_m \cos(m\theta) \beta_m^q \quad (47)$$

$$h_{01} = 2 \sum_m \sin(m\theta) \Delta_m. \quad (48)$$

In terms of the primitive basis functions, one has

$$\alpha_q = \langle \phi_{qm} | H | \phi_{qm} \rangle \quad (49)$$

$$\beta_m^q = \langle \phi_{0m} | H | \phi_{qm} \rangle \quad (50)$$

$$\Delta_m = \langle \phi_{00} | H | \phi_{1m} \rangle \quad (51)$$

and S_m^{qp} is the usual overlap integral.

One sees from the functional dependence of the energy, eq (44), on trigonometric functions that the energies of the bands are of the standard form. Because of the one-dimensional nature of the problem and the 2×2 dimensions of the secular matrix, it is possible simply to include account of overlap automatically without specifically using methods of orthonormalization, such as Löwdin's method [42-44].

The first (nominally ground state) primitive function for the ion at a site m is

$$\phi_0(z_m) = (a/\pi)^{1/4} \exp(-az_m^2/2) \quad (52)$$

with

$$z_m = z - m\zeta. \quad (53)$$

The second primitive basis function is

$$\phi_1(z_m) = \sqrt{2a} (a/\pi)^{1/4} z_m \exp(-az_m^2/2). \quad (54)$$

In practice, because of the vanishing of the overlap integral with m^2q^2 , summation only needs to be considered over a few wells. For the first s -band, the quantity α_0 is simply

$$\alpha_0 = \langle 00 | H | 00 \rangle = \frac{1}{4} \hbar\omega + \langle 00 | V | 00 \rangle. \quad (55)$$

The specific expressions for matrix elements of the truncated harmonic well potential system are given in Appendix E. The quantity β_m^0 is a two-center integral given by

$$\beta_m^0 = \langle 00|H|0m\rangle = \frac{1}{4} \hbar\omega(1 - 4m^2q^2)S_m^{00} + \langle 00|V|0m\rangle. \quad (56)$$

In this expression S_m^{00} is the overlap integral given by

$$S_m^{00} = \exp(-m^2q^2) \quad (57)$$

where the overlap is always measured relative to the zeroth cell.

The quantities associated with the first excited state are listed as follows. The diagonal matrix element is

$$\alpha_1 = \langle 10|H|10\rangle = \frac{3}{4}\hbar\omega + \langle 10|V|10\rangle. \quad (58)$$

The off-diagonal matrix element β_m^1 is given by

$$\begin{aligned} \beta_m^1 &= \langle 10|H|1m\rangle \\ &= \frac{3}{4}\hbar\omega \left(1 - 4m^2q^2 + \frac{16}{3}m^4q^4 \right) S_m^{00} + \langle 10|V|1m\rangle. \end{aligned} \quad (59)$$

Finally, the matrix element $\langle 00|H|1m\rangle$ is

$$\Delta_m = \langle 00|H|1m\rangle = -\frac{\hbar\omega}{2\sqrt{2}}mq(3 - 2m^2q^2)S_m^{00} + \langle 00|V|1m\rangle. \quad (60)$$

The evaluation of the energy based on these formulae is straightforward. One obtains an estimate of the harmonic frequency *via* the classical harmonic analysis. The quantity $a = m\omega/\hbar$ depends on the effective mass of the ion and the axial frequency. It is possible, of course, to get an immediate estimate of the extent of the band width for

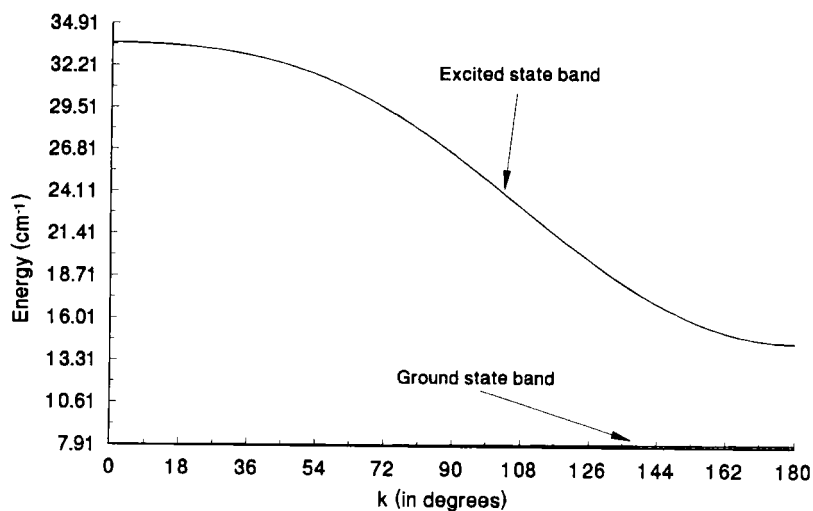


FIGURE 8. Axial channel bands for the ion in the one-dimensional harmonic well representation.

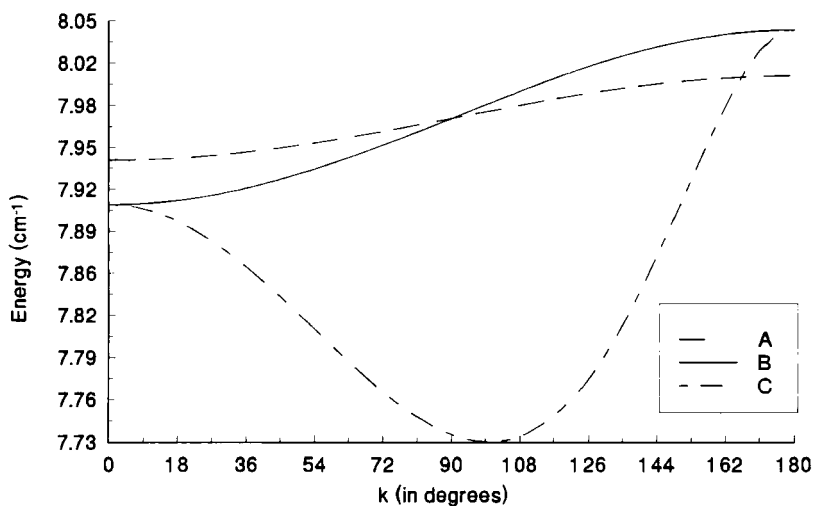


FIGURE 9. An enlargement of the (very narrow) ground state band. Curve A, complete neglect of overlap; B orthogonalized basis functions; and, C normalized with the overlap in the first zone.

any particular system from the quantity S_{01} , the overlap between adjacent wells. For a frequency of 10 cm^{-1} , a mass of 23 amu, and a distance from the minimum to the barrier-cusp of 0.75 \AA one finds $S_1^{00} = 0.022$ which is a reasonable quantity that suggests a sizable bandwidth. The band width of the ground state is very narrow, as seen in Figure 8. The excited state bandwidth is somewhat larger, as might be expected from overlap. Apart from the narrowness of the ground state band, the shape of the bands is that expected from general considerations of band theory.

It is well-known, since the original work of Wall and Glickler [41] on the inversion spectrum of ammonia in 1937, that the intersecting harmonic oscillator representation over-estimates the height of the barrier. Nevertheless, for ammonia it was found that an adjustment of the parameters could be found that yields reasonable agreement with the observed splittings. Thus, one expects that this intersecting oscillator model is a reasonable one, probably slightly better than the *ad hoc* square or triangular box barrier. It is easy to use and may provide adequate estimates of the band widths one can expect for tunnel systems.

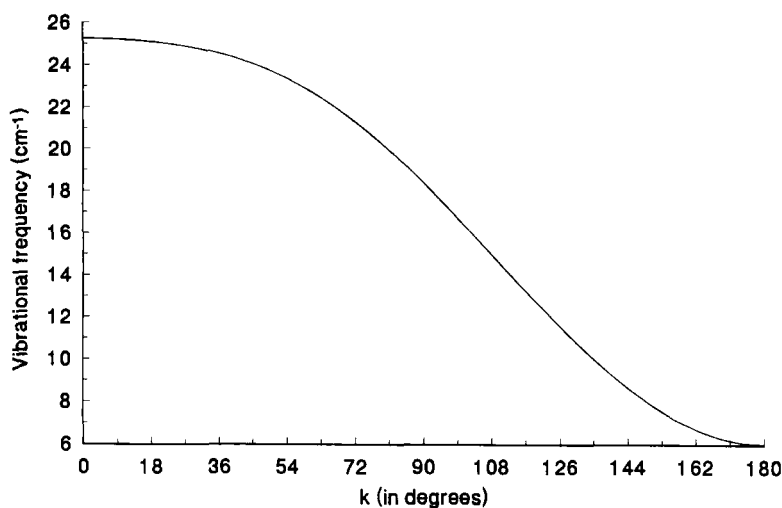


FIGURE 10. Plot of the frequency dispersion curve for the interband transition. This dispersion effectively follows the band shape of the excited state.

The work of Schmidt and Korzeniewski [11] proposed that supraionic conductivity might be seen in those systems for which the band width exceeds the height of the barriers. The height of the barrier is determined in this model by computing the energy distance between the minimum and the cusp

$$E_{\text{barrier}} = \frac{1}{2} \hbar \omega q^2. \quad (61)$$

For the frequencies chosen in that work [11], it was possible to see this effect. For the frequencies associated with the potentials examined here, and the associated frequencies of oscillation in the harmonic limit, we see that supraionic conduction is unlikely to exist in the ground state band. On the other hand, in the excited state, the width of the band is much greater. Indeed, because of the enhanced overlap associated with the larger amplitude axial vibrations in the excited state, this is to be expected. In essence, a *p*-type vibration is bond polarized along the axis and hence the band width ought to be greater than it is in the *s*-type state. For $\omega = 10 \text{ cm}^{-1}$ and $\zeta = 0.75 \text{ \AA}$, the barrier height for intersecting wells using sodium is about 30 cm^{-1} . This is a quantity greater than the height of the barrier for an actual interaction.

It is not reasonable, however, to rule out the possibility of supraionic conductivity merely because the criterion posed above is not satisfied. Another important factor ignored so far is the fact that the helical channels do not exist in a vacuum. Although for each helical unit cell used here there is a wide space in the *x,z* plane that does not contain a wall source, it is certainly the case that in real systems there are atoms of the underlying solvent structure in this region with which the ion can interact. Thus, in fact, a more realistic model of the conduction system consists of a triple or quadruple helical system of molecular species. The strands need not be identical. The result therefore of a dense medium that surrounds the channel is a further smoothing of the axial potential along the conduction channel. The smoothing results further in a lowering of the barriers to the point that, although the band width is probably small, the barriers are of the same dimension (in energy). Thus, supraionic conductivity is possible. The remaining issue has to do with scattering or activated transfer, as noted above.

5.2 IMPROVEMENT OF ACCURACY: ANHARMONIC POTENTIALS

The harmonic analysis can be improved greatly, even in one dimension, through the inclusion of higher order terms in the Taylor series expansion of the actual potential along the channel axis. Alternatively, it is possible also to consider the use of spherical Gaussian basis functions together with expansions of the Morse and other potentials in forms that are relatively easily integrated. This section summarizes a one-dimensional treatment with anharmonic potentials developed to high order. In the next section an alternative treatment is considered in three-dimensional Gaussian basis functions. The extension of this treatment to the case of band states would follow the general outline

given above, but will not be done here. The discussion of the limited two-well example serves adequately to illustrate the method.

The high order one-dimensional approach was first used to consider the inversion spectrum of ammonia along the v_2 inversion axis [45]. A similar calculation has been used to examine the influence of solvent on the splitting due to tunneling when a hydrogen atom, for example, can move between two positions of equilibrium [46].

The Taylor series restricted to the z -axis can be expressed as [28,45,46]

$$G(z+\mathbf{R}) = N_s \sum_{n=1}^{\infty} \frac{z^n}{n!} \sum_l A_{nl} P_l(\cos \theta_R) I_{nl}(R). \quad (62)$$

The variable θ_R is the angle between the vector drawn from the origin to the source of the external field and the z -axis. For the Coulomb (including the dipole), Morse, and three-body fields discussed it is a relatively straightforward, although not always an algebraically simple matter, to obtain the functions $I_{nl}(R)$. Expressing the potential simply as [45]

$$V(z) = \sum_{n=0}^{\infty} C_n z^n, \quad (63)$$

it is necessary only to evaluate matrix elements of z^n in the one-dimensional basis set. The coefficients C_n , from eq (44), are easily identified as

$$C_n = \frac{N_s}{n!} \sum_l A_{nl} P_l(\cos \theta_R) I_{nl}(R). \quad (64)$$

The Morse potential, for example, gives [45]

$$I_{nl}(R) = \alpha^n D \exp(\alpha R_0) \left\{ 2^n \exp(\alpha R_0) \left[2\alpha R k_{l-1}(2\alpha R) - (n-l)k_l(\alpha R) \right] \right. \\ \left. - 2 \exp(\alpha R_0) \left[\alpha R k_{l-1}(\alpha R) - (n-l)k_l(\alpha R) \right] \right\} \quad (65)$$

in which $k_l(x)$ is the modified spherical Bessel function of the third kind [30]. Similar expressions can be derived for other potentials. For the zeroth and first vibrational states, the matrix elements of z^n are given by the following expressions:

$$\langle a0|z^n|b0\rangle = 2^{1/2} \frac{(ab)^{1/4}(n-1)!!}{(a+b)^{(n+1)/2}} \exp(-Q^2) \quad (66)$$

with

$$Q^2 = \frac{1}{2} \frac{ab}{a+b} R^2 \quad (67)$$

and a and b are the Gaussian exponents of the two centers. For single center terms, $R = 0$ and $a = b$; the expression reduces to the correct form. The expansion coefficients $C_n(\mathbf{R})$ are evaluated for the two center contributions at the Gaussian center-of-gravity \mathbf{P}

$$\mathbf{P} = \frac{a\mathbf{A} + b\mathbf{B}}{a+b}. \quad (68)$$

For the first excited state, the matrix element is more complicated; it is

$$\begin{aligned} \langle a1|z^n|b1\rangle = & \frac{2^{3/2}(ab)^{3/4}(n-1)!!}{(a+b)^{(n+3)/2}} \left\{ \left[n + 1 - \frac{abR^2}{a+b} \right] \right|_{n(\text{even})} \\ & - \frac{(a-b)n!!R}{(a+b)^{1/2}(n-1)!!} \right|_{n(\text{odd})} \left. \right\} \exp(-Q^2). \quad (69) \end{aligned}$$

The general matrix elements of the kinetic energy operators are given by

$$\langle a0|ke|b0\rangle = \frac{\hbar}{2^{1/2}} (\omega_a \omega_b)^{1/2} \frac{(ab)^{3/4}}{(a+b)^{3/2}} \left[1 - \frac{abR^2}{a+b} \right] \exp(-Q^2) \quad (70)$$

and

$$\langle a1|ke|b1\rangle = \hbar(2\omega_a\omega_b)^{1/2} \frac{(ab)^{1/4}}{(a+b)^{3/2}} \left\{ \frac{3ab}{(a+b)^2} \right. \quad (71)$$

$$\left. + \frac{abR^2}{(a+b)^3} (a^2 + b^2 - 4ab) - \frac{abR^2}{a+b} \left[1 - \frac{(ab)^2 R}{(a+b)^3} \right] \right\} \exp(-Q^2). \quad (72)$$

As noted, this approach has been used to evaluate the inversion doubling in ammonia [45]. The vibrational spectrum of ammonia is a good point of reference for any calculation because it is so completely characterized. Thus, using a two center expansion to represent the mirror-equivalent positions of the nitrogen atom with respect to the plane defined by the triangle of hydrogen atoms in the molecule, it was possible to obtain an excellent fit of the parameters of the Morse potential to the data. It must be remembered, of course, that the number of parameters fit equaled the number of frequencies used; thus, the optimization was ideal. Details of that, and a related single center treatment that was inspired by the Swalen-Ibers calculation two decades before, are given in ref. [47].

One can be reasonably confident that if it is necessary only to focus on the energy states of the ion on the axis, and if a one-dimensional calculation is acceptable, then high accuracy ought to be possible with the easy use of high orders in the Taylor expansion of the potential as indicated above. If all degrees of freedom along the channel are of interest, it is necessary to move to a fully three-dimensional calculation, as discussed next. Although not discussed here, for systems with strong axial symmetry it is possible to combine the simplicity of the one-dimensional treatment to high orders in the Taylor expansion with radial motions normal to the axis. Because the utility of such an approach over the direct use of a Cartesian basis, as illustrated next, is not clear, I think it is not useful to consider it here.

5.3 ANALYSIS IN THREE-DIMENSIONAL GAUSSIAN BASIS: USE OF THE FULL POTENTIAL

The following sections present a treatment of the channel ion problem in three dimensions. The basis functions, although limited, now allow for the mixing of ionic vibrational modes along the channel axis in the axial and transverse directions. Moreover, it is possible to consider in detail the nature of the mobile ion/channel wall source phonon interaction, as will be discussed later.

The basis set is a simple orthonormal one. In fact, there is automatic orthonormality between the p -functions and between the s and p

functions in the x,y -plane. Along the z -axis, however, due to s and p_z overlap, orthonormality can be established *via* the Gram-Schmidt or Löwdin [42-44] methods for example. The explicit calculation of overlap integrals easily demonstrates that, for the most part, overlap for ionic systems is small. It is frequently the case, therefore, that one can get reasonably good approximations of the vibrational states of the system by ignoring overlap. The simplicity of the system is sufficient, however, to enable one to carry out complete calculations without neglect of overlap. Indeed, it is almost possible to hand-diagonalize the eigenvalue problem as the overlap matrix is block diagonal. Although many of the off-diagonal terms of the Hamiltonian operator are small, they do not vanish. Thus, the full eigenvalue problem [which is of the order of four in the first Brillouin zone for the basis set used] requires machine calculation.

The Hamiltonian operator is the same as used above for the *nearly free ion* case; here, however, because of the nature of the system and the spacing of sources, traps arise. It is no longer possible simply to consider a plane wave along the z -axis.

The channel is not closed, which means the following. If the channel were defined by three or more helical chains such that any point on the axis and at the center of a planar distribution of sources the ion "sees" a regular triangular or higher order planar polygonal array, then the system is closed; vibrations in the x,y -plane normal to the channel axis would be degenerate. In the double chain system we are using as a model, the vibrations in the plane normal to the channel axis at any point are non-degenerate. It is not possible, therefore, simply to consider the x,y -components of the ion wavefunction at any point to be the same as at any other point. In fact, a helical translation operator, similar to those used to generate the channel walls, eqs (1 & 2), operates to generate the ionic wavefunctions as an extension of the Bloch theorem.

In the following, $\chi_q(\mathbf{r})$ stands for a primitive basis function: $q = s, x, y, z$; a non-normalized element of the extended Bloch set is ϕ_q ; and, finally, ψ_q is an orthonormal band function of the type q . Let the primitive basis set consist of one s -function and three p -functions for the ion expanded at each position of equilibrium along the channel. Thus,

$$\chi_s(\mathbf{r}_n) = \exp[-a(x^2 + y^2 + (z - Z_n)^2)] \quad (73)$$

$$\chi_q(\mathbf{r}) = (q_n - Q_n) \exp[-a(x^2 + y^2 + (z - Z_n)^2)] \quad (74)$$

with $(q, Q) = (x, X; y, Y; z, Z)$. At any point Z_n along the channel that is a function of n displacements ζ , the x - and y -components of the wavefunction will be defined with reference to the coordinate system of the first cell by n rotations of the twist angle ϑ . Thus, the basis set at a point Z_n is

$$\phi_s(n) = e^{ink\zeta} \chi_s(n) \quad (75)$$

$$\phi_z(n) = e^{ink\zeta} \chi_z(n) \quad (76)$$

$$\phi_x(n) = e^{ink\zeta} \left\{ \chi_x(n) \cos(n\vartheta) - \chi_y(n) \sin(n\vartheta) \right\} \quad (77)$$

$$\phi_y(n) = e^{ink\zeta} \left\{ \chi_x(n) \sin(n\vartheta) + \chi_y(n) \cos(n\vartheta) \right\}. \quad (78)$$

In these expressions, the argument (n) implies $(\rho, z - n\zeta)$ and ρ , as before, is the radial coordinate measured from the axis, and x and y are defined with reference to the coordinate system of the zeroth cell. The basis set for the band calculation is simply the sum of these individual terms: viz.,

$$\Phi_q(\mathbf{r}) = \sum_{n=0}^N \phi_q(n). \quad (79)$$

We first consider the nature of the overlap matrix for the first Brillouin zone. Elements of the s - and z -functions are straightforward. One finds

$$\langle s \| s \rangle = N_0 \left[\langle s(0) \| s(0) \rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \langle s(0) \| s(m) \rangle \right] \quad (80)$$

$$\langle z \| z \rangle = N_0 \left[\langle z(0) \| z(0) \rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \langle z(0) \| z(m) \rangle \right] \quad (81)$$

and

$$\langle s \| z \rangle = N_0 \left[\langle s(0) \| z(0) \rangle + 2i \sum_{m=1}^{\infty} \sin(mk\zeta) \langle s(0) \| z(m) \rangle \right] \quad (82)$$

where the notation is $\langle f(m) | \mathcal{O} | g(n) \rangle = \langle \chi_f(m) | \mathcal{O} | \chi_g(n) \rangle$ for an operator \mathcal{O} including the overlap (unity) operator. In these expressions, ζ is the lattice spacing.

Because of the axial location of the radial origin,

$$\langle x(m) \| s(n) \rangle = \langle y(m) \| s(n) \rangle = 0 \quad (83)$$

for all m, n .

Overlap matrix elements in the x, y -plane involve the additional complication of the rotations of coordinates to allow the alignment of the primitive function with the locations of the sources in the unit cell n . Thus, one finds for the diagonal elements $\langle x \| x \rangle$ and $\langle y \| y \rangle$

$$\begin{aligned} \langle x \| x \rangle = N_0 \left\{ \langle x(0) \| x(0) \rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \left\{ \cos(m\vartheta) \langle x(0) \| x(m) \rangle \right. \right. \\ \left. \left. - \sin(m\vartheta) \langle x(0) \| y(m) \rangle \right\} \right\}. \end{aligned} \quad (84)$$

For all coordinate origins located on the axis, however, $\langle x(m) \| y(n) \rangle = 0$. Thus,

$$\langle x \| x \rangle = N_0 \left[\langle x(0) \| x(0) \rangle + 2 \sum_{m=1}^{\infty} \cos(m\vartheta) \cos(mk\zeta) \langle x(0) \| x(m) \rangle \right] \quad (85)$$

with a similar expression for $\langle y \| y \rangle$. It is possible likewise to show that all of the off-diagonal overlap matrix elements $\langle x \| y \rangle = 0$. Therefore, the overlap matrix in the first Brillouin zone is block diagonal:

$$S = \begin{pmatrix} \langle s \| s \rangle & \langle s \| z \rangle & 0 & 0 \\ \langle z \| s \rangle & \langle z \| z \rangle & 0 & 0 \\ 0 & 0 & \langle x \| x \rangle & 0 \\ 0 & 0 & 0 & \langle y \| y \rangle \end{pmatrix}. \quad (86)$$

A virtue of the simplicity of this example is the fact that Löwdin orthogonalization [42-44,48] can be applied *by hand*. Let \mathbb{L}_{\pm} stand for the eigenvalues of the upper left hand block of S

$$\mathbb{L}_{\pm} = \frac{1}{2} (\langle s \| s \rangle + \langle z \| z \rangle) \pm \frac{1}{2} \left[(\langle s \| s \rangle - \langle z \| z \rangle)^2 + 4 \langle s \| z \rangle^2 \right]^{1/2} \quad (87)$$

the remaining eigenvalues are trivially 1,1. The diagonalization of the overlap matrix yields a unitary matrix U

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix}. \quad (88)$$

The Löwdin matrix is constructed using the diagonal elements of the overlap matrix:

$$\begin{aligned} L &= UL^{-1/2}U^\dagger \\ &= \begin{pmatrix} a & b & 0 & 0 \\ b & a & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (89)$$

with

$$a = \frac{1}{2}(\mathbb{L}_+^{-1/2} + \mathbb{L}_-^{-1/2}) \quad \text{and} \quad b = \frac{1}{2}(\mathbb{L}_+^{-1/2} - \mathbb{L}_-^{-1/2}). \quad (90)$$

The orthonormal basis set, therefore, is

$$\hat{\phi} = \phi L. \quad (91)$$

As a result, with H the matrix of elements of the Hamiltonian operator in terms of the basis set of the first Brillouin zone, the matrix \hat{H} in the orthonormal bases is

$$\hat{H} = L^\dagger H L. \quad (92)$$

We turn next to examine the general structure of the matrix elements of the Hamiltonian operator H in terms of the basis functions of the first zone. A discussion of the matrix elements of H in terms of the primitive basis functions is found in Appendix F.

The matrix elements of H in the s - and z -functions are similar to those of the overlap operator:

$$\langle s|H|s\rangle = N_0 \left[\langle s(0)|H|s(0)\rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \langle s(0)|H|s(m)\rangle \right] \quad (93)$$

and

$$\langle z|H|z\rangle = N_0 \left[\langle z(0)|H|z(0)\rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \langle z(0)|H|z(m)\rangle \right]. \quad (94)$$

The off-diagonal sz -matrix element is

$$\langle s|H|z\rangle = N_0 \left[\langle s(0)|H|z(0)\rangle + 2i \sum_{m=1}^{\infty} \sin(mk\zeta) \langle s(0)|H|z(m)\rangle \right]. \quad (95)$$

Note, $\langle z(-n)|H|s\rangle = \langle z(n)|H|s\rangle^*$, as is required for the Hermitian properties of the treatment.

At this point, if one were to consider only the axial properties of the system, the analysis would be equivalent to the extended one-dimensional analysis discussed above. The energies of the bands—now effectively one dimensional—are the same as given by eq (44) for the simple collection of harmonic wells above. The full analysis, as can be seen by considering the consequences of Löwdin orthonormalization, mixes character of the x and y coordinates into that of the axial modes, although the smallness of the off-diagonal matrix elements essentially preserves the block-diagonal character of the problem.

The matrix elements in x and y are the following. The diagonal element $\langle x|H|x\rangle$ is

$$\begin{aligned} \langle x|H|x\rangle = N_0 \left\{ \langle x(0)|H|x(0)\rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \left\{ \cos(m\vartheta) \langle x(0)|H|x(m)\rangle \right. \right. \\ \left. \left. - \sin(n\vartheta) \langle y(0)| \sum_{N \geq 1} V_N |x(m)\rangle \right\} \right\} \quad (96) \end{aligned}$$

and

$$\begin{aligned} \langle y|H|y\rangle = N_0 & \left\{ \langle y(0)|H|y(0)\rangle + 2 \sum_{m=1} \cos(mk\zeta) \left\{ \cos(m\vartheta) \langle y(0)|H|y(m)\rangle \right. \right. \\ & \left. \left. + \sin(m\vartheta) \langle x(0)| \sum_{N \geq 1} V_N |y(m)\rangle \right\} \right\} \quad (97) \end{aligned}$$

The summation involving V_N , in which N is the source index, arises from the fact that the helical *source* molecule at $-z$ reflected from z also has coordinates that transform as $x \rightarrow -x$ and $y \rightarrow y$. Because of the manner in which the helix is constructed beginning with the zeroth cell, the matrix elements of the first Brillouin zone involving p_x and p_y functions vanish. Matrix elements involving sources outside the zone do not vanish.

The off-diagonal matrix elements are now given. The s -basis functions are coupled to the x and y functions in different cells.

$$\begin{aligned} \langle s|H|x\rangle = N_0 & \left\{ \langle s(0)|H|x(0)\rangle + 2i \sum_{m=1}^{\infty} \sin(mk\zeta) [\cos(m\vartheta) \langle s(0)|H|x(m)\rangle \right. \\ & \left. - \sin(m\vartheta) \langle s(0)|H|y(m)\rangle] \right\} \quad (98) \end{aligned}$$

and

$$\langle s|H|y\rangle = N_0 \left\{ \langle s(0)|H|y(0)\rangle + 2i \sum_{m=1}^{\infty} \sin(mk\zeta) [\cos(m\vartheta) \langle s(0)|H|y(m)\rangle \right.$$

$$+ \sin(m\vartheta) \langle s(0) | H | x(m) \rangle \Big] \Big\}. \quad (99)$$

The imaginary component arises because of the coordinate rotation by $m\vartheta$ at site m relative to the origin. The xy -off-diagonal matrix elements are given by

$$\begin{aligned} \langle x | H | y \rangle = N_0 \Bigg\{ & \langle x(0) | H | y(0) \rangle + 2i \sum_{m=1}^{\infty} \sin(mk\zeta) [\cos(m\vartheta) \langle y(0) | H | x(m) \rangle \\ & + \sin(m\vartheta) \langle y(0) | H | y(m) \rangle] \Big\}. \end{aligned} \quad (100)$$

The matrix element $\langle x | H | z \rangle$ is the following:

$$\langle x | H | z \rangle = N_0 \left[\langle x(0) | H | z(0) \rangle + 2 \sum_{m=1}^{\infty} \cos(mk\zeta) \langle x(0) | H | z(m) \rangle \right]. \quad (101)$$

The construction of the Cartesian matrix element $\langle x | H | z \rangle$ yields an (accidentally) even function. The matrix element $\langle y | H | z \rangle$, on the other hand, is complex:

$$\langle y | H | z \rangle = N_0 \left[\langle y(0) | H | z(0) \rangle + 2i \sum_{m=1}^{\infty} \sin(mk\zeta) \langle y(0) | H | z(m) \rangle \right]. \quad (102)$$

It is evident from the structure of the matrix elements, and the structure of the associated eigenvalue problem, that the band states that arise are of standard form. The widths of these bands are small for small overlap. The diagonal matrix elements $\langle x | H | x \rangle$ and $\langle y | H | y \rangle$ are nondegenerate and show an energy difference that is a result of the fact that the interaction of the central ion with the sources in the zeroth cell is anisotropic. Because sources are placed on the y -axis, vibrations of the ion along the line that joins the sources have higher energy than the vibrations in the x or z directions perpendicular to y . This fact was clear in the classical analysis earlier. One needs to remember, however, that the channel is defined as a double helix. The presence of two strands of identical sources means that for some of the

matrix elements there is a cancellation of contributions. As a result, the off-diagonal matrix elements can be very small. This can be seen from the structure of the primitive matrix elements that are discussed in more detail in Appendix F.

The principal new feature to arise from this formalism is the band structure that attaches to the transverse vibrations of an ion normal to the channel axis. It is certainly evident that, in view of the smallness of the overlap matrix elements, the widths of the bands will also be small. The principal contribution of band structure for these transverse vibrations would be in spectral band broadening. It is possible that the overlap could contribute as much as 10 cm^{-1} to the width of any vibrational band arising from the transverse motions. For real systems, immersed in a polar solvent such as water, this contribution is probably not be significant in comparison to the inhomogeneous broadening that arises from the folding of environmental vibrations into the fundamental modes of the channel ion. The structure of the axial vibrations, however, continues to be essentially the same as that uncovered in the simpler analysis of the overlapping, truncated harmonic oscillators.

5.4 MATRIX ELEMENTS OF THE SCREENED POTENTIAL

Section 2.2 introduced a Gaussian screened Coulomb potential energy to account for the fact that on close approach of the channel ion to a wall source, some penetration of the outer soft electronic screening charge should give rise to an electrostatic repulsion (assuming that the ion in the channel is anionic). Thus, to some extent, a hard/soft model of the charge accounts for polarization effects. The evaluation of matrix elements of such a potential energy operator is the subject of this section.

In Section 5.3, and in Appendix F, consideration is given both to the general form of the matrix elements of the Hamiltonian operator H and specifically to the kinetic and simple exponential forms of the potential energy. The consideration of the exponential form arises in connection with the Morse potential and its expansion in terms of Gaussian wavefunctions [49]. This section presents the evaluation of matrix elements of the screened electrostatic interaction in the form used in earlier sections.

The composite potential translates to the form of eq (5) discussed in Section 2.2. The evaluation of matrix elements in terms of the Gaussian primitives, however, yields closed forms that are remarkably similar to the matrix elements of the bare electrostatic Coulomb interaction $1/r$. Matrix elements of the Coulomb operator are obtained in terms of matrix elements of the three center overlap integral using the well-known transformation [15,16]

$$V_C = \frac{1}{r} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} du \exp(-r^2 u^2). \quad (103)$$

The matrix element of this operator in terms of s -functions is [16,50]

$$\langle s | V_C | s \rangle = 2 \left(\frac{a+b}{\pi} \right)^{1/2} \langle s || s \rangle F_0(U) \quad (104)$$

with

$$U = (a + b)(\mathbf{P} - \mathbf{C})^2 \quad (105)$$

where \mathbf{C} is the vector location of the wall source and \mathbf{P} again is

$$\mathbf{P} = \frac{a\mathbf{A} + b\mathbf{B}}{a + b} \quad (106)$$

with \mathbf{A} and \mathbf{B} the vector locations of the two points of expansion of the basis functions. The function $F_0(U)$ is

$$F_0(U) = \frac{1}{2} \sqrt{\frac{\pi}{U}} \operatorname{erf}(U^{1/2}) \quad (107)$$

and the general function $F_n(U)$ is [16]

$$F_n(U) = \int_0^{\infty} dt \, t^{2n} \exp(-Ut^2). \quad (108)$$

The Obara-Saika [50] recursion relations for the three-center overlap matrix elements can be used to derive the recursion relations for the ion-wall source interaction. Specific formulae will be listed below after we consider the interaction with the soft Gaussian part of the charge distribution.

Matrix elements of the soft part of the composite electrostatic interaction are evaluated according to the following scheme. First of all, for the s -type soft charge distribution, the Fourier transform, as noted in Appendix A, is

$$\rho(\mathbf{k}) = \exp(-k^2/4b) \quad (109)$$

where b is the reciprocal of the Gaussian screening length. Given a product of s -functions over two centers, the Fourier transform of that charge distribution is the following:

$$\begin{aligned}\rho_{ss}(\mathbf{k}) &= \left[\frac{\pi}{a+b} \right]^{3/2} \exp \left[-\frac{ab}{a+b} (\mathbf{A} - \mathbf{B})^2 \right] \exp \left[-k^2/4(a+b) \right] \\ &= \langle s \| s \rangle \exp \left[-\frac{k^2}{4(a+b)} \right].\end{aligned}\quad (110)$$

The matrix element $\langle s | V_s | s \rangle$ is

$$\begin{aligned}\langle s | V_s | s \rangle &= \frac{Ze^2}{2\pi^2} \langle s \| s \rangle \int_0^\infty d^3k \, k^{-2} \exp \left[-\frac{k^2}{4} \left(\frac{1}{b} + \frac{1}{a+b} \right) \right] j_0(k|\mathbf{C}-\mathbf{P}|) \\ &= \frac{Ze^2}{|\mathbf{C}-\mathbf{P}|} \langle s \| s \rangle \operatorname{erf} \left[\left(\frac{(a+b)b}{a+b+b} \right)^{1/2} |\mathbf{C}-\mathbf{P}| \right]\end{aligned}\quad (111)$$

in which $j_0(kR)$ is the modified spherical Bessel function of the first kind [30]. In the following, define T to be

$$T = \left(\frac{(a+b)b}{a+b+b} \right) (\mathbf{C} - \mathbf{P})^2. \quad (112)$$

Then, in view of the definition of $F_0(T)$, we find that

$$\langle s | V_s | s \rangle = Ze^2 \frac{2}{\sqrt{\pi}} \left(\frac{(a+b)b}{a+b+b} \right)^{1/2} \langle s \| s \rangle F_0(T). \quad (113)$$

Higher order matrix elements can be obtained from this expression by differentiation with respect to $2aA_x$, $2bB_x$, and so on. It is possible, however, to treat the composite potential, and, thus derive recursion relations for that operator. The similarity between the bare Coulomb matrix element and that of the soft potential enables us to write

$$\langle s | V_C - V_S | s \rangle^{(0)} = 2 \sqrt{\frac{a+b}{\pi}} \langle s || s \rangle \left[F_0(U) - \left(\frac{b}{a+b} \right)^{1/2} F_0(T) \right]. \quad (114)$$

In this expression, the superscript (0) indicates evaluation with respect to the $F_0(X)$ functions. The general expression is

$$\langle s | V_C - V_S | s \rangle^{(n)} = 2 \sqrt{\frac{a+b}{\pi}} \langle s || s \rangle \left[F_n(U) - \left(\frac{b}{a+b} \right)^{1/2} F_n(T) \right] \quad (115)$$

which is needed in the evaluation of the higher order terms. The matrix element $\langle x | V_S | s \rangle^{(0)}$ is found as follows:

$$\langle x | V_S | s \rangle^{(0)} = \frac{1}{2a} \frac{d}{dA_x} \langle s | V_S | s \rangle^{(0)}. \quad (116)$$

The evaluation of this uses the following easily proved relation

$$F_{m+1}(T) = - \frac{d}{dT} F_m(T). \quad (117)$$

Thus, considering the differentiation of $\langle s || s \rangle$ with respect to A_x , one finds

$$\begin{aligned} \langle x | V_S | s \rangle^{(0)} = \frac{2Ze^2}{\sqrt{\pi}} \left[\frac{(a+b)b}{a+b+b} \right]^{1/2} \langle s || s \rangle & \left\{ \frac{b}{a+b} (B_x - A_x) F_0(T) \right. \\ & \left. + \frac{b}{a+b+b} (C_x - P_x) F_1(T) \right\}. \end{aligned} \quad (118)$$

Now, this can be written as

$$\langle x | V_S | s \rangle^{(0)} = (P_x - A_x) \langle s | V_S | s \rangle^{(0)} - \frac{b}{a+b+b} (P_x - C_x) \langle s | V_S | s \rangle^{(1)}. \quad (119)$$

As a result, for the composite potential, we now can write

$$\begin{aligned}
\langle x | V_C - V_S | s \rangle^{(0)} &= (P_x - A_x) \left(\langle s | V_C | s \rangle^{(0)} - \langle s | V_S | s \rangle^{(0)} \right) \\
&- (P_x - C_x) \left(\langle s | V_C | s \rangle^{(1)} - \frac{b}{a+b+b} \langle s | V_S | s \rangle^{(1)} \right). \quad (120)
\end{aligned}$$

The derivative of this expression with respect to $2bB_x$ yields the xx -diagonal matrix element, while differentiation with respect to $2bB_y$ or $2bB_z$ gives the Cartesian off-diagonal elements: we find

$$\begin{aligned}
\langle x | V_C - V_S | y \rangle^{(0)} &= (P_x - A_x) \left(\langle s | V_C | y \rangle^{(0)} - \langle s | V_S | y \rangle^{(0)} \right) \\
&- (P_x - C_x) \left(\langle s | V_C | y \rangle^{(1)} - \frac{b}{a+b+b} \langle s | V_S | y \rangle^{(1)} \right) \quad (121)
\end{aligned}$$

and for the diagonal term

$$\begin{aligned}
\langle x | V_C - V_S | x \rangle^{(0)} &= (P_x - A_x) \left(\langle s | V_C | x \rangle^{(0)} - \langle s | V_S | x \rangle^{(0)} \right) \\
&- (P_x - C_x) \left(\langle s | V_C | x \rangle^{(1)} - \frac{b}{a+b+b} \langle s | V_S | x \rangle^{(1)} \right) \\
&+ \frac{1}{2(a+b)} \left(\langle s | V_C | s \rangle^{(0)} - \langle s | V_S | s \rangle^{(0)} \right. \\
&\left. - \langle s | V_C | s \rangle^{(1)} + \frac{b}{a+b+b} \langle s | V_S | s \rangle^{(1)} \right). \quad (122)
\end{aligned}$$

In the limit of tight screening, $b \gg a+b$ and the electrostatic interaction arising from the imbalance between hard and soft charges vanishes. In terms of the quantities used here, a , b , b , and the distances involved, these contributions to the band states are usually small. Thus, the major contribution to the band states of an ion—in the absence of the consideration of induced dipolar polarization effects—arises from the Morse potential implying dispersion interaction effects.

5.5 MATRIX ELEMENTS OF THE EXPONENTIAL THREE-BODY POTENTIAL

The exponential form of the three-body potential discussed above admits matrix elements in a Gaussian basis in a straightforward manner. We consider the operator

$$V_{3b}(\mathbf{r}) = \mathfrak{A} \exp \left[-g(r_G - R_G^0) - h(r_H - R_H^0) \right] \quad (123)$$

in which the individual distances r_G and r_H are $r_G = \mathbf{r} - \mathbf{G}$, $r_H = \mathbf{r} - \mathbf{H}$ and where \mathbf{G} and \mathbf{H} are the vector locations of the wall sources with which the test particle, the channel ion, interacts simultaneously. In terms of the transformation of the exponential function to an optimized sum of Gaussian functions *via* the Hehre-Stewart-Pople [51] formula [see Appendix C], we write

$$V_{3b} = \mathfrak{A} \exp(gR_G^0 + hR_H^0) \sum_{i,j} d_i d_j \exp \left[u_{iG}(r-G)^2 - u_{jH}(r-H)^2 \right] \quad (124)$$

where the coefficient u_{iG} (see eq (C6)) is

$$u_{iG} = g^2 \alpha_{1s,i} \quad (125)$$

and a similar term applies to u_{jH} . The matrix element $\langle s | V_{3b} | s \rangle$ depends on a sum of terms of the form

$$\begin{aligned} \langle s | E_{ij}^{3b} | s \rangle &= \left(\frac{a+b}{a+b+u_{iG}+u_{jH}} \right)^{3/2} \langle s || s \rangle \exp \left[-\frac{u_{iG}u_{jH}}{u_{iG}+u_{jH}} (\mathbf{G} - \mathbf{H})^2 \right] \\ &\times \exp \left[-\frac{(a+b)(u_{iG}+u_{jH})}{a+b+u_{iG}+u_{jH}} (\mathbf{P} - \mathbf{L}_{(ij)})^2 \right] \end{aligned} \quad (126)$$

in which \mathbf{P} is the same as given before, and $\mathbf{L}_{(ij)}$ is

$$\mathbf{L}_{(ij)} = \frac{u_{iG}\mathbf{G} + u_{jH}\mathbf{H}}{u_{iG} + u_{jH}}. \quad (127)$$

Matrix elements of the other primitive basis functions are readily obtained by differentiation. It is possible to demonstrate that the form of the matrix elements is very much the same as for the simpler matrix elements that arise with the Morse potential. Thus, one finds for $\langle x | E_{ij}^{3b} | s \rangle$,

$$\langle x | E_{ij}^{3b} | s \rangle = (M_{(ij)x} - A_x) \langle s | E_{ij}^{3b} | s \rangle. \quad (128)$$

Similarly,

$$\langle x | E_{ij}^{3b} | y \rangle = (M_{(ij)x} - A_x) \langle s | E_{ij}^{3b} | y \rangle \quad (129)$$

and

$$\begin{aligned} \langle x | E_{ij}^{3b} | x \rangle &= (M_{(ij)x} - A_x) \langle s | E_{ij}^{3b} | x \rangle \\ &+ \frac{1}{2(a + b + u_{iG} + u_{jH})} \langle s | E_{ij}^{3b} | s \rangle \end{aligned} \quad (130)$$

where \mathbf{M} is given by

$$\begin{aligned} \mathbf{M}_{(ij)} &= \frac{(a+b)\mathbf{P} + (u_{iG} + u_{jH})\mathbf{L}_{(ij)}}{a + b + u_{iG} + u_{jH}} \\ &= \frac{a\mathbf{A} + b\mathbf{B} + u_{iG}\mathbf{G} + u_{jH}\mathbf{H}}{a + b + u_{iG} + u_{jH}}. \end{aligned} \quad (131)$$

The full expression for the matrix elements of the three-body potential are obtained by summation:

$$\langle f | V_{3b} | g \rangle = \mathfrak{A} \exp(gR_G^0 + hR_H^0) \sum_{ij} d_i d_j \langle f | E_{ij}^{3b} | g \rangle. \quad (132)$$

Numerical experiments with this form of three-body potential in the classical harmonic limit indicate that its effect on the simple, local vibrational structure is small. It needs to be noted that this is a three-body effect involving wall sources and the ion. A similar three-body term operating between several ions in a multiply occupied channel may have a substantial effect on the energy states of the ions in the channel as well as on the mobility of the ion.

With reference to hydrogen bonding, I have recently found that the three-body term can have an important and profound influence on the effective *AB*-bond due to hydrogen in the *AHB* system. The sensitivity of the hydrogen bond to parameters in the *AH* and *BH* potentials is greater than one would expect on the basis of the influence of the term on classical limit properties [unpublished].

5.6 MOBILE ION-CHANNEL PHONON COUPLING

The issue of the coupling between the mobile ion and the vibrating channel wall sources is an extremely important one if the mobility of the ion is to be characterized accurately. The question of the self-trapping of the ion in channel deformations needs to be resolved for any particular model of the channel; whether or not self-trapping occurs can have a crucial bearing on the nature of the ionic mobility. It is, however, a relatively straightforward matter to develop the important interactions. There are two limits to consider: weak coupling in which the distortion of the channel walls is minimal and strong coupling in which deformation of the channel walls from the original regular radial spacing is sufficiently great to lead to deep traps. One can argue that there are no intermediate cases for the reason that as soon as there is any permanent distortion of the wall, the nature of the vibrational problem changes from one of an essentially free ion moving in a regular, rigid channel to one of the migration of the ion together with its distortion along the axis.

We consider the nature of the interaction between the ion and the wall source for the purpose of determining first the effect of the ion on vibrations of the wall. One properly needs to carry out a Born-Oppenheimer separation of the wall vibrations from the channel ion motions to be complete. It is probably adequate, however, in view of the approximate character of the model representation, to examine the wall-coordinate dependence of the total ion band energy $E(\mathbf{k})$ alone. As is well known [52,53] for molecular and solid state electronic systems, contributions to the vibrational wave equation arise in the form of terms

$$\langle \phi_i(\mathbf{r}, \mathbf{Q}) | \partial^2 / \partial \mathbf{Q}^2 | \phi_j(\mathbf{r}, \mathbf{Q}) \rangle$$

etc., in which ϕ is an electronic basis function (in the band picture) of the electronic coordinates \mathbf{r} and a parametric function of the vibrational coordinates \mathbf{Q} for metals and molecules. Similar terms arise for the case of the ion-wall interaction. The contribution of such terms to the equation of motion that determines the vibrational wavefunctions is frequently fairly small, although not always so. At this stage, we ignore terms of this form seeking instead to concentrate attention on the derivatives of the total band energy in the parametric representation on \mathbf{Q} .

Given a coordinate of a source C in the wall, with coordinates \mathbf{C} , we expand the band energy as a Taylor series in order to determine eventually the effect of the presence of the ion on the vibrations of the wall. Thus,

$$E(\mathbf{k}) = E_0(\mathbf{k}) + \mathbf{Q} \cdot \nabla_{\mathbf{Q}} E(\mathbf{k}) + \frac{1}{2} \mathbf{Q} \cdot [\nabla_{\mathbf{Q}} \nabla_{\mathbf{Q}'} E(\mathbf{K})] \cdot \mathbf{Q}' + \dots \quad (133)$$

Derivatives with respect to normal modes \mathbf{Q} are implied. The energy $E(\mathbf{k})$, however, is also a function of the Cartesian coordinates of the wall sources as well as the position of the ion. Derivatives with respect to the Cartesian coordinates can be converted to normal modes. Thus, we need to consider at least the gradient of the band energy with respect to Cartesian position. As one might expect, the derivatives of the energies involved here can be represented analytically. The problem differs somewhat from the molecular electronic case in that the positions of expansion of the ion wavefunctions are not associated necessarily with the center of gravity of any channel source species. The points of expansion generally lie on the channel axis, which remains a fixed coordinate throughout the analysis; both the channel ion and the wall source species can be considered to move relative to the channel axis without any loss of generality or restriction in manipulation. In the following paragraphs, we illustrate this with the consideration of three examples, the ss matrix elements of the Coulomb, screened electrostatic, and Morse potentials.

Consider the derivative of the ss -matrix element of the Coulomb operator:

$$\begin{aligned} \frac{\partial}{\partial C_x} \langle s | V_C | s \rangle^{(0)} &= 2 \left(\frac{a+b}{\pi} \right)^{1/2} \langle s || s \rangle \frac{\partial}{\partial C_x} F_0(U) \\ &= 2(a+b)(P_x - C_x) \langle s | V_C | s \rangle^{(1)}. \end{aligned} \quad (134)$$

This is related to the expression for the ion-dipole interaction, eq (D1). Therefore, one should consider the Taylor expansion of the Coulomb potential with respect to the coordinates at the wall source. In particular,

$$\sum_{n=0}^{\infty} \frac{C^n}{n!} \nabla_C^n \langle s | V_C | s \rangle = \sum_{n=0}^{\infty} \frac{C^n}{n!} \langle s | \nabla_C^n V_C | s \rangle \quad (135)$$

where, for reference, the coordinates of V_C are

$$V_c = \frac{Ze^2}{|\mathbf{r} - \mathbf{C}|} \quad (136)$$

and $\langle s| = \langle s(\mathbf{r})|$ is a function of \mathbf{r} in the integral. The functions $\langle s(\mathbf{r})|$ do not depend explicitly on the coordinates of C .

As matrix elements of the original potential V_c , for example, can be obtained from the ss -matrix element by differentiation with respect to $2aA_x$ etc., the elements of the Taylor series also needed in the vibrational problem for the wall sources can be obtained by differentiation. Further, it is possible to see the emergence of additional recursion relations; in fact, these have already have been presented to n th order by Obara and Saika [cf., Appendix, ref. 50, eqs (A22)-(A26)].

Turning to the soft part of the composite electrostatic interaction, it is equally clear that one has

$$\begin{aligned} \frac{\partial}{\partial C_x} \langle s| V_s |s \rangle^{(0)} &= \frac{2}{\sqrt{\pi}} \left[\frac{(a+b)b}{a+b+b} \right]^{1/2} \langle s||s \rangle \frac{\partial}{\partial C_x} F_0(T) \\ &= 2 \frac{(a+b)b}{a+b+b} (P_x - C_x) \langle s| V_s |s \rangle^{(1)}. \end{aligned} \quad (137)$$

Thus, one can write for the soft part also

$$\sum_{n=0}^{\infty} \frac{c^n}{n!} \nabla_C^n \langle s| V_s |s \rangle = \sum_{n=0}^{\infty} \frac{c^n}{n!} \langle s| \nabla_C^n V_s |s \rangle. \quad (138)$$

Finally, for the Morse potential, one has to consider the derivatives of the exponential operator in the matrix elements:

$$\langle s| \exp[-u(\mathbf{r} - \mathbf{C})^2] |s \rangle$$

From the specific form of the ss matrix element of the exponential operator, differentiation gives

$$\begin{aligned} \frac{\partial}{\partial C_x} \langle s| \exp[-u(\mathbf{r} - \mathbf{C})^2] |s \rangle \\ = \frac{(a+b)u}{a+b+u} (P_x - C_x) \langle s| \exp[-u(\mathbf{r} - \mathbf{C})^2] |s \rangle \end{aligned} \quad (139)$$

which is closely related to the result obtained for the electrostatic interactions already considered; in fact, as Obara and Saika show [50], the nuclear attraction integrals are derived with the use of the recursion relations of the three center overlap integrals.

If one is interested in carrying out a three dimensional harmonic limit analysis, extending the analysis in Section 5.1, it is possible to consider a two-center Taylor series expansion simultaneously about \mathbf{r} and \mathbf{C} . A quantum based calculation is needed for the ion in the channel because of the overlap between translationally equivalent unit cells. The sources in the wall of the channel, on the other hand, are fixed and a classical normal mode treatment suffices. The two systems are closely connected. Consequently, it is necessary to consider some form of self consistent calculation, if one is to take this approach.

A self-consistent treatment can be constructed in the following manner. (1) Given the ion in a rigid channel, determine the energies of the band states. (2) Expand these energies as a Taylor series about Cartesian points in the wall sources, as done above. (3) Calculate the vibrational states of the helical chains in the presence of [at least] linear interactions due to the presence of the ion. (4) Use the solutions to determine the displacements due to the presence of the ion, and carry out a new determination of the energy states for the ion. At this stage, the translational symmetry of the channel system is, strictly speaking, broken. However, the weak coupling limit is characterized by treating the interaction of the distorted chain on the ion in the same manner as one would handle the effect of an applied external field.

Strong coupling is implied when the following occurs. If, in the process of solving the self-consistent problem, one successively shifts states of the ion and positions of the wall sources to the point that the implied permanent distortions of the wall give rise to deep local wells, a band state in the extended rigid lattice is not possible. At this point, the treatment focuses on the use of local functions to handle the local states of the effectively massive ion. The ion has reduced mobility only with respect to quantum mechanical displacement or delocalization on a time scale comparable to the frequencies of vibration of the trapped ion. It is not necessarily immobile with respect to diffusional time scales. In addition, the wall-phonon dressed ion is still translationally invariant. The ion *plus* its distortion of the nearby wall moves along the channel.

Finally, a consequence of deep trapping can be a considerable blue shifting of the axial vibration of the ion. The deeper the trapping, the more the structure of the channel begins to assume the character of a uniform solvation or chelation of the ion. As a result, one can expect the axial vibrations to be much larger than those of the ion in a free channel state. The mobility of the ion should also decrease substantially due to the very large effective mass. A sufficiently sensitive spectroscopy may be able to resolve these issues.

5.7 IONIC MOBILITY IN THE SMALL POLARON-LIKE LIMIT

Several approaches to the issue of tight binding ionic mobility in the channel system can be explored. In the past decade, quantum transition state theory has matured to the point that it is possible to consider sophisticated treatments that include formally accurate accounts of quantum mechanical effects. For some well-defined and relatively ideal systems (*viz.*, single crystal surfaces) accurate calculations can be carried out [see, for example, Metiu *et al.*, ref. 55]. The same is true for the channel conductor, in principle. At this point, however, in the interest of presenting the notion of activated ion-migration in a channel, I opt for a simpler, less formally rigorous approach [56,57]. Although detailed calculations have not yet been carried out for a channel conductor, the application of the methods to be described is possible.

The small polaron limit for the channel conductor can be invoked with the observation that due to the likely strong interaction between the ion and its elastically constrained wall atoms, significant self-trapping should occur. Thus, the effective mass of the mobile ion increases as a result of the fact that activation is needed to move the ion from one site to another equivalent site. The activation requires the rearrangement distortion of the wall of the conduction channel to accommodate the transition state. In essence, although the system still has translational symmetry, extreme localization through self-trapping makes a local point of reference adequate for the evaluation of the transport process. One can consider an initial state as any equivalent point of equilibrium along the channel. Nearest final states are also all identical along the chain. This point of view for the ion is equivalent to the small polaron model for the electron originally developed by Holstein [58].

To begin, we consider an adiabatic transfer which corresponds to a strong coupling limit between the initial and final channel ion states. The ion is located in a well at 0 on the chain axis. We need to find the rate of transfer of the ion to a neighboring site one well away. In formal terms, the rate constant for the transfer is [59,60]

$$k = \int_0^\infty d\Omega \Omega^{6n-1} \dot{q}_r P(p, q) \quad (140)$$

where $q = dq/dt$. In this equation, P is the probability function

$$P(p, q) = \exp[-\beta(H-F)] \quad (141)$$

with the free energy F defined by

$$\exp(-\beta F) = \int d\Omega^{6n} \exp(-\beta H) \quad (142)$$

and, $\beta = 1/k_B T$ with k_B the Boltzmann constant. The absolute temperature is T . The volume element $d\Omega^{6n}$ is given by

$$d\Omega^{6n} = \prod_{i=1}^n d^3 p_i d^3 q_i. \quad (143)$$

In the transition state theory, one assumes that only migrations across the top of the saddle point in the (positive) reactive direction contribute to the rate constant [59]. Marcus [60] and Miller [61] have formalized this notion in a rigorous fashion for the quantum transition state theory. For our purposes here, it is sufficient, as well as illustrative, to consider an approach due to Glyde [59]. To obtain the rate constant, the free energy integral needs to be evaluated. It is reasonable to assume that near to the initial state the potential expands in a Taylor series to second order as [59]

$$V \approx V_0 + \frac{1}{2} \sum A_{q_i q_j} q_i q_j \quad (144)$$

where $A_{q_i q_j}$ is the matrix of derivatives of the potential evaluated with the initial state coordinates and V_0 is the value of the potential in the initial state. A similar Taylor expansion in the saddle point makes it possible to write the rate constant as

$$k = \frac{|A|_{3n}^{1/2}}{\left[(2\pi m)|A^\ddagger|_{3n-1}\right]^{1/2}} \exp\left[-\beta(V^\ddagger - V_0)\right] \quad (145)$$

and V^\ddagger is the value of the potential in the transition state. The determinant of the force constants, $|A^\ddagger|_{3n-1}$ is found by taking the derivatives of the potential evaluated at the transition state less the row and column corresponding to the single coordinate q_r along which the ion moves in making the transfer. For our purposes, this coordinate is the collinear with the channel axis.

As we have not considered specific models for the channel wall vibrations, we do not provide numbers for this transfer rate constant at this time. The transfer constant can be related to a transport coefficient and thence to the conductivity of the ion in the channel. It

is sufficient to note at this point that for shallow wells one has $V^* - V_0 \approx 0$, or at least this quantity is very small. Thus, even though the localization of the ion in any individual well may be sufficiently great to obviate the free or nearly free ion limit as a model, high ionic mobility is still possible in the channel system by virtue of low barriers to activation. In such cases, the ionic conductivity is essentially equal to the frequency of ionic motion in the shallow well. For a system that has an intrinsic frequency in the range of 10 cm^{-1} , the transport coefficient, $k\zeta$ is approximately $4,500 \text{ cm/s}$.

Transfers can also take place in the diabatic limit in which there is a weak splitting of the initial and final state potential energy curves [56,57]. The following is a simple treatment of the transfer in a diabatic limit. The rate constant depends on the transition probability

$$w = \frac{2\pi}{\hbar Q} \sum_{i,f} \exp(-\beta E_i) |L_{if}|^2 \delta(E_i - E_f) \quad (146)$$

which is a Boltzmann weighted average over all possible initial states, with energy E_i and all possible final states E_f . The matrix element L_{if} couples the initial and final states and Q is the sum-over-states

$$Q = \sum_{i=0}^{\infty} \exp(-\beta E_i). \quad (147)$$

Completeness relations and operator manipulations enable w to be expressed as [56]

$$w = \frac{2\pi}{\hbar Q} \int d\mathbf{v} d\mathbf{r} \exp(-\beta H) |L_{if}|^2 \delta(V_i - V_f). \quad (148)$$

Further manipulation, details of which can be found in ref. [57], yields

$$k_d = w = \frac{\sqrt{2}\pi}{\hbar} \frac{|A|_{3n}^{1/2}}{|A^*|_{3n-1}^{1/2}} \frac{|L_{if}|^2}{|\Delta F|} \exp\left(-\beta(V^* - V_{i0})\right) \quad (149)$$

with

$$|\Delta F| = |\partial(V_i - V_f)/\partial x| \quad (150)$$

which is effectively the difference in the slopes of the initial and final potential energy states at the point of intersection in the transition state. In this expression, x is the reaction coordinate. The determinants of the force constants, $|A|$ and so on, are the same as given above in the adiabatic case. It is possible to bring the result into the same form as in the adiabatic case by defining an effective mass of the form [57]

$$\sqrt{m_{\text{eff}}} = \frac{\hbar |\Delta F|}{2\pi\sqrt{\beta} |L_{\text{if}}|^2}. \quad (151)$$

As the interaction that leads to the migration of the ion becomes smaller, the effective mass increases. In essence, because the ion is trapped, its effective mass with respect to transfer to another well becomes very large.

Formalism similar to that just presented for the migration of the ion along the channel has been specialized to transfer at surfaces [55,62]; for the electron, the rates involved are associated with electrochemical electron transfer. We will return to consider the consequences of tunneling of ionic charge across surfaces of membranes in Section 7.

6. ESTIMATES FOR TWO OR MORE IONS PER CHANNEL

When there is more than one ion in the channel, the ion-ion interaction changes the nature of the problem. If the concentration is sufficiently great, the ions are closer together and this closeness greatly increases the values for the vibrational frequencies polarized along the channel axis. Thus, frequency shifts for axially polarized ionic vibrations may be a measure of the density of ionic concentration in the channel.

While it is eventually appropriate to carry out a fully quantum mechanical calculation of the multiple ion problem, especially in view of the persistence of multiple site occupancy in the channel and tunneling, it does not seem necessary at this point to do so. In the first place, the most significant effect of the ion-ion interaction can be seen in a purely classical harmonic analysis; quantum mechanical refinement will not significantly alter the facts.

The classical harmonic analysis is the following. First of all, the only additional interaction considered is the Coulomb repulsion operating between ions. It is easy to show that the form of the force constant operating along the channel axis is

$$k_{zz} = \frac{2e^2}{R^3} \quad (152)$$

where R is the distance between ions. The usual kind of one-dimensional analysis [33] gives frequencies

$$\omega = [4k_{zz}/M]^{1/2} \sin(\kappa\zeta/2) \quad (153)$$

in which κ comes from the travelling wave solution. The maximum value corresponding to the upper vibrational band edge is simply $(4k_{zz}/M)^{1/2}$. The dependence of the channel polarized frequency on distance between ions therefore follows a $3/2$ power law. The linear density of ions in the channel is $\Pi = N/L$ where N is the total number of ions in the channel and L is the total length. The effective distance between ions, therefore, is $1/\Pi$. The number of lattice sites between ions is

$$n = \text{int}(1/\Pi\zeta). \quad (154)$$

Figure 11 shows the dependence of the channel polarized frequency as a function of the number of lattice sites between ions. If the ionic density in the channel is sufficiently great to force ions to occupy adjacent lattice positions, the axial frequency will be observable in the infrared. As the distance between ions increases, however, the

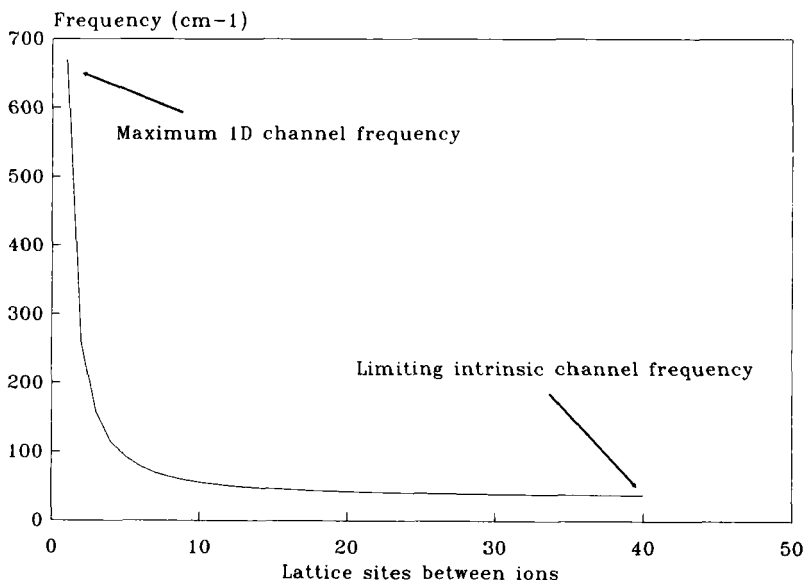


FIGURE 11. Frequency of the channel ion vibration as a function of linear density.

frequency drops quickly. For large spacing, corresponding to few ions per channel, the frequency drops essentially to the value predicted for the single ion per channel.

In sparsely occupied channels, there is the possibility that mobile ions may interact and the vibrational states exhibit excitonic structure. Such behavior, however, probably would only be observed in systems at low temperatures. As a result of the direct Coulomb interaction between ions, the splitting should be strongly distance-dependent. Thus, the extent of the splitting, if observed, should be another measure of the density of ions in the channel.

The major effect of multiple ion occupancy of a channel is the necessity of considering an appropriate many-body treatment of the states of the ionic system alone, even before considering the complications of interaction with the phonon modes of the channel walls. Although a number of issues are unresolved, a Hartree [49] approach seems likely to be appropriate. Whether or not the several ions that occupy the channel need to be treated as fermions or bosons, or mixed species, remains to be examined.

It seems reasonable to me to expect that in biological systems in which the resting ionic density is held at some steady state value, as for example in channels across nerves and muscle membranes, there can be a measurable change in frequency once these types of system change to a non-equilibrium, active state. If the opening of channel pores or gates has a cascade depletion effect, then this ought to be seen. Diffusive processes are very much slower than species vibrations. Thus, provided one can collect sufficient numbers of species to study and one has an instrument with a sufficiently quick acquisition time—in, for example, an interferometric mode—it may be possible to map the depletion dynamics of the channel in real time by following the changes in ionic frequency. The difficulty in this type of experiment, for a biological system, is the temperature at which the effect would have to be observed. With low ionic density in a channel, the frequencies are low, especially in a depleted state. Thus, one may not be able to observe any change unless the initial state were sufficiently densely populated to ensure an associated high frequency. As Figure 11 indicates, the frequency drops quickly for even relatively few lattice distances between ions. On the other hand, it may be possible to assemble artificial channel systems that can be held under the proper conditions to show an effect of this type.

Finally, a cautionary note must be added. Wall distortion has not been considered. It is possible that the interaction between wall-dressed ions is much different than the bare interaction. In view of the blue-shifting dressing with wall phonons may yield, the ultimate interpretation of channel ionic vibrations in multiply occupied systems may be a challenge.

7. A MODEL OF ION-GATING INTO BIOLOGICAL CONDUCTION CHANNELS

In this section I make use of the notions concerning ionic conduction in channels discussed in the preceding sections to build a model of a surface-gated ion transfer across a bilayer membrane system. A subtitle for this section might therefore be, *What to expect of ion channels and how to measure it*. The model of the incorporation of a transmembrane conduction channel advanced here is, to my knowledge, new. It incorporates easily into the existing model of a phospholipid membrane in, for example, an excitable structure such as an axon.

The facts are well-known [2,4]. The inactive, resting state of an unmyelinated axon is characterized by the restricted fluxes of ions across the surface of the structure. I argue that simple configurations of ionic aggregates in the surface of the axonal membrane account for the barrier to the free migration of sodium. In addition, I argue that *slight* alterations of the possible configurations of the surface account for the effective disappearance of barriers to ionic conduction in the active state. The flux of sodium that is associated with the active, excited state is explosively gated and the axon exhibits the phenomenon of the 'action potential'. Other ions migrate in response to the flux of sodium, but it is the sodium flux that governs the phenomenon.

7.1 THE MEMBRANE MODEL AND MECHANISM OF IONIC FLUX GATING

Biological membranes that encapsulate and give strength to cells consist primarily of bilayers of phospholipids. In the bilayer structure, the hydrophobic organic portions of the molecules are in juxtaposition enabling the polar phosphate portions of the molecules to face the water-based electrolyte of the intra- and extracellular fluids. Transport of materials to and from the cell occurs by several processes. In excitable tissue, nerves and muscles, it is now generally accepted that particular protein-based channels provide pathways for ionic transport by diffusion [3]. In addition, in excitable cells some of the transmembrane pathways are blocked and concentration gradients of sodium and potassium in particular, and other ionic species, exist across the cell membrane. Several of these concentration gradients relax locally and quickly with the passage of an impulse along the surface of the cell. The mechanism by which ionic flux across the membrane is controlled into and out of the conduction channels has been the subject of considerable interest for decades. It is possible to construct a relatively simple model of this process that, at the same time, suggests a possible way to demonstrate its applicability.

To begin, it is essential to note that the calcium cation is now recognized [63] as crucial to the functioning of biological processes. The role of calcium in membrane structure *can* be the following. Because

of the divalent character of the ion, it is possible for a calcium ion to associate with several of the phosphate groups of the lipid molecules at once. Thus, through coordination, the calcium ion binds the phosphate end-groups on the lipids into a two-dimensional calcium phosphate lattice. In addition, ethanolamine and other groups esterified to the phosphate can act as cross-linking agents to impart varying degrees of stiffness to the membrane surface. As a consequence of this hypothesis, information, in the form of nerve impulses, for example, migrate along the membrane surface as two-dimensional *lattice dilation waves*.

The next essential feature of the model is the incorporation of the ionic conduction channels into the lattice structure. Radioactive ion tracer studies long ago [64] established a relative surface density of ion channel pores. It is possible to use these data to estimate that an average distance between channels should be about 100 Å. As this distance is large, it is not easy to envisage a genuine, fast ionic conduction mechanism along the surface that contributes to the message

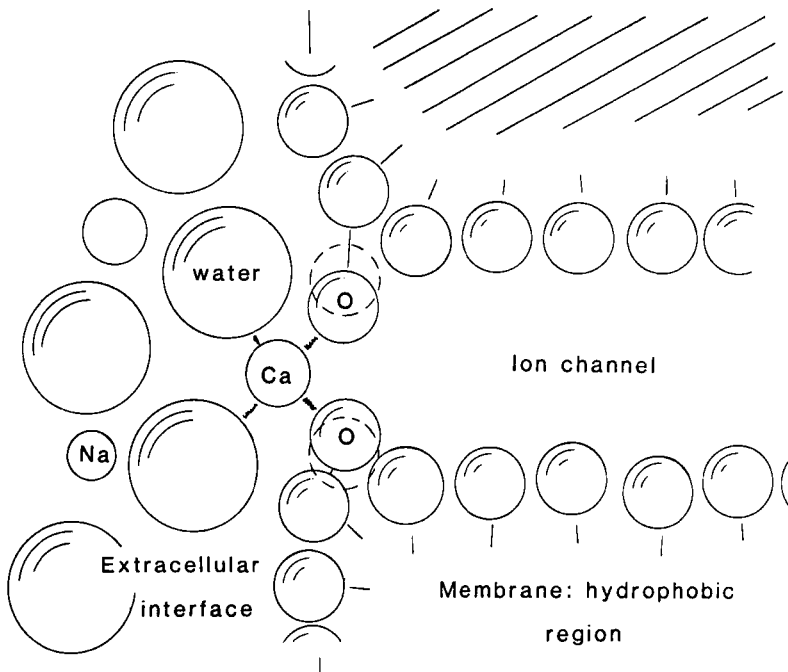


FIGURE 12. A cross-sectional view of the interface, surface and proteinaceous channel for conduction. An underlying channel distorts the surface, and this is indicated. Calcium is shown associated with oxygen atoms (in turn attached to phosphorus). The dilation of the surface (that would accompany the passage of an impulse along the surface of the membrane) is shown by the dotted distortion of the two bordering oxygen atoms.

transport in the sense that Ohm's law, for example, applies to conduction in metals. It is nevertheless true that a disturbance, which constitutes the nerve impulse message, does propagate. There is in fact another physical effect that can be used to account for the travelling disturbance. In particular, I propose that the mouth of an ionic conduction channel lies either incorporated within or immediately below the surface of the phosphate layer. In fact, if this assumption is reasonable, it is also reasonable to expect that there is a strong association between the molecular groups that form the mouth of the channel, the cations, and the phosphate heads of the lipid. This assumption about structure, which to my knowledge is not contravened by any experimental evidence, is reasonable from the point of view that if a channel and pore are to interact effectively with the travelling disturbance along the membrane, they *ought* to be an integral part of the mechanical system in order that gating be effectively triggered. Finally, in the resting state, ionic passage into the ionic channel is *blocked* by the calcium cations that bind and help to define the membrane surface. This situation is illustrated in Figure 12. In the resting state, one sees that there is a high barrier to ionic transport across

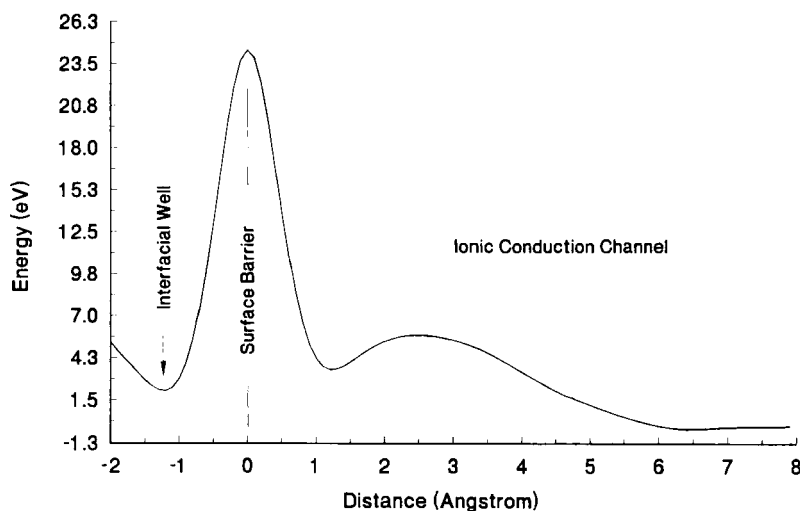


FIGURE 13. A single potential energy profile is illustrated together with labels that indicate the relationship of the structures in the interface, the surface, and the channel to the profile.

the mouth of the channel. I note additionally that the structure of the membrane energetically favors calcium because of the multidentate association. The cellular system, however, has excess concentrations of monovalent cations in the form of sodium and potassium, in particular, in the double layer regions of the membrane surface. Moreover, the cell metabolically maintains an electrolyte imbalance with sodium outside the complete membrane higher in concentration than potassium. The balance is such that the concentration of calcium in the structure of the membrane surface is sufficient to inhibit sodium or potassium flow across the membrane into the cell. The hydrophobic portions within the bilayer membrane present no thermodynamic driving force to equalize the sodium or potassium ion concentrations across the phosphate surface. The potential profile is shown in figure 13.

In the neighborhood of an (assumed) incorporated ionic channel, on the other hand, ion transfer is possible if the surface modifies appropriately. The mechanism of ionic gating therefore is now quite simple.

The model depends on the assumption that a nerve impulse is a propagating disturbance in the form of surface dilation wave; note that

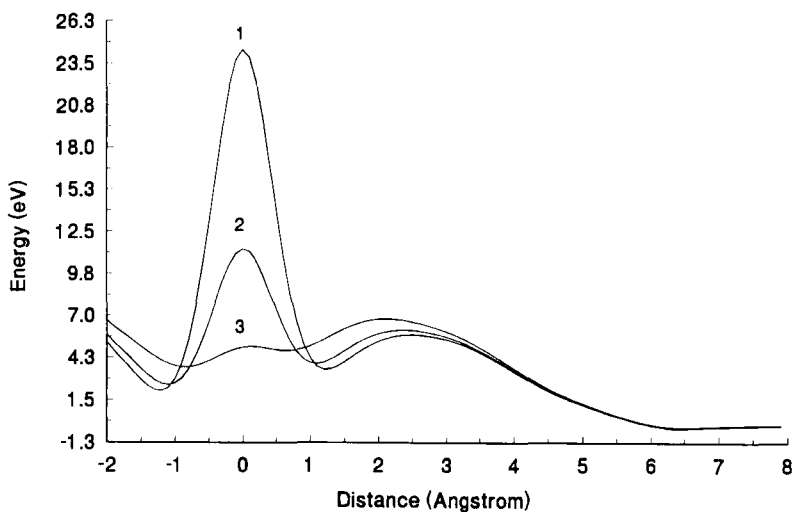


FIGURE 14. Three potential energy profiles: (1) the initial state for which the calcium ion and the configuration of phosphate groups on the lipids present a large barrier across the membrane surface; (2) the case in which the effective phosphate-ion radius is dilated by about 0.2\AA ; (3) the case in which the dilation by an additional 0.2\AA effectively eliminates the barrier.

the migration of a dilation wave can be handled with mathematical machinery that is much the same as Ohm's law or Fourier heat conduction [65]. This wave is confined almost entirely to the calcium phosphate surface. As the dilation wave enters the region near the mouth of a particular channel, a local expansion of the distances between the calcium and the neighboring phosphate groups *lowers* the potential barrier for calcium flux into the channel, as seen in Figure 14. This much is also clear from the figure. Because there is an excess of monovalent cations over the divalent calcium, replacement of the calcium by sodium or potassium further forces the dilation of the region of the membrane surface over the mouth of the channel. This follows as a result of the loss of compensating charge to the repulsion of one phosphate group against another. Thus, the barrier, as seen in the figure, disappears and sodium or potassium in the vicinity diffuses readily into the channel. Moreover, the opening over the channel mouth is driven, by electrostatic repulsion, even wider. This event should occur quickly; in view of the short distances required to precipitate the catastrophic replacement of calcium with sodium over the channel mouth, one can anticipate a gate opening time of the order of several vibrations.

Repair of the calcium gate/block due to diffusion occurs equally simply because of ionic depletion and competitive association. In particular, relaxation of the sodium or potassium ionic gradient in the vicinity of the *open* channel mouth quickly brings the levels of these ions locally to that of calcium. At that point, calcium competes effectively to bind to the phosphate. The system *repairs* to its resting state. The monovalent cations then can build up the concentration gradients observed through diffusion.

7.2 TUNNELING PROBES OF MEMBRANE SURFACES AND CHANNELS

At this point, it is possible to propose a scanning tunneling microscopic experiment based on the theory of the conduction in the channel and the flux of ions across the channel barrier. First of all, it is necessary to assume that there is a source of ions, such as protons or small group Ia and IIa cations, that can be sufficiently focused to be able to probe the ionic flux through pores in the surface. Several possibilities exist: one, it may be possible to form crack-tips in various glasses, and, two, it may be possible to harvest and incorporate single tubular fullerene fibers [6,7] (if they prove eventually to be collectible) in an ion source/holder that can be manipulated over the surface of a membrane. The fullerene may then act as a focused and efficient ionic conductor.

It is necessary to note, before considering the nature of the possible probe, that what follows differs from the scanning ion-conductance microscope of Hansma *et al.* [66] and the scanning electrochemical probe of Bard *et al.* [67]. First of all, I propose a

truly microscopic probe on the scale of several diameters of the ions involved. And second, I propose that actual tunneling across a well-defined quantal barrier takes place, not simply classical ionic conductance or electrochemistry in solution (that indeed can involve tunnel processes, but process that differ in form and action from those considered here).

The current one measures can be obtained in a variety of ways, but all generally originate with Bardeen's [68] transfer-Hamiltonian formalism. In view of the similarity of the ion transfer case to the electron transfer that has been investigated extensively in electrochemistry [69-72] and in connection with Scanning Tunneling Microscopy involving electrons and adsorbed molecular species, the tunnel current is given as [71,73]

$$j = \frac{2\pi e^2 v}{\hbar} \sum_{i,f} |M_{if}|^2 \delta(E_i - E_F) \delta(E_f - E_F) \quad (155)$$

where E_F is an effective Fermi level for the channel system and v is the small applied external potential. This formula applies to the case of low temperature (actually a temperature at which the local vibrational states are not significantly populated) and small applied external potential. The formalism that leads to the above expression for the current will, in the general case, reproduce the Tafel form that is well known in electrochemical systems.

An ion probe can involve the transfer of a single ion to an acceptor site through a barrier or it can involve the displacement of an ion in the structure. In the second case, the current will depend on (at least) a two-particle formalism to represent the exchange. To begin, we consider the first, single ion transfer. Formulae for the case of two (or more) ion transfers also contain terms for the single ion interactions, energies, and transfer as well as two-particle, many-center contributions. Aspects of the two-ion transfer will also be discussed.

The issue is the determination of the form of the transfer matrix element M_{if} . From an analysis [69,70,74] that has its origin in the Holstein treatment of the small polaron problem [58], the transfer matrix element is found to be

$$M_{if} = \frac{\langle i | H | f \rangle - S_{if} \langle f | H | f \rangle}{1 - |S_{if}|^2}; \quad (156)$$

although below we will use a slightly different form for the single proton transfer probe. S_{if} is the state overlap matrix.

Even before carrying out detailed calculations, it is possible to anticipate the outcome of an experiment in qualitative terms. To do so, one needs to bear in mind the operation of the electron-based scanning tunneling microscope. The electron is much smaller than any of the

atomic and molecular species it probes *via* tunneling. Moreover, there are many states to which the electron has access. Thus, even though the electron tunnels to the metal substrate, molecular species on a surface interact with, accept (in both real and virtual states), and transfer the electron. The electron is a fine-tipped probe. The atomic structure it can reveal, therefore, is great.

An ion, on the other hand, is of the same general size as parts of the structure to which it migrates. The proton is perhaps a limiting case in that there may be a number of sites to which it can migrate, as the facility with which aqueous systems conduct protons shows [75,76]. Nevertheless, a structure, a membrane surface for example, that admits ion transfer across the boundary at particular points can indeed reveal sites that are of the size of the probe. The interpretation of the results will simply require some care. Thus, although the electronic STM probe would in principle yield an image of the whole, and variegated, structure of the surface, an ion probe, particularly an ion-selective probe, would only show structure in the immediate vicinity of the channel pore. The membrane surface would be revealed as essentially flat except at the locations of the pores. The structure itself would generally appear only as a bump. In spite of the simplicity of the overall image, there is considerable information available. One would expect that in the case of an excitable structure, a nerve or muscle membrane, at low enough applied currents, only exchange tunneling would be involved and the current would be effectively ohmic. Thus, the size of the pore would not alter greatly. At higher applied potentials, it should be possible to trigger the catastrophic opening of the channel pore, as discussed above. Although it probably is not possible to resolve the opening, because of the extremely short times involved (*viz.* femtoseconds), it may be possible to see the open state and to observe the relaxation back to the resting state.

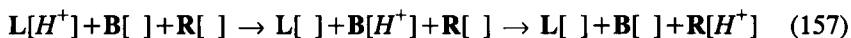
The following paragraphs sketch the theory of the *Ion Tunneling Microscopy* proposed here. The treatment, although different in starting form, has much in common with Marcus's recent treatment [77] of electron and ion transfer at the junction between two polar but immiscible liquids.

It is possible to use surface modified plane wavefunctions to attempt to account for the effect of a surface on states of an ion in a channel conductor. Instead, we consider simply an abundance of ion states that can arise in a large, but finite channel system that is the analog of the molecular electronic problem for the large, finite linear polyene [40]. This approach easily could have been adopted above to treat the ion-in-channel problem, the subject of this chapter. Given a sufficiently long finite polyene, it is well known that the central states mimic reasonably well the band states of a metal or similar delocalized system. The final states for the treatment, in which the migratory ion occupies the transmembrane channel, are likewise assumed to arise in a long, but finite, channel system. The membrane surface, which acts as a *gate* for certain ions such as sodium, potassium, and perhaps divalent cations, is treated as locally different from either the initial or final channels. In the absence of interaction with the

channels, one sees the states of the membrane surface as discrete ion vibrations within a planar framework of phosphate groups; the predominant interaction is probably between the ion and oxygen atoms on the phosphate.

Much of the standard work on tunneling makes use of parametric barriers of arbitrary geometry. It is possible, nevertheless, to treat tunneling in the cases developed here in terms of explicit configurations of molecules in the barrier region. The analysis used here employs a treatment explored earlier for the general problem of proton transfer [46]. In particular, basis functions are developed for the initial and final states to reflect the fact that the ion interacts both with the channels and the barrier. It is the case that in both the initial and final states, when the barrier region associated with the channel mouth is distant from the tip of the probe channel, the density of mobile probe-ion in the region of the barrier will be very small. Even so, it makes sense to construct initial and final state functions as linear combinations of basis functions for the ion in the channels (initial and final) and the barrier. The following analysis summarizes this approach.

To begin, we consider a proton transfer probe. The proton, we assume, is able to find unoccupied local sites within the membrane to occupy. The atomic-scale mechanistic picture of this is the following:



where $L[I]$, $B[I]$, and $R[I]$ stand for the *left*, probe channel, the *membrane* (possible *blocking*) site occupied by the proton, and the *right* trans-membrane channel or acceptor site. The notation $B[I]$, *etc.* indicates occupation; $[\]$ is an empty channel or block site.

We now consider the Hamiltonian operator that applies to the entire system of an proton in the probe channel, in the membrane surface phosphate complex, and in the transmembrane channel. A treatment of the states of this system [75] that has roots in Bardeen's transfer-Hamiltonian formalism [68] is the following. Consider a set of basis functions $u_0(\mathbf{r})$ for the proton in the probe channel on the *left* (L) side of the membrane surface, the proton in the surface membrane phosphate *barrier* complex (B), and the *right* (R) transmembrane channel. Two orthogonal functions can be constructed for the proton in its initial state in the probe:

$$\phi_{L1} = u_0(\mathbf{r}_L) \quad (158)$$

and

$$\phi_{L2} = \frac{1}{(1-S_{LB}^2)^{1/2}} \left[u_0(\mathbf{r}_B) - S_{LB} u_0(\mathbf{r}_L) \right] \quad (159)$$

with similar basis functions for the right channel acceptor side. The second function allows for the accumulation of proton on the site *B*. Complete right and left basis functions are now constructed as

$$\begin{aligned}\Psi_L &= a_{L1}\phi_{L1} + a_{L2}\phi_{L2} \\ \Psi_R &= a_{R1}\phi_{R1} + a_{R2}\phi_{R2}\end{aligned}\quad (160)$$

with which one finds the energy ϵ_L

$$\epsilon_L = \frac{1}{2} \left(H_{L1,L1} + H_{L2,L2} \right) \pm \frac{1}{2} \left((H_{L1,L1} - H_{L2,L2}) + 4H_{L1,L2}^2 \right)^{1/2}. \quad (161)$$

A similar term is found for ϵ_R . Let the eigenfunctions associated with the lower eigenvalues be Ψ_{L0} and Ψ_{R0} ; these functions are linear combinations of the basis functions ϕ_{R1} , *etc.* A new pair of orthonormal *left-right* basis functions for the whole system can be constructed from Ψ_{L0} and Ψ_{R0} . Let

$$\chi_L = \Psi_{L0} \quad (162)$$

and

$$\chi_R = \frac{1}{(1-S_{LR}^2)^{1/2}} \left(\Psi_{R0} - S_{LR} \Psi_{L0} \right) \quad (163)$$

in which S_{LR} is the left-right overlap that can be obtained in terms of the primitive channel basis functions. The energies associated with these states can also be found directly

$$E = \frac{1}{2} \left(H_{L,L} + H_{R,R} \right) \pm \frac{1}{2} \left((H_{L,L} - H_{R,R}) + 4H_{L,R}^2 \right)^{1/2}. \quad (164)$$

The energies involved, together with the associated state functions, are the essential ingredients in the formulation of the transfer-Hamiltonian matrix element needed to evaluate the transfer probability and, from that, the current one can associate with the process.

The matrix element H_{LR} can be obtained in terms of the matrix elements of the primitive basis functions $u(\mathbf{r})$. In particular, let a_{L1} ,

a_{L2} and a_{R1} , a_{R2} now be the coefficients in the expansion of Ψ_{L0} and Ψ_{R0} in terms of the functions ϕ for the lowest eigenvalues $\epsilon_{L(R)}^0$. The structure of the matrix element H_{LR} is the following:

$$\begin{aligned}
 H_{LR} = \frac{1}{(1-S_{LR}^2)^{1/2}} & \left\{ a_{L1} a_{R1} H_{lr} + \frac{a_{L1} a_{R2}}{(1-S_{lb}^2)^{1/2}} \left(H_{lb} - S_{br} H_{lr} \right) \right. \\
 & + \frac{a_{L2} a_{R1}}{(1-S_{lb}^2)^{1/2}} \left(H_{rb} - S_{lb} H_{lr} \right) + \frac{a_{L2} a_{R2}}{\left[(1-S_{lb}^2)(1-S_{br}^2) \right]^{1/2}} \\
 & \left. \times \left(H_{bb} - S_{lb} H_{lb} - S_{br} H_{rb} + S_{lb} S_{br} H_{lr} \right) - \epsilon_{L}^0 S_{LR} \right\}. \quad (165)
 \end{aligned}$$

Eq (165) shows a considerable and interesting structure. As is evident, there are terms in both the direct and step-wise transfer of an ion across the boundary. The index b indicates the barrier element, as noted. If, for example, the distance between the *right* and *left* initial and final states is physically large, both the overlap and Hamiltonian matrix elements will be vanishingly small. The matrix element (165) therefore depends primarily on the lb and br terms, and represents a step-wise transfer of the probe-ion.

The somewhat complicated function, Eq (165), is still only a single ion matrix element. The actual mechanism envisaged for the the larger ion transfer is more complicated for the reason that the membrane pore is closed in the normal state by the blocking ion and the whole process involves two or more ions. The process can be represented schematically as



In other words, the incoming ion in the probe channel, I^{m+} , encounters and replaces the ion J^{n+} originally blocking the transmembrane channel in the membrane surface. In the final state, the ion I^{m+} occupies the barrier site in the membrane. Of course, in the presence of an externally applied potential, Ψ , ion flow provides a source of additional ions to repeat the process to replace ions in the blocking site. If the barrier in the blocking site drops, as postulated, that

drop will be manifested in a different Hamiltonian operator H due to changes in the atomic distances in the membrane in the excited state of the nerve impulse. Such a situation represents a changing tunnel factor that is non-linear with any (small) applied potential.

TABLE 5.
Possible basis function combinations to use in constructing the transfer state function.

	L	B	R
A	1	2	
B		1	2
C	1		2
D	2	1	
E		2	1
F	2		1

An approach to the treatment of this replacement transfer problem is the analog of the simple valence bond treatment in the molecular electronic problem. The possibilities can be listed in tabular form; see Table 5. In the table, **L**, **B**, and **R** stand for the left probe channel, the membrane gate (block), and the right transmembrane channel, as before. The labels **A...F** indicate wavefunction product combinations $\Psi_L(1)\Psi_B(2)$ for the ions 1 and 2 in the channel and block sites respectively. Label **A** corresponds to an initial state in which the probe ion is in the probe channel and the block ion (nominally 2) is in the membrane surface. Case **B** corresponds to a final state in which the probe ion has displaced the block ion. Such a transition would result in a current. The next case, **C**, also represents a current-generating transition, but one that is a smaller contributor than case **B**, due to overlap. The last three cases represent exchange possibilities. In the following discussion, I will only concentrate of cases **A..C** ignoring the exchange contributions for the simple reason that, in view of the size of the ions involved, the size of the channels, and the large electrostatic repulsions, exchange due to indeterminacy is probably only a small contributor to the overall transfer current.

Let the state function now be constructed as a linear combination of the first three cases:

$$\Xi = a\Psi_L(1)\Psi_B(2) + b\Psi_B(1)\Psi_R(2) + c\Psi_L(1)\Psi_R(2). \quad (167)$$

The general overlap matrix element is

$$S_{AB} = \langle L(1)B(2) \| B(1)R(2) \rangle$$

$$= \int d^3r_1 \Psi_L(1)\Psi_B(1) \int d^3r_2 \Psi_B(2)\Psi_R(2) = S_{lr}S_{br}. \quad (168)$$

The overlap integrals can be used to construct the matrix to carry out Löwdin orthonormalization [42-44]. The last contribution, C of the table, is probably not as important as the first two. For simplicity, therefore, it is reasonable to consider the current due to the first two terms. The problem then reduces to one analogous to the electron transfer theory [69,70]. The transfer matrix element needed in the expression for the current is

$$\begin{aligned} H_{if} &= \langle L(1)B(2) | H | B(1)R(2) \rangle \\ &= \langle L(1) | H_1 | B(1) \rangle + \langle B(2) | H_2 | R(2) \rangle \\ &\quad + \langle L(1)B(2) | V(1,2) | B(1)R(2) \rangle \end{aligned} \quad (169)$$

where

$$H = H_1 + H_2 + V(1,2) \quad (170)$$

and the operators H_1 and H_2 operate in the spaces of the ions 1 and 2 respectively. Thus the first two single ion matrix elements above contain terms that are the same as considered for the single proton transfer case above. The interaction $V(1,2)$ represents a two-body term that has not explicitly been discussed. The origin of the interaction is clear. Direct Coulomb interaction between the ions applies. The nature of the matrix elements for the Coulomb operator between ions is the same as for the electron case (Obara and Saika, for example, [51] have worked out the general case of the two-particle four-center integrals). In addition, there is a weak attraction that arises from dispersion interactions. It is possible altogether to represent this part as a Morse potential. Matrix elements for the two-particle, four-center cases have been reported for the Morse potential elsewhere [49]. The remaining term needed in M_{if} is

$$\begin{aligned} H_{ff} &= \langle B(1)R(2) | H | B(1)R(2) \rangle \\ &= \langle B(1) | H_1 | B(1) \rangle + \langle R(2) | H_2 | R(2) \rangle \\ &\quad + \langle B(1)R(2) | V(1,2) | B(1)R(2) \rangle. \end{aligned} \quad (171)$$

The overlap matrix needed in the expression (156) for M_{if} is simply S_{AB} above.

The assumption so far is that once the ion transfer has taken place from the probe to the block sites, displacing the original block ion 2, subsequent transitions occur to sustain the current. If this is not the case, if the configuration $B(1)R(2)$ is sufficiently long lived, then the current-determining step may be the subsequent transfer to the transmembrane channel. The formulation of that case is similar to that already considered. One can assume an equilibrium population of $B(1)R(2)$ states leading to the final state $R(1,2)$; because of the multiple site character of the final state in the membrane channel, $R(1,2)$, not previously mentioned as a final state, is indeed possible. Finally, with respect to the current-generating molecular mechanisms, there is the possibility that $B(1)R(2)$ is a virtual state with a lifetime effectively shorter than the uncertainty. In this case, it is necessary to invoke higher orders of time-dependent perturbation theory [78], or to make use of super-exchange formalism [79], to account for the bridging effect of the intermediate state.

At this point, it is possible to anticipate with some more certainty the form the tunneling images will have. For the case of calcium as a probe and calcium as a channel blocker, the probe ion will simply displace an identical ion in the pore. The image can be generated in two ways. First, a plot can be made of the distance the probe must be raised above the surface (*via* a feedback loop) to maintain constant current over the pore. Thus, as a function of vertical displacement over the xy -membrane plane, a pore appears as a large vertical structure rising above the plane of the apparently flat membrane. In fact, the location of the channel gate ought to be viewed as a *hole*. Thus, one should probably plot the negative image of the scan of the ion over the surface.

When, in fact, the probe ion is monovalent, for example, sodium, and the gate ion is divalent, *e.g.*, calcium, the possibility arises that the applied potential may drive the membrane to the active state. For subcritical applied potentials, the contour of the surface and the ion gated ion pores should appear more or less the same as would be the case for calcium discussed above. With rising subcritical potentials the magnitude of the displacement of the probe over the pore ought to appear effectively *ohmic*. As critical action potential levels are reached, however, the image peak should grow grossly out of proportion in comparison to the equivalent probe case just outlined above. Alternately, if one plots the negative of the image, an exaggerated hole grows in the active state. It may be possible to locate the probe over the pore applying only subcritical potentials. Holding the probe in place in the xy -plane it is possible to increase the applied potential to the point the system makes its catastrophic transition. The increase in current could then be measured with the probe in place; the measured current ought to follow the well-known course of the action potential.

The discussion so far has concentrated on the issue of ion tunneling microscopy to generate some form of image of the membrane surface particularly in the vicinity of a transmembrane blocking site. There is the additional possibility that the ion tunnel effect can be used to generate a tunneling spectroscopy of the states in the surface

block site. The formalism involved in this case is similar to the treatment given to account for the role of molecular vibrations on the electron tunnel current [80].

In conclusion, it seems to me to be possible to construct a tunneling microscope that is site specific to probe the surface of the membrane. In the case of an excitable system, nerve or muscle, the probe should be able graphically to reveal the transition to the active state by giving an effective image of an extremely large peak (or hole for the image complement) when the membrane gate makes the transition to the active state. The proton probe, I suspect, will be more like an electron probe able to sample many more available sites on a membrane surface. As a result, the proton probe may show more fine structure. The resulting images, however, may be harder to interpret for the presence of surface ion-pores over membrane channels because of the additional complexity of the proton-generated images.

8. SUMMARY

A model of an ion in a simple channel has been explored in this study. The channel consists of entwined helical strands of source species, molecules, that present to a mobile ion an essentially regular sequence of potential wells into and out of which to move along the axis. The degrees of freedom available indicate that vibrational modes perpendicular to the axis ought to be observed in the range common to similar ions in a solvated condition. Degrees of freedom available to the ion along the channel axis should be observed in some cases as vibrations, possibly band broadened, of the order of a few tens of wavenumbers. Nevertheless, if a sequence of wells exist, the nearly free to tightly bound limits may apply. When, in the course of interacting with the walls of the channel, severe displacements take place, the ion may become deeply trapped. In all of these cases, an approach to the treatment of the theoretical model has been presented. Moreover, in several instances, these treatments have indicated the nature of the vibrational spectroscopy one might observe. Finally, a new model of the excitable membrane, consisting of the channel mouth incorporated as an integral part of the phospholipid structure, suggests a *scanning tunneling microscopy* that might be used to observe directly the pore structure on the membrane surface that is important for controlling and sustaining the action potential that migrates along the nerve and muscle fiber.

Ich bin Dr. jur. Ernst Lemberger für seine Gastfreundschaft in Wien, wo diese Arbeit angefangen war, zu Dank verpflichtet. I am also indebted to John Mintmire at NRL for the inspiration of his work on helical polymers and for his help and encouragement. Numerous other colleagues have endured the discussion of parts of this work. Hal Guard

and Bob Nowak have been patient and helpful. Mike Sevilla was particularly helpful several years ago at the time I formulated the model of nerve and muscle membrane surfaces. Mark Severson has been steadfast in offering sound advice on many matters concerning vibrational spectroscopy. Finally, I wish to acknowledge the encouragement given to me by Per-Olov Löwdin. The work was supported in part by the Office of Naval Research.

9. APPENDICES

Appendix A: Origin of the Gaussian screened Coulombic form of the interaction between the ion and the wall source

The general form of the Coulomb interaction between a test charge distribution ρ and a source distribution ρ_s is

$$V = \int d^3r_1 d^3r_2 \frac{\rho(\mathbf{r}_1)\rho_s(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (\text{A1})$$

The reference distance between the coordinate origins of the two charge distributions is \mathbf{R} . The interaction is therefore conveniently expressed as the inverse Fourier transform of the convolution of the charges and the Coulomb interaction

$$V = \frac{1}{2\pi^2} \int d^3k \, k^{-2} \rho(\mathbf{k}) \rho_s(\mathbf{k}) \exp(-i\mathbf{k} \cdot \mathbf{R}). \quad (\text{A2})$$

For a point charge, the Fourier transform of the charge density is $\rho(\mathbf{k}) = Ze$, the charge itself. The Fourier transform of a Gaussian charge distribution is

$$\rho_{\text{Gauss}}(\mathbf{k}) = Z_G e \exp(-k^2/4b) \quad (\text{A3})$$

which derives from the r -space form,

$$\rho_{\text{Gauss}}(\mathbf{r}) = Z_G (b/\pi)^{3/2} \exp(-br^2). \quad (\text{A4})$$

The evaluation of the integrals in eq (A2), using the Rayleigh formula [30] for $e^{-i\mathbf{k} \cdot \mathbf{R}}$, gives

$$V = \frac{Z_G Z_s e^2}{R} \operatorname{erf}(\sqrt{b}R) \quad (\text{A5})$$

where $\text{erf}(x)$ is the error function.

We assume that the Gaussian contribution to the interaction potential energy arises from the continuous, soft charge distribution due to the atomic and molecular electrons. The effective core charge is positive. The model potential, therefore, is assumed to consist of a single, positive core charge with an equal and compensating outer electronic charge. Thus, the complete potential is

$$V = \frac{e^2}{R} \left(1 - \text{erf}(\sqrt{5}R) \right) = \frac{e^2}{R} \text{erfc}(\sqrt{5}R). \quad (\text{A6})$$

Appendix B: Cartesian forms of the force constants for the vibration of an ion along a channel

This appendix contains formulae necessary to compute the classical harmonic vibrations of an ion at a position of equilibrium along the helical channel axis. For potential energy functions that depend explicitly on distances, and not angles, the Cartesian force constants are easy to find. From the Taylor series for the expansion of a function at a particular point, one finds the Cartesian diagonal elements from the formula [26,28,29]

$$k_{xx} = \frac{1}{3} \sum_i \left\{ I_{20}(R_i) + (2X_i^2 - Y_i^2 - Z_i^2) I_{22}(R_i)/R_i^2 \right\} \quad (\text{B1})$$

with similar terms for the other diagonal elements [obtained by permutation of the Cartesian variables]

$$k_{xy} = \sum_i I_{22}(R_i)/R_i^2 \quad (\text{B2})$$

for the off-Cartesian-diagonal elements. In these expressions, I_{20} and I_{22} are [81]

$$I_{20}(R) = \frac{d^2 V}{dR^2} + \frac{2}{R} \frac{dV}{dR} \quad (\text{B3})$$

and

$$I_{22}(R) = \frac{d^2 V}{dR^2} - \frac{1}{R} \frac{dV}{dR}. \quad (\text{B4})$$

For the Morse part of the potential, one readily finds [81]

$$I_{20}(R) = 2\alpha^2 D \exp[-\alpha(R-R_0)] \left[2(1-1/\alpha R) \exp[-\alpha(R-R_0)] - 1 + 2/\alpha R \right] \quad (\text{B5})$$

and

$$I_{22}(R) = 2\alpha^2 D \exp[-\alpha(R-R_0)] \left[(2+1/\alpha R) \exp[-\alpha(R-R_0)] - 1 - 1/\alpha R \right] \quad (\text{B6})$$

For the screened Coulomb part, one finds

$$I_{20}(R) = - \frac{4Ze^2}{\pi^{1/2}} b^{3/2} \exp(-bR^2) \quad (\text{B7})$$

and

$$I_{22}(R) = \frac{Ze^2}{R^3} \left\{ 3[1 - \text{erf}(\sqrt{b}R)] - \frac{2}{\sqrt{\pi}} \sqrt{b}R \exp(-bR^2) \right\}. \quad (\text{B8})$$

The expressions I_{20} and I_{22} for the three-body function are derived easily for eq (5). There is a slight difference in that the three-body term contains the simultaneous interaction of the channel ion with two sources in the wall. Thus, the ion force constants involve two terms for each nominal interaction:

$$k_{xx} = \frac{1}{3} \sum_{I,J} \left[I_{20}(R_I[R_J]) + (2X_I^2 - Y_I^2 - Z_I^2) I_{22}(R_I[R_J])/R_I^2 \right] \quad (\text{B9})$$

and

$$k_{xy} = \sum_{I,J} \left[X_I Y_I I_{22}(R_I[R_J])/R_I^2 \right] \quad (\text{B10})$$

in which the I_{20} and I_{22} for the three-body contribution are given by

$$I_{20}(R_I[R_J]) = \alpha^2 A \left[1 - \frac{2}{\alpha R_I} \right] \exp[-\alpha(R_I-R_0) - \alpha(R_J-R_0)] \quad (\text{B11})$$

and

$$I_{22}(R_I[R_J]) = \alpha^2 A \left[1 + \frac{1}{\alpha R_I} \right] \exp[-\alpha(R_I-R_0) - \alpha(R_J-R_0)]. \quad (\text{B12})$$

The form of the three-body potential examined here is but one of many that are possible. At the harmonic level, the most common form of many-body interaction is angle-dependent. The conversion of angle dependent force constants to Cartesian forms has been given by Pariseau, Suzuki and Overend [82]; details of the transformation can be found in their work.

Finally, for the (possible) dipole contribution, one obtains expressions for forces and force constants with the use of the Carlson-Rushbrooke formula [83] specialized to the ion-dipole interaction:

$$V = \mu \sum_{lm} \frac{(-r)^l}{R^{l+2}} B_{1l}^{mm'} Y_{1+l, -m-m'}(\hat{\mathbf{R}}) Y_{1m}(\hat{\mathbf{a}}) Y_{lm'}(\hat{\mathbf{r}}), \quad (\text{B13})$$

where the Y_{lm} are the spherical harmonic functions, μ is the magnitude of the dipole moment, the argument $\hat{\mathbf{R}}$ is the unit vector corresponding to \mathbf{R} and $\hat{\mathbf{a}}$ is the dipole unit vector. \mathbf{R} is the reference vector from the point of expansion—for the ion—to the wall source. $rY_{lm}(\hat{\mathbf{r}})$ gives the components of the displacement of the ion from the point of expansion. The coefficients $B_{1l}^{mm'}$ are given by [83]

$$B_{1l}^{mm'} = \frac{(4\pi)^{3/2} (-1)^{m+m'}}{[(2l+1)(2l'+1)(2l+2l'+1)]^{1/2}} \times \left\{ \frac{(l+l'+m+m')!(l+l'-m-m')!}{(l+m)!(l-m)!(l'+m')!(l'-m')!} \right\}^{1/2} \quad (\text{B14})$$

The specific Cartesian contributions to the force constants are given by

$$k_{xx} = \frac{6e\mu}{R^4} \left\{ X(2X^2 - 3Y^2 - 3Z^2)a_x + Y(4X^2 - Y^2 - Z^2)a_y + Z(4X^2 - Y^2 - Z^2)a_z \right\}, \quad (\text{B15})$$

and the quantities a_q are the direction cosines for the dipole located at \mathbf{R} . The remaining Cartesian diagonal force constants are obtained by cyclic permutation. The Cartesian off-diagonal matrix elements are

$$k_{xy} = 6 \frac{e\mu}{R^7} \left[Y(4X^2 - Y^2 - Z^2)a_x + X(4X^2 - Y^2 - Z^2)a_y + 5XYZa_z \right] \quad (\text{B16})$$

where again the remaining elements are obtained by cyclic permutation.

Appendix C: Matrix elements of the Coulomb and Morse operators in basis functions of the nearly free ion wavefunctions

The simple treatment of the nearly free ion was based on the use of a fast Fourier transform of the effective potential measured along the helical axis. It is possible, however, to evaluate matrix elements of both the Coulomb interaction and the Morse potential in terms of basis functions that are bound in the x,y -plane and a plane wave along the helical axis. The purpose of this appendix is to outline these evaluations.

The interaction of a mobile ion located on the helical channel axis with the sources of molecules in the channel wall takes place with the use of potential eq (A5) of Appendix A. The complementary error function can be written as an integral transform in the form

$$\frac{1}{r} \operatorname{erfc}(\sqrt{6}r) = \frac{2}{\sqrt{\pi}} \int_{\sqrt{6}}^{\infty} du \exp(-u^2 r^2). \quad (\text{C1})$$

Likewise, the Morse potential can be recast [49] into a representation in terms of a Gaussian operator similar to $\exp(-u^2 r^2)$. In particular, following Pople, Hehre, and Stewart [51], the exponential function can be expanded as

$$\exp(-\xi r) = (2^{3/4}/\pi^{1/4}) \sum_{k=1}^6 \delta_{1s,k} \alpha_{1s,k}^{3/4} \exp\left[-\alpha_{1s,k} (\xi r)^2\right]. \quad (\text{C2})$$

Therefore, the Morse potential,

$$V_M(r) = D \exp[-\alpha(r - r_0)] \left[\exp[-\alpha(r - r_0)] - 2 \right], \quad (\text{C3})$$

expands as

$$V_M(r) = D \exp(\alpha r_0) \sum_{k=1}^6 d_k \left\{ \exp(\alpha r_0 - 4u_k r^2) - 2 \exp(-u_k r^2) \right\} \quad (\text{C4})$$

and the coefficients d_k and u_k are given by

$$d_k = \frac{2^{3/4}}{\pi^{1/4}} \delta_{1s,k} \alpha_{1s,k}^{3/4} \quad (C5)$$

and

$$u_k = \alpha^2 \alpha_{1s,k}. \quad (C6)$$

The quantities $\alpha_{1s,k}$ and $\delta_{1s,k}$ are found in the original work by Pople *et al.*, while the derivative quantities, α_k and δ_k are given in Table C.1.

TABLE C.1

*Coefficients for the Gaussian expansion
of the exponential function*

	$\delta_{1s,k}$	$\alpha_{1s,k}$
1	1.31334e-01	6.51095e-02
2	4.16492e-01	1.58088e-01
3	3.70563e-01	4.07099e-01
4	1.68538e-01	1.18506
5	4.93615e-02	4.23592
6	9.16360e-03	23.1030

The notation is that of Hehre, Stewart, and Pople [51]. The label '1s' refers to the 1s Slater orbital. The coefficients apply to the case where $a = 1$.

The generic exponential interaction is denoted as \mathbf{e} , viz.,

$$\mathbf{e} = \exp(-fr^2); \quad (C7)$$

note that in this operator, $r = \{(x-X)^2 + (y-Y)^2 + (z-Z)^2\}^{1/2}$. Matrix elements of the operator \mathbf{e} are standard; one finds

$$\langle s | \mathbf{e} | s' \rangle = \frac{2\pi^{1/2}a}{N\zeta(2a+f)\sqrt{f}} \exp\left[-\frac{2af}{2a+f} R^2 - \frac{g^2}{4f} - igZ\right] \quad (C8)$$

with

$$R^2 = X^2 + Y^2. \quad (C9)$$

Rewrite eq (C8) as

$$\langle s | e | s' \rangle = \frac{1}{N} E_{xp}(R, g, f) \exp(-igZ) \quad (C10)$$

with

$$E_{xp}(R, g, f) = \frac{2\pi^{1/2}a}{\zeta(2a+f)\sqrt{f}} \exp\left[-\frac{2af}{2a+f} R^2 - \frac{g^2}{4f}\right] \quad (C11)$$

Matrix elements involve a summation over all the ionic interactions with the wall sources. Thus, it is necessary to consider terms of the form

$$\sum_I \langle s | e_I | s' \rangle = \frac{1}{N} E_{xp}(R, g, f) \sum_I \exp(-igZ_I) \quad (C12)$$

The radial distance R is constant for each interaction; the ion is located on the channel axis and each wall source is the same distance from the axis. Only the reflection of the wall source on the Z -axis changes. Thus, when the modulus square of the matrix element is considered, as it is in the variational calculation, one has

$$\begin{aligned} \left| \sum_I \langle s | e_I | s' \rangle \right|^2 &= E_{xp}(R, g, f)^2 \frac{1}{N^2} \sum_{I,J} \exp[-ig(Z_I - Z_J)] \\ &= E_{xp}(R, g, f)^2 \end{aligned} \quad (C13)$$

because $Z_I - Z_J = n\zeta$, (n =integer) and $\sum_{I,J} \exp[-ig(Z_I - Z_J)] = N^2$. For the Morse potential contribution, f in $E_{xp}(R, g, f)$ is simply replaced by u_k . Thus,

$$\begin{aligned} \langle s | V_{\text{Morse}} | s' \rangle &= D \exp(ar_0) \sum_{k=1}^6 d_k \left\{ \exp(ar_0) E_{xp}(R, g, 4u_k) \right. \\ &\quad \left. - 2 E_{xp}(R, g, u_k) \right\}. \end{aligned} \quad (C14)$$

On the other hand, for the screened Coulomb interaction, replacing f by u^2 in eq (C1) followed by a change of variable and simplification yields

$$\langle s | V_{\text{Coul}} | s' \rangle = \frac{2\pi e^2}{N\zeta} \int_L^1 dt \, t^{-1} \exp\left[-\frac{g^2}{4} \frac{1-t^2}{2at^2} - 2aR^2 t^2\right] \sum_{I,J} e^{-igZ}$$

$$= \frac{2\pi e^2}{\zeta} \int_b^\infty ds \, E_{xp}(R, g, s^2). \quad (C15)$$

Define the following function

$$\tilde{\gamma}_n(R, g) = \frac{2}{\zeta} \int_L^1 dt \, t^n \exp \left[-\frac{g^2}{4} \frac{1-t^2}{2at^2} - 2aR^2t^2 \right] \quad (C16)$$

with $-1 \leq n$ and the lower limit on the integration is given by

$$L = \sqrt{\frac{b}{2a+b}}. \quad (C17)$$

Evaluation of the integral is best carried out with the use of a Gaussian quadrature. Thus, the matrix element can be expressed as

$$\langle s | V_{Coul} | s' \rangle = e^2 \tilde{\gamma}_{-1}(R, g) \quad (C17)$$

The remaining matrix elements over p_x and p_y functions are expressed as follows:

$$\langle s | V_C | x' \rangle = e^2 \langle X \rangle \tilde{\gamma}_1(R, g) \quad (C18)$$

$$\langle x | V_C | y' \rangle = 4a e^2 \left[\langle XY \rangle \tilde{\gamma}_3(R, g) + \frac{1}{4a} \left(\tilde{\gamma}_{-1}(R, g) - \tilde{\gamma}_1(R, g) \right) \delta_{xy} \right]. \quad (C19)$$

In these expressions for the Coulomb and Morse potentials, it should be noted that the completely diagonal elements, viz., $\langle s | V | s \rangle$, follow for $g = 0$; thus,

$$\langle s | V_C | s \rangle = e^2 \tilde{\gamma}_{-1}(R, 0) \quad (C20)$$

and so on.

Appendix D: Evaluation of matrix elements involving the dipole interaction

The charge-dipole interaction can be derived, as is well known, by taking the negative gradient of $1/|\mathbf{r}-\mathbf{C}|$ as follows

$$V_{\text{dipole}} = -\underline{\mu} \cdot \nabla_C \frac{1}{|\mathbf{r}-\mathbf{C}|}; \quad (D1)$$

the negative gradient is used as the potential is defined as that from the dipole to the test charge. The components of the dipole moment $\underline{\mu}$ are defined simply as el_x and so on. With the use of this relationship, therefore, one can define the matrix elements of the dipole moment operator as [84]

$$\langle a | V_{\text{dipole}} | b \rangle = -\underline{\mu} \cdot \nabla_C \langle a | \frac{1}{|\mathbf{r}-\mathbf{C}|} | b \rangle \quad (\text{D2})$$

where now \mathbf{C} is naturally the location of the dipole in the wall of the channel.

With the use of this formula for the dipolar interaction, the matrix elements of the ion-dipole interaction in the basis of the nearly free wavefunctions are listed below.

$$\langle s | V_D | s \rangle = 4a(\langle \mu_x X \rangle + \langle \mu_y Y \rangle) \tilde{\gamma}_1(R, g) \quad (\text{D3})$$

$$\langle x | V_D | s \rangle =$$

$$-2a^{1/2} \left\{ \langle \mu_x \rangle \tilde{\gamma}_1(R, g) - 4a(\langle \mu_x X^2 \rangle + \langle \mu_y XY \rangle) \tilde{\gamma}_3(R, g) \right\} \quad (\text{D4})$$

$$\begin{aligned} \langle y | V_D | s \rangle = & -2a^{1/2} \left\{ -4a(\langle \mu_x XY \rangle \right. \\ & \left. + \langle \mu_y X^2 \rangle) \tilde{\gamma}_3(R, g) + \langle \mu_y \rangle \tilde{\gamma}_1(R, g) \right\} \end{aligned} \quad (\text{D5})$$

$$\begin{aligned} \langle x | V_D | y \rangle = & -4a \left\{ (\langle \mu_x Y \rangle + \langle \mu_y X \rangle) \tilde{\gamma}_3(R, g) \right. \\ & \left. - 4a(\langle \mu_x X^2 Y \rangle + \langle \mu_y XY^2 \rangle) \tilde{\gamma}_5(R, g) \right\} \end{aligned} \quad (\text{D6})$$

$$\begin{aligned} \langle x | V_D | x \rangle = & -4a \left\{ \langle \mu_x X \rangle \left[3\tilde{\gamma}_3(R, g) - \tilde{\gamma}_1(R, g) \right] \right. \\ & \left. + \langle \mu_y Y \rangle \left[\tilde{\gamma}_3(R, g) - \tilde{\gamma}_1(R, g) \right] - 4a \left[\langle \mu_x X^3 \rangle + \langle \mu_y X^2 Y \rangle \right] \tilde{\gamma}_5(R, g) \right\} \end{aligned} \quad (\text{D7})$$

$$\langle y | V_D | y \rangle = -4a \left\{ \langle \mu_x X \rangle \left[\tilde{\gamma}_3(R, g) - \tilde{\gamma}_1(R, g) \right] + \langle \mu_y Y \rangle \left[3\tilde{\gamma}_3(R, g) - \tilde{\gamma}_1(R, g) \right] \right\}$$

$$- 4a \left\{ \langle \mu_x XY^2 \rangle + \langle \mu_y Y^3 \rangle \right\} \mathfrak{F}_5(R, g) \}. \quad (D8)$$

The evaluation of matrix elements of the dipole moment operator in the full Cartesian basis makes use of the formulae given by Ōbara and Saika, [50 , eqs (A22) - (A26)].

Appendix E: Matrix elements of the truncated harmonic potential

Define a function

$$G(i, m, q) = \gamma(i/2, (m+3)^2 q^2) - \gamma(i/2, (m+1)^2 q^2). \quad (E1)$$

in which $\gamma(i, x)$ is the incomplete gamma function [30] of order i and $q = \sqrt{a}\zeta$. Then, in terms of the ground state primitive basis functions, the single center matrix element is given by

$$\begin{aligned} \langle 00 | V | 00 \rangle &= \frac{1}{2\sqrt{\pi}} \hbar\omega \left\{ \gamma(3/2, q^2) \right. \\ &+ \lim_{m=0} \sum \left\{ G(3, 2m, q) - 2(2m+2)qG(2, 2m, q) \right. \\ &\quad \left. \left. + (2m+2)^2 q^2 G(1, 2m, q) \right\} \right\} \quad (E1) \end{aligned}$$

The cell that contains the origin of coordinates for the whole system is used as the reference cell. The two-center matrix element that involves only the ground state primitive basis functions ϕ_0 is

$$\begin{aligned} \langle 00 | V | 0m \rangle &= \frac{1}{4\sqrt{\pi}} \hbar\omega S_{0m} \left\{ \gamma[3/2, (m+1)^2 q^2] + \gamma[3/2, (m-1)^2 q^2] \right. \\ &\quad \left. + 2mq \left[\gamma[1, (m+1)^2 q^2] - \gamma[1, (m-1)^2 q^2] \right] \right. \\ &\quad \left. + m^2 q^2 \left[\gamma[1/2, (m+1)^2 q^2] + \gamma[1/2, (m-1)^2 q^2] \right] \right\} \end{aligned}$$

$$\begin{aligned}
& \lim_{n=0} \left\{ G(3, 2n+m, q) + G(3, 2n-m, q) \right. \\
& - 2(2n+m+2) q G(2, 2n+m, q) - 2(2n-m+2) q G(2, 2n-m, q) \\
& \left. + 2(2n+m+2)^2 q^2 G(1, 2n+m, q) + 2(2n-m+2)^2 q^2 G(1, 2n-m, q) \right\} \quad (E3)
\end{aligned}$$

The single-center matrix element for the second primitive (first nominal excited state) basis function ϕ_1 is

$$\begin{aligned}
\langle 10 | V | 10 \rangle &= \frac{\hbar\omega}{\sqrt{\pi}} \left\{ \gamma(5/2, q^2) + \sum_{m=0}^{\infty} \left\{ \gamma(5/2, (2m+3)^2 q^2) \right. \right. \\
& - \gamma(5/2, (2m+1)^2 q^2) - 2(2m+2)q \left[\gamma(2, (2m+3)^2 q^2) - \gamma(2, (2m+1)^2 q^2) \right] \\
& \left. \left. + (2m+2)^2 q^2 \left[\gamma(3/2, (2m+3)^2 q^2) - \gamma(3/2, (2m+1)^2 q^2) \right] \right\} \right\}. \quad (E4)
\end{aligned}$$

The two-center matrix elements in the second primitive basis functions are

$$\begin{aligned}
\langle 10 | V | 1m \rangle &= \frac{\hbar\omega}{2\sqrt{\pi}} \left\{ \gamma(5/2, (m-1)^2 q^2) + \gamma(5/2, (m+1)^2 q^2) \right. \\
& + 2mq \left[\gamma(2, (m-1)^2 q^2) - \gamma(2, (m+1)^2 q^2) \right] \\
& - 2m^3 q^3 \left[\gamma(1, (m-1)^2 q^2) - \gamma(1, (m+1)^2 q^2) \right] \\
& \left. - m^4 q^4 \left[\gamma(1/2, (m-1)^2 q^2) - \gamma(1/2, (m+1)^2 q^2) \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& + \sum_{n=0}^{\infty} \left\{ G(5, 2n+m, q) - G(5, 2n-m, q) \right. \\
& - 2q \left[(2n+m+2)G(4, 2n+m, q) + (2n-m+2)G(4, 2n-m, q) \right] \\
& + q^2 \left[(2n+m+2)G(3, 2n+m, q) + (2n-m+2)G(3, 2n-m, q) \right] \\
& + 2m^2 q^3 \left[(2n+m+2)G(2, 2n+m, q) + (2n-m+2)G(2, 2n-m, q) \right] \\
& \left. - 2m^2 q^4 \left[(2n+m+2)G(1, 2n+m, q) + (2n-m+2)G(1, 2n-m, q) \right] \right\}. \quad (E5)
\end{aligned}$$

Finally, the off-diagonal matrix elements in the two primitive basis functions are

$$\begin{aligned}
\langle 00 | V | 1m \rangle &= \frac{\hbar\omega}{2\sqrt{2\pi}} \left\{ \gamma(2, (1-m)^2 q^2) - \gamma(2, (1+m)^2 q^2) \right. \\
& + mq \left[\gamma(3/2, (1-m)^2 q^2) + \gamma(3/2, (1+m)^2 q^2) \right] \\
& + (mq)^2 \left[\gamma(1, (1-m)^2 q^2) - \gamma(1, (1+m)^2 q^2) \right] \\
& - (mq)^3 \left[\gamma(1/2, (1-m)^2 q^2) + \gamma(1/2, (1+m)^2 q^2) \right] \\
& + \sum_{n=0}^{\infty} \left\{ G(4, 2n-m, q) - G(4, 2n-m, q) \right. \\
& + (4n+m+4)qG(3, 2n+m, q) - (4n-m+4)qG(3, 2n-m, q) \\
& + (2n+m+2)q^2G(2, 2n+m, q) - (2n-m+2)q^2G(2, 2n-m, q) \\
& \left. + m(2n+m+2)q^3G(1, 2n+m, q) - m(2n-m+2)q^3G(1, 2n-m, q) \right\} \Big\} e^{-m^2 q^2}. \quad (E6)
\end{aligned}$$

F: Matrix elements in Gaussian primitive basis functions

Three dimensional Cartesian Gaussian primitive functions have been used as basis functions for the molecular electronic problem for some time now [16]. Gaussian functions, of course, are natural for treating vibrational problems [37]. As has been noted in previous publications [49,85], advantage can be taken of the developments in the evaluation of molecular electronic integrals to apply to the vibrational problem. In this section, matrix elements are listed that find use in the general vibrational band problem. Considerable advantage can be taken of the Obara-Saika [50] recursion relations in simplifying the evaluation of specific matrix elements.

We begin by listing the overlap matrix elements for the s - and p -functions used in the basis set. The ss overlap integral is fundamental to all the following integrals. It is

$$\langle s \| s \rangle = \int d^3r \exp(-ar_A^2 - br_B^2) \quad (F1)$$

with $r_A = |\mathbf{r} - \mathbf{R}_A|$. Thus,

$$\begin{aligned} \langle s \| s \rangle &= \prod_{q=x,y,z} \int_{-\infty}^{\infty} dq \exp\left[-a(q - A_q)^2 - b(q - B_q)^2\right] \\ &= \left(\frac{\pi}{a+b}\right)^{3/2} \exp\left[-\frac{ab}{a+b} R_{AB}^2\right]. \end{aligned} \quad (F2)$$

The remaining overlap matrix elements follow from the Obara-Saika recursion relations [50]. Thus,

$$\langle x \| s \rangle = (P_x - A_x) \langle s \| s \rangle \quad (F3)$$

$$\langle x \| y \rangle = (P_x - A_x) \langle s \| y \rangle \quad (F4)$$

and, finally,

$$\langle x \| x \rangle = (P_x - A_x) \langle s \| x \rangle + \frac{1}{2(a+b)} \langle s \| s \rangle. \quad (F5)$$

The kinetic energy matrix elements are closely related. Define the general matrix element as

$$\langle q | \text{ke} | p \rangle = - \frac{\hbar^2}{2M} \langle q | \nabla^2 | p \rangle. \quad (\text{F6})$$

Then, with direct evaluation, one finds first

$$\langle s | \text{ke} | s \rangle = \frac{\hbar^2}{M} \frac{ab}{a+b} \left[3 - 2 \frac{ab}{a+b} R_{AB}^2 \right] \langle s || s \rangle \quad (\text{F7})$$

The successive matrix elements are then the following:

$$\langle x | \text{ke} | s \rangle = (P_x - A_x) \langle s | \text{ke} | s \rangle + 2 \frac{\hbar^2}{M} \frac{ab}{a+b} \langle x || s \rangle \quad (\text{F8})$$

$$\langle x | \text{ke} | y \rangle = (P_x - A_x) \langle s | \text{ke} | y \rangle + 2 \frac{\hbar^2}{M} \frac{ab}{a+b} \langle x || y \rangle \quad (\text{F9})$$

and, finally,

$$\begin{aligned} \langle x | \text{ke} | x \rangle &= (P_x - A_x) \langle s | \text{ke} | x \rangle + 2 \frac{\hbar^2}{M} \frac{ab}{a+b} \langle x || x \rangle \\ &+ \frac{1}{2(a+b)} \langle s | \text{ke} | s \rangle. \end{aligned} \quad (\text{F10})$$

An expansion of the Morse potential, for example, in a set of Gaussian functions is given by eq (C4) in Appendix C. Matrix elements of the Morse potential in terms of the Gaussian primitive basis functions are therefore simply three center overlap integrals [49]. These matrix elements can be evaluated for each term in the sum and then converted to the final expression in a straightforward manner.

Define an exponential operator E by

$$E = \exp(-ur_C^2) \quad (\text{F11})$$

which is referred to the point C. The quantities u assume the values needed in the expansion of the Morse potential, eq (C4), for example. By direct evaluation of the integrals, one finds

$$\langle s|E|s\rangle = \left(\frac{a+b}{a+b+u} \right)^{3/2} \langle s||s\rangle \exp \left\{ -\frac{(a+b)u}{a+b+u} (\mathbf{P} - \mathbf{C})^2 \right\}. \quad (\text{F12})$$

Next, $\langle x|E|s\rangle$ is

$$\langle x|E|s\rangle = (G_x - A_x) \langle s|E|s\rangle \quad (\text{F13})$$

in which

$$\mathbf{G} = \frac{(a+b)\mathbf{P} + u\mathbf{C}}{a+b+u}. \quad (\text{F14})$$

The off-diagonal pp' matrix elements are

$$\langle x|E|y\rangle = (G_x - A_x) \langle s|E|y\rangle. \quad (\text{F15})$$

The diagonal pp matrix elements are finally

$$\langle x|E|x\rangle = (G_x - A_x) \langle s|E|x\rangle + \frac{1}{2(a+b+u)} \langle s|E|s\rangle. \quad (\text{F16})$$

The behavior of the matrix elements of the primitive functions with respect to symmetry operations of the helical chain is important in determining the character of the matrix elements in the first Brillouin zone. The ss -matrix element is completely symmetric, of course. Consider next, however, the $\langle x(m)|E|s(n)\rangle$ matrix element, eq (F13). $G_x - A_x$ can be rewritten as

$$G_x - A_x = P_x - A_x + \frac{u}{a+b+u} (C_x - P_x). \quad (\text{F17})$$

Similar terms apply to the matrix elements of the y - and z -coordinates. For the helix, $A_x = B_x = A_y = B_y = 0$. As a result,

$$\langle x|E|s\rangle = \frac{u}{a+b+u} C_x \langle s|E|s\rangle. \quad (\text{F18})$$

A similar term applies to the sp_y matrix element. The remaining distinct matrix elements are

$$\langle x|E|x\rangle = \frac{1}{2(a+b+u)} \langle s|E|s\rangle + \left(\frac{u}{a+b+u} \right)^2 C_x^2 \langle s|E|s\rangle \quad (\text{F19})$$

$$\langle x|E|y\rangle = \left[\frac{u}{(a+b+u)} \right]^2 C_x C_y \langle s|E|s\rangle \quad (\text{F20})$$

$$\langle x|E|z\rangle = \frac{u}{(a+b+u)} C_x \left[P_z - A_z + \frac{u}{a+b+u} (C_z - P_z) \right] \langle s|E|s\rangle. \quad (\text{F21})$$

The transformation properties of the coordinates of the sources with reference to the zeroth cell are easy to determine. The coordinates of the two source atoms in the zeroth cell have been determined to be $(0, \pm p, 0)$ where p is the radial distance of the source from the z -axis. It is easy to determine that

$$\begin{aligned} x(-n) &= -x(n) \\ y(-n) &= y(n) \\ z(-n) &= -z(n). \end{aligned} \quad (\text{F22})$$

This property is important in determining the parity of the matrix elements. The evaluation of the interaction of the ion with the sources can be carried out by summing along one chain in the positive axial direction. Reflection in the negative direction yields a second term the sign of which is determined by (F22). When these summations are carried out at the same time as the summations over cells, the character of the matrix elements in the various Brillouin zones emerges.

10. REFERENCES

1. Pullman, A., *Chem. Rev.*, **1991**, 91, 793; this large and complete review should be consulted for individual citations of computational methods and results used to investigate the energy states available to the ion in model channels. I make no effort here to attempt to duplicate citations.
2. An early review covering experimental and theoretical results on gates and pores is given by Armstrong, C. M., *Quart. Rev. Biophys.*, **1975**, 7, 179
3. Work by Parsegian, A., *Nature*, **1969**, 223, 844, some time ago showed how dielectric considerations implicit to the intramembrane region, e.g., the channels, lower the energy barriers to the ion transfer.
4. The original impetus to consider channels and pores comes from the work of Hodgkin, A. L., and Huxley, A. F., *J. Physiol.* **1952**, 117, 500
5. Whittingham, M. S. and Jacobson, A. J. *Intercalation Chemistry* (Academic Press, NY, 1982); particular reference I make is to the examples of the Group Ia- and IIa-doped β -aluminas and the fast ionic conduction observed. Chapters in this reference discuss

conduction models and provide further references to the original literature.

6. Iijima, S., *Nature*, **1991**, 354, 56
7. Ebbeson, J. W. and Ajayan, P. M., *Nature*, **1992**, 358, 220
8. Stamp P. C. E. and Zhang, C., *Phys. Rev. Letters*, **1991**, 66, 1902
9. Kronig, R. de L. and Penney, W. G., *Proc. Roy. Soc. (London)*, **1931**, A130, 499
10. Dougherty, D. A. and Stauffer, D. A., *Science*, **1990**, 250, 1558; it has been regarded for some time as necessary to have significant amounts of water incorporated into the structure of the channel. This work suggests that π -bonding involving aromatic species is sufficiently strong to act as a substitute for water within the channel and to make the transfer more or less isoenergetic. See also the recent publication of Suzuki, S., Green, P. G., Bumgarner, R. E., Dasgupta, S., Goddard III, W. A. and Blake, G. A., *Science*, **1992**, 257, 942 that discusses the formation of hydrogen bonds between benzene and water.
11. Schmidt, P. P. and Korzeniewski, C., *J. Chem. Soc. Far. 2*, **1984**, 80, 2017
12. Fraser, A., and Frey, H., *Biophys. J.*, **1968**, 8, 731
13. Mintmire, J. W. in *Density Functional Methods in Chemistry*, Labenowski, J., and Andezelm, Eds. (Springer Verlag, NY, 1990)
14. Painter, P. C., Coleman, M. M., and Koenig, J. L., *The Theory of Vibrational Spectroscopy and its Application to Polymeric Materials* (John Wiley and Sons, NY, 1982)
15. Gradshteyn I. S. and Ryzhik, I. M., *Table of Integrals, Series, and Products* (Academic Press, NY, 1965)
16. Shavitt, I., *Methods of Comput. Phys.*, **1963**, 2, 1
17. Murrell, J. N., Carter, S., Farantos, S. C., Huxley, P. and Varandas, A. J. C., *Molecular Potential Energy Functions* (John Wiley and Sons, NY, 1984)
18. Löwdin, P.-O., *Rev. Mod. Phys.*, **1963**, 35, 724
19. Löwdin, P.-O., *Adv. Quantum Chem.*, **1965**, 2, 216
20. Hall, G. G., *Theoret. Chim. Acta*, **1985**, 67, 439
21. Maxey B. W. and Popov, A. I., *J. Am. Chem. Soc.*, **1969**, 91, 20
22. McKinney, W. J. and Popov, A. I., *J. Phys. Chem.*, **1970**, 74, 535
23. Wuepper, J. L. and Popov, A. I., *J. Am. Chem. Soc.*, **1970**, 92, 1493
24. Wong, M. K., McKinney, W. J. and Popov, A. I., *J. Phys. Chem.*, **1971**, 75, 56
25. Edgell, W. F., Lyford, J., Wright, R., Risen, W., and Watts, R., *J. Am. Chem. Soc.*, **1970**, 2240
26. Schmidt, P. P., Pons, B. S. and McKinley, J. M., *J. Chem. Soc. Far 2*, **1980**, 76, 979
27. Born, M. and Landé, A., *Verh. d. D. Phys. Ges.*, **1918**, 20, 210
28. McKinley, J. M. and Schmidt, P. P., *J. Chem. Soc. Far. 2*, **1982**, 78, 867
29. McKinley, J. M. and Schmidt, P. P., *Chem. Phys. Letters*, **1984**, 110, 379
30. Arfken, G., *Mathematical Methods for Physicists* (Academic Press, NY, 2nd Ed., 1970)

31. Morse, P. M. and Feshbach, H. *Methods of Mathematical Physics* (McGraw-Hill Book Co., NY, 1953)
32. See, for example, Schmidt, P. P. and Pons, B. S., *Electrochim. Acta*, **1982**, 27, 867, 875 for an application to questions of the stability of solvation of ions at interfaces.
33. Ziman, J. M., *Electrons and Phonons, The Theory of Transport Phenomena in Solids* (Oxford University Press, London, 1960)
34. Kittel, C., *Quantum Theory of Solids* (John Wiley & Sons, NY, 1963)
35. Harrison, W. A., *Solid State Theory* (McGraw-Hill Book Co., NY, 1970)
36. Robertson, D. H., Brenner, D. W. and Mintmire, J. W., *Phys. Rev. B*, **1992**, 45, 12592
37. Davydov, A. S., *Quantum Mechanics* (Pergamon Press, London, 1965)
38. Mott, N. S. and Jones, H., *Theory of the Properties of Metals and Alloys* (Cambridge University Press, Cambridge, 1936, Dover Publications, NY, 1958)
39. Press, W. H., Flannery, B. P. Teukolsky, S. H. and Vetterling, W. J., *Numerical Recipes, The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986)
40. Salem, L., *Molecular Orbital Theory of Conjugated Systems* (W. H. Benjamin, NY, 1966)
41. Wall, F. T. and Glockler, Wm., *J. Chem. Phys.*, **1937**, 5, 314
42. Löwdin, P.-O., *J. Chem. Phys.*, **1950**, 18, 365
43. Löwdin, P.-O., *Adv. in Phys.*, **1956**, 5, 1
44. Löwdin, P.-O., *Adv. Quantum Chem.*, **1970**, 5, 185
45. Schmidt, P. P. and Chang, S. S., *J. Phys. Chem.*, **1986**, 90, 4945
46. Schmidt, P. P. *J. Phys. Chem.*, **1989**, 93, 6610
47. Swalen, J. D. and Ibers, J. A., *J. Chem. Phys.*, **1962**, 36, 1914
48. Pilar, F. L. *Elementary Quantum Chemistry* (McGraw-Hill book Co., 1968)
49. Schmidt, P. P. *Chem. Phys. Letters*, **1989**, 159, 511
50. Obara, S. and Saika, A., *J. Chem. Phys.*, **1986**, 84, 3963
51. Hehre, W. J., Stewart, R. F. and Pople, J. A., *J. Chem. Phys.*, **1969**, 51, 2657
52. Born, M. and Huang, K., *Dynamical Theory of Crystal Lattices* (Oxford University Press, London, 1954)
53. Longuet-Higgins, H. C., *Adv. Spect.*, **1961**, 2, 429
54. Jørgensen, P. and Simons, J., Eds., *Geometrical Derivatives of Energy Surfaces and Molecular Properties* (D. Reidel, Dordrecht, 1986)
55. Engel, V., Schinke, R., Hennig, S., and Metiu, H., *J. Chem. Phys.*, **1990**, 92, 1
56. Schmidt, P. P. *J. Chem. Soc. Far. 2*, **1984**, 80, 181; it should be noted, as Marcus [Marcus, R. A., *J. Chem. Phys.*, **1984**, 81, 4494] has pointed out, that my original treatment [Schmidt, P. P., *J. Chem. Phys.*, **1973**, 58, 4384] of the semiclassical limit for the rate constant assumes that all environmental coordinates are classical. Marcus, on the other hand, has derived a more complete and accurate expression that accounts for the fact that only some of the vibrational coordinates are classical. It is probably safe to say, however, that for the problem I consider in this paper, all the

modes are classical except for the migrating ion. That they may not be well represented as harmonic is another matter, one that is better addressed with Marcus's formalism.

57. Schmidt, P. P. *J. Chem. Soc. Far. 2*, **1986**, 82, 1399
58. Holstein, T. *Adv. Phys. (NY)*, **1959**, 8, 325, 343
59. Glyde, H. R., *Rev. Mod. Phys.*, **1967**, 39, 373
60. Marcus, R. A., *J. Chem. Phys.*, **1964**, 41, 2624; **1965**, 43, 679
61. Miller, W., Schwartz, S. D. and Tromp, J. W., *J. Chem. Phys.*, **1983**, 79, 4889
62. Voter, A. F. and Doll, J. D., *J. Chem. Phys.*, **1984**, 80, 5814
63. see, for example, Requena, J., Whitembury, J., Tiffert, T., Eisner, D. A., and Mullins, L. J., *J. Gen. Physiol*, **1985**, 85, 789
64. see ref. 2, p. 202
65. Schmidt, P. P., *Biochim. Biophys. Acta*, **1971**, 233, 765
66. Hansma, P. K., Drake, B., Marti, O., Gould, S. A. C. and Prater, C. B., *Science*, **1989**, 243, 641; also, Prater, C. B., Hansma, P. K., Tortonese, M and Quate, C. F., *Rev. Sci. Instrum.*, **1991**, 62, 2634
67. Bard, A. J., Fan, F.-R. F., Kwak, J. and Lev, O., *Anal. Chem.*, **1989**, 61, 132
68. Bardeen J., *Phys. Rev. Letters*, **1961**, 6, 57
69. Kestner, N. R., Logan, J. and Jortner, J., *J. Phys. Chem.*, **1974**, 78, 2148
70. Schmidt, P. P., *Specialist Periodical Reports, Electrochemistry, Vol. 5* (The Chemical Society, London, 1975)
71. Bennett, A. J., Duke, C. D. and Silverstein, S. D., *Phys. Rev.*, **1968**, 176, 696
72. Tersoff, J. and Hamann, D. R., *Phys. Rev. B*, **1985**, 31, 805
73. West, P., Kramar, J., Baxter, D. V., Cave, R. J. and Baldeschwieler, J. D., *IBM J. Res. Develop.*, **1986**, 30, 484
74. Schmidt, P. P., *Proc. Indian Acad. Sci. (Chem. Sci.)*, **1986**, 97, 233
75. Bell, R. P., *The Proton in Chemistry* (Cornell University Press, Ithaca, 1959)
76. Bell, R. P., *The Tunnel Effect in Chemistry* (Chapman and Hall, London, 1980)
77. Marcus, R. A., *J. Phys. Chem.*, **1989**, 94, 4125
78. See, for example, Schmidt, P. P. in *Specialist Periodical Reports, Electrochemistry, Vol. 6* (The Chemical Society, London, 1976)
79. Hu, Y. and Mukamel, S., *Chem. Phys. Letters*, **1989**, 160, 410
80. Kirtley, J. Scalapino, D. J. and Hansma, P. K., *Phys. Rev. B*, **1976**, 14, 3177
81. Schmidt, P. P., Severson, M. W. and McKinley, J. M., *J. Chem. Phys.*, **1987**, 86, 5392
82. Pariseau, M. A., Suzuki, I. and Overend, J., *J. Chem. Phys.*, **1965**, 4211, 2335
83. Carlson, B. C. and Rushbrooke, G. S., *Proc. Camb. Phil. Soc.*, **1950**, 46, 626
84. Slater, J. C. and Frank, N. H., *Electromagnetism* (McGraw-Hill Book Co., NY, 1947)
85. Schmidt, P. P., Dunlap, B. I., and White, C. T., *J. Phys. Chem.*, **1991**, 91, 10537

MOLECULAR INTEGRALS
OVER GAUSSIAN BASIS FUNCTIONS

Peter M.W. Gill

Department of Chemistry
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Table of Contents

1. Quantum Chemical Procedures

2. Basis Functions

- 2.1 Slater Functions
- 2.2 Gaussian Functions
- 2.3 Contracted Gaussian Functions
- 2.4 Gaussian Lobe Functions
- 2.5 Delta Functions

3. A Survey of Gaussian Integral Algorithms

- 3.1 Performance Measures
 - 3.1.1 Flop-Cost
 - 3.1.2 Mop-Cost
 - 3.1.3 CPU-Time
- 3.2 Fundamental Integrals
 - 3.2.1 The Overlap Integral
 - 3.2.2 The Kinetic-Energy Integral
 - 3.2.3 The Electron-Repulsion Integral
 - 3.2.4 The Nuclear-Attraction Integral
 - 3.2.5 The Anti-Coulomb Integral
- 3.3 The Boys Algorithm
- 3.4 The Contraction Problem
- 3.5 The Pople-Hehre Algorithm
- 3.6 Bras, Kets and Brackets
- 3.7 The McMurchie-Davidson Algorithm
- 3.8 The Obara-Saika-Schlegel Algorithm
- 3.9 The Head-Gordon-Pople Algorithm
- 3.10 Variations on the HGP Theme

4. The PRISM Algorithm

- 4.1 Shell-Pair Data
- 4.2 Selection of Shell-Quartets
- 4.3 Generation of the $[0]^{(m)}$ Integrals
- 4.4 Contraction Steps
- 4.5 Transformation Steps
 - 4.5.1 Two-Electron Transformations on the MD PRISM
 - 4.5.2 One-Electron Transformations on the MD PRISM
 - 4.5.3 Two-Electron Transformations on the HGP PRISM
 - 4.5.4 One-Electron Transformations on the HGP PRISM
- 4.6 Loop Structure of PRISM in Gaussian 92
- 4.7 Performance of PRISM in Gaussian 92

5. Prospects for the Future

6. Acknowledgments

7. References

1. QUANTUM CHEMICAL PROCEDURES

The major goal of Quantum Chemistry is to obtain solutions to atomic and molecular Schrödinger equations [1]. To be useful to chemists, such solutions must be obtainable at a tolerable computational cost and must be reasonably accurate, yet devising solution methods which meet both of these requirements has proven remarkably difficult. Indeed, although very many variations have been developed over the years, almost every currently existing method can be traced to a prototype introduced within ten years of Schrödinger's seminal paper.

The most uniformly successful family of methods begins with the simplest possible n -electron wavefunction satisfying the Pauli antisymmetry principle – a Slater determinant [2] of one-electron functions $\chi_i(\mathbf{r},\omega)$ called spinorbitals. Each spinorbital is a product of a molecular orbital $\psi_i(\mathbf{r})$ and a spinfunction $\alpha(\omega)$ or $\beta(\omega)$. The $\psi_i(\mathbf{r})$ are found by the self-consistent-field (SCF) procedure introduced [3] into quantum chemistry by Hartree. The Hartree-Fock (HF) [4] and Kohn-Sham density functional (KS) [5,6] theories are both of this type, as are their many simplified variants [7–16].

In SCF methods, the n -electron Schrödinger equation is replaced by a set of n coupled integro-differential one-electron equations. The HF equations are a well-defined approximation to the Schrödinger equation and constitute the starting point for a variety of subsequent treatments [17]: the equations of KS theory are formally equivalent [18] to the original Schrödinger equation for the ground state. In both cases, the equations are highly non-linear and require iterative techniques for their solution.

Commonly, initial guesses for the molecular orbitals are obtained and these are then used to compute the potential felt by an electron in the field of the nuclei and the other electrons. The corresponding one-electron Schrödinger equation is then solved to determine another set of orbitals and the process is continued until successive sets of orbitals differ negligibly, at which point self-consistency is said to have been achieved. The most time-consuming part of this procedure is the evaluation of the *potential* which, within a basis set (Section 2), is represented by various types of integrals (Section 3). Moreover, even if we proceed beyond the HF SCF level, to correlated levels of theory, these integrals remain central to the problem of determining the energy and wavefunction [17].

The result of any quantum chemical procedure is the molecular energy, parametrically determined by the nuclear geometry. To locate equilibrium and transition structures, we usually compute the first derivatives of the energy with respect to nuclear motion [19]; harmonic vibrational frequencies can be obtained if second derivatives are available [20]; third and higher derivatives are needed [21] for higher-level studies of potential surfaces. Not surprisingly, n th-derivatives of the integrals are required to compute n th-derivatives of the energy and the efficient generation of integrals and their n th-derivatives is the focus of this Review.

2. BASIS FUNCTIONS

Because computers can represent numbers, but not functions, the molecular orbitals at each stage of the SCF procedure have to be represented by an expansion in a finite set of basis functions $\phi_i(\mathbf{r})$, $i = 1, 2, \dots N$. If the set is mathematically complete, the result of the SCF procedure is termed the HF or KS limit: otherwise the result is dependent on the basis set used. Many types of basis function have been explored, and several are currently used in routine applications. However, their interrelationships and relative strengths and weaknesses are not often clarified and it may be instructive to do so here.

2.1 Slater Functions

Until the 1960's, Slater basis functions [22]

$$\phi_a^{\text{Slater}}(\mathbf{r}) \equiv (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} \exp[-\alpha|\mathbf{r} - \mathbf{A}|] \quad (1)$$

were very popular. Like exact wavefunctions, they exhibit cusps at the nuclei and decay exponentially but their use necessitates the evaluation of integrals which are very time-consuming to compute. Although several groups have made useful progress in developing efficient algorithms for the evaluation of such integrals, explicit use of Slater basis functions is presently restricted to rather small molecules. It should be noted, however, that any Slater function can be approximated, to any desired accuracy, by a sum of Gaussian functions and the difficult Slater integrals then become relatively easy contracted Gaussian integrals (see below). This is the philosophy of the STO-nG basis sets [23]. In a similar vein, the *product* of a pair of Slater functions can also be approximated to any accuracy by a sum of Gaussians and this approach has been suggested and explored by Harris and Monkhorst [24].

2.2 Gaussian Functions

A *primitive* Gaussian function

$$\phi_a^{\text{PGF}}(\mathbf{r}) \equiv (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} \exp[-\alpha |\mathbf{r} - \mathbf{A}|^2] \quad (2)$$

has *center* $\mathbf{A} = (A_x, A_y, A_z)$, *angular momentum* $\mathbf{a} = (a_x, a_y, a_z)$, and *exponent* α . The suggestion by Boys [25] to use Gaussians as basis functions was a crucial step in the progression of quantum chemistry from a largely qualitative, to a largely quantitative, discipline. The use of a Gaussian basis set in a HF or KS calculation leads to very much simpler integrals (see below) than those which arise within a Slater basis and, although it is known [26] that more Gaussian than Slater functions are required to achieve a given basis set quality, the simplicity of Gaussian integrals more than compensates for this.

A set of primitive basis functions with the same center and exponent are known as a *primitive shell*. For example, a set of *p*-functions $\{p_x, p_y, p_z\}$ on an atom is termed a primitive *p*-shell and, if an *s*-function (with the same exponent) is added, the shell becomes a primitive *sp*-shell. The most commonly occurring shells in modern computational chemistry are *s*, *p*, *sp*, *d* and *f*.

2.3 Contracted Gaussian Functions

It is found that *contracted* Gaussian functions (CGFs) [27]

$$\phi_a^{\text{CGF}}(\mathbf{r}) \equiv \sum_{k=1}^{K_A} D_{ak} (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} \exp[-\alpha_k |\mathbf{r} - \mathbf{A}|^2] \quad (3)$$

where K_A is the *degree of contraction* and the D_{ak} are *contraction coefficients*, are even more computationally effective [26] than Slater functions. It is crucial to note that, although they have different contraction coefficients and exponents, all of the primitive functions in a contracted function share the same center \mathbf{A} and angular momentum \mathbf{a} . A set of CGFs with the same center and the same set of exponents is termed a *contracted shell* by analogy with a primitive shell (defined above).

Over the years, many contracted Gaussian basis sets have been constructed and the interested reader will find the excellent review by Davidson and Feller [26] very illuminating. As a rule, one or two CGFs are used to model each of the core atomic orbitals ($1s$ for lithium to neon; $1s$, $2s$ and $2p$ for sodium to argon; *etc.*) and the CGFs are often highly contracted (a typical K value is 6). Each valence atomic orbital ($1s$ for hydrogen and helium; $2s$ and $2p$ for lithium to neon; *etc.*) is generally more weakly contracted (K less than about 4). Finally, high-quality basis sets contain functions whose angular momenta are higher than that of the valence orbitals (*e.g.* p for hydrogen and helium, d for lithium to argon, *etc.*) and, in most cases, these functions are uncontracted ($K = 1$).

Two distinct classes of contracted Gaussian functions are in common use. In *general* contraction schemes, different contracted functions share the same primitive exponents (with different contraction coefficients) while, in *segmented* schemes, different contracted functions are constructed from primitive functions with different exponents. As a rule, basis functions of the former type tend to have higher degrees of contraction but the higher computational cost implied by this can be partially ameliorated by the use of algorithms which are carefully constructed to take maximum advantage of the exponent sharing. In this Review, we will confine our attention to the efficient treatment of segmented basis sets: we will extend our analysis to the generally contracted case in a future paper.

2.4 Gaussian Lobe Functions

Many of the programming complexities which arise when general contracted Gaussian functions are used disappear if *all* of the functions are constrained to be s -functions, *i.e.*

$$\phi_a^{\text{Lobe}}(\mathbf{r}) \equiv \sum_{k=1}^{K_A} D_{ak} \exp[-\alpha_k |\mathbf{r} - \mathbf{A}|^2] \quad (4)$$

Such basis functions were advocated by a number of authors [28] on the basis of their manifest simplicity and because an array of variously centered s functions can mimic functions of higher angular momentum (p , d , f , *etc.*). However, for obvious reasons, Gaussian Lobe basis sets have to be rather large to yield useful results and become unwieldy in high angular momentum cases. They are rarely used nowadays because of the availability of highly efficient algorithms and programs which can handle CGFs of arbitrary angular momentum.

2.5 Delta Functions

Still more of the programming complexities vanish if another constraint is applied to the Gaussian basis set, namely that its exponents be infinite, which yields a basis composed entirely of Dirac delta functions

$$\phi_a^{\text{Delta}}(\mathbf{r}) \equiv \delta(\mathbf{r} - \mathbf{A}) \quad (5)$$

The set of delta functions at all points in space is mathematically complete and procedures based on these simplest of basis functions have been devised and implemented, first [29–33] for diatomics and, more recently, [34–36] for arbitrary polyatomic systems.

The manifest simplicity of delta functions is both their strength and weakness: computer programs based on them are refreshingly straightforward but, to yield results of chemical significance, delta basis sets must be large, typically thousands of functions per atom. The construction of efficient delta basis sets (*i.e.* 3-dimensional grids) remains an active area of research but, most commonly, they consist of points on concentric spheres about each atom. Most workers use the results of Lebedev [37] who has found optimal quadrature formulae for the surface of a sphere. However, agreement has not yet been reached on which spherical radii are best [38–42].

3. SURVEY OF GAUSSIAN INTEGRAL ALGORITHMS

What are these "integrals" to which we have referred? From the fact that the Schrödinger Hamiltonian contains only one- and two-electron operators, it is straightforward to show [17] that most of the matrix elements [43] which arise in computing the SCF energy and its derivatives with respect to nuclear motion can be written in terms of integrals of the general form

$$(ab|cd) \equiv \iint \phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_1) f(|\mathbf{r}_1 - \mathbf{r}_2|) \phi_c(\mathbf{r}_2)\phi_d(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (6)$$

and their n th-derivatives with respect to displacement of the basis functions. Each of the integrations in (6) is over 3-dimensional space and, thus, the integral is 6-dimensional. The function $f(x)$ is normally very simple – for example, $f(x) \equiv 1/x$ in the familiar case of two-electron repulsion integrals – but it suffices for our present purposes to consider a general function.

The integral (6) is based on two pairs of basis functions, one describing electron 1 and the other describing electron 2. Since there are N functions in the basis set, there are $N(N+1)/2$ distinct basis function pairs and, similarly, there are

$$N_{\text{total}} = \frac{1}{2} \left[\frac{N(N+1)}{2} \right] \left[\frac{N(N+1)}{2} + 1 \right] = \frac{1}{8} N(N+1)(N^2 + N + 2) \quad (7)$$

distinct integrals of the form (6). A *class* of integrals (and/or their n th-derivatives) is defined as the set of all integrals associated with a *shell-quartet*. For example, a $(pplpp)$ class is the set of 81 $(pplpp)$ integrals associated with four p -shells each containing three p -functions. Because all of the integrals in a class share the same four centers and sets of exponents, their generation involves many common intermediate quantities. For this reason, it is always computationally expedient to compute integrals and their derivatives in classes rather than individually.

When large basis sets $N > 10^2$ are used, the generation of the integrals is a major computational task: in fact, in the most common quantum chemistry methods (such as direct SCF, either in the context of HF [44] or KS [45] calculations), it is rate-determining. Obviously, therefore, it is of paramount importance to devise and implement highly efficient generation algorithms. This realization has stimulated the development of a series of integral strategies [46–61] over the last two decades.

It is convenient to divide the evolution of two-electron integral methods into three generations which are distinguished from one another by the general goals which motivated their development. In the 1950's and early 1960's, the target was simply to be able to perform SCF calculations at the most primitive level. The first algorithm for Gaussian functions was outlined by Boys in his classic paper [25] on the use of such functions in SCF calculations and his methodology was subsequently developed and elaborated by a number of workers, including Shavitt [46], Taketa, Huzinaga and O-ohata [47] and Clementi and Davis [48]. Probably the most remarkable achievements of these First Generation methods were Clementi's minimal-basis SCF calculations on pyrrole, pyridine and pyrazine [49] which, while trivial by today's standards, must have used an astonishing amount of computer time in 1967. However, despite opening the door to computational quantum chemistry, these early algorithms and their implementations were both inefficient and slow.

The "axis-switch" method of Pople and Hehre (PH), which was the centerpiece of the Gaussian 70 program [50], revolutionized notions in the early 1970's of the range of chemical systems which could routinely be submitted to a HF calculation. It constitutes a Second Generation method because, unlike its predecessors, its objective was to enable quantum chemical SCF calculations to become a standard tool in the repertoire of practising chemists. To achieve this, the algorithm had to be fully optimized and the implementation carefully designed with contemporary computer architectures in mind. The PH algorithm, which was constructed principally to deal with basis functions of low angular momentum and high degree of contraction, becomes very inefficient under other conditions and this deficiency motivated the development of the Dupuis-Rys-King (DRK) [51] and McMurchie-Davidson (MD) [52] schemes in the mid-1970's. Until the mid-1980's, many computer programs included both PH for *s* and *p* functions and one of DRK and MD for higher functions (*d*, *f*, *etc.*). These Second Generation approaches enabled quantum chemists to perform high-quality SCF calculations on an enormous variety of molecular systems and paved the way for rigorous quantum chemistry to become established as a part of mainstream chemistry.

The success of the Second Generation methods soon spawned a number of new algorithms [53–61], each of which sought to improve upon the ones which had preceded it. The primary goal for these Third Generation methods has been maximum computational efficiency and the main driving force behind their development has been the desire to apply rigorous SCF methods to molecular systems with tens, hundreds, or even thousands, of atoms.

The most recent integral algorithms evolved with, and were influenced by, the advent of supercomputer technologies – if a new method cannot be "vectorized" and/or "parallelized", it faces a cool reception these days – and, of these, the Obara-Saika-Schlegel (OS) [53, 54], Head-Gordon-Pople (HGP) [55] and PRISM [61] algorithms are the most significant.

3.1 Performance Measures

Every integral method exhibits certain strengths and certain weaknesses and there is always a trade-off between the complexity of an algorithm (which is normally directly related to the difficulty of implementing it) and its computational performance. For example, the original Boys algorithm is not very difficult to understand or to code into a computer program but, as we have seen, its practical performance leaves plenty of room for improvement. At the other extreme, while the performance of the PRISM algorithm is much more satisfactory, it is also greatly more complicated conceptually and implementationally. However, since performance is ultimately more important to more people than is the difficulty of writing the underlying computer program, we will take a pragmatic stance and ignore the latter in this Review.

If we wish to compare a variety of integral algorithms, it is certainly highly desirable to be able to *quantify* their relative performances. Hitherto, three measures of performance have been proposed and used in the literature and, although related, they are distinct and it is instructive to describe each briefly at this point.

3.1.1 Flop-Cost

The ubiquitous measure of the *theoretical* performance of an integral algorithm is its Flop-cost. More precisely, this is the number of Flops which are required to form a specified class of integrals from a defined set of starting quantities and a Flop (Floating-Point Operation) is defined to be a floating-point add, subtract, multiply or divide. For example, (using values quoted by Head-Gordon and Pople [55]), the Flop-cost of computing a (*pplpp*) class, *i.e.* 81 (*pplpp*) integrals, where each of the *p*-functions is a sum of two primitive functions (*i.e.* $K = 2$), is 20,000 using MD but only 15,170 using HGP.

It should be re-emphasized that a Flop-cost is a *theoretical* measure and that comparisons like the one above are practically useful only if the two algorithms have been implemented equally well: a well-written MD program would undoubtedly generate a (*pplpp*) class much more rapidly than a poorly written HGP program. Thus, establishing that a given algorithm has a small Flop-cost for a certain integral class tells us only that the *potential* exists for an implementation of that algorithm to perform well. It is prudent to view with considerable skepticism claims which are based exclusively on a purely theoretical measure like a Flop-count.

3.1.2 Mop-Cost

A newer measure of an algorithm's *theoretical* performance is its Mop-Cost which is defined exactly as the Flop-cost except that Memory Operations (Mops) are counted instead of Floating-Point Operations (Flops). A Mop is a load from, or a store to, fast memory. There are sound theoretical reasons why Mops should be a better indicator of practical performance than Flops, especially on recent computers employing vector or RISC architectures, and this has been discussed in detail by Frisch *et al.* [62]: to cut a long story short, the Mops measure is useful because, on modern computers and in contrast to older ones, memory traffic generally presents a tighter bottleneck than floating-point arithmetic.

The Mop-cost of forming an integral class depends on the intrinsic expense of the algorithm *and* on implementational detail. Consider, for example, the following two versions of a subroutine:

```
Version #1      DO 10 I = 1,N
                  A(I) = B(I) + C(I) * D(I)
10 CONTINUE
                  DO 20 I = 1,N
                    E(I) = A(I) + F(I)
20 CONTINUE
```

```
Version #2      DO 10 I = 1,N
                  A(I) = B(I) + C(I) * D(I)
                  E(I) = A(I) + F(I)
10 CONTINUE
```


The two versions achieve the same results but the second will run faster on many computers. Both versions have Flop-costs of $3N$ but their Mop-costs are different, $7N$ and $6N$, respectively. This is because, in the second version, $A(I)$ does not need to be loaded from fast memory since, having just been produced, it will already reside in a register. This example is certainly a very simplistic one but it serves to illustrate the principle of Mop reduction.

Mop comparisons between integral programs have not yet begun to appear in the literature but these may be anticipated in the near future. As a design rule for the future, an algorithm which minimizes Mops is a better target than one which minimizes Flops. A particularly striking example of the usefulness of this approach is the recent discovery by Johnson *et al.* that a certain family of novel recurrence relations, which appear uncompetitive with older recurrence relations on the basis of Flop-cost, are exceedingly competitive in Mop-cost and in actual timings [63].

3.1.3 CPU-Time

The most popular measure of the *practical* performance of an algorithm is the amount of CPU time which a specified computer requires to complete a specified task, commonly a single iteration of a specified SCF procedure. While this is certainly a very appealing measure (it directly reflects an expense which is uppermost in the minds of most computational chemists), it will be extremely difficult for another worker to reproduce the timing unless the specifications are very complete.

It is important not only to specify precisely the computer whose CPU time has been measured, but also to record the compiler version, any non-default compiler options that were used and the operating system. Specification of the task which was performed is even more demanding. To write, for example, that the timing pertains to one cycle of a HF calculation is far from complete. It is necessary also to state clearly exactly what was timed: Was it the integral generation only? Was Fock matrix formation included? Was the diagonalization included? And, even more important, which integrals were actually computed? As we shall shortly see, most modern integral programs use sophisticated "screening" techniques to avoid computing very small integrals. Unless a clear description of the screening procedure used is given, including the cutoff threshold, any quoted CPU-time is obviously meaningless. However, provided that all of the various requirements above are met, a CPU-time measurement can be a very useful indication of the performance of an integral algorithm.

The optimal strategy to adopt when comparing algorithms is to employ as many measures as possible in the comparison. Previous reviews of integral algorithms [64–67] have been helpful in this regard, that of Hegarty and van der Velde [66] being particularly noteworthy.

3.2 Fundamental Integrals

One of the problems which plagues the two-electron integral literature is that of notation and we certainly hope that this Review does not add to the confusion. We will adopt a notation system which appears (slowly!) to be becoming the standard and will introduce it as the need arises. We will use **A**, **B**, **C** and **D** to represent the position vectors of the centers of the four basis functions in (6) and will use α , β , γ and δ to represent exponents of generic primitives within these functions.

All discussion of two-electron integrals ultimately begins by considering the special case of the integral (6) in which each of the four functions is a primitive *s*-Gaussian, *i.e.*

$$I \equiv \iint e^{-\alpha|\mathbf{r}_1-\mathbf{A}|^2} e^{-\beta|\mathbf{r}_1-\mathbf{B}|^2} f(|\mathbf{r}_1-\mathbf{r}_2|) e^{-\gamma|\mathbf{r}_2-\mathbf{C}|^2} e^{-\delta|\mathbf{r}_2-\mathbf{D}|^2} d\mathbf{r}_1 d\mathbf{r}_2 \quad (8)$$

and we will refer to this as the *Fundamental Integral*. The first step in its evaluation is to invoke the Gaussian Product Rule ([17], p. 411) which immediately reduces the integral to

$$I = G_{AB} G_{CD} \iint e^{-\zeta|\mathbf{r}_1-\mathbf{P}|^2} f(|\mathbf{r}_1-\mathbf{r}_2|) e^{-\eta|\mathbf{r}_2-\mathbf{Q}|^2} d\mathbf{r}_1 d\mathbf{r}_2 \quad (9)$$

where

$$G_{AB} = \exp\left[\frac{-\alpha\beta}{\alpha+\beta}|\mathbf{A}-\mathbf{B}|^2\right] \quad G_{CD} = \exp\left[\frac{-\gamma\delta}{\gamma+\delta}|\mathbf{C}-\mathbf{D}|^2\right] \quad (10)$$

$$\zeta = \alpha + \beta \quad \eta = \gamma + \delta \quad (11)$$

$$\mathbf{P} = \frac{\alpha\mathbf{A} + \beta\mathbf{B}}{\alpha + \beta} \quad \mathbf{Q} = \frac{\gamma\mathbf{C} + \delta\mathbf{D}}{\gamma + \delta} \quad (12)$$

The reduction from (8) to (9) is a crucial simplification because it transforms a four-center (**A**, **B**, **C**, **D**) problem into a two-center (**P**, **Q**) one. Most of the difficulties associated with the use of Slater functions can be traced to the fact that a corresponding reduction is not possible for these functions.

Next, we replace each of the three factors in the integrand of (9) by its Fourier representation

$$e^{-\zeta|\mathbf{r}_1-\mathbf{P}|^2} = (2\pi)^{-3} \int \left(\frac{\pi}{\zeta}\right)^{3/2} e^{-\mathbf{k}_1^2/4\zeta} e^{i\mathbf{k}_1\cdot(\mathbf{r}_1-\mathbf{P})} d\mathbf{k}_1 \quad (13)$$

$$f(|\mathbf{r}_1-\mathbf{r}_2|) = (2\pi)^{-3} \int \mathfrak{I}(\mathbf{k}_2) e^{i\mathbf{k}_2\cdot(\mathbf{r}_1-\mathbf{r}_2)} d\mathbf{k}_2 \quad (14)$$

$$e^{-\eta|\mathbf{r}_2-\mathbf{Q}|^2} = (2\pi)^{-3} \int \left(\frac{\pi}{\eta}\right)^{3/2} e^{-\mathbf{k}_3^2/4\eta} e^{i\mathbf{k}_3\cdot(\mathbf{r}_2-\mathbf{Q})} d\mathbf{k}_3 \quad (15)$$

Substituting (13) – (15) into (9) and re-ordering the integrations yields

$$\begin{aligned} I = & (2\pi)^{-3} G_{AB} G_{CD} \left(\frac{\pi^2}{\zeta\eta}\right)^{3/2} \iiint e^{-\mathbf{k}_1^2/4\zeta - \mathbf{k}_3^2/4\eta} e^{-i\mathbf{k}_1\cdot\mathbf{P} - i\mathbf{k}_3\cdot\mathbf{Q}} \mathfrak{I}(\mathbf{k}_2) \\ & (2\pi)^{-3} \int e^{i\mathbf{r}_1\cdot(\mathbf{k}_1+\mathbf{k}_2)} d\mathbf{r}_1 (2\pi)^{-3} \int e^{i\mathbf{r}_2\cdot(\mathbf{k}_3-\mathbf{k}_2)} d\mathbf{r}_2 d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3 \end{aligned} \quad (16)$$

The fourth and fifth integrals in (16) are Fourier representations of the three-dimensional Dirac delta function, whence

$$\begin{aligned} I = & \frac{G_{AB} G_{CD}}{8(\zeta\eta)^{3/2}} \\ & \iiint e^{-\mathbf{k}_1^2/4\zeta - \mathbf{k}_3^2/4\eta - i\mathbf{k}_1\cdot\mathbf{P} - i\mathbf{k}_3\cdot\mathbf{Q}} \mathfrak{I}(\mathbf{k}_2) \delta(\mathbf{k}_1 + \mathbf{k}_2) \delta(\mathbf{k}_3 - \mathbf{k}_2) d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3 \end{aligned} \quad (17)$$

By virtue of the "sampling" property of the delta function, the triple integral in (17) collapses to a single integral, yielding

$$I = \frac{\pi G_{AB} G_{CD}}{2(\zeta\eta)^{3/2} R^3} \int_0^\infty u \sin u e^{-u^2/4T} \mathfrak{Z}(u/R) du \quad (18)$$

where

$$\vartheta^2 = \frac{\zeta\eta}{\zeta + \eta} \quad (19)$$

$$\mathbf{R} = \mathbf{Q} - \mathbf{P} \quad (20)$$

$$T = \vartheta^2 R^2 \quad (21)$$

Our evaluation of the Fundamental Integral cannot proceed further than (18) unless we now specify the two-electron function $f(\mathbf{x})$. We are now in a position to consider some of the integral types which arise in quantum chemical calculations: overlap, kinetic-energy, electron-repulsion, nuclear-attraction and anti-coulomb.

3.2.1 The Overlap Integral

The Fundamental Overlap Integral is given by

$$I^{\text{overlap}} \equiv \int e^{-\alpha|\mathbf{r}-\mathbf{A}|^2} e^{-\gamma|\mathbf{r}-\mathbf{C}|^2} d\mathbf{r} \quad (22)$$

and the first task which arises is to write this in the form (8). To achieve this, we choose

$$\beta = \delta = 0 \quad (23)$$

$$f^{\text{overlap}}(\mathbf{r}) \equiv \delta(\mathbf{r}) \quad (24)$$

in order to convert a four-center problem into a two-center one and to convert a two-electron integral into a one-electron one. This simple trick enables us to treat overlap integrals on an equal footing with any other integrals which we may be able to cast in the form (8).

The Fourier transform of (24) is easily shown to be

$$\mathfrak{S}^{\text{overlap}}(\mathbf{k}) = 1 \quad (25)$$

and the expression which is obtained if (25) is substituted into (18) is readily integrable by parts, finally yielding

$$I^{\text{overlap}} = \left[\frac{\pi}{\alpha + \gamma} \right]^{3/2} e^{-\alpha\gamma|\mathbf{A}-\mathbf{C}|^2/(\alpha+\gamma)} \quad (26)$$

which is the familiar formula for the overlap of two *s*-Gaussians.

3.2.2 The Kinetic-Energy Integral

The Fundamental Kinetic-Energy Integral is given by

$$I^{\text{kinetic}} \equiv \int e^{-\alpha|\mathbf{r}-\mathbf{A}|^2} \left[-\frac{1}{2} \nabla^2 \right] e^{-\gamma|\mathbf{r}-\mathbf{C}|^2} d\mathbf{r} \quad (27)$$

where the Laplacian denotes differentiation with respect to \mathbf{r} . However, it is clearly equivalent to differentiate with respect to \mathbf{C} , from which it immediately follows that

$$I^{\text{kinetic}} = -\frac{1}{2} \left[\frac{\partial^2}{\partial C_x^2} + \frac{\partial^2}{\partial C_y^2} + \frac{\partial^2}{\partial C_z^2} \right] \int e^{-\alpha|\mathbf{r}-\mathbf{A}|^2} e^{-\gamma|\mathbf{r}-\mathbf{C}|^2} d\mathbf{r} \quad (28)$$

Thus, the Fundamental Kinetic-Energy Integral can be obtained from the second derivatives of the Fundamental Overlap Integral with respect to motion of the center \mathbf{C} . Since we are interested in algorithms which offer *n*th-derivatives of two-electron integrals, the problem of generating kinetic-energy integrals and their *n*th-derivatives is subsumed into the problem of generating overlap integrals and their *n*th-derivatives.

3.2.3 The Electron-Repulsion Integral

The Fundamental Electron-Repulsion Integral is given by

$$I^{EE} \equiv \iint e^{-\alpha|\mathbf{r}_1-\mathbf{A}|^2} e^{-\beta|\mathbf{r}_1-\mathbf{B}|^2} \frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} e^{-\gamma|\mathbf{r}_2-\mathbf{C}|^2} e^{-\delta|\mathbf{r}_2-\mathbf{D}|^2} d\mathbf{r}_1 d\mathbf{r}_2 \quad (29)$$

which is obviously already of the form (8) with

$$f^{EE}(\mathbf{r}) \equiv r^{-1} \quad (30)$$

The Fourier transform of (30) is

$$\mathfrak{F}^{EE}(\mathbf{k}) = 4\pi k^{-2} \quad (31)$$

and, upon substituting (31) into (18), we obtain

$$I^{EE} = G_{AB}G_{CD} \frac{2\pi^2}{(\zeta\eta)^{3/2}R} \int_0^\infty \frac{\sin u}{u} e^{-u^2/4T} du \quad (32)$$

which is related to the error function and is often re-written

$$I^{EE} = G_{AB}G_{CD} \frac{2\pi^{5/2}}{\zeta\eta(\zeta+\eta)^{1/2}} F_0^{EE}(T) \quad (33)$$

where

$$F_0^{EE}(T) = \int_0^1 e^{-Tu^2} du = 1 - T/3 + T^2/10 - \dots \quad (34)$$

The efficient computation of the function (34) has been discussed in a number of papers [46, 50–52, 54, 55, 61, 68–71] and it is generally agreed that a carefully constructed interpolation scheme, such as that described in [71], is the most effective approach.

3.2.4 The Nuclear-Attraction Integral

The Fundamental Nuclear-Attraction Integral is given by

$$I^{\text{NE}} \equiv \int e^{-\alpha|\mathbf{r}-\mathbf{A}|^2} e^{-\beta|\mathbf{r}-\mathbf{B}|^2} \frac{1}{|\mathbf{r}-\mathbf{C}|} d\mathbf{r} \quad (35)$$

and, as with the Fundamental Overlap Integral, we must begin by casting this in the form (8). This can be accomplished by taking

$$\gamma = \infty \quad (36)$$

$$\delta = 0 \quad (37)$$

$$f^{\text{NE}}(r) \equiv r^{-1} \quad (38)$$

which replaces the nuclear center by a Gaussian with an extremely large exponent and thereby transforms the problem into that of the Fundamental Electron-Repulsion Integral.

3.2.5 The Anti-Coulomb Integral

Recently, we have developed a straightforward least-squares method [72] for modeling the potential of a charge distribution using a second, simpler, distribution and Kutzelnigg and coworkers have developed a promising method [73] to enhance the rate of convergence of CI-type expansions. Curiously, both of these methods are ultimately based on the Fundamental Anti-Coulomb Integral

$$I^{\text{AC}} \equiv \iint e^{-\alpha|\mathbf{r}_1-\mathbf{A}|^2} e^{-\beta|\mathbf{r}_1-\mathbf{B}|^2} |\mathbf{r}_1-\mathbf{r}_2| e^{-\gamma|\mathbf{r}_2-\mathbf{C}|^2} e^{-\delta|\mathbf{r}_2-\mathbf{D}|^2} d\mathbf{r}_1 d\mathbf{r}_2 \quad (39)$$

We note that the only feature which distinguishes (39) from (29) is the two-electron function

$$f^{\text{AC}}(r) \equiv r^{+1} \quad (40)$$

and the evaluation of the Fundamental Anti-Coulomb Integral is similar [74–76] to that of the Fundamental Electron-Repulsion Integral.

3.3 The Boys Algorithm [25]

Boys noted in his original paper that, if one differentiates the Fundamental Integral with respect to one of the coordinates of one of the centers, one obtains a primitive integral over a p and three s functions. Moreover, further differentiations lead to primitive integrals over still higher angular momentum functions. Based on this observation, Boys proposed that formulae for general primitive integrals be obtained by the repeated differentiation of the formula for the Fundamental Integral.

In the event that we wish to compute an integral (6) in which one or more of the four Gaussian basis functions is *contracted*, the Boys algorithm expresses the contracted integral as a sum of primitive integrals and then computes each of the latter using the formulae described in the foregoing paragraph.

The Boys algorithm is pedagogically useful and also serves to emphasize the highly important connection between the derivatives of Gaussian integrals and Gaussian integrals over higher angular momentum functions. However, its practical usefulness is limited by three important considerations:

- (a) It rapidly becomes exceptionally tedious (and error-prone) to generate formulae by repeated differentiation of expressions such as (26) and (33);
- (b) The resulting formulae are inefficient because they fail to make use of the fact that integrals in the same class share many common intermediates;
- (c) The resulting formulae are inefficient because they fail to make use of the fact that the primitives in a contracted basis function all share the same center. The task of making the best possible use of this is called the Contraction Problem.

More recent integral algorithms have sought to ameliorate, to a greater or lesser degree, each of these three deficiencies. As we will see, all of these algorithms employ recurrence relations (RR's) to express integrals of high angular momentum in terms of integrals of lower angular momentum. This tactic automatically improves consideration (a) and (b) above. Addressing (c) is more difficult but the rewards are very significant: the greatest improvements in computational efficiency over the last decade have almost all resulted from new solutions to the Contraction Problem.

3.4 The Contraction Problem

Suppose that we wish to form a class of contracted integrals and that each of the basis functions is K -fold contracted, *i.e.* is a sum of K primitive functions. Then, in a straightforward method (such as the Boys algorithm), each contracted integral **(abcd)** is expressed as a sum of its component primitive integrals **[abcd]** which, in turn, are computed individually, *i.e.*

$$(\mathbf{abcd}) = \sum_{i=1}^K \sum_{j=1}^K \sum_{k=1}^K \sum_{l=1}^K D_{ai} D_{bj} D_{ck} D_{dl} [\mathbf{a}_i \mathbf{b}_j | \mathbf{c}_k \mathbf{d}_l] \quad (41)$$

It is clear from (41) that the computational effort to construct the desired class of **(abcd)** integrals will rise with the fourth power of K . In fact, it is easily shown [66] that the total Flop-cost of forming the class can always be expressed as

$$\text{Flop-cost} = xK^4 + yK^2 + z \quad (42)$$

Of course, an analogous expression for the Mop-cost also exists.

Table I: Flop-cost parameters for generating integral classes ^a				
Class	Parameter	PH	MD	HGP
<i>(pp pp)</i>	x	220	1,100	920
	y	2,300	600	30
	z	4,000	0	330
<i>(sp,sp sp,sp)</i>	x	220	1,500	1,400
	y	2,300	1,700	30
	z	4,000	0	800
<i>(dd dd)</i>	x	—	27,300	14,600
	y	—	24,000	30
	z	—	0	11,300
<i>(ff ff)</i>	x	—	342,000	108,000
	y	—	383,000	30
	z	—	0	135,000

(a) Taken from [55].

The x , y and z parameters for various algorithms and various integral classes have been tabulated in a number of papers [55, 57, 59, 61] and are valuable in rationalizing the observed performances of different integral methods. In Table I, for example, we list Flop-cost parameters for the PH, MD and HGP methods to generate various integral classes. Clearly, each method is characterized by unique sets of parameters and these yield important information about the method's theoretical performance behavior. For example, the remarkably small x parameter and remarkably large z parameter for PH argue that it should be a very powerful method for generating highly contracted ($pp|pp$) integrals but a wasteful one for forming uncontracted ones. Similar qualitative analyses for other methods are summarized in Table II.

Table II: Algorithmic costs as a function of degree of contraction					
	PH	HGP	OS	MD	DRK
Small K	High	Low	Moderate	Moderate	Moderate
Large K	Low	Moderate	High	High	High
Contraction ^a	Early	Midway	Late	Late	Late

(a) At what point in the algorithm, the contraction step occurs.

It is clear from Table II that none of the five algorithms included is the universal panacea for all integral problems. The best single method is HGP but, since typical SCF calculations on large molecules involve highly contracted, mildly contracted and weakly contracted integrals (see Section 2), a program which seeks to be near-optimal under all circumstances has to switch from one integral algorithm to another, basing its decision upon the type of integral under consideration at any moment. The Gaussian 82 [77] and Gaussian 86 [78] programs adopted this hybrid approach, employing a PH routine (Link 311) for all integrals involving only s and p functions and a DRK routine (Link 314) for any others. However, such strategies are not only arbitrary and artificial but also render the program complicated and difficult to improve.

The third row of Table II reveals that there is a very simple correlation between the performance behaviour of an algorithm and the point at which the primitive integrals are added together into contracted integrals: early-contraction methods are best suited to highly contracted integrals; late-contraction methods are best suited to weakly contracted integrals. This observation underlies the PRISM algorithm which we will discuss shortly.

3.5 The Pople-Hehre Algorithm [50]

We have seen that PH is exceptionally efficient for highly contracted classes, exhibiting x parameters (for classes involving only s and p functions) which are very much smaller than those of other, more recent, methods. How is this efficiency achieved?

Pople and Hehre showed that, given the position vectors **A**, **B**, **C** and **D** and two exponents γ and δ , there exists a unique Cartesian axis system [79] in which many primitive integrals vanish by symmetry. Moreover, because this axis system is independent of the exponents α and β , it can be used for all α, β pairs. After looping over these, the accumulated integral combinations are rotated into a second Cartesian system [80] which depends only on **A**, **B**, **C** and **D**, and the next γ, δ pair is then selected. When all γ, δ pairs have been treated, the desired integrals are finally obtained by rotating back to the original Cartesian system of the molecule.

The x parameter in (42) measures the work required to form and manipulate exclusively primitive quantities. Thus, of the steps in the PH method, only the computation and accumulation of the non-vanishing primitive integrals contributes to x and, since the unique axis system described in the foregoing paragraph was carefully designed to minimize the number of such integrals, x is correspondingly small.

The y and z parameters in (42) measure the computational effort to rotate from the first cartesian frame to the second, and from the second to the third, respectively, and to accumulate them therein. These are relatively substantial tasks and this explains the large y and z parameters in Table I for PH.

Only if the basis functions are sufficiently contracted (*i.e.* K is large enough), does the work saved by the use of special axis systems outweigh the effort which must be expended to perform the two rotations and it is interesting to determine the value of K at which PH becomes cheaper than HGP. For a $(pp|pp)$ class, using the parameters in table I, one finds that PH is competitive with HGP even when K is as small as 2.

Notwithstanding the impressive performance of PH on integral classes involving contracted basis functions with low angular momentum, it founders when applied to uncontracted classes with high angular momentum, for example $[dd|dd]$, because of the huge costs incurred in the two rotation steps [81]. For such classes, new techniques had to be developed and we will discuss some of these in the next few sections of this Review.

3.6 Bras, Kets and Brackets [61]

Before proceeding further, it is useful to introduce a simple but powerful notation for integrals and their n th-derivatives. All of the modern integral algorithms can be easily represented within this notation and, of course, uniform descriptions greatly facilitate any algorithmic comparisons which we may make.

In (6) we defined a general two-electron integral over the two-electron operator f . However, it has long been realized that such an equation defines an *inner product between two functions*

$$(\mathbf{ab}| \equiv \phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_1) \quad (43)$$

$$|\mathbf{cd}) \equiv \phi_c(\mathbf{r}_2)\phi_d(\mathbf{r}_2) \quad (44)$$

Thus, taking inspiration from the notation introduced by Dirac, we will refer to $(\mathbf{ab}|$ and $|\mathbf{cd})$ as a "bra" and "ket", respectively, and to (\mathbf{abcd}) as a "braket".

For the purposes of defining bras and kets, it is useful to generalize (43) and (44) substantially. Though the resulting definitions may appear complicated, it should be borne in mind that, like (43) and (44), bras and kets are simply functions of the positions of electrons 1 and 2, respectively.

We define a *primitive bra*

$$\begin{bmatrix} \bar{\mathbf{a}} & \bar{\mathbf{b}} \\ \mathbf{a} & \mathbf{b} & \mathbf{p} \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' \end{bmatrix} \equiv \frac{(2\alpha)^{a'}(2\beta)^{b'}}{(2\zeta)^{p'}} \cdot \frac{\partial^{\bar{\mathbf{a}}+\bar{\mathbf{b}}}}{\partial A_x^{\bar{a}_x} \partial A_y^{\bar{a}_y} \partial A_z^{\bar{a}_z} \partial B_x^{\bar{b}_x} \partial B_y^{\bar{b}_y} \partial B_z^{\bar{b}_z}} \begin{bmatrix} 0 & 0 \\ \mathbf{a} & \mathbf{b} & \mathbf{p} \\ 0 & 0 & 0 \end{bmatrix} \quad \dots (45)$$

where

$$\begin{bmatrix} 0 & 0 \\ \mathbf{a} & \mathbf{b} & \mathbf{p} \\ 0 & 0 & 0 \end{bmatrix} \equiv D_A D_B e^{-\alpha(\mathbf{r}-\mathbf{A})^2} e^{-\beta(\mathbf{r}-\mathbf{B})^2} \prod_{i=x,y,z} (i-A_i)^{a_i} (i-B_i)^{b_i} \zeta^{p_i/2} H_{p_i}[\zeta^{1/2}(i-P_i)] \quad \dots (46)$$

and H_n is the n th Hermite polynomial. The other symbols in (45) and (46) have been defined earlier in this Review. The Hermite polynomials are defined by $H_0(x) \equiv 1$, $H_1(x) \equiv 2x$ and

$$H_{n+1}(x) \equiv 2x H_n(x) - 2n H_{n-1}(x) \quad (47)$$

and possess the useful property that $dH_n(x)/dx = 2n H_{n-1}(x)$.

Certain special cases of (45) arise sufficiently often that it is useful to introduce more concise notations for them. In particular, a Hermite function on center \mathbf{P} is denoted by

$$[p] \equiv \begin{bmatrix} 0 & 0 & \\ 0 & 0 & p \\ 0 & 0 & 0 \end{bmatrix} \quad (48)$$

a product of Cartesian Gaussian primitives on \mathbf{A} and \mathbf{B} is

$$[ab] \equiv \begin{bmatrix} 0 & 0 & \\ \mathbf{a} & \mathbf{b} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (49)$$

and the product of Cartesian Gaussian primitives on \mathbf{A} and \mathbf{B} and a Hermite function on \mathbf{P} is

$$[abp] \equiv \begin{bmatrix} 0 & 0 & \\ \mathbf{a} & \mathbf{b} & p \\ 0 & 0 & 0 \end{bmatrix} \quad (50)$$

A bra in which $\bar{\mathbf{a}} = \bar{\mathbf{b}} = \mathbf{a} = \mathbf{b} = \mathbf{0}$ will be termed a *p-bra*.

Having defined a primitive bra, we define a *contracted bra* by

$$\begin{pmatrix} \bar{a} & \bar{b} \\ a & b & p \\ a' & b' & p' \end{pmatrix} \equiv \sum_{k_A=1}^{K_A} \sum_{k_B=1}^{K_B} \begin{pmatrix} \bar{a} & \bar{b} \\ a & b & p \\ a' & b' & p' \end{pmatrix} \quad (51)$$

and its various special cases by

$${}_{a'b'p'}(p| \equiv \begin{pmatrix} 0 & 0 \\ 0 & 0 & p \\ a' & b' & p' \end{pmatrix} \quad (52)$$

$$(ab| \equiv \begin{pmatrix} 0 & 0 \\ a & b & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (53)$$

$${}_{a'b'p'}(abp| \equiv \begin{pmatrix} 0 & 0 \\ a & b & p \\ a' & b' & p' \end{pmatrix} \quad (54)$$

The definitions of *primitive and contracted kets* are entirely analogous to (45)–(54). As previously discussed, a bracket is then an inner product between a bra and a ket

$$\begin{pmatrix} \bar{a} & \bar{b} \\ a & b & p \\ a' & b' & p' \end{pmatrix} \begin{pmatrix} \bar{c} & \bar{d} \\ c & d & q \\ c' & d' & q' \end{pmatrix} \equiv \iiint \begin{pmatrix} \bar{a} & \bar{b} \\ a & b & p \\ a' & b' & p' \end{pmatrix} f(|\mathbf{r}_1 - \mathbf{r}_2|) \begin{pmatrix} \bar{c} & \bar{d} \\ c & d & q \\ c' & d' & q' \end{pmatrix} d\mathbf{r}_1 d\mathbf{r}_2 \quad \dots (55)$$

and a *pq-bracket*, which results from a p-bra and a q-ket, is an important special case. It should be noted that the symbols which we use to represent bras, kets and brackets, although "matrix-like" in appearance, have no connection whatever to matrices. They are, we emphasize, nothing more than compact notations for very general one-electron functions and their inner products.

3.7 The McMurchie-Davidson Algorithm [52]

Whereas the principal concern expressed in the paper by Pople and Hehre [50] was extremely high efficiency for a limited set of integral classes, the main emphasis of McMurchie and Davidson was generality and extendability. Not surprisingly, then, their respective algorithms are, to a large degree, complementary.

The target $(\mathbf{ab}|\mathbf{cd})$ are made from bra-contracted $(\mathbf{ab}|\mathbf{cd})$ using

$$(\mathbf{ab}|\mathbf{cd}) = \sum_{k=1}^{K_C} \sum_{l=1}^{K_D} (\mathbf{ab}|\mathbf{c}_k\mathbf{d}_l) \quad (56)$$

MD used elementary properties of Hermite polynomials to derive a three-term RR

$$[(\mathbf{c} + \mathbf{1}_i)\mathbf{d}|\mathbf{q}] = q_i |(\mathbf{c}(\mathbf{q} - \mathbf{1}_i)| + (Q_i - C_i)|\mathbf{c}|\mathbf{q}] + (2\eta)^{-1} |(\mathbf{c}(\mathbf{q} + \mathbf{1}_i)| \quad (57)$$

by which the $(\mathbf{ab}|\mathbf{cd})$ can be formed from $(\mathbf{ab}|\mathbf{q})$. (Note that the subscript i represents a Cartesian direction (x , y or z) and $\mathbf{1}_i$ is the unit 3-vector in the i th direction.) Since it has no effect on the bras, we describe this step as a "ket-transformation".

The $(\mathbf{ab}|\mathbf{q})$ are formed from $[\mathbf{ab}|\mathbf{q}]$ using

$$(\mathbf{ab}|\mathbf{q}) = \sum_{k=1}^{K_A} \sum_{l=1}^{K_B} [\mathbf{a}_k\mathbf{b}_l|\mathbf{q}] \quad (58)$$

and, in the bra-transformation step, the $[\mathbf{ab}|\mathbf{q}]$ are formed from $[\mathbf{p}|\mathbf{q}]$ using the bra version

$$[(\mathbf{a} + \mathbf{1}_i)\mathbf{b}|\mathbf{p}] = p_i [\mathbf{a}(\mathbf{p} - \mathbf{1}_i)| + (P_i - A_i)[\mathbf{a}|\mathbf{b}|\mathbf{p}] + (2\zeta)^{-1} [\mathbf{a}(\mathbf{p} + \mathbf{1}_i)| \quad (59)$$

of the RR (57).

To evaluate the $[\mathbf{p}|\mathbf{q}]$, one can follow the procedure discussed in Section 3.2. This turns out to be straightforward because the Fourier Transform of a \mathbf{p} -bra is just a Cartesian Gaussian function and it is not difficult to show that

$$[\mathbf{p}|\mathbf{q}] = (-1)^q [\mathbf{p} + \mathbf{q}] \quad (60)$$

where, for repulsion integrals,

$$[\mathbf{r}] = D_A D_B D_C D_D G_{AB} G_{CD} \frac{2\pi^{5/2}}{\zeta \eta (\zeta + \eta)^{1/2}} \int_0^1 (\vartheta \mathbf{u})^T H_{r_x}(R_x \vartheta \mathbf{u}) H_{r_y}(R_y \vartheta \mathbf{u}) H_{r_z}(R_z \vartheta \mathbf{u}) e^{-T \mathbf{u}^2} d\mathbf{u} \quad (61)$$

which is a generalization of (33). By invoking the elementary RR for differentiation of Hermite polynomials, MD showed that $[\mathbf{r}] \equiv [\mathbf{r}]^{(0)}$ integrals can be generated from $[\mathbf{0}]^{(m)}$ integrals using the two-term RR

$$[\mathbf{r}]^{(m)} = R_i [\mathbf{r} - \mathbf{1}_i]^{(m+1)} - (r_i - 1) [\mathbf{r} - \mathbf{2}_i]^{(m+1)} \quad (62)$$

where $\mathbf{2}_i$ denotes twice the unit vector in the i th direction,

$$[\mathbf{0}]^{(m)} = D_A D_B D_C D_D G_{AB} G_{CD} \frac{2^{1/2} \pi^{5/2}}{(\zeta \eta)^{3/2}} (2\vartheta^2)^{m+1/2} F_m(T) \quad (63)$$

and

$$F_m(T) = \int_0^1 \mathbf{u}^{2m} e^{-T \mathbf{u}^2} d\mathbf{u} \quad (64)$$

MD advocated the use of seven-term Taylor interpolation to evaluate (64) for the largest value of m , followed by downward recursion using

$$F_m(T) = \frac{1}{2m+1} \left[2T F_{m+1}(T) + e^{-T} \right] \quad (65)$$

to compute (64) for smaller m values.

3.8 The Obara-Saika-Schlegel Algorithm [53, 54]

Four years before the appearance of the first paper by Obara and Saika, Schlegel published an important article concerning the rapid computation of *first derivatives* of two-electron integrals with respect to nuclear motion. However, although his method was in widespread use for some years, it was apparently not recognized that it is equivalent to a completely new algorithm for computing integrals themselves: only after Obara and Saika independently discovered the same algorithm did the penny drop.

The target $\langle \mathbf{ab}|\mathbf{cd} \rangle$ are generated from $[\mathbf{ab}|\mathbf{cd}]$ according to

$$\langle \mathbf{ab}|\mathbf{cd} \rangle = \sum_{i=1}^{K_A} \sum_{j=1}^{K_B} \sum_{k=1}^{K_C} \sum_{l=1}^{K_D} [\mathbf{a}_i \mathbf{b}_j | \mathbf{c}_k \mathbf{d}_l] \quad (66)$$

and $[\mathbf{ab}|\mathbf{cd}] \equiv \langle \mathbf{ab}|\mathbf{cd} \rangle^{(0)}$ are formed from $\langle 00|00 \rangle^{(m)}$ using the eight-term OS RR

$$\begin{aligned} \langle (\mathbf{a} + \mathbf{1}_i) \mathbf{b} | \mathbf{cd} \rangle^{(m)} &= (P_i - A_i) \langle \mathbf{ab} | \mathbf{cd} \rangle^{(m)} + \frac{\eta}{\zeta + \eta} R_i \langle \mathbf{ab} | \mathbf{cd} \rangle^{(m+1)} \\ &+ \frac{a_i}{2\zeta} \left[\langle (\mathbf{a} - \mathbf{1}_i) \mathbf{b} | \mathbf{cd} \rangle^{(m)} - \frac{\eta}{\zeta + \eta} \langle (\mathbf{a} - \mathbf{1}_i) \mathbf{b} | \mathbf{cd} \rangle^{(m+1)} \right] \\ &+ \frac{b_i}{2\zeta} \left[\langle \mathbf{a} (\mathbf{b} - \mathbf{1}_i) | \mathbf{cd} \rangle^{(m)} - \frac{\eta}{\zeta + \eta} \langle \mathbf{a} (\mathbf{b} - \mathbf{1}_i) | \mathbf{cd} \rangle^{(m+1)} \right] \\ &+ \frac{c_i}{2(\zeta + \eta)} \langle \mathbf{ab} | (\mathbf{c} - \mathbf{1}_i) \rangle^{(m+1)} + \frac{d_i}{2(\zeta + \eta)} \langle \mathbf{ab} | \mathbf{c} (\mathbf{d} - \mathbf{1}_i) \rangle^{(m+1)} \end{aligned} \quad (67)$$

and its analogues which increment \mathbf{b} , \mathbf{c} and \mathbf{d} . The $\langle 00|00 \rangle^{(m)}$, which are closely related to the $[0]^{(m)}$ in (63), are defined by

$$\langle 00|00 \rangle^{(m)} = D_A D_B D_C D_D G_{AB} G_{CD} \frac{2\pi^{5/2}}{\zeta \eta (\zeta + \eta)^{1/2}} F_m(T) \quad (68)$$

When generating a class of integrals with non-trivial angular momentum, there are usually very many sequences in which (67) can be applied and the task of determining the most efficient sequence can be discussed as a tree-search problem. OS did not provide a solution to this problem however and, as such, the OS algorithm is not completely well-defined.

3.9 The Head-Gordon-Pople Algorithm [55]

Although no clear prescription was given by OS for the use of their RR, it was clearly a significant advance and immediately received considerable attention. The next major step forward was taken by HGP who not only introduced a second RR but also gave a precise description of how the two RR's can be used in tandem to good effect.

The target $(\mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d})$ are made from $(\mathbf{a}|\mathbf{b}|\mathbf{n}|\mathbf{0})$ using a two-term RR

$$|c(\mathbf{d} + \mathbf{1}_i)| \equiv |(c + \mathbf{1}_i)\mathbf{d}| + (C_i - D_i)|c\mathbf{d}| \quad (69)$$

which HGP derived from the eight-term RR of OS and which they designated the horizontal recurrence relation (HRR). In fact, the same RR (termed a transfer relation) had been used for many years to transform the 2-dimensional integrals which arise in the DRK method [51] but, until the HGP paper, it had apparently not been applied to *contracted* integrals. This distinction is important, for it implies that the computational expense incurred by the use of (69) is independent of the contraction degree of the integrals, *i.e.* the work contributes only to the z parameter in (42).

The $(\mathbf{a}|\mathbf{b}|\mathbf{n}|\mathbf{0})$ are made from $(\mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0})$, again using the HRR, *i.e.*

$$(\mathbf{a}|\mathbf{b} + \mathbf{1}_i|) = ((\mathbf{a} + \mathbf{1}_i)\mathbf{b}| + (A_i - B_i)(\mathbf{a}|\mathbf{b}|) \quad (70)$$

and, as before, (70) contributes only to the z parameter in (42).

The $(\mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0})$ are made by simple contraction of $[\mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0}]$, *i.e.*

$$(\mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0}) = \sum_{i=1}^{K_A} \sum_{j=1}^{K_B} \sum_{k=1}^{K_C} \sum_{l=1}^{K_D} [\mathbf{m}_i|\mathbf{0}_j|\mathbf{n}_k|\mathbf{0}_l] \quad (71)$$

and the $[\mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0}] \equiv \langle \mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0} \rangle^{(0)}$ are generated using the OS RR. Because the second and fourth indices in $\langle \mathbf{m}|\mathbf{0}|\mathbf{n}|\mathbf{0} \rangle$ are $\mathbf{0}$, the 5th, 6th, and 8th terms of the OS RR always vanish and HGP called this five-term special case the vertical recurrence relation (VRR). A simple method was given by HGP for deciding whether to reduce at the first or at the third index when applying the VRR and this completes the specification of their algorithm.

3.10 Variations on the HGP Theme

Since the advent of HGP, there have been a number of attempts to improve upon it without changing its essential structure, viz.

$$\langle 00|00 \rangle^{(m)} \rightarrow [m0|n0] \rightarrow (m0|n0) \rightarrow (ab|cd) \quad (72)$$

Initially, it was suggested [82] that, instead of using the VRR, a DRK scheme (such as that of Saunders [65]) could be used to generate the $[m0|n0]$, after which the contraction and HRR steps could proceed as usual. This idea was explored by Lindh, Ryu and Liu (LRL) in a fascinating paper [59a] which convincingly argues that, for the construction of $[m0|n0]$, the HGP and DRK algorithms are essentially equivalent mathematically. This remarkable result was also supported by Flop-cost determinations, using each method, for $(pp|pp)$, $(dd|dd)$ and $(ff|ff)$ classes.

A second, and somewhat more successful, attempt to render the VRR obsolete was made independently by Hamilton and Schaefer (HS) [57] and LRL [59a]. By combining the translational invariance condition for first derivatives [83] with the HRR, HS discovered a six-term RR

$$\begin{aligned} [ab|(c+1_i)d] &= \frac{a_i}{2\eta}[(a-1_i)b|cd] + \frac{b_i}{2\eta}[a(b-1_i)|cd] + \frac{c_i}{2\eta}[ab|(c-1_i)d] \\ &+ \frac{d_i}{2\eta}[ab|c(d-1_i)] - \frac{2\zeta}{2\eta}[(a+1_i)b|cd] - \left[\frac{2\beta}{2\eta}(A_i - B_i) + \frac{2\delta}{2\eta}(C_i - D_i) \right] [ab|cd] \end{aligned} \quad \dots (73)$$

which, if applied to integrals in which $\mathbf{b} = \mathbf{d} = \mathbf{0}$, becomes

$$\begin{aligned} [m0|(n+1_i)0] &= \frac{m_i}{2\eta}[(m-1_i)0|n0] + \frac{n_i}{2\eta}[m0|(n-1_i)0] \\ &- \frac{2\zeta}{2\eta}[(m+1_i)0|n0] - \left[\frac{2\beta}{2\eta}(A_i - B_i) + \frac{2\delta}{2\eta}(C_i - D_i) \right] [m0|n0] \end{aligned} \quad \dots (74)$$

Both HS and LRL proposed that the OS RR be used only to generate $[\mathbf{m0|00}] \equiv \langle \mathbf{m0|00} \rangle^{(0)}$ and that (74) then be used to form $[\mathbf{m0|n0}]$ from these. It seems probable from the estimates made by both groups (they disagree somewhat) of the Flop-cost of this new algorithm that it constitutes a marginal improvement over the HGP algorithm.

In the same paper where they discuss an alternate derivation of (74), LRL also present the Reduced Multiplication Rys (RMR) scheme for computing $[\mathbf{m0|n0}]$ and compare it with the DRK, HGP and HS algorithms. RMR is found to require noticeably fewer Flops than DRK and HGP (which, as mentioned above, are equivalent). The major source of this improvement is the sophisticated treatment of reusable intermediate data by RMR and the interested reader is referred to [59a] for further details.

As a general rule, the construction of the $[\mathbf{m0|n0}]$ using the VRR is considerably more expensive than the subsequent contraction and HRR steps. However, if one is dealing with integral classes of high angular momentum and low contraction or with derivatives of such classes with respect to nuclear motion, the HRR step can become sufficiently expensive to warrant optimization and Ryu, Lee and Lindh [59b] have recently studied this problem. Recognizing that efficient application of the HRR involves a complicated tree-search problem, they devised a heuristic solution which eliminates 13%, 25%, 38% and 44%, respectively, of the Flops which previous HRR implementations had needed for (fff) , $(gg|gg)$, $(hhlhh)$ and $(iilii)$ classes.

Recently, however, Johnson *et al.* [63] have found that, if one wishes to minimize the number of *Mops* (as opposed to Flops) in the transformation of $(\mathbf{m0|n0})$ to (\mathbf{abcd}) , it is often preferable to dispense with the HRR entirely and, in lieu of it, employ RR's from a novel family which these authors term "nth-order transfer relations". We will say more about these later in the context of the PRISM algorithm.

4. THE PRISM ALGORITHM [61]

In addition to stimulating a number of variations on the HGP theme, the seminal paper by Head-Gordon and Pople [55] also served to catalyse the development of a completely different approach to the Contraction Problem called the PRISM algorithm.

We recall from our discussion of the Contraction Problem (Section 3.4) that none of the algorithms hitherto suggested has proven optimal under all circumstances: the Pople-Hehre method is highly efficient for highly contracted classes but is very poor for weakly contracted ones; the Hamilton-Schaefer-Lindh-Ryu-Liu method employing (74) is extremely effective for mildly contracted classes but is otherwise grossly inferior to Pople-Hehre. However, we observed that there is a simple connection between the behavior of an algorithm and the point at which the primitive integrals are combined to yield contracted integrals: early-contraction methods are best suited to highly contracted classes and late-contraction methods are best suited to weakly contracted classes.

Ideally, we would like an algorithm to *choose dynamically*, on the basis of the type of class being generated, the optimal point at which to perform the contraction step. The crucial insight that leads to the PRISM algorithm is the realization that some of the algorithms which we have already discussed can be generalized into forms in which such dynamic flexibility is possible. We call the generalizations of the MD and HGP methods the MD-PRISM and HGP-PRISM algorithms, respectively.

The MD-PRISM [61] was discovered and implemented long before the HGP-PRISM (which has not previously been discussed in the literature) but, for the purposes of this Review, it is convenient to develop them together in order to emphasize their similarities and differences.

We have already observed (Section 3.6) that any two-electron integral (6), and any n th-derivative of that integral with respect to motion of the basis functions, is an inner product between a function of the position of electron #1 (which we term a bra) and a function of the position of electron #2 (which we term a ket). Henceforth, rather than explicitly considering integrals and/or their n th-derivatives, we will examine the more general problem of computing contracted brackets.

After developing the mathematical foundations of the PRISM algorithm, we will discuss its implementation within the Gaussian 92 computer program [84].

The MD-PRISM and HGP-PRISM algorithms are most easily understood when presented in terms of diagrams which resemble rectangular prisms – whence their names. To simplify the discussion, we will confine our attention to the “front” face of each prism. The generalizations to the complete prisms are neither difficult nor especially interesting: they are outlined in [61c].

The front face of the MD PRISM is shown below. As the arrows indicate, the MD-PRISM algorithm consists of a set of highly interrelated pathways from shell-pair data to the desired brackets.

Shell-Pair Data

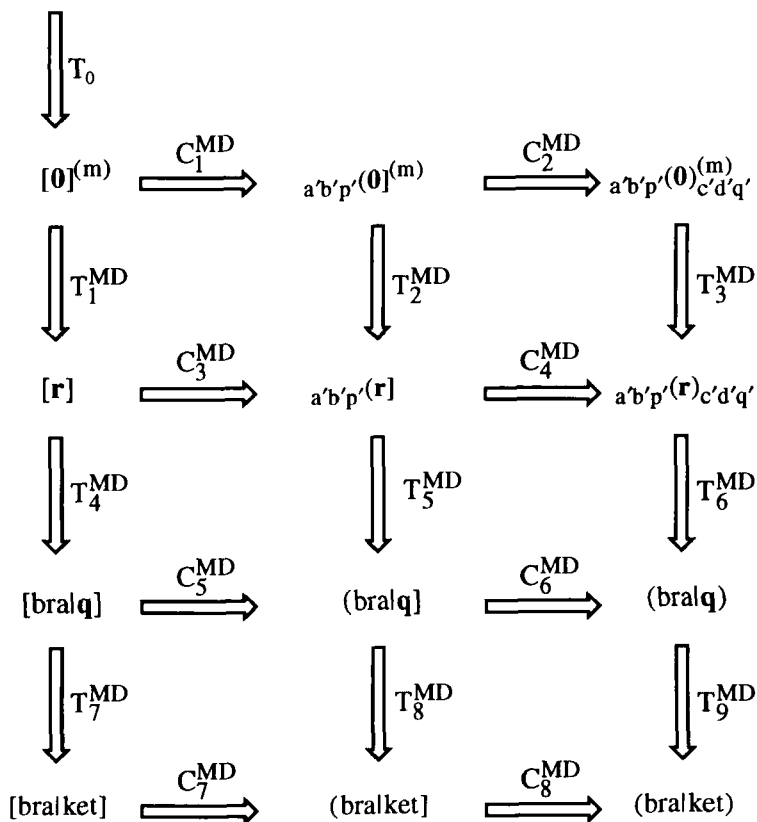


Figure 1. The front face of the MD PRISM.

The front face of the HGP PRISM is shown below. Like the MD-PRISM algorithm, the HGP-PRISM algorithm consists of a set of highly interrelated pathways from shell-pair data to the desired brackets.

Shell-Pair Data

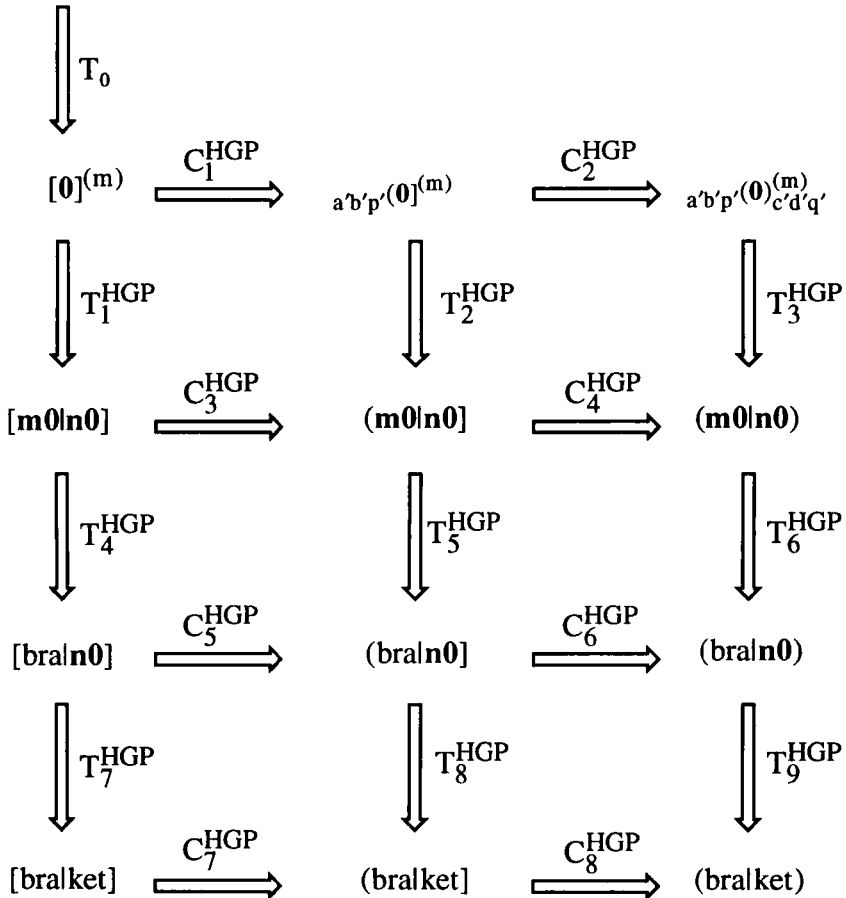


Figure 2. The front face of the HGP PRISM.

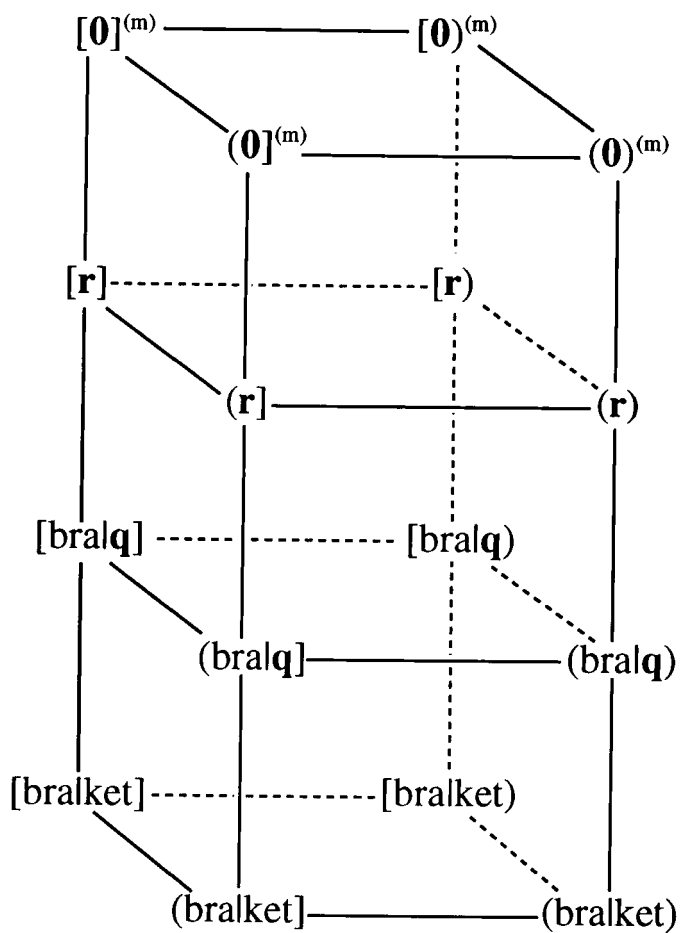


Figure 3. The Full MD PRISM.

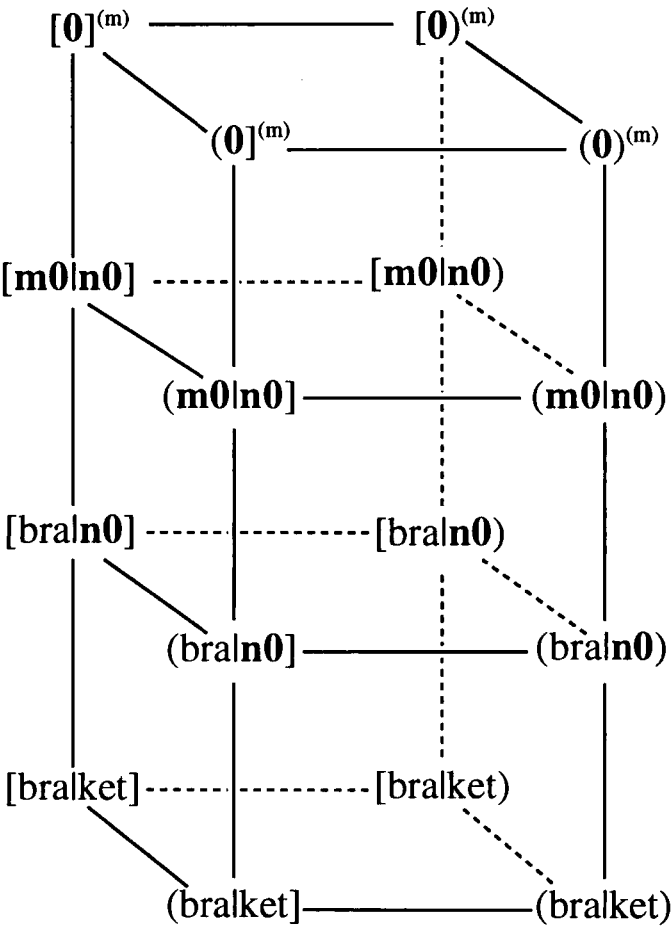


Figure 4. The Full HGP PRISM

In the subsections which follow, we will examine various aspects of the MD and HGP PRISMs but, before doing so, we may make some general observations.

- (1) The first step (T_0), the generation of $[0]^{(m)}$ integrals from shell-pair data, is common to both the MD- and HGP-PRISM algorithms. We will discuss T_0 in detail in Section 4.3.
- (2) In either algorithm, the metamorphosis of $[0]^{(m)}$ to brackets involves exactly two Contraction steps (C_i , denoted by horizontal arrows and discussed in Section 4.4) and three Transformation steps (T_i , denoted by vertical arrows and discussed in Section 4.5) and, since these five steps can be performed in *any* order, there are 10 paths from $[0]^{(m)}$ to brackets in Figure 1 and a further 10 in Figure 2. We will use appropriate permutations of two “C”s and three “T”s to label these paths. For example, one of the paths in Figure 1 is MD-CCTTT and one of those in Figure 2 is HGP-TCCTT.
- (3) By comparing Figures 1 and 2 with the descriptions in Sections 3.7 and 3.9, it becomes immediately apparent that the classical McMurchie–Davidson and Head–Gordon–Pople algorithms correspond to the MD-TTCTC and HGP-TCCTT paths, respectively.
- (4) As a result of the “perpendicular” separation of contraction and transformation steps in Figures 1 and 2, the steps which contribute to the x , y , and z cost-parameters of (42) are clearly delineated. Specifically, the T_0 , T_1 , T_4 , T_7 , C_1 , C_3 , C_5 and C_7 steps contribute to x ; the T_2 , T_5 , T_8 , C_2 , C_4 , C_6 and C_8 steps contribute to y ; and the T_3 , T_6 and T_9 steps contribute to z . This observation permits a computer program to compute x , y , and z parameters (for a given bracket class) for each of the paths on each of the prisms and, using (42), then very easily to determine the Flop- or Mop-cheapest path.

4.1 Shell-Pair Data

Inspection of (6) reveals that it describes the interaction of two charge distributions, $\phi_a(\mathbf{r})\phi_b(\mathbf{r})$ and $\phi_c(\mathbf{r})\phi_d(\mathbf{r})$, and our first task is to collect information about all such charge distributions in the molecule. Because brackets are formed in classes, rather than individually, it is convenient to compile data for shell-pairs (rather than basis-function-pairs) and this shell-pair dataset is central to any modern integral program. To generate all of the desired brackets, we will later loop over all pairs of shell-pairs, that is, over all shell-quartets.

We begin by considering all pairs [85] of shells in the basis set, categorizing each shell-pair as either *significant* or *negligible*. A shell-pair is negligible if the shells involved are so far apart (relative to their diffuseness) that their overlap is negligible; otherwise, it is significant. Because all of the basis functions which we have discussed (Section 2) decay at least exponentially, *most of the shell-pairs in a large molecule are negligible*. Indeed, the number of significant pairs grows only *linearly* with increasing molecular size and it therefore follows that the number of significant shell-quartets grows only *quadratically*. This is a very important point for it is this, more than anything else, which permits rigorous SCF calculations to be performed on very large systems.

Each time that we discover a significant shell-pair, we generate a *model* for that pair, *i.e.* a second shell-pair, *with fewer primitives than the first*, whose potential is as close as possible (in a least-squares sense) to that of the first. Only very recently has the theory necessary for such *potential*-fitting procedures been developed [72] and our methodology may be contrasted with previous approaches [86–92] in which many alternative modeling criteria have been employed. In particular, we note that it is common in Kohn-Sham calculations to expand the electronic density in an auxiliary basis set [89]. This is equivalent to modeling each charge distribution $\phi_a(\mathbf{r})\phi_b(\mathbf{r})$ by an expansion in the auxiliary basis. The modeling scheme employed to obtain the results discussed later in this Review was the simple one described by Head-Gordon and Pople [55]: all primitives with amplitudes below 10^{-10} are discarded while all others are kept. More elaborate modeling procedures are currently being developed [72b].

Once we have compiled a list of models for all of the significant shell-pairs, they are sorted by "type", *i.e.* by the angular momenta of the component shells and by the degree of contraction of the (modeled) shell-pair. Thus, all uncontracted *ss*-pairs are stored consecutively, followed by all doubly-contracted

(K=2) *ss*-pairs, and so forth. For each shell-pair, the parent shells and (**A** – **B**) vector are recorded; for each component primitive, we record (2 α), (2 β), 1/(2 ζ), **P** and

$$U_P = D_A D_B G_{AB} (\pi/\zeta)^{3/2} \left[\frac{1}{2\zeta} \right]^{a+b} \quad (75)$$

The last factor in (75), which scales it according to the angular momentum (*a* + *b*) of the shell-pair, is termed the *principal scaling* and is included only if the MD-PRISM is used. For reasons which will become clearer below, its presence reduces the Flop- and Mop-costs of the algorithm.

We have recently developed [93] an upper-bound on (6)

$$(\mathbf{ab}|\mathbf{cd}) \leq \text{Min} \left\{ I_{ab}^* I_{cd}^*, V_{ab}^* S_{cd}^*, S_{ab}^* V_{cd}^* \right\} \quad (76)$$

where

$$I_{ab}^* = (\mathbf{ab}|\mathbf{ab})^{1/2} \quad (77)$$

$$V_{ab}^* = \text{Max}_{\mathbf{G}} \left| \int \phi_a(\mathbf{r}) \phi_b(\mathbf{r}) f(|\mathbf{r} - \mathbf{G}|) d\mathbf{r} \right| \quad (78)$$

$$S_{ab}^* = \int |\phi_a(\mathbf{r}) \phi_b(\mathbf{r})| d\mathbf{r} \quad (79)$$

which is significantly stronger than the familiar [94–97] Schwarz bound (which is the first of the three braced quantities in (76)). Our next task is to compute shell-pair generalizations of (77)–(79) for each of the significant shell-pairs by evaluating (77)–(79) for each of their constituent basis-function-pairs and taking appropriate maxima [97].

The total computational effort involved in setting up the shell-pair data increases linearly with the size *N* of the basis. For tasks such as large Direct SCF calculations [44, 45], it is entirely negligible compared with the subsequent work; for less computationally demanding tasks, such as finding potential-derived atomic charges [98], it typically constitutes 10% of the job time.

4.2 Selection of Shell-Quartets

Given a sorted list of significant shell-pairs, we can construct all potentially important shell-quartets [99] by pairing the shell-pairs with one another. For the sake of vectorization, we deal with *batches* of shell-quartets of the same “type” and we utilize the memory which is available as effectively as possible in order to maximize the sizes of such batches [100].

Not every pair of shell-pairs, however, is necessarily accepted as a worthwhile shell-quartet. Although the shell-pair database has been carefully screened and contains no negligible shell-pairs, there are several ways in which a pair of significant shell-pairs may yield a shell-quartet which can be neglected...

- (1) The quartet may be equivalent, by point group symmetry, to another quartet which has already been treated.
- (2) The largest bracket associated with the quartet may be negligibly small. This can be anticipated by the upper-bound formula (76) if the cutoff parameters (77)–(79) have been precomputed as indicated in Section 4.1.
- (3) The largest density matrix (or delta-density matrix) elements which will multiply any of the brackets associated with the quartet may be negligibly small [44,97]. This is particularly common in late SCF cycles when incremental Fock matrix formation is being used.

Unfortunately, in general, it is not easy to vectorize the shell-quartet selection process because of the conditional nature of quartet acceptance. In the special case where one desires the electrostatic potential on a large grid, Johnson *et al.* have circumvented this problem by the so-called “fixed shell-pair” scheme [98] which is completely vectorizable. However, this approach is not directly applicable to the general case.

4.3 Generation of the $[0]^{(m)}$ Integrals

Given a batch of shell-quartets, the real computational work (denoted T_0 in Figures 1 and 2) can begin. In the first stage of this, the seven basic shell-quartet parameters

$$\mathbf{R} = \mathbf{Q} - \mathbf{P} \quad (80)$$

$$R^2 = R_x^2 + R_y^2 + R_z^2 \quad (81)$$

$$2\vartheta^2 = 1 / \left[\frac{1}{2\zeta} + \frac{1}{2\eta} \right] \quad (82)$$

$$2T = 2\vartheta^2 R^2 \quad (83)$$

$$\mathbf{U} = \mathbf{U}_P \mathbf{U}_Q \quad (84)$$

are constructed. Given the shell-pair data which were generated earlier (Section 4.1), (80)–(84) can be computed in just 12 Flops and 17 Mops and this accounts for a rather small fraction of the total CPU time in most Direct SCF calculations (less than 4% in the pentacene run described in Section 4.7).

In the second stage, the $[0]^{(m)}$ integrals ($0 \leq m \leq L$) are evaluated. If we desire two-electron repulsion integrals, *i.e.* $f(x) \equiv 1/x$ in (6), the relevant definition (Section 3.2.2) is

$$[0]^{(m)} = \mathbf{U} \left(2\vartheta^2 \right)^{m+1/2} \mathbf{G}_m(T) \quad (85)$$

where

$$\mathbf{G}_m(T) = (2/\pi)^{1/2} \int_0^1 t^{2m} \exp(-Tt^2) dt \quad (86)$$

For other types of integrals, (85) and (86) must be appropriately modified but retain the same general form.

If T (which measures the extent to which the bra and ket charge distributions overlap) is less than a critical value T_{crit} , $G_L(T)$ can be evaluated using an interpolation procedure. We follow Shipman and Christoffersen [68] in favoring Chebyshev interpolation as an effective means for computing such functions but we have adopted the approach of Elbert and Davidson [69] who prefer to employ approximations of lower degree. We have chosen cubic interpolation as our standard and have discussed our methodology in detail in [71]. Values of $G_m(T)$ for $0 \leq m < L$ can then be obtained, with numerical stability, by downward recursion (see (65), for example). The function $\exp(-T)$ is needed for this and, for speed, we also compute this by interpolation [71].

On the other hand, if T is greater than T_{crit} , the distributions overlap negligibly and (in the case of two-electron repulsion integrals) the $[0]^{(m)}$ then reduce [101] to classical multipole terms

$$[0]^{(m)} = \left(\frac{U}{R} \right) \left(\frac{1}{R^2} \right) \left(\frac{3}{R^2} \right) \dots \left(\frac{2m-1}{R^2} \right) \quad (87)$$

which can be computed recursively with great efficiency.

The subroutine in Gaussian 92 which is responsible for the evaluation of the $[0]^{(m)}$ in the context of Direct SCF is Calc0m. This routine accounts for a noticeable fraction of the total CPU time (7% in the pentacene run described in Section 4.7) and has been carefully optimized. It runs at roughly 160 MFlops on a single-processor Cray Y-MP (whose theoretical peak speed is 333 MFlops).

In computations of the electrostatic potential on a grid, Gaussian 92 calls the subroutine Calc0G to compute the $[0]^{(m)}$ and this routine accounts for roughly 50% of the total CPU time. As a result of the use of the “fixed shell pair” scheme [98] and very careful optimization, Calc0G runs at approximately 180 MFlops on the Cray Y-MP.

Given a batch of $[0]^{(m)}$ integrals, it “only” remains to traverse one of the prisms in order to obtain the brackets which we seek. As mentioned earlier, this involves some combination of two Contraction steps and three Transformation steps and we now focus on these in detail.

4.4 Contraction Steps

The horizontal arrows in Figures 1 and 2 correspond to contraction steps and, in typical Direct SCF calculations using PRISM, these account for a significant fraction of the total CPU time (15% in the pentacene run described in Section 4.7). In electrostatic grid calculations, the fraction is even higher. It is therefore very important that they be executed as efficiently as possible.

As was indicated in Section 3.6, contraction within the bracket framework involves the summation of primitive quantities which may, or may not, also need to be scaled before being added together. In the case where scaling is required, the contraction amounts to a dot product, *i.e.*

$$A_i = \sum_{j=1}^K S_{ij} B_{ij} \quad (88)$$

where i loops over the members of the batch and j loops over the length of the contraction. In Gaussian 92, (88) is implemented with i as the innermost loop and with the outer j loop unrolled sixfold. The resulting code, a representative kernel of which is

```

DO 20 J=Jbeg,Jend,6
  DO 10 I=1,N
    A(I) = A(I) + S(I,J ) *B(I,J ) + S(I,J+1)*B(I,J+1)
$          + S(I,J+2)*B(I,J+2) + S(I,J+3)*B(I,J+3)
$          + S(I,J+4)*B(I,J+4) + S(I,J+5)*B(I,J+5)
10    CONTINUE
20 CONTINUE
```

has a Flop/Mop ratio which approaches unity, possesses a good balance between adds and multiplies, is manifestly vectorizable and, not surprisingly, runs very fast on most platforms.

We note, too, that things only improve under the “fixed shell-pair” scheme which is used in electrostatic grid calculations because the scalings become loop-invariants [98].

4.5 Transformation Steps

As we indicated in Section 3.3, modern integral algorithms invariably employ recurrence relations to build complicated brackets from simple ones. Their use permits algorithms to deal (in principle) with brackets of arbitrarily high angular momentum and, additionally, to make good use of the intermediates that are shared by fraternal brackets. In the next four subsections, we will discuss the various recurrence relations which are used to move vertically on the prisms.

4.5.1 Two-Electron Transformations on the MD PRISM

The T_1^{MD} step in Figure 1 transforms $[0]^{(m)}$ to $[r]$ integrals and the T_2^{MD} and T_3^{MD} steps are the half-contracted and contracted variants of this. We will examine T_1^{MD} in detail because, although it is the simplest transformation step on either prism, it nonetheless shares many features with the more complicated transformation steps and, therefore, has useful pedagogical value.

The recurrence relation on which T_1^{MD} , T_2^{MD} and T_3^{MD} are based is the one-center RR (62)

$$[r]^{(m)} \equiv R_i [r - \mathbf{1}_i]^{(m+1)} - (r_i - 1) [r - \mathbf{2}_i]^{(m+1)} \quad (89)$$

which McMurchie and Davidson derived originally using the elementary properties of Hermite polynomials. It is easily shown that, if the total angular momentum of the desired bracket class is L , there are $(L+1)$ $[0]^{(m)}$ and $(L+3)!/L!/3!$ $[r]$ integrals and that, in general, the efficient generation of the latter from the former involves a complicated tree-search problem. Johnson *et al.* have carefully analysed this and have constructed highly optimized solution-trees [102]. In Gaussian 92, the subroutine MakMD1 employs these solutions in forming a "driver" for the T_1^{MD} step.

In Gaussian 92, a "driver" is an array of instructions (coded as integers) for the formation of one set of integrals from another. Subroutine MakMD1 first constructs a driver (the MD1 driver) to form the $[r]$ from the $[0]^{(m)}$ using (89). Given this driver, a set of $[0]^{(m)}$ and the corresponding set of \mathbf{R} vectors (80), subroutine DoMD1 then takes responsibility for the actual construction of the $[r]$. For the purposes of illustration, it is useful to work through a simple example, such as $L = 2$.

The explicit $L = 2$ solution-tree for (89) is [102]

$$\begin{aligned}
 [s]^{(1)} &= [\mathbf{0}]^{(1)} \\
 [p_z]^{(1)} &= R_z[\mathbf{0}]^{(2)} \\
 [p_y]^{(1)} &= R_y[\mathbf{0}]^{(2)} \\
 [p_x]^{(1)} &= R_x[\mathbf{0}]^{(2)} \\
 [s] &= [\mathbf{0}]^{(0)} \\
 [p_z] &= R_z[\mathbf{0}]^{(1)} \\
 [p_y] &= R_y[\mathbf{0}]^{(1)} \\
 [p_x] &= R_x[\mathbf{0}]^{(1)} \\
 [d_{zz}] &= R_z[p_z]^{(1)} - [s]^{(1)} \\
 [d_{yz}] &= R_y[p_z]^{(1)} \\
 [d_{yy}] &= R_y[p_y]^{(1)} - [s]^{(1)} \\
 [d_{xz}] &= R_z[p_x]^{(1)} \\
 [d_{xy}] &= R_x[p_y]^{(1)}
 \end{aligned} \tag{90}$$

Thus, from $[\mathbf{0}]^{(0)}$, $[\mathbf{0}]^{(1)}$, $[\mathbf{0}]^{(2)}$, we form ten $[\mathbf{r}]$ integrals. We will assume (correctly) that the former are stored in locations 1, 2 and 3, respectively. Subroutines such as DoMD1 are very simple, containing only a handful of basic DO-loops corresponding to all useful special cases of the RR (89). The driver is a two-dimensional integer array, each row of which (an *instruction*) specifies the DO-loop and the locations to use. The MD1 driver for $L = 2$ is shown in Table III and DoMD1 is reproduced immediately below this.

Table III: The MD1 driver for $L = 2$

4	3	0	1	0
5	3	0	2	0
3	3	0	3	0
6	4	2	1	1
7	5	0	1	0
4	4	0	3	0
5	5	2	2	1
8	3	0	2	0
3	3	2	3	1
9	2	0	1	0
10	2	0	2	0
2	2	0	3	0

```

SUBROUTINE DoMD1(W,R,n,Driver,nDrive)
IMPLICIT INTEGER (a-z)
INTEGER Driver(5,nDrive)
REAL*8 W(N,*),R(n,3)
DO 40 knt = 1,nDrive
  Z      = Driver(1,knt)
  P1     = Driver(2,knt)
  P2     = Driver(3,knt)
  Axis   = Driver(4,knt)
  j      = Driver(5,knt)
  IF (j.eq.0) THEN
    DO 10 i = 1,n
      W(i,Z) = R(i,Axis) * W(i,P1)
10    CONTINUE
    ELSE IF (j.eq.1) THEN
      DO 20 i = 1,n
        W(i,Z) = R(i,Axis) * W(i,P1) - W(i,P2)
20    CONTINUE
    ELSE
      DO 30 i = 1,n
        W(i,Z) = R(i,Axis) * W(i,P1) - W(i,P2) * j
30    CONTINUE
    END IF
40 CONTINUE
RETURN
END

```

Each instruction in the driver is encoded by five integers:

- (1) The location in which to store the result;
- (2) The location of the first term in (89);
- (3) The location of the second term in (89);
- (4) The value of i in (89) ($1 \Rightarrow x$, $2 \Rightarrow y$, $3 \Rightarrow z$);
- (5) The coefficient of the second term in (89).

For example, the first instruction (4, 3, 0, 1, 0) corresponds to the fourth equation in (90) and the fourth instruction (6, 4, 2, 1, 1) corresponds to the last equation in (90). We note the following:

- (a) DoMD1 uses the coefficient of the second term in (89) to determine which of the three special cases of (89) to use.
- (b) Each instruction results in the appropriate RR being applied to an entire *column* (n elements) of the W array, not just to a single element. That is, DoMD1 actually forms n sets of $[\mathbf{r}]$, not just a single set. Of course, n is the batch size described in Section 4.2 and this is the device by which PRISM is vectorized.
- (c) The driver is frugal in its use of locations: to the greatest extent possible, it re-uses locations by *overwriting* intermediates when they are no longer needed. For example, instruction #1 places $[p_x]^{(1)}$ in location 4 but, after this has been used in instruction #4 to generate $[d_{xx}]$, the $[p_x]^{(1)}$ is overwritten in instruction #6 by $[d_{xz}]$.

The other PRISM subroutines whose names begin with Do (see Section 4.6) operate very similarly to DoMD1. Each is handed the W array, a driver and, if necessary, some auxiliary arrays such as R in DoMD1 and then proceeds to generate columns of W by combining columns with one another and, possibly, with the auxiliary arrays.

To complete this subsection, we must derive two further RRs. In order to undertake the T_2^{MD} or T_3^{MD} steps, we require RRs involving half-contracted or fully-contracted $[\mathbf{r}]^{(m)}$, respectively. These are easily derived [61c] by replacing R_i in (89) by appropriate identities. Thus, if the easily verified identity

$$R_i \equiv \frac{2\alpha}{2\zeta}(B_i - A_i) + (Q_i - B_i) \quad (91)$$

is substituted into (89), we obtain the half-contracted RR

$$\begin{aligned} a'b'p'(\mathbf{r})^{(m)} &= (B_i - A_i)_{(a'+1)b'(p'+1)}(\mathbf{r} - \mathbf{1}_i)^{(m+1)} \\ &+ (Q_i - B_i)_{a'b'p'}(\mathbf{r} - \mathbf{1}_i)^{(m+1)} \\ &- (r_i - 1)_{a'b'p'}(\mathbf{r} - \mathbf{2}_i)^{(m+1)} \end{aligned} \quad (92)$$

which is suitable for use in the T_2^{MD} step. Likewise, if the identity

$$R_i \equiv \frac{2\alpha}{2\zeta}(B_i - A_i) + \frac{2\gamma}{2\eta}(C_i - D_i) + (D_i - B_i) \quad (93)$$

is substituted into (89), we obtain the fully-contracted RR

$$\begin{aligned} a'b'p'(\mathbf{r})_{c'd'q'}^{(m)} &= (B_i - A_i)_{(a'+1)b'(p'+1)}(\mathbf{r} - \mathbf{1}_i)_{c'd'q'}^{(m+1)} \\ &+ (C_i - D_i)_{a'b'p'}(\mathbf{r} - \mathbf{1}_i)_{(c'+1)d'(q'+1)}^{(m+1)} \\ &+ (D_i - B_i)_{a'b'p'}(\mathbf{r} - \mathbf{1}_i)_{c'd'q'}^{(m+1)} \\ &- (r_i - 1)_{a'b'p'}(\mathbf{r} - \mathbf{2}_i)_{c'd'q'}^{(m+1)} \end{aligned} \quad (94)$$

which is suitable for use in the T_3^{MD} step.

4.5.2 One-Electron Transformations on the MD PRISM

The T_4^{MD} , T_7^{MD} , and T_8^{MD} steps in Figure 1 transform *uncontracted* p-bras (or q-kets) to *uncontracted* bras (or kets). There is little to say about such transformations except that they are accomplished by (57) and (59) and form the backbone of the classical MD algorithm [52]. Their chief weakness is that, because they are uncontracted transformations, they contribute to the x and y cost-parameters in (42) and therefore become very expensive when applied to bracket classes of moderate or high degree of contraction. For such classes, it would obviously be more efficient to contract *before* transforming: this is precisely what the T_5^{MD} , T_6^{MD} and T_9^{MD} steps achieve. But how can this be done?

It is only the second terms in (57) and (59) which prevent the application of these RRs to contracted bras and kets because the prefactors $(P_i - A_i)$ and $(Q_i - C_i)$ vary from primitive to primitive. However, if we replace these by identities based on the definitions (12) of \mathbf{P} and \mathbf{Q} , as was first suggested in [61a], a new RR, which can be applied to *contracted* bras and kets, emerges. In bra notation, it can be expressed [61b] as

$$\begin{aligned} \left(\begin{array}{cc} 0 & 0 \\ \mathbf{a} + \mathbf{1}_i & \mathbf{b} \quad \mathbf{p} \\ \mathbf{a}' & \mathbf{b}' \quad \mathbf{p}' \end{array} \right| &= p_i \left(\begin{array}{ccc} 0 & 0 & \\ \mathbf{a} & \mathbf{b} & \mathbf{p} - \mathbf{1}_i \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' \end{array} \right| + (\mathbf{B}_i - \mathbf{A}_i) \left(\begin{array}{ccc} 0 & 0 & \\ \mathbf{a} & \mathbf{b} & \mathbf{p} \\ \mathbf{a}' & \mathbf{b}' + \mathbf{1} & \mathbf{p}' + \mathbf{1} \end{array} \right| \\ &+ \left(\begin{array}{ccc} 0 & 0 & \\ \mathbf{a} & \mathbf{b} & \mathbf{p} + \mathbf{1}_i \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' + \mathbf{1} \end{array} \right| \end{aligned} \quad (95)$$

Eqn. (95) is the contracted analogue of (59) and is much more efficient for highly contracted bracket classes. The corresponding contracted analogue of (57) is obvious.

Thus, by using (59), (89), (92), (94) and (95) in judicious combinations, one can traverse any of the 10 paths in Figure 1 (and, indeed, with only trivial extensions, any of the 20 paths in Figure 3). By choosing always to go by the cheapest (in a Flops or Mops sense) path, one gains the full benefit of the PRISM algorithm.

4.5.3 Two-Electron Transformations on the HGP PRISM

The T_1^{HGP} step in Figure 2 transforms $[0]^{(m)}$ to $[m0|n0]$ integrals and the T_2^{HGP} and T_3^{HGP} steps are the half-contracted and contracted variants of this. However, before we can even contemplate the HGP algorithm within a bracket framework, it is necessary to recast the OS RR (67), which involves $\langle \mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d} \rangle^{(m)}$ integrals, into a new form involving $[\mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m)}$ integrals. Although this is straightforward to do, it has not previously appeared in the literature. Upon substituting the connection formula

$$[\mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m)} \equiv \left(\frac{2\zeta\eta}{\zeta+\eta} \right)^m \langle \mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d} \rangle^{(m)} \quad (96)$$

which is easily derived by comparing (63) and (68), into (67) and simplifying, we discover a new RR

$$\begin{aligned} [(\mathbf{a} + \mathbf{1}_i)|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m)} &= (\mathbf{B}_i - \mathbf{A}_i) \left(\frac{2\beta}{2\zeta} \right) [\mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m)} + \mathbf{R}_i \left(\frac{1}{2\zeta} \right) [\mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m+1)} \\ &\quad + a_i \left\{ \left(\frac{1}{2\zeta} \right) [(\mathbf{a} - \mathbf{1}_i)|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m)} - \left(\frac{1}{2\zeta} \right)^2 [(\mathbf{a} - \mathbf{1}_i)|\mathbf{b}|\mathbf{c}|\mathbf{d}]^{(m+1)} \right\} \\ &\quad + b_i \left\{ \left(\frac{1}{2\zeta} \right) [\mathbf{a}(\mathbf{b} - \mathbf{1}_i)|\mathbf{c}|\mathbf{d}]^{(m)} - \left(\frac{1}{2\zeta} \right)^2 [\mathbf{a}(\mathbf{b} - \mathbf{1}_i)|\mathbf{c}|\mathbf{d}]^{(m+1)} \right\} \\ &\quad + c_i \left(\frac{1}{2\zeta} \right) \left(\frac{1}{2\eta} \right) [\mathbf{a}|\mathbf{b}(\mathbf{c} - \mathbf{1}_i)|\mathbf{d}]^{(m+1)} + d_i \left(\frac{1}{2\zeta} \right) \left(\frac{1}{2\eta} \right) [\mathbf{a}|\mathbf{b}|\mathbf{c}(\mathbf{d} - \mathbf{1}_i)]^{(m+1)} \end{aligned} \quad (97)$$

by which any $[\mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d}]$ can be reduced to the $[00|00]^{(m)} \equiv [0]^{(m)}$ discussed in Section 4.3. Through this sleight of hand, we free ourselves of the necessity to consider the OS $\langle \mathbf{a}|\mathbf{b}|\mathbf{c}|\mathbf{d} \rangle^{(m)}$ integrals any further and we greatly clarify the relationship between the MD and HGP algorithms. We note too that, in addition to being more aesthetically pleasing than (67), (97) is also computationally superior to it because all of the *four*-center exponent factors in (67) are replaced by *two*-center factors in (97). Thus, each of the terms in (97) is now clearly a bracket, *i.e.* of the form (55).

In the original HGP method (Section 3.9), a special case of the OS RR, termed the VRR, is used to compute $[\mathbf{m}\mathbf{0}|\mathbf{n}\mathbf{0}]$ from $\langle\mathbf{00}|\mathbf{00}\rangle^{(m)}$ integrals. If, instead, we use the analogous special case of (97), we obtain a new RR

$$\begin{aligned}
 [(\mathbf{a} + \mathbf{1}_i)\mathbf{0}|\mathbf{c}\mathbf{0}]^{(m)} &= (\mathbf{B}_i - \mathbf{A}_i)\left(\frac{2\beta}{2\zeta}\right)[\mathbf{a}\mathbf{0}|\mathbf{c}\mathbf{0}]^{(m)} + \mathbf{R}_i\left(\frac{1}{2\zeta}\right)[\mathbf{a}\mathbf{0}|\mathbf{c}\mathbf{0}]^{(m+1)} \\
 &+ \mathbf{a}_i\left\{\left(\frac{1}{2\zeta}\right)[(\mathbf{a} - \mathbf{1}_i)\mathbf{0}|\mathbf{c}\mathbf{0}]^{(m)} - \left(\frac{1}{2\zeta}\right)^2 [(\mathbf{a} - \mathbf{1}_i)\mathbf{0}|\mathbf{c}\mathbf{0}]^{(m+1)}\right\} \\
 &+ \mathbf{c}_i\left(\frac{1}{2\zeta}\right)\left(\frac{1}{2\eta}\right)[\mathbf{a}\mathbf{0}|\mathbf{c} - \mathbf{1}_i\mathbf{0}]^{(m+1)}
 \end{aligned}
 \tag{98}$$

by which we can compute $[\mathbf{m}\mathbf{0}|\mathbf{n}\mathbf{0}]$ from $[\mathbf{00}|\mathbf{00}]^{(m)} \equiv [\mathbf{0}]^{(m)}$. This, of course, is precisely what is needed to perform the $\mathbf{T}_1^{\text{HGP}}$ step in Figure 2.

Finally, just as we derived RRs for \mathbf{T}_2^{MD} and \mathbf{T}_3^{MD} from the RR for \mathbf{T}_1^{MD} in Section 4.5.1, we can derive RRs for $\mathbf{T}_2^{\text{HGP}}$ and $\mathbf{T}_3^{\text{HGP}}$ from (98) by replacing \mathbf{R}_i by the identities (91) and (93). In this way, we eventually obtain

$$\begin{aligned}
& \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a}+1_i & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' \end{array} \right)^{(m)} = \\
& (\mathbf{B}_i - \mathbf{A}_i) \left\{ \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' + 1 & \mathbf{p}' + 1 & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' \end{array} \right)^{(m)} + \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' + 1 & \mathbf{b}' & \mathbf{p}' + 2 & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' \end{array} \right)^{(m+1)} \right\} \\
& + (\mathbf{C}_i - \mathbf{D}_i) \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' + 1 & \mathbf{c}' + 1 \ \mathbf{d}' \ \mathbf{q}' + 1 \end{array} \right)^{(m+1)} \\
& + (\mathbf{D}_i - \mathbf{B}_i) \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' + 1 & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' \end{array} \right)^{(m+1)} \\
& + \mathbf{a}_i \left\{ \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} - 1_i & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' + 1 & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' \end{array} \right)^{(m)} - \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} - 1_i & \mathbf{b} & 0 & \mathbf{c} \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' + 2 & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' \end{array} \right)^{(m+1)} \right\} \\
& + \mathbf{c}_i \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \mathbf{a} & \mathbf{b} & 0 & \mathbf{c} - 1_i \ \mathbf{d} \ 0 \\ \mathbf{a}' & \mathbf{b}' & \mathbf{p}' + 1 & \mathbf{c}' \ \mathbf{d}' \ \mathbf{q}' + 1 \end{array} \right)^{(m+1)}
\end{aligned} \tag{99}$$

for $\mathbf{T}_3^{\text{HGP}}$. The derivation of the (slightly simpler) $\mathbf{T}_2^{\text{HGP}}$ recurrence relation is left as an exercise for the reader...

4.5.4 One-Electron Transformations on the HGP PRISM

The T_4^{HGP} , T_5^{HGP} , T_6^{HGP} , T_7^{HGP} , T_8^{HGP} and T_9^{HGP} steps in Figure 2 correspond to the transformation of $[\mathbf{m}0]$ and $(\mathbf{m}0|$ to $[\mathbf{br}]$ and $(\mathbf{br}|$, respectively, and the analogous ket transformations. As Head-Gordon and Pople emphasized [55], the RR (70) which achieves this for $(\mathbf{br}| = (\mathbf{a}b|$ in the uncontracted case is also applicable in the contracted case. This represents an important difference between the MD and HGP PRISMs because it is therefore *always* preferable to contract fully before the last two transformations on the HGP PRISM whereas this is not the case on the MD PRISM.

At first glance, there is not much more that can be said about this transformation. The RR (70) is extremely simple and is easy to use and it might appear that our analysis can probe no further. However, as Ryu, Lee and Lindh have shown [59b], if one wishes to apply (70) in a way that minimizes the number of Flops involved, a complicated tree-search problem must first be solved. These authors were unable to solve the general problem but gave heuristic solutions which clearly indicated that substantial savings were available. However, this is not the approach which is followed in the HGP-PRISM algorithm...

As Johnson *et al.* have recently found [63], if one seeks a transformation scheme which is *Mop-optimal*, rather than *Flop-optimal*, one is led to introduce an entire new *family* of RRs, which these authors terms “*n*th-order transfer relations”. The 1st-order transfer relation is simply (70); the 2nd-order transfer relations are obtained by applying (70) to itself; and so on. The interested reader is referred to the original literature for the explicit forms of the first few transfer relations.

One is immediately led to ask why the *Flop-optimal* and *Mop-optimal* solutions should be so different and the answer, surprisingly, lies in the extreme simplicity of (70). Although it contains two Flops, it involves four Mops and this is an unhealthy balance: in essence, not enough “real work” is done between loading the right-hand-side and saving the left-hand-side. Instead, by the use of comparatively long *n*th-order transfer relations, we reduce the amount of memory traffic dramatically and, on many modern machines, the CPU cost falls. This is particularly revealing because the *Mop-optimal* solutions frequently boast *Flop* costs of disconcertingly large proportions.

4.6 Loop Structure of PRISM in Gaussian 92

Call **ListS2** to form list of significant shell-pairs

Call **SortS2** to sort shell-pair data by type

Call **CutoS2** to compute cutoff parameters for each shell-pair

Loop over LTot values

Call **TabGmT** to set up appropriate interpolation tables

Call **MakMD1** to make $[0]^{(m)} \rightarrow [r]$ driver

Loop over bra angular momentum types

Call several routines to form bra-transformation drivers

Loop over ket angular momentum types

Call several routines to form ket-transformation drivers

Call **FillAv** to decide which paths to make available

Call **MkDrTp** to make transposition drivers

Call **MakVR1** to make $[0]^{(m)} \rightarrow [m0ln0]$ driver

Call **MkDrPQ** to make scatter drivers

Call several routines to make $(0)^{(m)} \rightarrow (r)$ drivers

Call several routines to make contraction drivers

Call **MkCost** to compute PRISM step-costs in Flops and Mops

Loop over KBra values

Form petite list of bra shell-pairs of current type

Loop over KKet values

Call **Choose** to select the cheapest path

Call **PthInf** to determine info about the chosen path

Call **CalcSF** to compute bra- and ket-scalings

Compute maximum number of quartets per batch

Loop over batches of quartets of current type

Call **PickS4** to select a batch of quartets

Call **LoadS4** to form scaling and distance arrays

Call **CalcS4** to compute R2, T, Theta and U values

Call **Calc0m** to compute $[0]^{(m)}$ integrals

Call **DoMD1** to transform $[0]^{(m)} \rightarrow [r]$

Call **DoCont** to contract $[r] \rightarrow (r)$

Call **DoCont** to contract $(r) \rightarrow (r)$

Call **DoShuf** to scatter $(r) \rightarrow (plq)$

Call **DoTran** to transform $(plq) \rightarrow (bralq)$

Call **DoShuf** to transpose $(bralq)$

Call **DoTran** to transform $(bralq) \rightarrow (braket)$

Call **Gobbbxx** to digest $(braket)$'s

Call **Scatxx** to scatter Fock contributions etc.

4.7 Performance of PRISM in Gaussian 92

Notwithstanding the apparently strong *theoretical* arguments in support of the utility of the PRISM philosophy, our efforts are in vain unless it can be demonstrated that a real computer program, employing the algorithm, runs fast. In their paper on the HGP algorithm [55], Head-Gordon and Pople presented CPU timings on various computers which suggested that the HGP method was significantly faster than any other method which existed at that time (1988), with the possible exception of the PH axis-switch technique (Section 3.5). On that basis, we have systematically used the HGP method as a performance target while developing the PRISM methods and have published results [61] which establish clearly that the MD-PRISM algorithm is generally superior to the HGP algorithm. However, it was noted in the conclusions of [61c] that the MD-PRISM “is substantially *inferior* to HGP for weakly contracted classes of high angular momentum” and that “a modified version of PRISM that does not suffer from this defect needs to be developed”. We believe that, by incorporating paths on the new *HGP* PRISM, we now have an algorithm which is *uniformly and significantly* superior to all existing methods.

As an indication of the performance of our implementation [103] of PRISM in Gaussian 92, we have measured the CPU time for a single Hartree-Fock SCF iteration on a number of polyacenes. Our aim in making these measurements is to establish some well-defined benchmarks against which other programs in the future can be tested.

The precise specifications underlying our timings are:

- | | | |
|-----|----------------|---|
| (1) | Computer: | Dedicated IBM RS/6000 Model 320 (AIX) |
| (2) | Procedure: | Direct RHF-SCF (1st cycle only) |
| (3) | Basis Set: | 6-31G* |
| (4) | Accuracy: | 10 ⁻¹⁰ |
| (5) | Symmetry Used: | D ₂ |
| (6) | Initial Guess: | Projected INDO |
| (7) | Geometry: | C-C bond lengths = 1.4 Å
C-H bond lengths = 1.1 Å
All angles = 120 degrees. |

Table IV: Timings ^a for One Direct HF-SCF Cycle on Some Polyacenes				
Molecule	N ^b	Integrals ^c	Fock ^d	Total ^e
C ₆ H ₆	102	38 (61)	13	52 (74)
C ₁₀ H ₈	166	163 (269)	67	230 (336)
C ₁₄ H ₁₀	230	403 (652)	191	594 (843)
C ₁₈ H ₁₂	294	747 (1215)	371	1118 (1585)
C ₂₂ H ₁₄	358	1192 (1956)	627	1819 (2583)

- (a) In CPU seconds with HGP values in parentheses.
See text for detailed specifications.
- (b) Number of basis functions.
- (c) Time to construct all needed integrals.
- (d) Time to digest integrals into an RHF Fock matrix.
- (e) Sum of (c) and (d).

We may infer from the data in Table IV that the cutoff scheme used in PRISM is working satisfactorily. Even though none of the molecules considered could be considered large, the CPU times are already increasing much less rapidly than N^4 : between C₁₈H₁₂ and C₂₂H₁₄ the functional dependence is close to $N^{2.5}$.

It is also interesting to note that the fraction of the Total time which is associated with Fock matrix construction becomes larger in the larger systems. Presumably, this is because (due to our modeling scheme) the average degree of contraction is less in the larger molecules. This renders the integrals cheaper to form but has no effect on the Fock construction time.

5. PROSPECTS FOR THE FUTURE

The theory and practice of Molecular Integrals over Gaussian Basis Functions have come a long way since Boys' 1950 proposal. However, notwithstanding the enormous progress that has been made, there is little justification for the notion that we have reached the end of the road. Today, one can undertake Quantum Chemical calculations which would have been inconceivable only a decade ago and, without a doubt, computational chemists will be making the same observation a decade from now.

Advances in computer hardware will continue to catalyse and fuel the construction of novel software strategies. The advent of vector machines during the 1980s revolutionized the way in which integral programs were constructed and a second revolution is now underway in response to the possibilities afforded by the new Massively Parallel Processing (MPP) machines. The problem of effectively implementing PRISM within an MPP framework, which is presently under investigation in a number of groups, is a difficult one and it may eventually turn out that only radical departures from the conventional wisdom will yield efficient codes on the teraflop computers of the future.

Another, equally important, avenue to future developments involves the importation of ideas from other disciplines. This type of cross-fertilization – a ubiquitous ingredient in the advance of science – is apparent in the papers of Feibelman [104], Yang [105], Panas and Almlöf [106], Galli and Parrinello [107] and others. Indeed, some of these workers provide tantalizing evidence that the Holy Grail of Quantum Chemistry – a practical scheme for SCF calculations whose expense increases asymptotically only *linearly* with the system size – may be achieved in the rather near future.

Speculatively combining such hardware and software advances leads to the conclusion that accurate *ab initio* calculations on molecular systems with thousands of atoms may soon be routine and Quantum Chemistry will have come of age.

6. ACKNOWLEDGMENTS

It is a pleasure to acknowledge the contributions which many others have made to my research effort at Carnegie Mellon University.

For more than four years, John Pople has been friend and mentor to me. Like others before me, I have been greatly influenced by the seasoned pragmatism which underlies his approach to Quantum Chemistry.

In a similar vein, it was a number of conversations with Martin Head-Gordon, shortly after his discovery of the algorithm which now bears his name, that sparked my early interest in the theory of Molecular Integrals.

More recently, the careful work of Benny Johnson has been instrumental in elucidating how best to apply PRISM in the context of the two- and three-center problems. The extraordinary speed of the Electrostatic Potential program in Gaussian 92 is compelling testimony to his successes in this endeavor.

The greater part of what I know about constructing effective computer implementations of abstract algorithms I have learned from Mike Frisch and I am indebted to him for his advice over the years.

The MD-PRISM was first presented at the Argonne National Laboratory's "Methodology of the Evaluation of Integrals in LCAO Calculations" workshop in August 1990. The HGP-PRISM was first presented at the "PRISM" workshop in Connecticut in August 1992. I am grateful to Argonne and to Lorentzian Inc., respectively, for invitations to speak at these meetings.

My work at Carnegie Mellon has been funded under various National Science Foundation grants.

Finally, I thank my wife, Elizabeth, for her loving constancy, her encouragement and her typing of this manuscript.

7. REFERENCES

- [1] E. Schrödinger, *Ann. Physik* 79 (1926) 361.
- [2] (a) J.C. Slater, *Phys. Rev.* 34 (1929) 1293.
(b) J.C. Slater, *Phys. Rev.* 35 (1930) 509.
- [3] (a) D.R. Hartree, *Proc. Cam. Phil. Soc.* 24 (1928) 89.
(b) D.R. Hartree, *Proc. Cam. Phil. Soc.* 24 (1928) 111.
(c) D.R. Hartree, *Proc. Cam. Phil. Soc.* 24 (1928) 426.
- [4] V. Fock, *Z. Physik* 61 (1930) 126.
- [5] P. Hohenberg and W. Kohn, *Phys. Rev.* B136 (1964) 864.
- [6] W. Kohn and L.J. Sham, *Phys. Rev.* A140 (1965) 1133.
- [7] L.H. Thomas, *Proc. Cam. Phil. Soc.* 23 (1927) 542.
- [8] E. Fermi, *Rend. Accad., Lincei* 6 (1927) 602.
- [9] P.A.M. Dirac, *Proc. Cam. Phil. Soc.* 26 (1930) 376.
- [10] J.C. Slater, *Phys. Rev.* 81 (1951) 385.
- [11] (a) R. Pariser and R.G. Parr, *J. Chem. Phys.* 21 (1953) 466.
(b) R. Pariser and R.G. Parr, *J. Chem. Phys.* 21 (1953) 767.
(c) J.A. Pople, *Trans. Faraday Soc.* 49 (1953) 1375.
- [12] J.A. Pople and D.L. Beveridge, *Approximate Molecular Orbital Theory* (McGraw-Hill, New York, 1970)
- [13] R.C. Bingham, M.J.S. Dewar and D.H. Lo, *J. Am. Chem. Soc.* 97 (1975) 1285.
- [14] M.J.S. Dewar and W. Thiel, *J. Am. Chem. Soc.* 99 (1977) 4899.
- [15] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart, *J. Am. Chem. Soc.* 107 (1985) 3902.
- [16] J.J.P. Stewart, *J. Comput. Chem.* 10 (1989) 221.
- [17] A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Quantum Chemistry* (McGraw-Hill, New York, 1989).
- [18] The Kohn-Sham equations would indeed be equivalent to the Schrödinger equation if the universal exchange-correlation functional were known. Since it has not yet been discovered, practical forms of Kohn-Sham theory are (like Hartree-Fock theory) well-defined approximations to Schrödinger theory.
- [19] (a) P. Pulay, *Adv. Chem. Phys.* 69 (1987) 241.
(b) H.B. Schlegel, *Adv. Chem. Phys.* 67 (1987) 249.
(c) R. Fournier, J. Andzelm and D.R. Salahub, *J. Chem. Phys.* 90 (1989) 6371.

- [20] (a) J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, *Int. J. Quantum Chem., Symp.* 13 (1979) 225.
(b) R. Fournier, *J. Chem. Phys.* 92 (1990) 5422.
- [21] (a) S.M. Colwell, D. Jayatilaka, P.E. Maslen, R.D. Amos and N.C. Handy, *Int. J. Quantum Chem.* 40 (1991) 179.
(b) P.E. Maslen, D. Jayatilaka, S.M. Colwell, R.D. Amos and N.C. Handy, *J. Chem. Phys.* 95 (1991) 7409.
(c) P.E. Maslen, N.C. Handy, R.D. Amos and D. Jayatilaka, *J. Chem. Phys.* 97 (1992) 4233.
- [22] J.C. Slater, *Phys. Rev.* 36 (1930) 57.
- [23] W.J. Hehre, R.F. Stewart and J.A. Pople, *J. Chem. Phys.* 51 (1969) 2657.
- [24] H.J. Monkhorst and F.E. Harris, *Chem. Phys. Letters* 3 (1969) 537.
- [25] S.F. Boys, *Proc. Roy. Soc. (London)* A200 (1950) 542.
- [26] E.R. Davidson and D. Feller, *Chem. Rev.* 86 (1988) 681.
- [27] E. Clementi, *IBM J. Res. and Dev.* 9 (1965) 2.
- [28] (a) H. Preuss, *Z. Naturforsch.* 11a (1956) 823.
(b) J.L. Whitten, *J. Chem. Phys.* 39 (1963) 349.
(c) H. Sambe, *J. Chem. Phys.* 42 (1965) 1732.
(d) J.F. Harrison, *J. Chem. Phys.* 46 (1967) 1115.
(e) A.A. Frost, *J. Chem. Phys.* 47 (1967) 3707.
- [29] (a) E.A. McCullough Jr., *Chem. Phys. Letters* 24 (1974) 55.
(b) E.A. McCullough Jr., *J. Chem. Phys.* 62 (1975) 3991.
(c) E.A. McCullough Jr., *Comp. Phys. Rep.* 4 (1986) 265.
- [30] (a) L. Adamowicz and E.A. McCullough Jr., *J. Chem. Phys.* 75 (1981) 2475.
(b) L. Adamowicz, R.J. Bartlett and E.A. McCullough Jr., *Phys. Rev. Lett.* 54 (1985) 426.
- [31] (a) A.D. Becke, *J. Chem. Phys.* 76 (1982) 6037.
(b) A.D. Becke, *J. Chem. Phys.* 78 (1983) 4787.
(c) A.D. Becke, *Phys. Rev. A* 33 (1986) 2786.
- [32] (a) L. Laaksonen, D. Sundholm and P. Pyykkö, *Int. J. Quantum Chem.* 27 (1985) 601.
(b) L. Laaksonen, P. Pyykkö and D. Sundholm, *Comp. Phys. Rep.* 4 (1986) 313.
- [33] D. Heinemann, B. Fricke and D. Kolb, *Chem. Phys. Letters* 145 (1988) 125.
- [34] M. Defranceschi, M. Suard and G. Berthier, *Int. J. Quantum Chem.* 25 (1984) 863.
- [35] S.A. Alexander, R.L. Coldwell and H.J. Monkhorst, *J. Comp. Phys.* 76 (1988) 263.

- [36] (a) A.D. Becke and R.M. Dickson, *J. Chem. Phys.* 89 (1988) 2993.
(b) A.D. Becke, *Int. J. Quantum Chem., Symp.* 23 (1989) 599.
(c) A.D. Becke, *J. Chem. Phys.* 92 (1990) 3610.
(d) A.D. Becke, *J. Chem. Phys.* 96 (1992) 2155.
- [37] (a) V.I. Lebedev, *Zh. Vychisl. Mat. Mat. Fiz.* 15 (1975) 48.
(b) V.I. Lebedev, *Zh. Vychisl. Mat. Mat. Fiz.* 16 (1976) 293.
(c) V.I. Lebedev, *Sibirsk. Mat. Zh.* 18 (1977) 132.
(d) V.I. Lebedev, *Proc. Conf. Novosibirsk (1978)*, ed. S.L. Sobolev (Nauka Sibirsk. Otdel., Novosibirsk, 1980).
- [38] A.D. Becke, *J. Chem. Phys.* 88 (1988) 2547.
- [39] B. Delley, *J. Chem. Phys.* 92 (1990) 508.
- [40] D.R. Salahub, in *Density Functional Methods in Chemistry*, eds. J.K. Labanowski and J.W. Andzelm (Springer-Verlag, New York, 1991).
- [41] C.W. Murray, N.C. Handy and G.J. Laming, *Mol. Phys.* in press.
- [42] P.M.W. Gill, B.G. Johnson and J.A. Pople, *Chem. Phys. Letters*, submitted.
- [43] Overlap, kinetic-energy, nuclear-attraction and two-electron-repulsion integrals can all be written in this general form. The matrix elements of the exchange-correlation contribution to the energy in Kohn-Sham theory cannot, however. They are so awkward that a second basis set (usually of the Gaussian or Delta type) must be introduced to permit the computation. This, however, lies beyond the scope of our present Review.
- [44] J. Almlöf, K. Faegri and K. Korsell, *J. Comput. Chem.* 3 (1982) 385.
- [45] J. Andzelm, in *Density Functional Methods in Chemistry*, eds. J.K. Labanowski and J.W. Andzelm (Springer-Verlag, New York, 1991).
- [46] I. Shavitt, in *Methods in Computational Physics, Vol. 2*, eds. B. Alder, S. Fernback and M. Rotenberg (Academic Press, New York, 1963).
- [47] H. Taketa, S. Huzinaga and K. O-ohata, *J. Phys. Soc. Japan* 21 (1966) 2313.
- [48] E. Clementi and D.R. Davis, *J. Comput. Phys.* 1 (1967) 223.
- [49] (a) E. Clementi, H. Clementi and D.R. Davis, *J. Chem. Phys.* 46 (1967) 4725.
(b) E. Clementi, *J. Chem. Phys.* 46 (1967) 4731.
(c) E. Clementi, *J. Chem. Phys.* 46 (1967) 4737.
- [50] (a) J.A. Pople and W.J. Hehre, *J. Comp. Phys.* 27 (1978) 161.
(b) W.J. Hehre, W.A. Lathan, M.D. Newton, R. Ditchfield and J.A. Pople, Gaussian 70, Program No. 136, QCPE, Indiana University, Bloomington, Indiana.

- [51] (a) M. Dupuis, J. Rys and H.F. King, *J. Chem. Phys.* 65 (1976) 111.
(b) H.F. King and M. Dupuis, *J. Comput. Phys.* 21 (1976) 44.
(c) J. Rys, M. Dupuis and H.F. King, *J. Comput. Chem.* 4 (1983) 154.
- [52] L.E. McMurchie and E.R. Davidson, *J. Comput. Phys.* 26 (1978) 218.
- [53] (a) H.B. Schlegel, *J. Chem. Phys.* 77 (1982) 3676.
(b) H.B. Schlegel, *J. Chem. Phys.* 90 (1989) 5630.
- [54] (a) S. Obara and A. Saika, *J. Chem. Phys.* 84 (1986) 3963.
(b) S. Obara and A. Saika, *J. Chem. Phys.* 89 (1988) 1540.
- [55] M. Head-Gordon and J.A. Pople, *J. Chem. Phys.* 89 (1988) 5777.
- [56] K. Ishida, *J. Chem. Phys.* 95 (1991) 5198.
- [57] T.P. Hamilton and H.F. Schafer III, *Chem. Phys.* 150 (1991) 163.
- [58] I. Panas, *Chem. Phys. Letters* 184 (1991) 86.
- [59] (a) R. Lindh, U. Ryu and B. Liu, *J. Chem. Phys.* 95 (1991) 5889.
(b) U. Ryu, Y.S. Lee and R. Lindh, *Chem. Phys. Letters* 185 (1991) 562.
- [60] T. Helgaker and P.R. Taylor, *Theor. Chim. Acta* 83 (1992) 177.
- [61] (a) P.M.W. Gill, M. Head-Gordon and J.A. Pople, *Int. J. Quantum Chem., Symp.* 23 (1989) 269.
(b) P.M.W. Gill, M. Head-Gordon and J.A. Pople, *J. Phys. Chem.* 94 (1990) 5564.
(c) P.M.W. Gill and J.A. Pople, *Int. J. Quantum Chem.* 40 (1991) 753.
- [62] M.J. Frisch, B.G. Johnson, P.M.W. Gill, D.J. Fox and R.H. Nobes, *Chem. Phys. Letters*, submitted.
- [63] B.G. Johnson, P.M.W. Gill and J.A. Pople, *Chem. Phys. Letters*, submitted.
- [64] V.R. Saunders, in *Computational Techniques in Quantum Chemistry and Molecular Physics*, eds. G.H.F. Diercksen, B.T. Sutcliffe and A. Veillard (Reidel, Dordrecht, 1975).
- [65] V.R. Saunders, in *Methods in Computational Physics*, eds. G.H.F. Diercksen, and S. Wilson (Reidel, Dordrecht, 1983).
- [66] D. Hegarty and G. van der Velde, *Int. J. Quantum Chem.* 23 (1983) 1135.
- [67] D. Hegarty, in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*, ed. C.E. Dykstra (Reidel, Dordrecht, 1984).
- [68] L.L. Shipman and R.E. Christoffersen, *Comp. Phys. Comm.* 2 (1971) 201.
- [69] S.T. Elbert and E.R. Davidson, *J. Comp. Phys.* 16 (1974) 391.
- [70] F.E. Harris, *Int. J. Quantum Chem.* 23 (1983) 1469.
- [71] P.M.W. Gill, B.G. Johnson and J.A. Pople, *Int. J. Quantum Chem.* 40 (1991) 745.

- [72] (a) P.M.W. Gill, B.G. Johnson, J.A. Pople and S.W. Taylor, *J. Chem. Phys.* 96 (1992) 7178.
(b) P.M.W. Gill and J.A. Pople, in preparation.
- [73] (a) W. Kutzelnigg, *Theor. Chim. Acta* 68 (1985) 445.
(b) W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* 94 (1991) 1985.
- [74] A. Preiskorn and B. Zurawski, *Int. J. Quantum Chem.* 27 (1985) 641.
- [75] M.J. Bearpart, N.C. Handy, R.D. Amos and P.E. Maslen, *Theor. Chim. Acta* 79 (1991) 361.
- [76] W. Klopper and R. Röhse, *Theor. Chim. Acta* 83 (1992) 441.
- [77] J.S. Binkley, M.J. Frisch, D. J. Defrees, K. Raghavachari, R.A. Whiteside, H.B. Schlegel, E.M. Fluder and J.A. Pople, Gaussian 82, Carnegie-Mellon University, Pittsburgh (1984).
- [78] M.J. Frisch, J.S. Binkley, H.B. Schlegel, K. Raghavachari, C.F. Melius, R.L. Martin, J.J.P. Stewart, F.W. Bobrowicz, R.A. Whiteside, D.J. Fox, E.M. Fleuder and J.A. Pople, Gaussian 86, Carnegie Mellon Quantum Chemistry Publishing Unit, Pittsburgh, PA (1987).
- [79] One axis coincides with \overline{AB} another passes through Q , and the third is their mutual perpendicular.
- [80] One axis coincides with \overline{AB} , another is mutually perpendicular to \overline{AB} and \overline{CD} , and the third is their mutual perpendicular.
- [81] P.C. Hariharan, PhD thesis, Carnegie Mellon University (1973).
- [82] Hamilton and Schaefer [57] ascribe this idea to C. Murray.
- [83] A. Komornicki, K. Ishida, M. Morokuma, R. Ditchfield and M. Conrad, *Chem. Phys. Letters* 45 (1977) 595.
- [84] Gaussian 92, M.J. Frisch, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J.P. Stewart and J.A. Pople, Gaussian Inc., Pittsburgh PA, 1992.
- [85] Pairs of shells are required for most PRISM calculations but, for example, if overlap integrals are being calculated as outlined in Section 3.2.1, the "shell-pairs" are constructed by pairing each shell with a "dummy" ($\beta = 0$) shell.
- [86] D.L. Wilhite and R.N. Euwema, *J. Chem. Phys.* 61 (1974) 375.
- [87] J. Rys, H.F. King and P. Coppens, *Chem. Phys. Letters* 41 (1976) 383.
- [88] M. Yanez, R.F. Stewart and J.A. Pople, *Acta. Cryst. A* 34 (1978) 641.

- [89] (a) B.I. Dunlap, J.W.D. Connolly and J.R. Sabin, *J. Chem. Phys.* 71 (1979) 3396.
(b) B.I. Dunlap, J.W.D. Connolly and J.R. Sabin, *J. Chem. Phys.* 71 (1979) 4993.
- [90] (a) G.G. Hall and D. Martin, *Isr. J. Chem.* 19 (1980) 255.
(b) G.G. Hall, *Theor. Chim. Acta* 63 (1983) 357.
(c) G.G. Hall and C.M. Smith, *Int. J. Quantum Chem.* 25 (1984) 881.
(d) C.M. Smith and G.G. Hall, *Theor. Chim. Acta* 69 (1986) 63.
- [91] C. Van Alsenoy, *J. Comput. Chem.* 9 (1988) 620.
- [92] (a) A. Fortunelli and O. Salvetti, *J. Comput. Chem* 12 (1991) 36.
(b) A. Fortunelli and O. Salvetti, *Chem. Phys. Letters* 186 (1991) 372.
- [93] P.M.W. Gill, B.G. Johnson and J.A. Pople, *Chem. Phys. Letters*, submitted.
- [94] J.C. Maxwell, *A Treatise on Electricity and Magnetism* (Clarendon, London, 1904).
- [95] J.L. Whitten, *J. Chem. Phys.* 58 (1973) 4496.
- [96] J. Power and R.M. Pitzer, *Chem. Phys. Letters* 24 (1974) 478.
- [97] (a) R. Ahlrichs, *Theor. Chim. Acta* 33 (1974) 157.
(b) M. Haser and R. Ahlrichs, *J. Comput. Chem.* 10 (1989) 104.
(c) H. Horn, H. Weiß, M. Haser, M. Ehrig and R. Ahlrichs, *J. Comput. Chem.* 12 (1991) 1058.
- [98] B.G. Johnson, P.M.W. Gill, J.A. Pople and D.J. Fox, *Chem. Phys. Letters*, submitted.
- [99] Pairs of shell-pairs are required for most PRISM calculations but, for example, if nuclear-attraction integrals are being calculated as outlined in Section 3.2.4, the "shell-quartets" are constructed by pairing the shell-pairs with the infinite-exponent Gaussian at each of the nuclei.
- [100] For example, we might begin by forming batches of (sslss) quartets with $K_{\text{bra}} = K_{\text{ket}} = 1$ and, once these have all been treated, we might proceed with batches of (sslss) quartets with $K_{\text{bra}} = 2$ and $K_{\text{ket}} = 1$, and so forth. This approach, which is known as *extrinsic* vectorization, has been discussed by a number of authors. Our version is described in detail in [61b].
- [101] Eq. (87) is readily derived from (86) by examining the asymptotic expansion of the latter and it is the attractive simplicity of (87) which motivates the definition (74).
- [102] B.G. Johnson, P.M.W. Gill and J.A. Pople, *Int. J. Quantum Chem.* 40 (1991) 809.

- [103] The following paths were made available in these calculations: MD-CCTTT, MD-CTCTT, MD-TCCTT, MD-TCTCT, MD-TTTCC and HGP-TCCTT.
- [104] P.J. Feibelman, *J. Chem. Phys.* 81 (1984) 5864.
- [105] W. Yang, *Phys. Rev. Lett.* 66 (1991) 1438.
- [106] I. Panas and J. Almlöf, *Int. J. Quantum Chem.* 42 (1992) 1073.
- [107] G. Galli and M. Parrinello, *Phys. Rev. Lett.* 69 (1992) 3547.

Soliton Dynamics in trans-Polyacetylene

Wolfgang Förner

Chair for Theoretical Chemistry
Friedrich-Alexander University Erlangen-Nürnberg
Egerlandstr. 3
D-91058 Erlangen, FRG.

Contents

- Abstract
- 1. Introduction
- 2. Ab initio Calculations on trans-Butadiene
- 3. The Pariser Parr Pople Model
- 4. Parametrization
- 5. Static Properties
- 6. Dynamics
 - 6.1. Neutral Polymethine Chains
 - 6.2. Charged Polymethine Chains
 - 6.3. Singly Charged Polyenes
 - 6.4. Doubly Charged Polyenes
 - 6.5. Excited Polyenes
- 7. Conclusion
- Acknowledgement
- References

Abstract

The Pariser-Parr-Pople Hamiltonian for the description of the π -electrons in trans-polyacetylene is reparametrized using ab initio Coupled Cluster Doubles calculations based on a Restricted Hartree Fock reference on trans-butadiene. To avoid the spin contaminations inherent in Unrestricted Hartree Fock (UHF) type calculations on polymethine chains in the doublet state the Annihilated Unrestricted Hartree Fock (AUHF) model is applied in our PPP calculations (tPA: $(\text{CH})_\infty$, polyenes: $\text{H}-(\text{CH})_{2N}-\text{H}$, polymethines: $\text{H}-(\text{CH})_{2N+1}-\text{H}$). In geometry optimizations on polymethine chains it is shown that in contrast to results from Hückel type models the width of neutral solitons is strongly

reduced, in agreement to semiempirical all valence electron calculations (MNDO method). Further it is shown that the UHF model predicts artificially a metallic groundstate for polymethines due to spin contaminations, corresponding to an "infinite" soliton width. Negatively charged solitons have a larger width in agreement with MNDO and *ab initio* Hartree Fock results. Positively charged solitons are found to have the same width as negatively charged ones in disagreement with MNDO calculations, but in agreement with *ab initio* results. The photoinduced low and high energy absorptions can qualitatively be assigned to charged and neutral solitons. Within the new parametrization simulations of the dynamics of neutral polymethine chains, charged solitons, excited states, and doped chains are discussed. This work was presented at the First Congress of the International Society for Theoretical Chemical Physics, June 28-July 3, 1993, in Girona.

1. Introduction

The Su-Schrieffer-Heeger Hamiltonian (SSH) for trans-polyacetylene (tPA) is of simple Hückel-type [1]. For a recent review see [2]. However, a series of experiments has been carried out [3] showing results which require explicit inclusion of electron-electron interactions on Pariser-Parr-Pople (PPP) level [4] for their interpretation. Consequently molecular dynamics applying a first neighbor extended Hubbard Hamiltonian [5] and also the full PPP model have been studied [6]. These applications have shown that properties of solitons in tPA like width and kinetic mass change rather drastically compared to SSH results, emphasizing the necessity of explicit electron-electron interactions in the Hamiltonian. First molecular dynamics studies have been performed by Su and Schrieffer [7]. With an improved method for the gradient calculations their procedure was used in several applications for tPA in our Laboratory [8]. Analytical energy gradients for the SSH and the PPP Hamiltonian have been derived [9] and applied [6]. It turned out that the SSH Hamiltonian had to be augmented by a linear potential term to avoid lattice shrinking [10]. This was done e.g. by using an additional Lagrange multiplier [10]. Guinea [11] called it a "renormalization of parameters" and avoided it by keeping the chain ends fixed. The necessity for this term also in PPP models [6] rules out that its origin could be the missing electron-electron interaction as it was supposed previously. In a previous paper [12] we have shown that the necessity for linear terms originates simply from inconsistent assumptions in the original ansatz of SSH [1]. The potential ansatz was reformulated in a consistent way such that it contains only terms with a direct physical interpretation [12]. In order to compare results obtained with the SSH and PPP model Hamiltonian, especially potential curves for bond alternation with *ab initio*

potentials, trans-butadiene has been chosen as model system. For this molecule ab initio Hartree Fock (HF) and correlation calculations have been performed within atomic basis sets of increasing quality. It should be mentioned, that for trans-butadiene, hexatriene, longer polyenes and tPA a number of force fields applying ab initio and other methods have been determined in order to fit IR data (see [13,14] and references therein for some examples; tPA: $(\text{CH})_\infty$, polyenes: $\text{H}-(\text{CH})_{2N}-\text{H}$, polymethines: $\text{H}-(\text{CH})_{2N+1}-\text{H}$). A more recent fit to extensive, high quality correlated ab initio data specially for butadiene was published by Guo and Karplus [15]. In the present paper the σ - π separability of the correlation energy is investigated in detail. Also the effects of different energy contributions (one-particle without electron-electron interactions, full HF, correlation) on the bond alternation are calculated separately on ab initio level. We decided to give a rather lengthy discussion, because especially in solid state physical papers quantum chemical concepts were misinterpreted many times. Comparisons with semi-empirical descriptions are given and both the SSH and PPP Hamiltonian in their usual parametrization are shown to be unable to reproduce the ab initio potential curves. Mean field PPP excitation energies are discussed as function of the parametrization and compared to experiment as well as to theoretical results published previously. The excitation spectrum of trans-butadiene is discussed in detail. In the next part of this paper we describe our reparametrization of the PPP model, ground state soliton dynamics in tPA and a method which avoids spin contaminations. The original Hückel parameters of the SSH model [1] have been revised already by the authors themselves [2a]. However, we want to go beyond the Hückel level and also we do not want to include additional terms violating the zero differential overlap (ZDO) approximation (see [2b]). Thus a careful reinvestigation of the parameters seems to be necessary (see [16,17] for more details). Our reparametrization is based on the calculated dimerization potential of trans-butadiene (ab initio SCF plus correlation) and on optimized ground state geometries (ab initio SCF [14], double zeta basis) for $\text{C}_{21}\text{H}_{23}^+$, $\text{C}_{22}\text{H}_{24}$ and $\text{C}_{22}\text{H}_{24}^{2-}$. The excitation energies were used only as a general guideline (see below). Procedures like that for parametrization of semiempirical models have been advocated e.g. by Freed and co-workers [18-20]. Especially [20] deals with parameters for trans-butadiene and its excitation spectrum. Using the annihilated unrestricted Hartree Fock (AUHF) method we performed dynamical studies on the system. The AUHF method is used because it avoids the spin contamination inherent in UHF models [21]. In recent work we have applied this method to trans-polyenes and found travelling solitons using the conventional parameter sets. Also against disorder in the system these solitons remained stable [22]. The AUHF method leads to pure spin states and works well around equilibrium structures. However, it was shown for small molecules that in dissociation regions it approaches the behaviour of restricted HF (RHF) and thus dissociates wrong ($\text{A}_2 \rightarrow \text{A}^+ + \text{A}^-$ instead of $\text{A}_2 \rightarrow 2\text{A}$) [23]. The atomic

displacements occurring in soliton dynamics in trans-polyenes are rather small and remain in the region where AUHF is superior over UHF. In the final part of this paper we present soliton dynamics in neutral and charged polymethine chains, in disordered chains, excited polyenes and in polyenes doped with one or two electrons or holes, respectively, applying the derived set of parameters. In [12] as well as in [22] we have shown that the soliton width in small chains is a function of the chain length in SSH and UHF-PPP models as it was proposed later on the basis of more sophisticated methods [24] and of experiments [25]. Therefore we also study in static calculations the soliton properties in chains of different length.

2. Ab Initio Calculations on trans-Butadiene

The ab initio HF calculations reported below have been performed with the GAUSSIAN 76 [26] program package. The atomic basis sets applied are a minimal (STO-3G [26]) one, a split valence (6-31G [26]) one, a split-valence one plus a set of five d-functions on carbon (6-31G* [26]), and one with an additional set of p-functions on hydrogen (6-31G** [26]). The correlation energy has been computed using Møller-Plesset many body perturbation theory of second order (MP2) [27], the linear approximation of Coupled Cluster Doubles theory (L-CCD) [28] and full CCD [28]. The program package used is described in the literature [29]. Since this package was designed for application to large molecules which exhibit in general no point group symmetry except C_1 and uses localized orbitals, it takes no advantage of molecular symmetry. Therefore, in the largest calculation reported here (trans-butadiene, 6-31G basis set, only 1s core-orbitals of carbon frozen) ≈ 66000 independent excitations (≈ 30000 for the σ - σ and 78 for the π - π correlation) had to be treated. References for other implementations and the numerous applications of the theory (also going beyond CCD level) can be found in Ref. [29].

The geometrical model applied for trans-butadiene in this work is the same as the one applied usually in studies on polyene chains [1]. In dynamic simulations the coordinates perpendicular to the chain axis ($\pm y_o$) of the CH units are usually kept constant, and only the coordinates u_i parallel to the axis are varied. In this way for an ideally dimerized chain with the bond alternation $u_o = 0.037 \text{ \AA}$ [1] one obtains 1.336 \AA for the short CC-bond (r_1), 1.465 \AA for the long CC-bond (r_2), and 119.9° for the CCC-bond angle φ . This is in reasonable agreement with experiment for trans-butadiene ($r_1 = 1.35 \pm 0.02 \text{ \AA}$, $r_2 = 1.46 \pm 0.03 \text{ \AA}$, $\varphi = 122^\circ$ [30]).

First of all we want to discuss some general trends, concerning certain contributions to the bond alternation in polyenes. Trans-butadiene is a very small molecule, compared to long polyenes or even infinite tPA. However, our calculations, as well as those on tPA published previously (see below) exhibit the same general trends in both systems, although in butadiene we certainly have to face large end effects. Due to the small conjugated segments in tPA and also to

the relatively short polyenes studied so far, it seems to be somewhat questionable to introduce cyclic boundary conditions. Moreover, studies on benzene and larger $4n+2$ rings like in [31] introduce aromaticity effects in addition which are not present in linear polyenes or in tPA. Thus despite end effects we prefer to concentrate on a linear system, and as pointed out above our results on butadiene show the same general trends as published previously on tPA and longer polyenes.

In tPA since the works of Suhai [32] and Liegener [33] who used an improved Greens function approach instead of the direct perturbational one it is well known that the ab initio HF-approximation in its basis set limit overestimates the bond alternation. Inclusion of $\approx 70\%$ of the correlation energy reduced the dimerization to its experimental value [32]. Correlation energy in this sense is defined as the energy difference between the exact solution of the Schrödinger equation and the exact HF (basis set limit) solution, both in Born-Oppenheimer (adiabatic) approximation. Semiempirical PPP configuration interaction (CI) studies [34] lead to the conclusion that with increasing size of polyenes, the order of excitations (double, triple, quadruple, etc.) necessary to describe the system increases. However this behaviour is well known as an artefact of limited CI-expansions, called size-inconsistency. Thus the conclusions drawn in Ref. [34] are due rather to an artefact of the method applied than to a real physical importance of high order excitations in these systems. Although FCI results discussed in [34] are exact solutions of the PPP Hamiltonian, comparisons of truncated CI results like CISD to FCI do not tell too much about the physical importance of higher order excitations. Due to the size inconsistency of truncated CI expansions FCI contains e.g. certain unphysical contributions from double and from quadruple excitations which cancel each other exactly. However, in CISD the size inconsistent part of the doubles contribution cannot be canceled, because the quadruples are missing, while in FCI the cancelation occurs. The two unphysical size inconsistent contributions from doubles and quadruples, which in fact cancel each other, cannot be separated from the remaining size consistent parts in FCI or in CISD. Thus comparisons may lead to the conclusion that higher order excitations are even important for a correct description of the ground state [34], what is only true in CI expansions and an artefact of the method. On the other hand, in MP2 for example (and also in CC expansions) the size inconsistent contributions from the double excitations are excluded from the beginning by virtue of the linked cluster theorem, as well as in higher orders of MP the corresponding ones from quadruple excitations which only would cancel with the former exactly. In fact MP is size consistent in any individual order. Therefore in the MP series, as well as in CC expansions no artificial necessity of higher order excitations appears, as it happens in CI expansions. The MP2 [32] and MP3 [33] results on the gap of infinite tPA are a nice example of this fact. Indeed Suhai [32] found the experimental value for the dimerization in his calculations and

there is no reason why higher order excitations should lead to another result which would contradict experimental findings. Thus the conclusion reached by Suhai [32] and emphasized by Wendel et al. [35] that the dimerized state is the HF-ground state of tPA and that correlation reduces the alternation is certainly justified.

The opposite conclusion that correlation increases the dimerization reached in PPP valence bond (VB) theory [36] holds only in comparison to Hückel theory which is not the semiempirical equivalent to ab initio HF. In fact Hückel theory would correspond to diagonalization of the one-electron part of the ab initio Fock matrix rather than to HF itself although in Hückel theory correlation may appear implicitly via the parameters. And indeed, if one considers only the one-electron part of the ab initio Fock matrix (core diagonalization) in a 6-31G** basis set the equidistant geometry for trans-butadiene is favored. Inclusion of electron-electron interactions on one-particle level (HF), which corresponds to PPP theory yields an increase of the bond-alternation beyond its experimental value. The inclusion of π - π correlations (CCD-level) decreases this value again to the experimental one [16].

Therefore the effects of the different energy contributions are the same in trans-butadiene as in tPA [32]. The corresponding semiempirical models show exactly the same qualitative behaviour. In these types of models through adjustment of the parameters relevant for the σ bonds the experimental bond alternation can be fitted already on Hückel-level. This is discussed in our previous paper [16] on SSH, PPP and PPP/CCD level. Note that in these calculations we used a parametrization for the resonance integrals which differs from the SSH one. Namely, we applied an exponential ansatz for these integrals $B_n = \beta_0 \exp(-AR_n)$ [37-39] (M/PPP). On Hückel level (corresponding to ab initio core diagonalization) the energy-minimum occurs trivially for $u/u_0 = 1$. In PPP theory (corresponding to ab initio HF) the dimerization is strongly enhanced and again reduced by inclusion of correlation. Therefore the conclusion reached by Kivelson et al. [40] that electron-electron interactions on one-particle level decreases dimerization is certainly not valid on ab initio level and an artefact of their inconsistent inclusion of non-ZDO types of integrals. If the constants for the σ bonds as calculated on PPP level are used for all three models the above described behaviour is even more obvious. Therefore there is no discrepancy concerning the qualitative influence of electron-electron interactions between ab initio and semiempirical models, provided that the really corresponding models are compared. Hückel theory is certainly not the semiempirical counterpart to ab initio HF.

Suhai [32] found that only 15-20% of the correlation energy he obtained for tPA is due to π - π correlations. On this basis he concluded that σ - π separation does not work in tPA. However in order to investigate the reliability of σ - π separation in molecular dynamics studies on polyenes the influences of the

different types of correlation involved (σ - σ , σ - π , π - π) on the total energy as function of geometry must be studied. To be able to use reasonably large basis sets in ab initio CCD calculations, trans-butadiene has been chosen as model system. Concerning the importance of the different kinds of excitations on potential hypersurfaces end effects should be unimportant. The σ - σ and π - π contributions to the correlation energy have been computed directly by freezing the orbitals of other types. The σ - π contribution is then obtained by subtracting these quantities from the result of a frozen core calculation.

Table 1: Parameters of a parabola $V(x)=\frac{1}{2}k(x-x_0)^2+E_0$ ($x=u/u_0$, $u_0=0.037$ Å) adjusted to the three calculated values of the energy ($x=0.5, 1, 1.5$) relative to its value at $x=0$ ($V(0)$ as calculated from the parabola, the exact value is $V(0)=0$) for different methods.

Method	x_0	E_0 (eV)	$V(0)$ (eV)
HF/CCD/STO-3G ¹⁾	1.126	-0.246	-0.0004
HF/CCD/STO-3G ²⁾	1.145	-0.249	+0.0009
HF/CCD/6-31G ¹⁾	1.012	-0.178	-0.001
HF/CCD/6-31G ²⁾	1.009	-0.173	-0.002
HF ³⁾	1.386	-0.339	-0.001
HF/MP2 ³⁾	1.250	-0.274	+0.001
HF/L-CCD ³⁾	0.997	-0.155	+0.002
HF/CCD ³⁾	1.098	-0.199	-0.001

1) frozen core calculations

2) only π - π correlations considered

3) 6-31G** basis set, only π - π correlations considered

The different energy contributions from the three methods (MP2, L-CCD, and CCD) were reported by us [16] for trans-butadiene as function of u/u_0 in a minimal basis and a split valence basis. The qualitative features turned out to be the same for the two different basis sets. Surprisingly there is a qualitative difference between MP2 on one hand and L-CCD and CCD on the other. In both cases the σ - σ and σ - π correlation are responsible for most of the total correlation energy. But in L-CCD and CCD the variation of the correlation energy with u/u_0 is almost exclusively due to π - π correlations. On the contrary in MP2 the variation of the σ - π contribution is even larger than that of the π - π correlation energy. Most probably this is an artefact of the low order of MP theory applied. The π - π correlation energies obtained in 6-31G* and 6-31G** basis set are almost identical implying that p-functions on the hydrogens do not influence the π - π

correlation. The results obtained with a PPP reference function are qualitatively similar to the ab initio ones, although the π - π correlation energies are smaller in the semiempirical case [16]. We have adjusted parabolas to the three energy values actually calculated (relative to the equidistant molecule) with the different methods.

In Table 1 we show the parameters of these adjusted parabolas. Clearly enough for a minimal basis set the σ - π separation is not justified. However, the parameters for the split valence basis are practically the same for frozen core

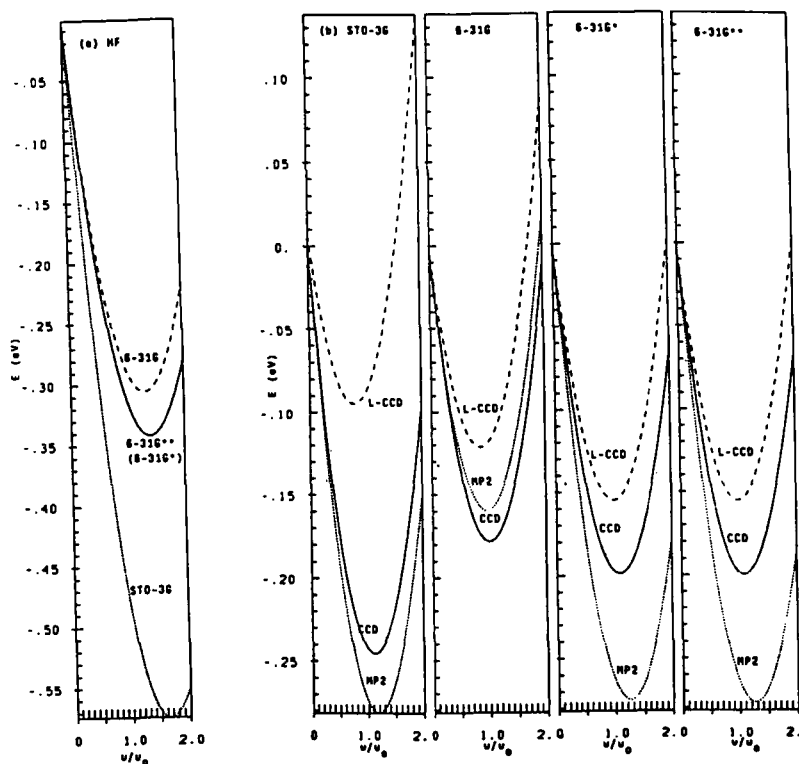


Fig. 1: Total ab initio energies (relative to the equidistant geometry) of trans-butadiene calculated with different basis sets as function of the dimerization

- (a) Hartree Fock
- (b) Hartree Fock plus correlation energy (MP2, L-CCD, CCD; in STO-3G and 6-31G basis only core orbitals are frozen; in 6-31G* and 6-31G** basis also σ orbitals are frozen).

calculations and for those considering only π - π excitations. Though σ - σ and σ - π correlations give rise to the largest contributions to the correlation energy in agreement with Suhai [32] their variation with bond alternation is negligible.

In Fig. 1 the total energies obtained with the different methods and basis sets are shown as function of u/u_0 , relative to the energies at $u=0$. The minimal basis set overestimates the energy difference between $u=u_0$ and $u=0$ (Fig. 1b). In the split valence case (Fig. 1b) the minimum for CCD occurs exactly at $u/u_0=1$, while the HF-minimum (Fig. 1a) occurs for all three basis sets beyond $u/u_0=1.5$. The results for 6-31G* and 6-31G** are almost identical (Fig. 1b) and the energy minimum occurs at ≈ -0.20 eV (CCD) and $u/u_0=1$, energetically somewhat below the split valence basis. The L-CCD results seem to be rather unreasonable being energetically as low as the corresponding minimal basis ones. MP2 shows the same minimum as CCD but at somewhat higher energies. The curves shown in Fig. 1 are drawn by adjusting a parabola to the actually calculated points ($u/u_0=0.5, 1, 1.5$) as above and calculating additional points from it. With the usually used parameters the corresponding PPP curves do not match the ab initio one (valence split plus polarization basis set, CCD level) and thus a reparametrization is necessary.

For the first four valence excitation energies due to $\pi \rightarrow \pi^*$ excitations, values obtained with different methods and by experiment for trans-butadiene are given in Table 2. For the two triplet states ab initio results agree reasonably well with experiment. The PPP-FCI [34] (Full CI) results are too low by roughly 0.3-0.4 eV. SSH/PPP in HF approximation places them too high. For the 1B_u state all PPP results are within the experimental range. The older ab initio results in this case are all too high compared to the accuracy with which other states are reproduced. Buenker et al. [45] discuss a possible non vertical effect in the experimental spectrum. The lowering of the excitation energy upon non vertical excitation obtained by Nascimento et al. [47] points also into this direction. If this is true, then the success of PPP-methods for this state might be accidental since PPP assumes planarity of the excited states. A matter of controversy is the 2^1A_g state. PPP-SCF and ab initio calculations [45-47], as well as an assignment by McDiarmid [43] place it around 7 eV. However, more recent measurements by Chadwick et al. [44] show a state of 1A_g symmetry below 1B_u at 5.4-5.8 eV. PPP-FCI [34] yields valence 2^1A_g below 1B_u while older high quality ab initio calculations put the state above 1B_u [45-47]. However, for polyenes larger than hexatriene fluorescence from the 1B_u state indicates 2^1A_g below 1B_u . Unfortunately high quality ab initio calculations have not been performed for these large systems. PPP-FCI and -CISD [34,49] reproduces this state ordering but results in excitation energies of about 1 eV too small compared to recent multireference CI calculations [50]. However, for butadiene and hexatriene no fluorescence from 1B_u has been found yet [51]. For octatetraene and higher polyenes there is agreement, that 2^1A_g is below 1B_u [51]. Octatetraene in alkanes is planar in its excited states

[52]. However, for butadiene and trans-hexatriene even the ordering of the excited states is still not established [53-55]. For cis-hexatriene the 2^1A_1 excitation was observed below 1B_2 [54] and the molecular geometry in the excited 2^1A_1 state was

Table 2: Singlet and triplet valence π - π^* excitation energies in trans-butadiene (in eV).

Method	3B_u	3A_g	1B_u	2^1A_g
Experiments	3.22-3.3	4.8-4.93	5.76-6.29	7.06
References:	[41]	[41]	[42]	[43]
Ref. [44]				5.4-5.8
PPP-HF ¹⁾	3.65	6.23	5.97	7.92
HF-CI [45]	3.31	4.92	7.67	7.02
MCSCF [46]	3.36	5.06	6.88	6.77
GVB-CI [47] ²⁾	3.35	5.08	6.79	7.06
GVB-CI [47] ³⁾			6.47	
PPP-FCI [34] ⁴⁾	2.92	4.59	6.01	5.82
RPA [48] ⁵⁾			6.38	
PPP-CISD [49] ⁶⁾			5.96	5.65
ab initio Multireference CI [50]				
A ⁷⁾			6.74	6.82
B ⁸⁾			6.50	5.55
C ⁹⁾			4.97	5.26

- 1) SSH parameters, $\gamma_0=11.08$ eV
- 2) Generalized Valence Bond Model, vertical excitation
- 3) as 2) but 1B_u calculated non-vertical
- 4) full CI, Ohno formula, $\gamma_0=11.13$ eV
- 5) Random Phase Approximation, ab initio including diffuse spd function sets at the center of mass.
- 6) Doubles CI, vertical excitation
- 7) vertical planar to planar excitation
- 8) minimum to minimum planar excitation
- 9) minimum to minimum non-planar excitation

found to deviate slightly from planarity due to rotations of the terminal CH_2 groups [55]. In trans-butadiene the question of excited state geometries is unsolved. A Frank-Condon analysis of absorption and Raman spectra for the 1B_u excitation lead to the conclusion that the state is essentially planar [56,57]. Absorption spectroscopy results (in jet-cooled samples) provided no evidence for a substantially more distorted 1B_u state in butadiene than in octatetraene [58].

However, McDiarmid and Sheybani [59] found indirect support for a twisting of the terminal CH_2 groups in trans-butadiene. Recently, high quality multireference CI (MRCI) results have been published [50] (see Table 2). There it was found that for planar structures 2^1A_g is below 1B_u . However, both excited states are shown to prefer a distorted geometry with rotated terminal CH_2 groups. In case of non planar structures the ionic excited state (1B_u if the molecule is planar) is below the covalent one (2^1A_g if planar). Since the measured excitation energies of 1B_u are in the range 5.76-6.29 eV [42] one might assume that a more or less planar to planar excitation occurs (6.50-6.74 eV [50]) with subsequent geometry relaxation, since direct excitation to the distorted $1^1A'$ state would require only ≈ 5 eV [50]. If the forbidden excitation at 5.4-5.8 eV is due to the 2^1A_g state [44] the finding of this state below 1B_u would also point into this direction. However, from the above discussion it is clear that the problem of the excited states of butadiene is still unsolved, and for a parametrization of HF-PPP models the excitation energies are only of limited use, especially that for 2^1A_g which is a true more-determinantal state [50]. However, keeping these limitations in mind, we discussed in our previous paper the excitation energies for the four states obtained by the HF-PPP method as function of γ_0 , the parameter which determines the strength of the electron-electron interaction. We could not find a unique parameter set which would place all the states qualitatively into the correct regions.

3. The Pariser Parr Pople Model

In simulations on longer chains instead of the above mentioned physical displacements u_i (relative to the equidistant chain) usually the staggered coordinates ψ_i are used to construct the Hamiltonian [1]. These are defined as:

$$\psi_i = (-1)^{i+1} u_i \quad (1)$$

The ideally dimerized chain is then described by

$$u_i = (-1)^{i+1} u_0 \Rightarrow \psi_i = u_0 \quad (2)$$

The SSH [1] as well as the PPP-Hamiltonian [6] for tPA are described in the literature. However, since we discuss parameter dependences a short review seems to be adequate to define all parameters. The total energy of a chain is given as:

$$E = E_\sigma + E_\pi \quad (3)$$

where the energy due to the σ electrons enters through a potential and E_π is calculated on SSH (Hückel type) or PPP level. The Hückel matrix is defined by the resonance integrals β_{rs}^{SSH} :

$$H_{rs}^{SSH} = \beta_{rs}^{SSH} (1 - \delta_{rs}) \quad (4)$$

where r and s run over all CH-sites. The diagonal elements are shifted to zero since only CH groups are involved. Then the total π -electron energy in Unrestricted Hartree Fock (UHF) approximation can be written as

$$E_{\pi} = \sum_{r,s=1}^N H_{rs}^{SSH} (P_{rs}^{\alpha} + P_{rs}^{\beta}) \quad (5)$$

where N denotes the number of CH-units considered, α and β the two spin orientations and \underline{P}^{σ} are the charge density bond order matrices ($\sigma = \alpha$ or β):

$$P_{rs}^{\sigma} = \sum_{i=1}^N o_i^{\sigma} c_{ri}^{\sigma} (c_{si}^{\sigma})^* \quad (6)$$

Here c_n^{σ} is the coefficient of the basis-function at site r in the σ -spin MO i and o_i^{σ} is the occupation number of this MO (0 or 1). These are obtained by diagonalization of \underline{H}^{SSH} :

$$\underline{H}^{SSH} \underline{c}_i^{\sigma} = e_i^{\sigma} \underline{c}_i^{\sigma} \quad (7)$$

In the Hückel-case the \underline{c}_i^{σ} , e_i^{σ} are identical for the two spin orientations, only the o_i^{σ} differ in doublet, triplet, etc. cases.

In SSH models [1] the β_n^{SSH} are non-vanishing only for nearest neighbors ($s=r\pm 1$) and expanded up to first order around the equidistant chain values.

$$\beta_{rs}^{SSH} = \beta_{r,r+1}^{SSH} (1 - \delta_{rN}) \delta_{s,r+1} + \beta_{r,r-1}^{SSH} (1 - \delta_{r1}) \delta_{s,r-1} \quad (8)$$

for open chains. Further [1]

$$\beta_{r,r\pm 1}^{SSH} = \beta_0 \pm (-1)^r (\psi_r + \psi_{r\pm 1}) \alpha \quad (9)$$

Therefore the electronic part of the SSH Hamiltonian contains two adjustable parameters, namely β_0 and the so-called electron-phonon coupling constant α . In original SSH-theory they are chosen such, that for an infinite ideally dimerized chain the π -band width and the single particle gap of tPA are reproduced, leading to $\beta_0 = -2.5$ eV and $\alpha = 4.1$ eV/Å [1].

In PPP-theory [4,60a,61] electron-electron interaction is introduced via the two-electron integrals γ_n applying the zero differential overlap (ZDO) approximation. In ZDO approximation all integrals containing products of basis-functions which are not centered at the same site are assumed to vanish. This

holds only if the basis functions are assumed to be orthogonalized linear combinations of the $2p_{\pi}$ -AO at the sites [60b]. The Fock matrix for spin σ (with $\lambda=\alpha$ if $\sigma=\beta$ and vice versa) is then given by [4,60a]:

$$F_{rs}^{\sigma} = [-I + \gamma_0 P_{rr}^{\lambda} + \sum_{t=1}^N (P_{tt}^{\alpha} + P_{tt}^{\beta} - z_t) \gamma_{rt} (1 - \delta_{rt})] \delta_{rs} + H_{rs}^{SSH} - P_{rs}^{\sigma} \gamma_{rs} (1 - \delta_{rs}) \quad (10)$$

The actual value of I is unimportant for our purpose but usually the ionization potential (I_C) of carbon is used for I . γ_0 is the on-site Coulomb repulsion. In the usual parametrization [60a] γ_0 is given by $I_C - A_C$, where A_C is the electron affinity of carbon. This leads to $\gamma_0 = 11.08$ eV [60a]. Upon variation of γ_0 we will simply scale I according to γ_0 ($I = I_C \gamma_0 / 11.08$ eV). Thus we do not consider I as independent parameter since it leads only to a geometry independent shift of total energy and the energy level spectrum. z_i is the charge of the ionic cores after removing all π -electrons from a site (+1 in case of carbon). The two-electron integrals are approximated by the Ohno formula [62]

$$\gamma_{rs} = b / \sqrt{(b/\gamma_0)^2 + R_{rs}^2}; \quad b = e^2 / 4\pi\epsilon_0 = 14.3989 \text{ eV}\text{\AA} \quad (11)$$

where e is the elementary charge, ϵ_0 the dielectric constant of the vacuum and R_{rs} the distance between sites r and s .

The \underline{c}_i^{σ} , necessary to compute \underline{P}^{σ} are obtained by self-consistent solution (SCF) of the (unrestricted PPP) coupled eigenvalue equations

$$\underline{F}^{\sigma} \underline{c}_i^{\sigma} = \epsilon_i^{\sigma} \underline{c}_i^{\sigma} \quad (12)$$

With the one-electron Hamiltonian matrix

$$H_{rs}^N = [-I - \sum_{t=1}^N z_t \gamma_{rt} (1 - \delta_{rt})] \delta_{rs} + H_{rs}^{SSH} \quad (13)$$

the total energy is given by

$$E_{\pi} = \frac{1}{2} \sum_{\sigma} \sum_{r,s=1}^N \langle F_{rs}^{\sigma} + H_{rs}^N \rangle P_{rs}^{\sigma} + \frac{1}{2} \sum_{r,s=1}^N z_r z_s \gamma_{rs} (1 - \delta_{rs}) \quad (14)$$

where the second term represents the Coulomb repulsion of the ionic cores. Thus the electronic part of the PPP-Hamiltonian contains three adjustable parameters, namely β_0 , α , and γ_0 . The energy due to the σ -electrons is given by a harmonic potential for each CC-bond

$$E_{\sigma} = \frac{1}{2} K \sum_{r=1}^{N-1} (R_{r,r+1} - R_0)^2 \quad (15)$$

where R_0 is the equilibrium σ -bond length. K and R_0 are determined by requiring that the ideally dimerized chain is the global minimum of the total energy $E_{\pi} + E_{\sigma}$. The formulas involved are given in detail in Ref. [12].

An extension of the PPP-Hamiltonian was suggested by Kivelson et al. [40]. They studied the influence of two-electron integrals which are neglected in the ZDO-approximation. However, as mentioned above Löwdin has shown (see e.g. Ref. [60b]) that these integrals are indeed negligible if the basis functions are considered as orthogonalized linear combinations of the $2p_{\pi}$ functions of carbon. Therefore a straightforward inclusion of integrals of this kind having a magnitude larger than 1/3 of the usually considered integrals (γ_{π}) as done in Ref. [40] should lead to inconsistencies. In this case first of all the γ_{π} should be changed due to a basis transformation and secondly inclusion of these integrals would require explicit introduction of basis set overlap (since ZDO cannot be considered as a valid approximation). Further these new correction terms [40] had been already discussed and found to be negligible in the early literature on PPP and related (Hubbard) models [61]. Also in more recent literature the suggestion to include such terms is criticized (see [63] for some examples). Recently, Vogl and Campbell [64] found on the basis of first principles LDA calculations that no charge builds up between two carbons linked by a double bond and thus concluded, that terms like bond charge repulsion are insignificant in agreement with the ZDO approximation [61]. Indeed Kivelson et al. [40] found a decrease of Peierls distortion compared to SSH upon inclusion of electron-electron interactions in this way. As ab initio results discussed above show, this finding seems to be unreasonable.

The UHF ansatz is necessary because in case of neutral solitons one has to deal with a doublet state. Thus a DODS (different orbitals for different spins) ansatz, as the UHF one, is necessary to describe the system. However, in the UHF method described so far, one Slater determinant with different spatial orbitals for electrons of different spins is applied, which is not an eigenfunction of \hat{S}^2 , i.e. $\langle \hat{S}^2 \rangle \neq S(S+1)\hbar^2$. The best way to overcome this difficulty would be to use the PHF (Projected Hartree Fock) method, also called EHF method (Extended Hartree Fock) where before the variation the correct spin eigenfunction is projected out of the DODS ansatz Slater determinant [66,67a]. Unfortunately numerical solution of the rather complicated EHF equations in each time step seems to be too tedious at present. Moreover for large systems the EHF wavefunction approaches the UHF one [68], however, this might be due to the approximations used in [67a]. Another possibility is to apply the projection after the variation using again Löwdin's projection operator [66]. Projection and annihilation techniques were

also applied to the correlation problem in the MP framework [67b], where in the work of Handy and Knowles (see [67b]) also the idea of annihilating just the first contaminating component was developed. The general idea to recover correlation partially by a DODS ansatz, e.g. also with the alternant molecular orbital (AMO) method goes back to Löwdin and was developed further by others [67c]. As a review on these methods and a source of further references the paper by Mayer [67d] is recommended. However, for our purpose, i.e. the simulation of the dynamics of larger polymethine and polyene chains, these techniques appear to be computationally too tedious and we had to look for a simpler, but still accurate approach, especially since our PPP model contains correlation effects implicitly already via the parametrization. Thus we decided to apply the AUHF method [21] (Annihilated Unrestricted Hartree Fock). In this method the lowest contaminating spin multiplicity is annihilated from the density matrix in each SCF iteration cycle. In this way one obtains a wave function with $\langle \hat{S}^2 \rangle$ close to the correct value (e.g. $0.7501\hbar^2$ instead of $(3/4)\hbar^2$ for a doublet state [22]). To study the dissociation behaviour of AUHF we considered a system of two units [17]. The resonance integral β between the units is chosen as

$$\beta = \beta' e^{-a(R-R_0)} \quad (16)$$

With $\beta' = -2.5$ eV, $R_0 = 1.4$ Å and $a = 1.46$ Å⁻¹ we obtain around R_0 roughly the same (linearized) behaviour as in the SSH parametrization. The two electron integral γ was as usually calculated using Ohno's formula. The on-site term is $\gamma_0 = 11.08$ eV and the ionization potential $I = 11.5$ eV. First of all we obtain a symmetry adapted RHF solution with a density matrix $P_{\alpha\beta} = 1$. In this case the total energy is (including repulsion γ of the ionic cores)

$$E_{RHF} = -2I + \frac{1}{2}\gamma_0 + 2\beta - \frac{1}{2}\gamma \quad (17)$$

For $R \rightarrow \infty$ this leads to $E_{RHF}(\infty) = -2I + (1/2)\gamma_0 = -17.46$ eV. Using a symmetry breaking UHF solution where we have an α spin on one atom and a β spin on the other we find that UHF (although the guess density matrix is asymmetric) converges back to (17) up to $R = 1.54$ Å corresponding to $2.5 u_0$ ($u_0 = 0.025$ Å). Such a large value never occurred in our PPP simulations. Increasing R further from 1.54 Å the UHF solution differs from the RHF one and converges to $E_{UHF}(\infty) = -2I = -23$ eV. In this case we have $\langle \hat{S}^2 \rangle = 1\hbar^2$ instead of zero as required for a singlet.

At $R = 5.14$ Å also in RHF a symmetry breaking occurs where on one site a charge of e and one of $-e$ on the other occurs. Linear combination of the two possible solutions of this kind leads to the symmetry adapted HF wave function. As calculations up to $R = 100$ Å show this solution converges also to -23 eV. The

AUHF curve cannot be distinguished from the symmetry adapted RHF curve. This coincides with ab initio results on LiH [23] where it was found that around equilibrium AUHF follows UHF but it dissociates in the wrong way as RHF ($A_2 \rightarrow A^+ + A^-$). In our PPP case we obtain no level splitting around R_0 and thus around equilibrium RHF, UHF and AUHF are identical. Note that in the corresponding potential curves [17] the minimum would occur only if the σ -electron potential would be included. However, the wrong behaviour of AUHF appears in bond length regions which do not occur in our dynamical simulations. In case of a three site system (symmetric stretch, doublet state) up to roughly 1.6-1.7 Å AUHF is much closer to UHF than to RHF. In this case for construction of the RHF density matrix occupation numbers were 2 for the lowest bonding orbital and 1 for the non-bonding one. However, in contrast to UHF the AUHF method gives correctly $\langle \hat{S}^2 \rangle = 0.7500 \hbar^2$ while in UHF $\langle \hat{S}^2 \rangle$ increases from 0.76 \hbar^2 to 1.75 \hbar^2 with increasing bond stretching. Thus in the bond length regions occurring in soliton dynamics the AUHF energy is only slightly above the UHF one and thus superior over RHF (lower energy in AUHF) and UHF (correct $\langle \hat{S}^2 \rangle$ in AUHF).

Our time iteration scheme is based on the classical equations of motion [1] for the staggered coordinates (ψ_i) and velocities ($\omega_i = d\psi_i/dt$) for CH-unit i (m is the mass of a CH group):

$$F_i = m\dot{\omega}_i = -\frac{\partial E}{\partial \psi_i} \quad (18)$$

Then from $\psi_i(l)$ and $\omega_i(l)$ at a given time lt ($\tau=1.0$ fs) their values at time $(l+1)\tau$ are obtained by [1]

$$\begin{aligned} \omega_i(l+1) &= \omega_i(l) + F_i(l) \tau / m \\ \psi_i(l+1) &= \psi_i(l) + \omega_i(l+1) \tau \end{aligned} \quad (19)$$

We have shown that this one-step procedure works very well for tPA [65]. The gradient of the total energy with respect to the geometrical variables is computed by a simple and exact formula [9] in UHF and RHF cases in contrast to the very time consuming numerical procedure introduced by SSH [1]. In case of the AUHF approximation the formula for the gradients is no longer exact, but is numerically still a good approximation to the exact gradient [22]. Geometry optimization with this procedure can be done by setting $\omega_i=0$ (or by reducing the velocities by a factor of e.g. 0.9) after each time step. In this case the procedure corresponds to the gradient method. Excitation energies for excitation from occupied orbital i into the virtual orbital a are given by

$${}^{1,3}\Delta E_{i-a} = (e_a - e_i) - J_{ia} + {}^{1,3}K_{ia} \quad (20)$$

in our simple HF approximation, where $^1k=2$ (singlet excitations), $^3k=0$ (triplet excitations) and J_{ia} (K_{ia}) is the Coulomb (exchange) integral between the orbitals i and a . However, for increasing number (N) of electrons

$$\lim_{N \rightarrow \infty} J_{ia} \sim \lim_{N \rightarrow \infty} K_{ia} \sim \lim_{N \rightarrow \infty} \left(\frac{1}{N} \right) \rightarrow 0 \quad (21)$$

holds, and thus the correction terms to the difference of the HF eigenvalues vanishes. For this reason the excitation energies for polymers are usually calculated far too large in HF approximations and have to be corrected for correlation (see section V).

4. Parametrization

In the PPP case we have to deal with a three-dimensional parameter space, namely the resonance integral β_0 , the electron-phonon coupling constant α , and the on-site Coulomb repulsion integral γ_0 . Thus we varied the values of β_0 and α and calculated the energy of butadiene in the model at $u_p/u_0 = 0$ ($p=1$), 0.5 ($p=2$), 1.0 ($p=3$) and 1.5 ($p=4$). Note, that p is not a site index here. The values obtained ($E_{PPP}(u_p)$) are compared with the corresponding $E_{CCD}(u_i)$ energies. The energies are computed relative to $u_p=0$. Then we calculated

$$\Delta_4 = \left[\sum_{p=1}^4 (E_{PPP}(u_p) - E_{CCD}(u_p))^2 \right]^{1/2} \quad (22)$$

as function of γ_0 for each α , β_0 pair. γ_0 was varied from $\gamma_0=0$ to $\gamma_0=17.25$ eV in steps of 0.25 eV. In all cases we collected the excitation energies 1B_u , 3B_u , 2^1A_g , 3A_g which can also be qualitatively compared to experimental values. We found [17] that there is no unique set of parameters with minimal error Δ_4 . In addition (what is very interesting) in all (α, β_0) cases where $\gamma_0=0$ is found as minimum of $\Delta_4(\gamma_0)$ the optimal Δ_4 values are much larger than those where $\gamma_0>0$ is optimal. Thus a Hückel-type SSH treatment of trans-butadiene is not appropriate and inclusion of electron-electron interactions ($\gamma_0 \neq 0$) is essential. Only for $\gamma_0>0$ the bond alternation potential can be described with a reasonably small error. The correct description of such potentials is crucial for soliton dynamics.

We have collected 12 parameter sets with $\gamma_0>0$ where the error Δ_4 is optimal between 0.7 and 1.3 meV. An additional criterion for the choice of these sets was also that the excitation energies especially for 1B_u and 3B_u are not too far from experiment. For the 2^1A_g state also the experimental situation is not clear, and in addition this state is a true more-determinantal one. Thus one cannot expect that HF excitation energies could be correct. The 1B_u state is most probably not planar (see the above discussion) and it is not clear whether the excitation occurs

vertical. Therefore, for our parametrization we take comparisons of excitation energies only as a very qualitative guideline and concentrate on total groundstate energies as function of geometry. In addition to the PPP parameter sets we selected also two SSH type sets ($\gamma_0=0$), where Δ_4 is optimal but is still larger than 10 meV. Using these sets of parameters we performed geometry optimizations for $C_{21}H_{23}^+$, $C_{22}H_{24}$, and $C_{22}H_{24}^{2+}$ in the RHF approach (closed shell systems), because for these three polyene chains ab initio geometry optimizations exist in the literature [14] for comparison.

Villar et al. [14] published for the three systems mentioned above the bond lengths R_i^{HF} between sites i and $i+1$ obtained from ab initio calculations. For the 14 parameter sets mentioned above we performed optimizations on these three systems and computed for each one

$$\Delta_N = \left[\sum_{i=1}^{N-1} (R_i^{HF} - R_i^{PPP})^2 \right]^{1/2} \quad (23)$$

These values are given in Table 2 of Ref. [17].

Looking at all the data it is obvious that one of the parameter sets (called III in [17]) is the optimal one: the energy error for trans-butadiene is in a medium region (0.82 meV), the bond length error is optimal ($\Delta_{21}=2.26$ pm, $\Delta_{22}=1.27$ pm, $\Delta_{22}^{2+}=2.22$ pm) and the soliton half width (5.9 sites) fits best to the value of Villars et al. (6.0 sites [14]). The excitation energies in trans-butadiene of 5.55 eV for 1B_u (exp. 5.76-6.29 eV) and 3.09 eV for 3B_u (exp. 3.22-3.3 eV) agree reasonably well with experiment (keeping the above mentioned limitations in mind), at least better than for most of the other parameter sets. Our value for the 1B_u excitation energy is smaller than recent Multireference CI results of 6.7 eV for the vertical excitation [50] due to the simple HF scheme used for our calculation (see discussion above). The parameter values found seem to be reasonable in themselves. $\beta_0=-2.4$ eV is in the region where most PPP parametrizations for carbon put it (2.3-2.5 eV, e.g. [60a]). The on-site Coulomb repulsion $\gamma_0=11.25$ eV is very close to its most reasonable value (I_c-A_c) of 11.08 eV [60a]. The electron-phonon coupling of $\alpha=1.9$ eV/Å is much smaller than the SSH value (4.1 eV/Å [1]). This agrees with findings from ab initio calculations on infinite tPA [69]. Thus the basic assumption of the SSH Hamiltonian, that the properties of tPA are governed mostly by electron-phonon interactions and that electron-electron interactions are negligible does not hold. On the contrary electron-electron interactions are of utmost importance as also concluded in [69]. We conclude that the above described set of parameters ($\beta_0=-2.4$ eV, $\alpha=1.9$ eV/Å, $\gamma_0=11.25$ eV, $I=11.72$ eV) is appropriate to describe properties of polyenes on mean field PPP level in reasonable agreement with ab initio, ab initio plus correlation calculations, and certain experiments. Correlation effects on the potential surface are implicitly included via the parametrization.

5. Static Properties

We have shown in Table 3 of Ref. [17] how some properties of polyenes behave with increasing chain length in PPP calculations. The spring constant K of the σ -electron potential and the equilibrium σ -bond length R_0 are calculated such that for each chain (C_NH_{N+2}) the geometry $\psi/u_0=1$ with $u_0=0.0238$ Å is the equilibrium structure. This u_0 value leads to the optimized geometry of tPA obtained by Suhai [70] (CC bond lengths of 1.366 Å and 1.450 Å, CCC bond angle of 123.9°). This value of u_0 is also close to recent experimental results (0.026 Å [71]) and recent ab initio calculations including electronic correlation (0.025 Å [69]). We found [17] that the spring constant converges to a value of ≈ 38.4 eV/Å² and the equilibrium σ -bond length is 1.50 Å almost independent of the number of units N as one expects. The energies of highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals are practically converged from $N=80$ -100. The "band gap" (HOMO-LUMO separation) approaches ≈ 5.7 eV for $N \rightarrow \infty$. As expected this value is much larger than ab initio (including correlation) and experimental values for the gap (around 2 eV [32]). The actually calculated gap of 5.7 eV is a typical artefact of HF calculations. Suhai [32] found a gap of roughly 3 eV using the ab initio MP2 method and Liegener [33] obtained 2.03 eV using MP3, a value which agrees with the peak in the optical absorption spectrum of tPA. The low energy edge he assigns to excitons from his results on exciton binding energies. The four lowest excitation energies 1B_u , 3B_u , 2^1A_g , and 3A_g converge very slowly to the gap value. Even for $N=160$ they are 0.5-1.0 eV below the gap. This is due to the fact that all integrals in the HF approximation which correct the energy level difference to the actual excitation energy are approaching zero very slowly for $N \rightarrow \infty$ (see above) [72]. The energy gap can be corrected to realistic values only by explicit inclusion of correlation corrections [32,33,73]. The A_g energies approach the B_u energies because in the infinite polymer the two excitations are separated in reciprocal (k) space only by an infinitesimal dk .

König and Stollhoff [69] calculated the bond alternation potential ($E(u)$ per C_2H_2 unit) of infinite tPA using the ab initio L-CCD method and a double zeta atomic basis set. We computed the corresponding potential for chains of length $N=40$ and 50 using our new parametrization for the mean field PPP method. The potential curves are practically identical for the two chain lengths. In Fig.2 we show these results in comparison with those estimated from Fig. 1 in Ref. [69]. In Fig. 2 the potentials per CH unit are shown. The third curve is the corresponding one for trans-butadiene (CCD method, 6-31G** basis set). The lines are drawn just for guiding the eye, the symbols mark the actually calculated points. Our PPP curve results from a fit to CCD values computed with an atomic basis set containing polarization functions while the curve from Ref. [69] results from an L-CCD calculation using a double zeta basis set without such functions. We know from the results discussed above that for trans-butadiene both the size of the

basis set and the method applied (CCD or L-CCD) influence the potential curves

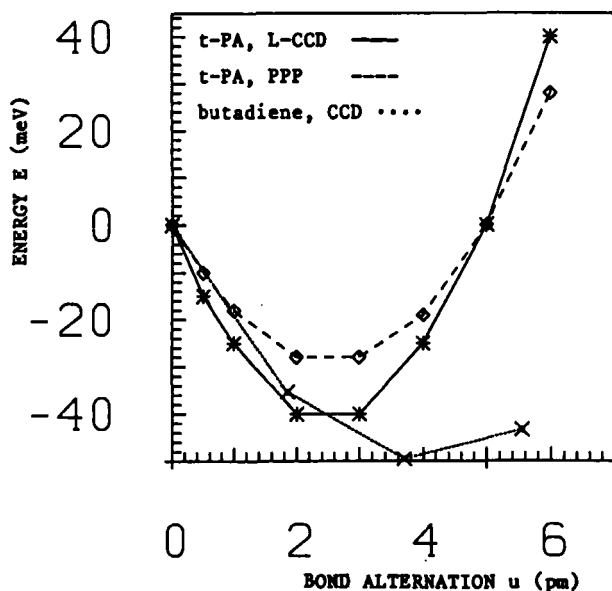


Fig. 2: The total energy E (in meV) per CH unit as function of the bond alternation u (in pm) for infinite tPA using ab initio L-CCD values from Ref. [5] (solid line), the PPP method (dashed line, $N=50$) and for trans-butadiene using the ab initio CCD method (pointed line).

considerably. Keeping this in mind we feel that the agreement between the two curves is satisfactory. From the PPP potential curve in Fig. 2 an effective force constant (including σ -electron effects) of $25.0 \text{ eV}/\text{\AA}^2$ can be deduced. This agrees fairly well with the experimental value of $26.0 \text{ eV}/\text{\AA}^2$ [74-76]. Vanderbilt and Mele [74] found that α has to be increased to $8 \text{ eV}/\text{\AA}$ in order to fit this force constant. However, they did not use a PPP model but an SSH type one, and thus effects which are due to electron-electron interactions can become incorporated in the value of α , simulating a larger electron-phonon interaction. For small rings Hayden and Soos [77] found $\alpha=3.21 \text{ eV}/\text{\AA}$ and the other parameters similar to ours. This difference might be due to their exact solution of the PPP problem, while we parametrize the HF-PPP model. In addition our ansatz for the lattice potential differs from that in [77]. Takahashi and Paldus [78] parametrize also HF-PPP models to fit data of benzene, and then computed larger rings. They found $\alpha=4.04 \text{ eV}/\text{\AA}$ close to SSH. However, they did not use the Ohno formula

for the γ_n and also an anharmonic lattice which leads to parameter values different from ours. Further it is possible that the description of open chains requires other sets of parameters than that of aromatic rings. Since the conjugation length in tPA seems to be rather small the application of periodic boundary conditions might be dangerous. However, we want to emphasize the agreement of our reduced value for α with the conclusions of König and Stollhoff [69] from comparisons to ab initio L-CCD results on infinite tPA.

Before turning to the dynamics we want to discuss some soliton properties as functions of chain length. For this purpose we used the ideal soliton shape [1]

$$\psi_i = u_0 \tanh\left(\frac{N_0 - i}{L}\right) \quad (24)$$

where $N_0 = (N+1)/2$, i.e. the center of the chain. K and R_0 are determined for $\psi_i = u_0$ as usual and then L is optimized. The results are shown in detail in Table 4 of Ref. [17]. As shown there in AUHF the half width of the neutral soliton converges to roughly 2.5 lattice sites while in UHF it is much larger. Also the spin density waves appearing in UHF are suppressed in the AUHF model which leads to a correct value of $\langle \hat{S}^2 \rangle$. The average spin densities on odd and even numbered sites are too small but still much closer to experiment than the very large UHF values. The band gap converges to roughly the same value as in case of even numbered chains. The neutral soliton levels are shifted to the band edges as expected. The spring constant K converges to a slightly higher value than for even numbered chains. This is most probably due to the end kink present in odd numbered chains with geometry $\psi_i/u_0 = 1$. The energy difference between an end kink and a soliton in the middle of the chain converges to roughly -0.25 eV. The corresponding value for charged solitons is much larger (-0.89 eV). Also the half width of charged solitons is larger than that of neutral ones (roughly 10 sites). The UHF values of the neutral soliton half width are also quite large and do not converge with chain size. These results are in qualitative agreement with those of Ref. [22] for another parameter set and also with results obtained with other methods found in the literature. The small S° width in AUHF calculations agrees with MNDO results where $L=3$ was found [79]. The charged solitons both have the same width in our PPP calculations which disagrees with the MNDO results of [79] where for the cation a very small width was found. In contrast to the semiempirical MNDO results ab initio calculations on $C_{21}H_{23}^\pm$ show the same width of roughly six sites for cation and anion [14] in agreement with our PPP results of about five sites in chains of this length. In the UHF model the soliton width diverges with increasing chain size because in an infinite chain UHF results artificially in the equidistant geometry as equilibrium structure, which corresponds to an infinite soliton width. Also the spin density wave (SDW) in UHF reaches

nearly undamped through the chain and the ratio of the absolute values of the spin densities on neighboring carbons approaches 1 with $\langle \hat{S}^2 \rangle$ values of order 2-10 \hbar^2 . Even for smaller chains the extent of the SDW is far too large to match any experimental findings of SDW half-widths of 12 [80] or 6-8 sites [81,82]. Monte Carlo calculations were performed with rather small values of the electron-electron interaction parameters [83,84] and lead to SDW's comparable to experiment.

Table 3: Spin densities in equidistant polymethines computed with the UHF-PPP and the AUHF-PPP method (own calculation) in comparison to exact PPP solutions for the same system from Ref. [85] and the expectation value $\langle \hat{S}^2 \rangle$ (in \hbar^2).

Sites	UHF-PPP	AUHF-PPP	Exact PPP
<hr/> C ₇ H ₉ <hr/>			
1, 7	0.499	0.183	0.332
2, 6	-0.299	-0.026	-0.105
3, 5	0.482	0.364	0.353
4	-0.363	-0.043	-0.139
$\langle \hat{S}^2 \rangle$	1.1062	0.7506	3/4
<hr/> C ₉ H ₁₁ <hr/>			
1, 9	0.329	0.118	0.266
2, 8	-0.216	-0.018	-0.086
3, 7	0.383	0.257	0.293
4, 6	-0.322	-0.038	-0.125
5	0.422	0.361	0.303
$\langle \hat{S}^2 \rangle$	1.2107	0.7509	3/4
<hr/> C ₁₁ H ₁₃ <hr/>			
1, 11	0.475	0.078	0.227
2, 10	-0.292	-0.012	-0.071
3, 9	0.454	0.177	0.251
4, 8	-0.376	-0.030	-0.110
5, 7	0.436	0.308	0.264
6	-0.393	-0.041	-0.121
$\langle \hat{S}^2 \rangle$	1.5006	0.7511	3/4

However, together with exact solutions of the PPP model for smaller systems [85]

these results imply that correct SDW amplitudes can only be obtained by explicit inclusion of correlation. Since such exact solutions cannot be computed in time simulations on chains of reasonable size, one has to use a mean field approximation. As the soliton widths show, UHF gives the wrong dependence of the energy on geometry, while AUHF soliton widths are comparable to *ab initio* results. Thus, although AUHF underestimates the spin polarization from the point of view of energy hypersurfaces it has to be preferred. In Table 3 we compare UHF and AUHF spin densities in polyenes in equidistant geometry with the corresponding exact results from [85]. The underestimation of SDW amplitudes by AUHF is obvious. In UHF for this small system the spin contamination is not too large, but the overestimation of SDW amplitudes can already be recognized. Sasai and Fukutome [86] found a smaller SDW width in their UHF calculations. This is probably due to their use of an exponentially screened Ohno potential for the γ_n , which leads to a much faster decay of the integrals for larger separations, making the model more similar to an extended Hubbard one.

The charged soliton levels appear closer to mid gap than the neutral ones. Also the gap in charged chains is by roughly 1.5 eV smaller than in neutral ones. In Fig. 3a we show the energy level scheme for the different solitons. The energies are scaled such that the gap is 2 eV in all cases which corresponds to the experimental value of the gap. From the energy levels one could assign the photo induced high energy absorption in tPA (1.35 eV) to an excitation from the S^- level to the conduction band or from the valence band to the S^+ level. However, this assignment cannot explain the strong temperature dependence of the high energy peak. In our level scheme the low energy peak would correspond to an $S^+S^- \rightarrow 2S^0$ excitation (0.8 eV). Recently, Tanaka et al. [87] came to the same conclusion using another parametrization for the HF-PPP model. However, since it is not obvious if states in the gap can be scaled as the gap itself, we computed the shifts in energy of the bottom of the "conduction band" states (C), the S^- level (width $L=8$) and of the top of the "valence band" states (V) in $C_{81}H_{83}^-$ using our RHF-PPP reference Slater determinant and the MP2 quasiparticle scheme in the form used by Liegener [88]. The results are shown in Fig. 3b (unscaled). Obviously the gap of ≈ 4.4 eV in RHF-PPP is reduced to ≈ 3.2 eV upon inclusion of correlation. This compares well to Suhai's *ab initio* result within his best basis set of ≈ 3 eV in infinite tPA [32]. The S^- state is shifted above midgap to $\approx +0.4$ eV, while a similar calculation on S^+ shifts it below midgap (≈ -0.4 eV). Since one has to expect that the inclusion of higher orders of the MP series would shift C further down and S^- further up, the low energy absorption (exp. 0.45 eV) should be assigned to $S^- \rightarrow C$ and $V \rightarrow S^+$ excitations. The present value of 1.2 eV for this energy appears to be too large, but is the only excitation below midgap in our level scheme. Scaling again the gap to 2 eV and using this factor of $2/3.2$ also for the $S^- \rightarrow C$ excitation energy leads to ≈ 0.7 eV. Since states in the gap as shown in Fig. 3b experience larger shifts than the band edges, one can expect that the

inclusion of higher order correlations, as well as dynamic effects could well reduce the excitation energy to that of the low energy feature in the photoinduced spectrum of tPA. Moreover, $S^- \rightarrow C$ and $V \rightarrow S^+$ would be temperature independent in contrast to $S^- \rightarrow S^+$ excitations which require a bound soliton pair of not too large separation in order to have a large enough Frank-Condon factor. From our AUHF results an $S^0 \rightarrow C$ excitation would require roughly 5 eV. Correlation will shift C downward and S^0 upward.

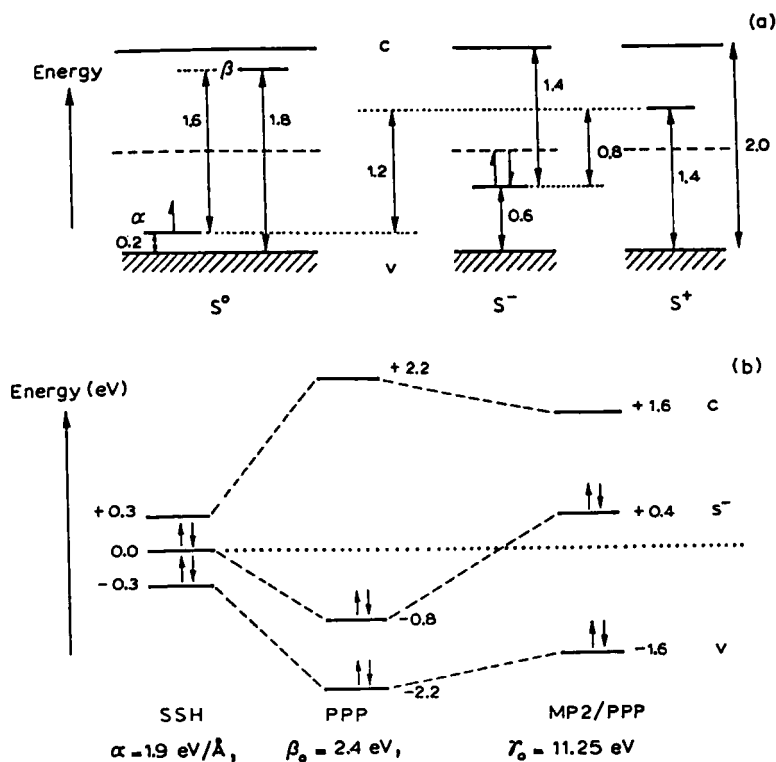


Fig. 3: (a) Energy level diagram for neutral (S^0 , AUHF) and charged (S^+ , RHF) solitons (all numbers in eV, C denotes bottom of the conduction band, V top of the valence band). The energies ($N=121$) are scaled such that the gap is 2.0 eV in all cases.
 (b) Top of the valence band (V), S^- level, and bottom of the conduction band (C) in a $C_{81}H_{83}^-$ chain with a soliton in the center (width $L=8$), calculated with the SSH, the RHF-PPP and the PPP/MP2 method.

Assuming again for the $S^0 \rightarrow C$ excitation a similar scaling as for the gap one would expect the excitation around 1.8 eV. Thus one is tempted to assign the high energy feature in the photoinduced absorption spectrum of tPA (exp. 1.35 eV) to neutral solitons (see [89,90] and references therein). Therefore we reach at the same conclusion as experimental assignments [91], namely that charged solitons are responsible for the low energy absorption and neutral solitons probably trapped in unconjugated chain segments for the high energy absorption [91]. It is concluded in [91] that the neutral solitons do not form triplet pairs and are also not produced by triplet excitation. Another possibility for their production would be the reaction $S^+ + S^- \rightarrow 2S^0$ which according to [92] is allowed if charge conjugation symmetry breaking terms like second neighbor hopping appear in the Hamiltonian even with otherwise negligible parameter values. Trapping would then prevent the neutral solitons from recombination. The temperature dependence of the high energy absorption could then be due to the fact that the above mentioned reaction should be most effective in a bound pair of charged solitons or that with increasing temperature the solitons would be able to leave their traps.

6. Dynamics

6.1. Neutral Polymethine Chains

All simulations are started from a geometry $\psi/u_0=1$ which correspond besides some end effects to the equilibrium ground state geometry in polyene chains (even number of carbons) and to an end-kink in polymethine chains (odd number of carbons). Since the equilibrium geometry corresponds to a soliton of tanh shape in the center of the chain, in the latter case from this start during the simulation a soliton starts to evolve and travels through the chain. We have chosen chain lengths of 31 carbons for polymethines and of 32 carbons for polyenes. This chain length seems to be sufficient since in the usually synthesized polyacetylene the conjugation length is quite small, being 30-50 carbons at maximum. In our calculations on polymethines we also studied the influence of disorder on the soliton properties. For this purpose we used in the simulations for each parameter A a modified value $A(1+x_i)$ where x_i is a random number with $-x < x_i < x$, i being the site index. x is the maximal degree of disorder in a simulation and is varied ($x=0.05, 0.10, 0.15, 0.20$). Thus e.g. $x=0.05$ corresponds to 5% disorder in the parameters. As disordered parameters we have chosen the resonance integral β_0 , the electron-phonon coupling constant α , the on-site Coulomb repulsion γ_0 , and the ionization potential I . These parameters are all influenced by the disordered environment of a chain in tPA, while the masses and core charges of the CH units are kept constant. As time interval for simulation 1 fs turned out to be sufficiently small to keep the total energy reasonably constant (within less than 10%) as compared to the kinetic energy.

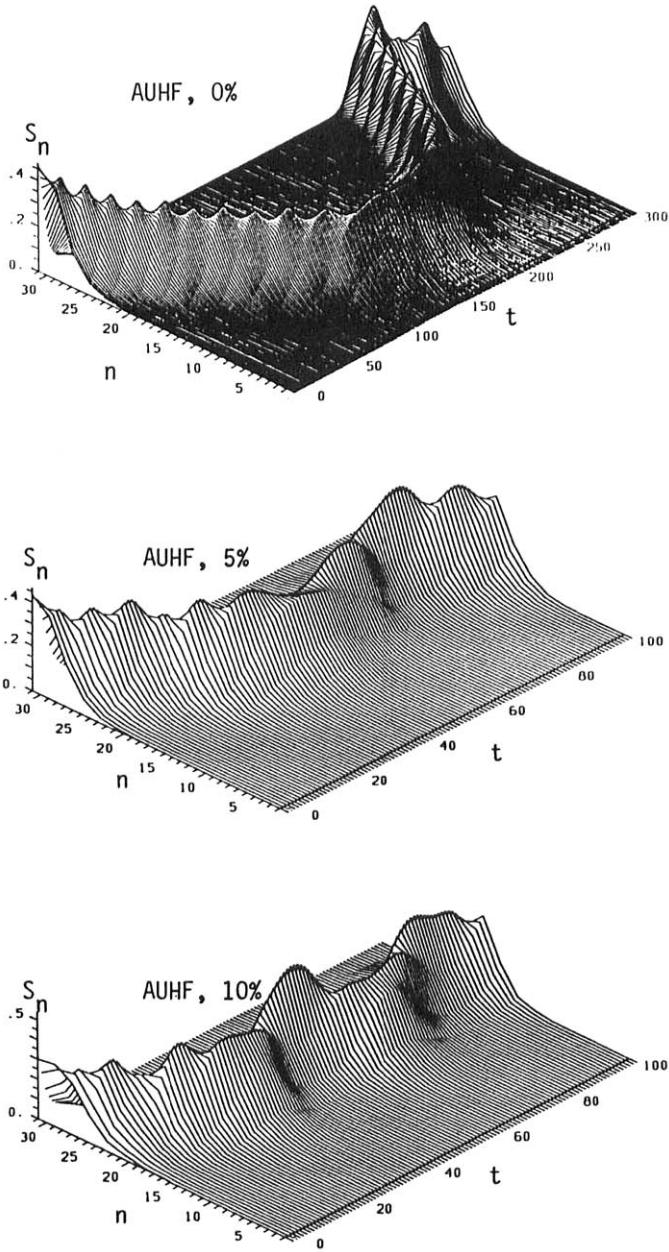


Fig. 4: Continued on next page.

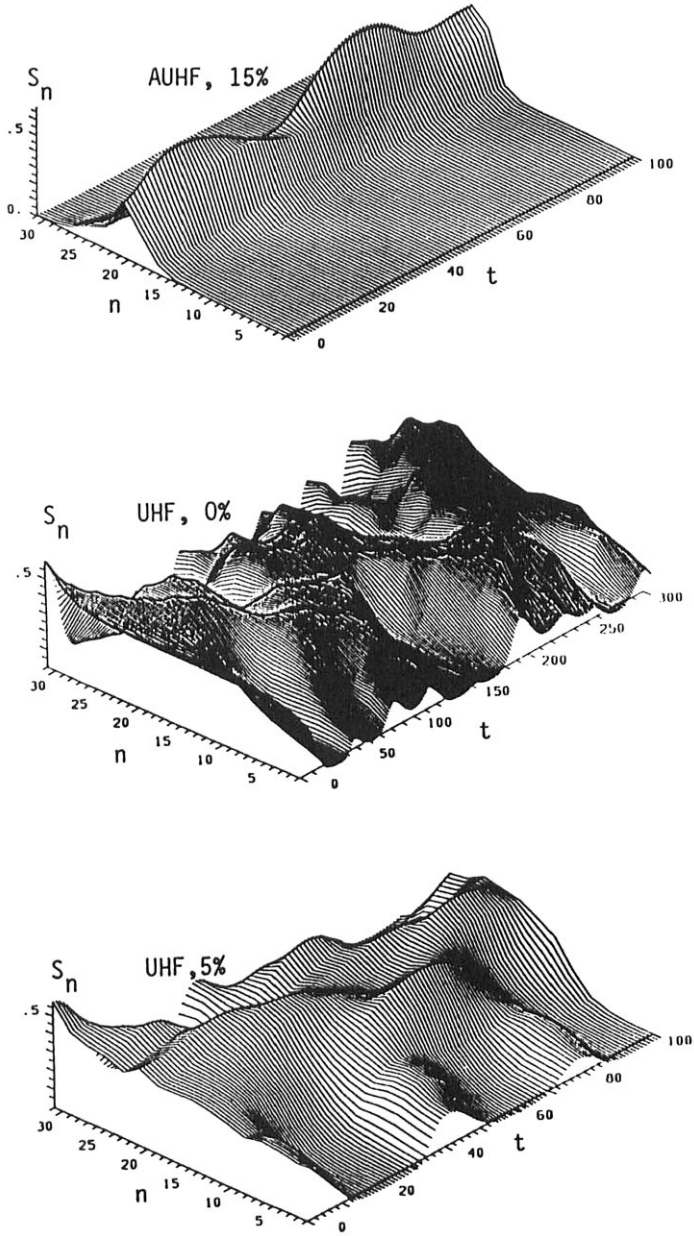


Fig. 4: Continued on next page.

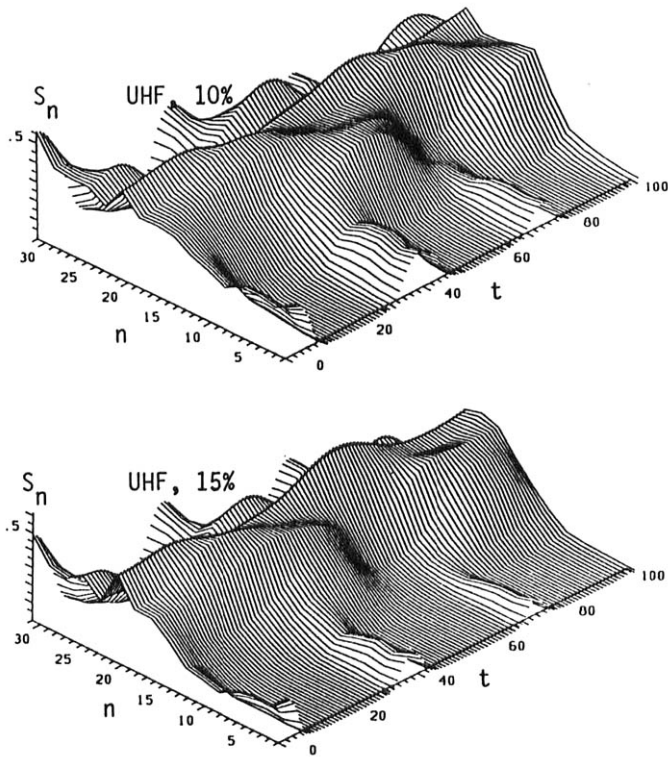


Fig. 4: Spin density $S_n(t)$ at odd numbered sites only as function of site n and time t (in fs) as obtained from AUHF and UHF simulations starting from an end-kink in a neutral polymethine chain of 31 sites with different disorder strengths of 0%, 5%, 10%, and 15%.

In Fig. 4 we show the results of such simulations using the above discussed AUHF and UHF approximations for comparison and applying different disorder strengths. In the Figure we display the spin densities S_n as function of site n and time t on odd numbered sites only, because the spin densities oscillate between positive values on odd numbered sites and negative values on even numbered sites. In case of completely ordered chains we observe the formation and propagation of a soliton of rather small width in the AUHF approximation in agreement with the static results reported above. Its velocity is roughly 2.9 km/s, much smaller than the values obtained with the usually applied parameters [6,7] in UHF or SSH models. The soliton obviously survives reflections at the chain ends with-

out change of shape or velocity. In the corresponding UHF calculation an extended spin density wave of large amplitude is seen. Thus the soliton has a very large width and correspondingly a higher velocity. However, as discussed above this large soliton width has to be interpreted as an artefact of the UHF approximation. When the disorder is increased to 5% and 10% we see that the soliton propagation becomes erratic and the soliton is not able to reach the other chain end. Increasing the disorder further the spin becomes localized around one site from the beginning due to Anderson localization. However, we do not expect that the effects of the environment on the parameters would exceed 2-3%. Thus we conclude that soliton movement should be possible within a reasonable disorder strength. In UHF approximation we observe the same behaviour: with increasing disorder the spin becomes more and more localized. Optimization runs had lead to solitons in the middle of the chains with widths in agreement with the static results discussed above. In the disordered cases we did not perform optimizations. We finally want to note, that due to the rather deep potential wells at each site in AUHF the optimization for 0% disorder had to be performed with a very slow "cooling" by reduction of the velocities by a factor of 0.99 after each time step. Otherwise the optimization does not reach at equilibrium with the soliton in the middle of the chain, but stops with the soliton close to the chain end. This is avoided with the slow "cooling" method, because in this case the system has in the beginning enough kinetic energy to pass the local minima and to reach the true equilibrium. These rather deep potential wells at the sites originate from the increasing tendency to lattice pinning, i.e. decreasing soliton velocities, with decreasing width of the solitons. This is a very general feature of soliton dynamics in different non-linear systems.

6.2. Charged Polymethine Chains

In this case we used the restricted HF approximation since we have closed shell systems with either a doubly occupied or an unoccupied soliton level. The results for negatively and positively charged solitons are, as expected from the discussion given above more or less identical. Therefore we show in Fig. 5 only the dynamics of positively charged solitons. Here we show instead of the spin densities the charge densities Q_n at odd numbered sites as function of site and time. In an ordered chain we observe the formation of a charged soliton of rather large width, which travels nearly 5 times through the chain in 100 fs, corresponding to an average velocity of roughly 18 km/s. However, as the second plot shows in the case of charged solitons already at 5% disorder strength the Anderson localization of the soliton level plays a crucial role. However, since the charge is more delocalized than the spin in the neutral case, fractions of it are localized on several sites and oscillate somewhat between them. However, as discussed above in the neutral case we assume that 5% disorder strength is

already larger than what one should expect, since the tPA chains in the disordered material are rather far apart from each other and thus their influence on the effective set of parameters simulating the environment should not exceed 2-3% disorder strength.

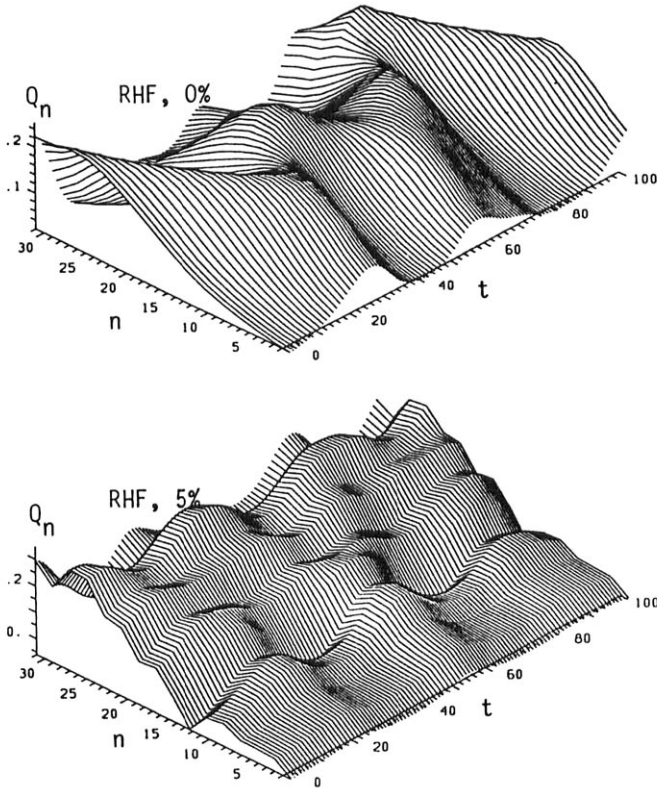


Fig. 5: Charge density $Q_n(t)$ at odd numbered sites only as function of site n and time t (in fs) as obtained from RHF simulations starting from an end-kink in a positively charged polymethine chain of 31 sites with different disorder strengths of 0% and 5%.

At these values soliton movement is still possible as a calculation with 2.5% disorder shows. In that case a small fraction of the charge becomes trapped, while still a charged soliton moves through the chain and can be identified from the

time evolution of the geometry of the system. Optimization runs result again in a soliton in the middle of the chain in ordered systems, and in the case of disorder we again did not perform optimizations.

6.3. Singly Charged Polyenes

In this case we have used polyene chains of 32 sites with a positive or a negative net charge which might occur from doping. In their early simulation using their Hückel-type model, SSH [7] found that in this case a polaron is formed from the ideal alternating chain. We have used the AUHF and also the UHF approximations again for comparisons and have performed optimization runs as well as simulations. Here again the results for positively and negatively charged polyenes are nearly identical so that we can restrict the discussion here to the positively charged case. The results are shown in Fig. 6, where Fig. 6 a-d shows the AUHF and Fig. 6 e-h the UHF results. Let us first discuss the more realistic AUHF case. In Fig. 6a the normalized staggered coordinate $\psi_n' = \psi_n / u_0$ obtained in an optimization run is shown. Obviously from the ideally dimerized charged polyene a polaron of relatively small width is formed in the middle of the chain. It is clearly a polaron, because the minimum ψ' in the middle of the chain is 0, as it is typical for a polaron. A pair of a soliton and an anti-soliton would have the reversed dimerization, i.e. $\psi_i' = -1$ in its center. Fig. 6b shows the normalized staggered coordinate from a simulation. We see that the polaron is formed within roughly 10 fs and the system after that bounces back to the dimerized structure due to conservation of energy. The minimum value of the staggered coordinate in the polaron of ideally 0 is overshooted to roughly -0.5 since the system has kinetic energy from the polaron formation. Fig. 6c shows the corresponding spin density on odd numbered sites which obviously is localized in the right hand side of the polaron center. The spin density on even numbered sites can be obtained from that as its mirror image relative to the polaron center. Similarly the charge density (Fig. 6d) on odd numbered sites is mainly localized at the same site as the spin density with smaller amount of negative charge on the other side of the polaron center. As above the charge density on even numbered sites is the mirror image of that on odd numbered sites relative to the polaron center. The polaron is stabilized with respect to the dimerized chain by roughly 0.23 eV.

In the UHF approximation (Fig. 6 e-h) the situation is somewhat different. In the optimization an extremely broad polaron is obtained which is practically an equidistant chain some sites apart of the two chain ends. This is obviously an artefact of the well known tendency of UHF to produce metallic polyenes as equilibrium structures [39]. In the simulation we observe also the formation of this polaron like structure after roughly 10 fs (Fig. 6f), however this is destroyed after two oscillations. The spin density in UHF (6g) looks very similar to the

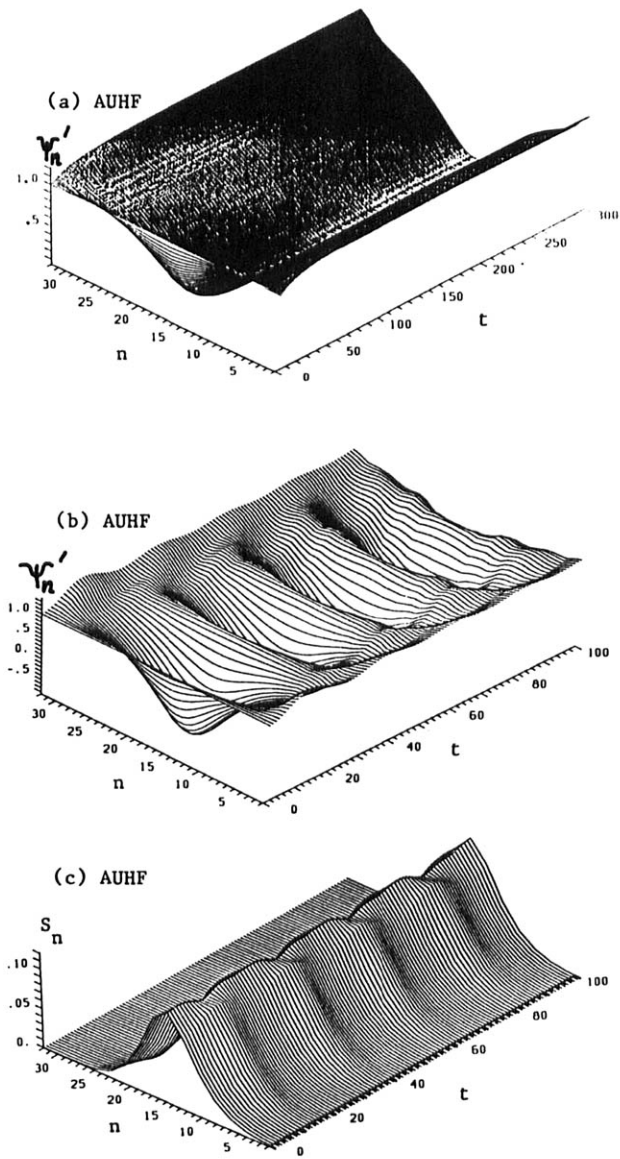


Fig. 6: Continued on next page.

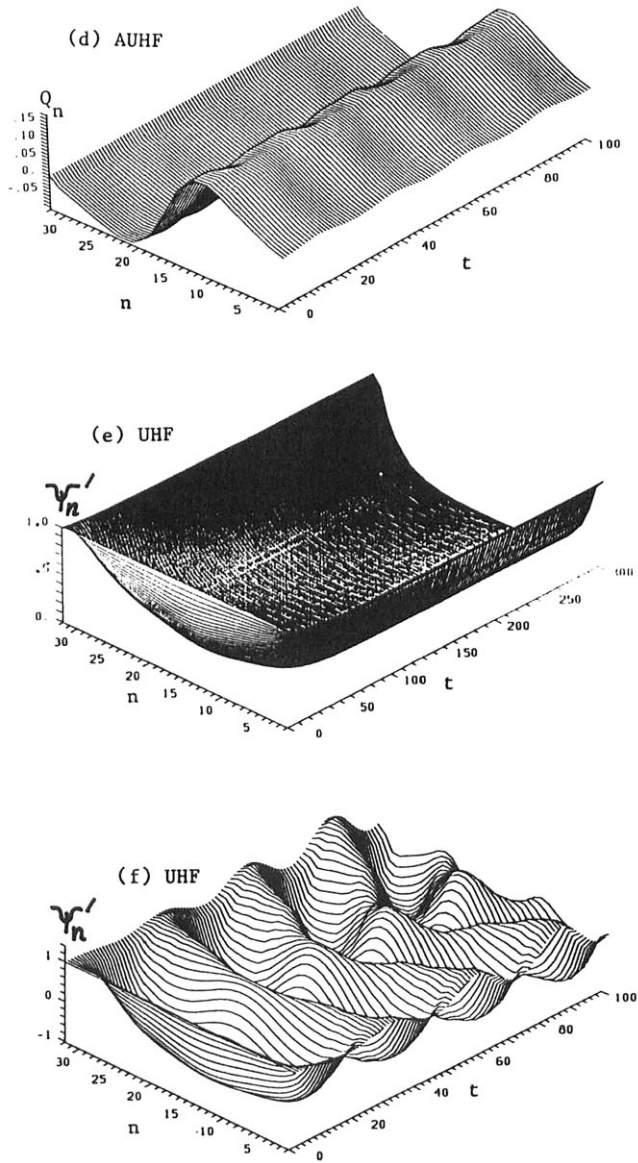


Fig. 6: Continued on next page.

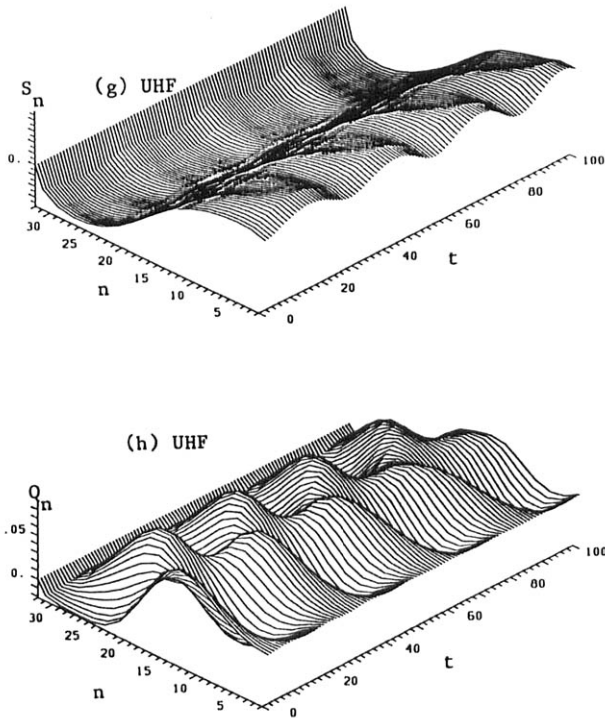


Fig. 6: The normalized staggered coordinate $\psi_n'(t) = \psi_n(t)/u_0$ as obtained from optimization (a,e, note that here time t stands symbolically for the iteration step) and simulation (b,f) runs, together with the spin density $S_n(t)$ (c,g) at odd numbered sites and the charge density $Q_n(t)$ (d,h) on odd numbered sites as obtained from the simulations, using the AUHF (a-d) and the UHF (e-h) approximation as functions of site n and time t (in fs) starting from ideally dimerized positively charged polyenes with 32 sites.

charge density in AUHF (6d), with negative values on the left side of the polaron center. Together with the corresponding spin density on even numbered sites we see a spin density wave of large amplitude spread over the whole chain which reverses its phase at the polaron center. In the charge density we can observe a soliton like structure oscillating around the center of the polaron. The charge

density at even numbered sites is again the mirror image of Fig. 6h which respect to the polaron center. This is a good example to show how the spin contaminations in UHF can result in misleading conclusions. One has to note that again in UHF the expectation values $\langle \hat{S}^2 \rangle$ are much larger than the correct one of $0.75 \hbar^2$ for a doublet state, while the AUHF results are typically between the correct value and $0.76 \hbar^2$.

6.4. Doubly Charged Polyenes

Here we have again to deal with a singlet state and thus RHF can be used. Note that in the singlet cases discussed in this work AUHF calculations give the same results as RHF ones, while UHF sometimes gives a splitting of α - and β -spin states, accompanied by large spin contaminations showing the unreliability of such results. Therefore we discuss here only RHF calculations for singlet states. In case of doubly charged polyene chains in principle two different non-linear quasiparticle systems can be formed. First of all the formation of a pair of a charged soliton and an equally charged antisoliton is possible. On the other hand these two solitons can be closely bound forming a bipolaron, which is nothing but a doubly charged polaron. Using the SSH Hamiltonian we found [93] that the former case occurs and this behaviour is also suggested by experiments as discussed above, namely that bipolarons are not stable in polyenes but decay to free solitons and antisolitons. The results of our calculations on doubly charged polarons are shown in Fig. 7, where Fig. 7a shows again the normalized staggered coordinates of an optimization run. Obviously as required for a pair of solitons the coordinates decrease nearly to -1 in the center of the chain. However, for a clear cut separation of the two solitons the chain is too small. We have a soliton, changing the coordinates from +1 to -1 at the right end of the chain and an antisoliton, changing the coordinates again from -1 to +1 at the left hand side of the chain. In the simulation (Fig. 7b) the same structure is formed within about 10 fs and the system bounces back to the dimerized geometry. This cycle is repeated twice and after that enough energy is radiated into phonons, so that the two equally charged solitons are able to remain at opposite chain ends. A breather vibration shows up in the center of the chain. The charge density at odd numbered sites (Fig. 7d) resembles the same behaviour. We see clearly a positively charged soliton at the right end of the chain, which oscillates somewhat around a position close to the chain end. We have here a competition between two effects, namely the tendency of solitons to be in the center of a chain and the repulsion between the two positive charges. The charge density at the left hand side of the chain shows again up in the corresponding plot for even numbered sites. The stabilization energy of the soliton pair with respect to an ideally dimerized chain is 1.37 eV, somewhat less than twice the corresponding stabilization energy of a single charged soliton in the middle of a chain (roughly 0.8 eV [17]).

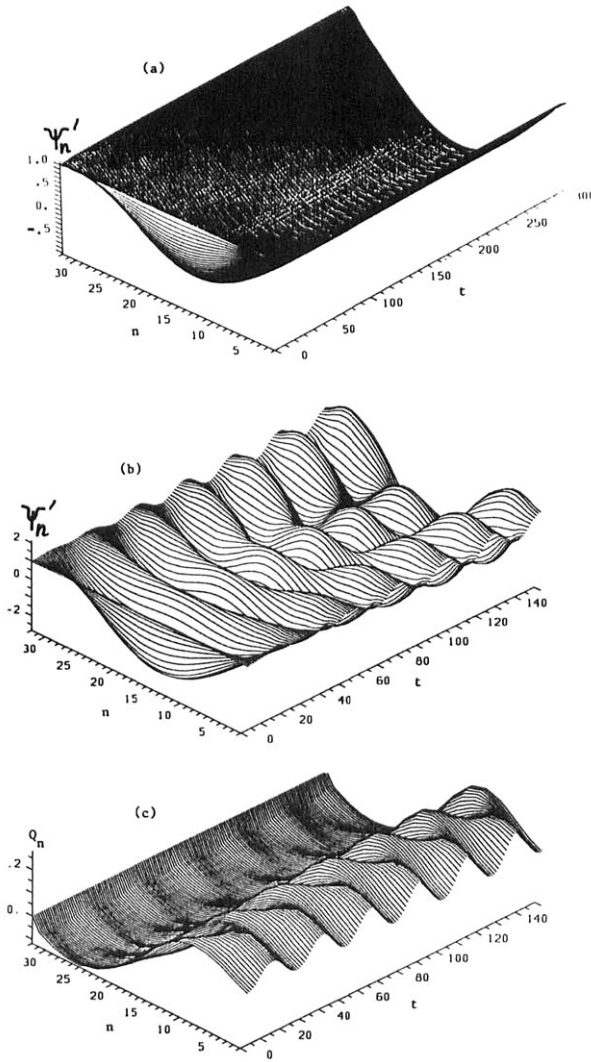


Fig. 7: The normalized staggered coordinate $\psi'_n(t)=\psi_n(t)/u_0$ as obtained from optimization (a, note that here time t stands symbolically for the iteration step) and simulation (b) runs, together with the charge density $Q_n(t)$ (c) at odd numbered sites as obtained from the simulations, using the RHF approximation as functions of site n and time t (in fs) starting from ideally dimerized doubly positively charged polyenes with 32 sites.

6.5. Excited Polyenes

Recently [94] we performed simulations on excited triplet and singlet states of polyenes using equ. (20) together with the ground state energy to determine the total energy of the excited system. The necessary gradient of the excitation energy was computed by determination of the gradient of energy eigenvalues and the occurring Coulomb and exchange integrals with the help of perturbation theory. In that work we found as expected two neutral solitons in the triplet state (3B_u) and two solitons with the characteristics of charged solitons in the singlet excited state (1B_u). Thus the possibility of formation of a neutral polaron, which is a close bound state of the two solitons is ruled out. The charges themselves cannot show up in such calculations because first of all we did not determine the wave function of the excited states and further in principle the HF wave function of an excited singlet state is a superposition of an S^+S^- and an S^-S^+ state and thus does not show charged solitons explicitly. However the solitons could be identified from the time evolution of the geometry of the system [94]. Here we want to discuss shortly the first excited triplet state of polyenes, where we simply changed the occupation numbers in the density matrices, i.e. the highest occupied (in the ground state) β -spin orbital was not populated, and the electron was put into the lowest unoccupied (in the ground state) α -spin orbital to obtain the triplet state. Then we followed the dynamics of the system again with the AUHF and the UHF method. The results are shown in Fig. 8. Note, that in the UHF case we had to use a time step of 0.5 fs to obtain consistent results, while in all other calculations reported in this paper 1 fs was sufficiently small. The optimization shows clearly that with the AUHF method (8a) and also with UHF (8d) a pair of solitons is formed. However, in UHF their width is again much larger and the UHF wave function contains large spin contaminations. In the time simulation we see that in the case of the AUHF method (8b) the solitons are formed quite fast, separate and then oscillate between one of the chain ends and the center of the chain where a breather evolves. More clearly this behaviour can be realized from the plot of the spin density at odd numbered sites (8c). The other soliton again has its spin density on even numbered sites and performs the same movement in the left half of the chain. From this it is not quite clear whether the solitons, both neutral and triplet coupled attract or repel each other and would separate in long chains. The picture of the dynamics points into the direction that the two solitons just pass each other, however, a clear distinction is not possible. The stabilization energy of the two solitons compared to the ideally dimerized chain is roughly 0.75 eV which is about three times the corresponding stabilization energy of a single neutral soliton in a polymethine chain (0.25 eV [17]). Also this result seems to indicate that there should be some binding between the solitons. In the UHF case the pictures of the dynamics (8e,f) are not that clear. Obviously soliton formation occurs also, but (8f) the movement of the soliton seems to be rather erratic.

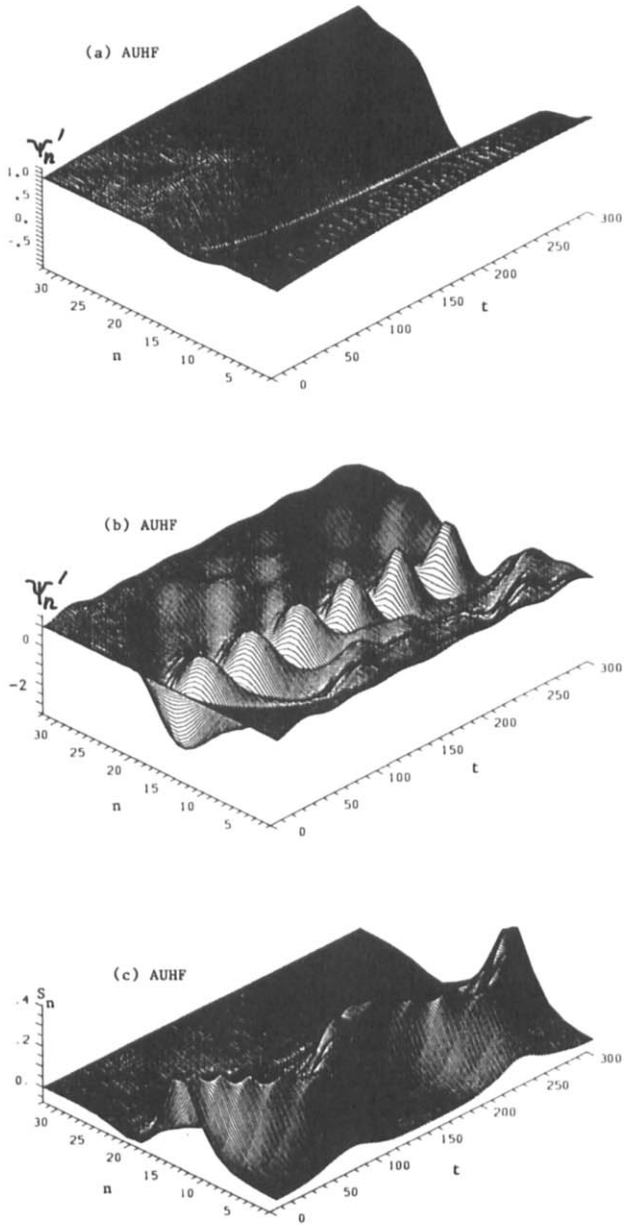


Fig. 8: Continued on next page.

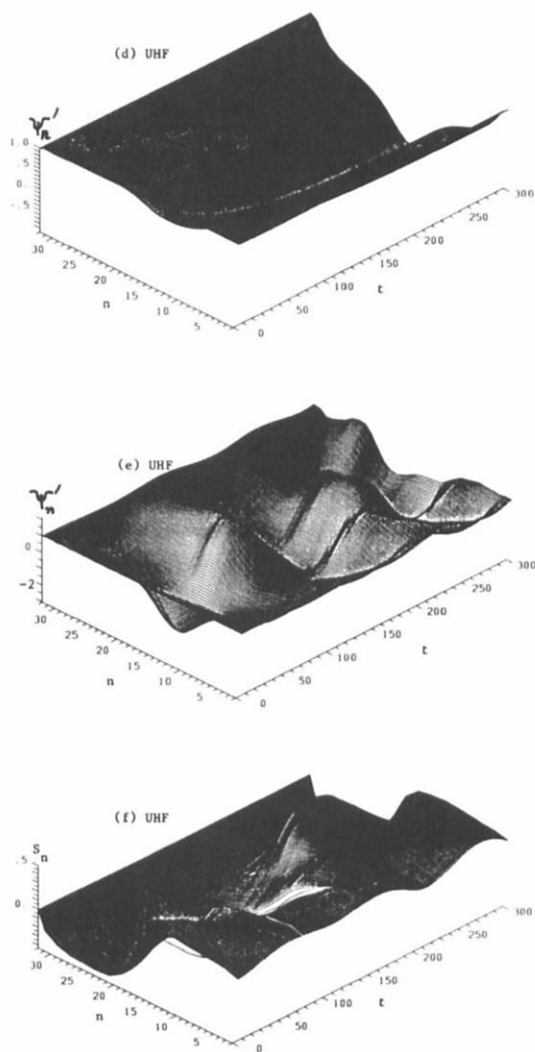


Fig. 8: The normalized staggered coordinate $\psi'_n(t) = \psi_n(t)/u_0$ as obtained from optimization (a,d, note that here time t stands symbolically for the iteration step) and simulation (b,e) runs, together with the spin density $S_n(t)$ (c,f) at odd numbered sites as obtained from the simulations, using the AUHF (a-c) and the UHF (d-f) approximation as functions of site n and time t (in fs for a-c and in 0.5 fs for d-f) starting from ideally dimerized triplet excited polyenes with 32 sites.

We do not want to discuss the singlet excited states here, because they require at least a two-determinantal wave function for their description and most probably also an MCSCF (Multiconfiguration SCF) treatment including all single excitations is necessary. The derivation of the necessary formulae has been already completed [95] and the numerical application of them is in progress in our Laboratory.

7. Conclusion

We have performed a reparametrization of the PPP Hamiltonian on the basis of ab initio calculations including correlation effects on CCD level on the bond alternation potential of trans-butadiene. We have shown that the new parameters which differ significantly from the SSH [1] ones, especially concerning the electron-phonon coupling constant, which is reduced to roughly half of its SSH value, lead to reasonable soliton properties in static calculations. Also soliton levels in the gap are predicted qualitatively correct. The force constant of a longitudinal vibration is calculated in fair agreement with its experimental value. Soliton dynamics using this new parameter set were presented. We found that in the system mobile neutral and charged solitons exist also for disorder of reasonable strength. Note, that solitons do not exist only for this specific set of parameters but also in a wider range around the actually considered values. In singly charged polyene chains the development of charged polarons is observed in computer experiments in agreement with the early results of SSH based on Hückel theory [7]. In doubly charged polyenes we found the bipolaron to be instable and instead pairs of charged solitons are formed which repel each other. In addition in the center of the chain breather formation occurs. In agreement with experiments we find that in triplet excited chains pairs of neutral solitons are formed together with a breather vibration in the center of the chain.

In order to improve the model further we are currently taking quantum effects in the lattice into account, i.e. treating the CH units not classically but on quantum mechanical basis. To this end we use an ansatz state similar to Davydov's so-called $|D_1\rangle$ state [96] developed for the description of solitons in proteins. However, there vibrations are coupled to lattice phonons, while in tPA fermions (electrons) are coupled to the lattice phonons. The results of this study will be the subject of a forthcoming paper. Further we want to improve the description of the electrons by going to semiempirical all valence electron methods or even to density functional theories. Further we introduce temperature effects into the theory which can be done with the help of a Langevin equation (random force and dissipation terms) or by a thermal population of the lattice phonons. Starting then the simulations with an optimized soliton geometry in the center of the chain (equilibrium position) one can study the soliton mobility as function of temperature. Further in the same way the mobility of polarons can be

investigated. Finally we want to extend our studies to other polymers with degenerate ground state like pernigraniline and also to those without degenerate ground state like polyparaphenylene or polythiophene where no solitons but polarons exist.

Acknowledgement

First of all the author wants to thank Professor Ramon Carbo and Professor Janos Ladik for giving him the opportunity to present this work at the First Congress of the International Society for Theoretical Chemical Physics held in Girona, Spain on June 28 - July 3, 1993. Especially I want to thank Professor P.-O. Löwdin for the invitation to publish this paper in *Advances in Quantum Chemistry* and for putting the author's attention to important references missing in the preliminary version of the manuscript. The financial support of the "Deutsche Forschungsgemeinschaft" (Project No. Fo 175/2-3) and the "Fond der Chemischen Industrie" is gratefully acknowledged.

References

- [1] W.P. Su, J.R. Schrieffer, and A.J. Heeger, *Phys. Rev. Lett.* **42**, 1698 (1979); W.P. Su, *Solid State Comm.* **35**, 899 (1980); W.P. Su, J.R. Schrieffer, and A.J. Heeger, *Phys. Rev.* **B22**, 2099 (1980).
- [2] (a) A. J. Heeger, S. Kivelson, J. R. Schrieffer, and W. P. Su, *Rev. Mod. Phys.* **60**, 781 (1988).
(b) S. Kivelson, W.-P. Su, J.R. Schrieffer, and A.J. Heeger, *Phys. Rev. Lett.* **58**, 1899 (1987).
- [3] H. Thomann, L.R. Dalton, Y. Tomkiewicz, N.S. Shiren, and T.C. Clarke, *Phys. Rev. Lett.* **50**, 533 (1983); H. Thomson, H. Kim, A. Morobel-Sosa, L.R. Dalton, M.T. Jones, B.H. Robinson, T.C. Clarke, and Y. Tomkiewicz, *Synth. Met.* **9**, 255 (1984); H. Thomann, J.E. Cline, B.M. Hofmann, H. Kim, A. Morobel-Sosa, R.H. Robinson, and L.R. Dalton, *J. Phys. Chem.* **89**, 1994 (1985); A.J. Heeger and J.R. Schrieffer, *Solid State Comm.* **48**, 207 (1983); Z.G. Soos and S. Ramasesha, *Phys. Rev. Lett.* **51**, 2374 (1983); M. Sasai and H. Fukutome, *Synth. Met.* **9**, 295 (1984); J. Orenstein and G.L. Baker, *Phys. Rev. Lett.* **49**, 1043 (1982); B.R. Weinberger, *Phys. Rev. Lett.* **50**, 1693 (1983); G. Blanchet, C.R. Fincher, and A.J. Heeger, *Phys. Rev. Lett.* **51**, 2132 (1983); G.B. Blanchet, C.P. Fincher, T.C. Chung, and A.J. Heeger, *Phys. Rev. Lett.* **50**, 1938 (1983); A.R. Bishop, D.K. Campbell, P.S. Lomdahl, B. Horowitz, and S.R. Phillpot, *Phys. Rev. Lett.* **52**, 671 (1984); W.P. Su, *Phys. Rev.* **B34**, 2988 (1986); S. Kivelson and W.-K. Wu, *Phys. Rev.* **B34**, 5423 (1986).
- [4] R. Pariser and P.G. Parr, *J. Chem. Phys.* **21**, 660, 707 (1953); J.A. Pople,

- Trans. Far. Soc. 49, 1375 (1953); J. Ladik, Acta Phys. Acad. Sci. Hung. 18, 185 (1965); J. Ladik, D.K. Rai, and K. Appel, J. Mol. Spectr. 27, 72 (1968).
- [5] C.L. Wang and F. Martino, Phys. Rev. B34, 5540 (1986).
- [6] W. Förner, C.L. Wang, F. Martino and J. Ladik, Phys. Rev. B 37, 4567 (1988).
- [7] W.P. Su and J.R. Schrieffer, Proc. Natl. Acad. Sci. USA 77, 5626 (1980).
- [8] W. Förner, M. Seel, and J. Ladik, Solid State Comm. 57, 463 (1986); J. Chem. Phys. 84, 5910 (1986); C.-M. Liegener, W. Förner, and J. Ladik, Solid State Comm. 61, 203 (1987); A. Godzik, M. Seel, W. Förner, and J. Ladik, Solid State Comm. 60, 209 (1986).
- [9] W. Förner, Solid State Comm. 63, 941 (1987).
- [10] S. Kivelson and D.E. Heim, Phys. Rev. B26, 4278 (1982).
- [11] F. Guinea, Phys. Rev. B30, 1884 (1984).
- [12] W. Förner, Synth. Metals 30, 135 (1989).
- [13] R. J. Hemley, B. R. Brooks and M. Karplus, J. Chem. Phys. 85, 6550 (1986); F. Zerbetto, M. Z. Zgierski, F. Negri, and G. Orlandi, J. Chem. Phys. 89, 3681 (1988); J. T. L. Navarrette and G. Zerbi, Synth. Met. 32, 151 (1989); G. Zerbi and G. Zannoni, J. Phys. C44, 3273 (1983); G. Zannoni and G. Zerbi, J. Mol. Struct. 100, 485 (1983); A. Peluso, M. Seel and J. Ladik, Can. J. Chem. 63, 1553 (1985); B. Horovitz, Phys. Rev. Lett. 47, 1491 (1981).
- [14] H. O. Villar, M. Dupuis, J. D. Watts, G. J. B. Hurst, and E. Clementi, J. Chem. Phys. 88, 1003 (1988); H. O. Villar, M. Dupuis and E. Clementi, J. Chem. Phys. 88, 5252 (1988); Phys. Rev. B37, 2520 (1988).
- [15] H. Guo and M. Karplus, J. Chem. Phys. 94, 3679 (1991).
- [16] W. Förner, Chem. Phys. 160, 173 (1992).
- [17] W. Förner, Chem. Phys. 160, 189 (1992).
- [18] K. F. Freed, J. Chem. Phys. 60, 1765 (1974).
- [19] S. Iwata and K. F. Freed, J. Chem. Phys. 61, 1500 (1974).
- [20] Y. S. Lee, K. F. Freed, H. Sun, and D. L. Yeager, J. Chem. Phys. 79, 3862 (1983).
- [21] T. Kovar, Master Thesis, University Erlangen-Nürnberg, FRG (1986).
- [22] W. Förner, Phys. Rev. B44, 11743 (1991); J. Mol. Struct. (Theochem) 282, 235 (1993).
- [23] P. Saalfrank, Ph. D. Thesis, University Erlangen-Nürnberg, FRG (1991).
- [24] J. L. Bredas and A. J. Heeger, Chem. Phys. Lett. 154, 56 (1989).
- [25] M. Löglund, P. Dannetun, S. Stafström, W. R. Salaneck, M. G. Ramsey, C. W. Spangler, C. Frederiksson, and J. L. Bredas, Phys. Rev. Lett. 70, 970 (1993).
- [26] J.S. Binkley, R.A. Whiteside, P.C. Hariharan, R. Seeger, and J.A. Pople, GAUSSIAN 76, QCPE No. 368 (1976).

- [27] C. Møller and M.S. Plesset, *Phys. Rev.* 46, 618 (1934).
- [28] J. Čížek, *J. Chem. Phys.* 45, 4256 (1966).
- [29] W. Förner, J. Čížek, P. Otto, J. Ladik, and E.O. Steinborn, *Chem. Phys.* 97, 235 (1985); W. Förner, J. Ladik, P. Otto, and J. Čížek, *Chem. Phys.* 97, 251 (1985); W. Förner, L. Pylypow, and J. Čížek, *Chem. Phys.* 110, 355 (1986); W. Förner, *Chem. Phys.* 114, 21 (1987).
- [30] D.G. Jenkins, A.D. Mitchell, and L.C. Cross, eds., "Tables of interatomic distances and configurations in molecules and ions", Special Publication No. 11 Chem. Soc., (London, 1956).
- [31] M. Takahashi and J. Paldus, *Can. J. Phys.* 62, 1226 (1984).
- [32] S. Suhai, *Phys. Rev.* B27, 3506 (1983).
- [33] C.-M. Liegener, *J. Chem. Phys.* 88, 6999 (1988).
- [34] K. Shulten, I. Ohmine, and M. Karplus, *J. Chem. Phys.* 64, 4422 (1976), and references therein.
- [35] H. Wendel, M. Seel, and J. Ladik, *Solid State Comm.* 54, 551 (1983).
- [36] S.N. Dixit and S. Mazumdar, *Phys. Rev.* B29, 1824 (1984).
- [37] R.S. Mulliken, *J. Chem. Phys.* 46, 675 (1949).
- [38] J. Ladik, *Acta Phys. Acad. Sci. Hung.* 18, 173 (1965).
- [39] J. Paldus and E. Chin, *Int. J. Quant. Chem.* 24, 373 (1983); J. Paldus, E. Chin and M.G. Grey, *Int. J. Quant. Chem.* 24, 395 (1983).
- [40] S. Kivelson, W.-P. Su, J.R. Schrieffer, and A.J. Heeger, *Phys. Rev. Lett.* 58, 1899 (1987).
- [41] H. Brongersma, Ph. D. thesis, Leyden (1968); J.H. Moore Jr., *J. Phys. Chem.* 76, 1130 (1972); O.A. Mosher, W.M. Flocker, and A. Kuppermann, *J. Chem. Phys.* 59, 6502 (1973).
- [42] G. Herzberg, "Electronic Spectra of Polyatomic Molecules" (van Nostrand, Princeton, 1966), p. 656.
- [43] R. McDiarmid, *J. Chem. Phys.* 64, 514 (1976).
- [44] R.R. Chadwick, D.P. Gerrity, and B.S. Hudson, *Chem. Phys. Lett.* 115, 24 (1985).
- [45] R.J. Buenker, S. Shih, and S.D. Peyerimhoff, *Chem. Phys. Lett.* 44, 385 (1976).
- [46] M. Aoyagi, Y. Osamura, and S. Iwata, *J. Chem. Phys.* 83, 1140 (1985).
- [47] M.A.C. Nascimento and W.A. Goddard III, *Chem. Phys.* 36, 147 (1979).
- [48] V. Galasso, *J. Chem. Phys.* 89, 4529 (1988).
- [49] A. C. Lasaga, R. J. Aerni and M. Karplus, *J. Chem. Phys.* 73, 5230 (1980).
- [50] P. G. Szalay, A. Karpfen and H. Lischka, *Chem. Phys.* 130, 219 (1989).
- [51] B. Hudson, B. E. Kohler and K. Schulten, *Excited States* 6, 1 (1982); B. S. Hudson and B. E. Kohler, *Chem. Phys. Lett.* 14, 299 (1972); *J. Chem. Phys.* 59, 4984 (1973); K. Schulten and M. Karplus, *Chem. Phys. Lett.* 14, 305 (1972); M. F. Granville, G. R. Holtom and B. E. Kohler, *J. Chem.*

- Phys. 72, 4671 (1980); R. Snijder, E. Arvidson, C. Foote, L. Harrigan, and R. L. Christensen, J. Amer. Chem. Soc. 107, 4117 (1985); J. H. Simpson, J. McLaughlin, D. S. Smith, and R. L. Christensen, J. Chem. Phys. 87, 3360 (1987); B. E. Kohler in "Conjugated Polymers: The Novel Science and Technology of Conducting and Nonlinear Optically Active Materials", eds. J. L. Bredas and R. Silbey (Kluwer, Amsterdam, to be published).
- [52] B. E. Kohler and J. B. Snow, J. Chem. Phys. 79, 2134 (1983).
- [53] P. M. Johnson, J. Chem. Phys. 64, 4638 (1976); L. J. Rothberg, D. P. Gerrity and V. Vaida, J. Chem. Phys. 73, 5508 (1980); J. P. Doering and R. McDiarmid, J. Chem. Phys. 75, 2477 (1981); R. R. Chadwick, D. P. Garrity and B. S. Hudson, Chem. Phys. Lett. 115, 24 (1985); T. Fuji, A. Kamati, M. Shimuru, Y. Aadachi, and S. Maeda, Chem. Phys. Lett. 115, 369 (1985); R. J. Cave and E. R. Davidson, J. Phys. Chem. 91, 4481 (1987); *ibidem* 92, 614 (1988); Chem. Phys. Lett. 148, 190 (1988).
- [54] W. J. Buma, B. E. Kohler and K. Song, J. Chem. Phys. 92, 4622 (1990).
- [55] W. J. Buma, B. E. Kohler and K. Song, J. Chem. Phys. 94, 6367 (1991).
- [56] R. J. Hemley, J. I. Dawson and V. Vaida, J. Chem. Phys. 78, 2915 (1983).
- [57] M. F. Granville, B. E. Kohler and J. B. Snow, J. Chem. Phys. 75, 3765 (1981).
- [58] D. G. Leopold, R. D. Pendley, J. L. Roebber, R. J. Hemley, and V. Vaida, J. Chem. Phys. 81, 4218 (1984).
- [59] R. McDiarmid and A.-H. Sheybani, J. Chem. Phys. 89, 1255 (1988).
- [60] (a) J. Ladik, "Quantenchemie", F. Enke Verlag (Stuttgart, 1973).
(b) P.-O. Löwdin, unpublished result (1954) according to Ref. [151] in [60a]; R. G. Parr, J. Chem. Phys. 33, 1184 (1960).
- [61] J. Hubbard, Proc. Roy. Soc. London A276, 238 (1963); R. G. Parr, J. Chem. Phys. 20, 1499 (1952); R. Pariser and R. G. Parr, *ibidem* 21, 767 (1953); J. A. Pople, Proc. Roy. Soc. London A68, 81 (1955); L. Salem, "Molecular Orbital Theory of Conjugated Systems", W. A. Benjamin, New York 1966.
- [62] K. Ohno, Theor. Chim. Acta 2, 219 (1964).
- [63] See for example J. T. Gammel and D. K. Campbell, Phys. Rev. Lett. 60, 70 (1988); A. Painelli and A. Girlando, Solid State Commun. 66, 273 (1988); A. Painelli, Synth. Met. 27, A15 (1988); D. K. Campbell, J. T. Gammel and E. Y. Loh, Synth. Met. 27, A9 (1988).
- [64] P. Vogl and D. K. Campbell, Phys. Rev. B41, 12797 (1990).
- [65] W. Förner, J. Ladik, D. Hofmann, M. Seel, A. Godzik, Godzik, and F. Martino, J. Mol. Struct. (Theochem) 188, 231 (1989).
- [66] P.-O. Löwdin, Phys. Rev. 97, 1509 (1955); J. Appl. Phys. Suppl. 33, 251 (1962); Ann. Acad. Reg. Sci. Upsaliensis 2, 127 (1958); Rev. Mod. Phys. 32, 328 (1960); P.-O. Löwdin (ed.), "Quantum Theory of Atoms, Molecules, and Solid State; a Tribute to John. C. Slater" (Academic Press,

- New York 1966).
- [67] (a) F. Martino and J. Ladik, *J. Chem. Phys.* 52, 2262 (1970); I. Mayer, J. Ladik and G. Biczó, *Intern. J. Quantum Chem.* 7, 583 (1973).
- (b) H. B. Schlegel, *J. Chem. Phys.* 84, 4530 (1986); *J. Phys. Chem.* 92, 3075 (1988); P. J. Knowles and N. C. Handy, *J. Phys. Chem.* 92, 3097 (1988); *J. Chem. Phys.* 88, 6991 (1988); P. M. W. Gill, M. W. Wong, R. H. Nobes, and L. Radom, *Chem. Phys. Lett.* 148, 541 (1988).
- (c) P.-O. Löwdin, *Proc. Symp. Mol. Phys. at Nikko, Japan, 1953*, p. 13, (Maruzen, 1954); *Proc. 10th Solvay Conf., "Les Electrons dans le Metaux"*, p. 71 (Bruxelles 1955); *Svensk Kemisk Tidskr.* 67, 365 (1955); T. Itoh and H. Yoshizumi, *J. Phys. Soc. Japan* 10, 201 (1955); *J. Chem. Phys.* 23, 412 (1955); *Busseiron Kenkyu* 83, 13 (1955); P.-O. Löwdin, *Texas J. Sci.* 8, 163 (1956); *Ed. du Centre Nat. Rech. Sci* 82 (1958); R. Lefebvre, H. H. Dearman and H. M. McConnell; *J. Chem. Phys.* 32, 176 (1960); P.-O. Löwdin, R. Pauncz and J. de Heer, *J. Chem. Phys.* 36, 2247, 2257 (1962); R. Pauncz, *J. Chem. Phys.* 37, 2739 (1962); J. de Heer, *Rev. Mod. Phys.* 35, 631 (1963); R. Pauncz, in "Molecular Orbitals in Chemistry, Physics and Biology", P.-O. Löwdin (Ed.), Academic Press (New York, 1964), p. 433; *Tetrahedron* 19, Suppl. 2, 43 (1963); *J. Chem. Phys.* 43, S69 (1965); O. Goscinski and J. L. Calais, *Arkiv Fysik* 29, 135 (1965); J. de Heer and R. Pauncz, *J. Chem. Phys.* 39, 2314 (1963); R. Pauncz, "Alternant Molecular Orbital Method (W. B. Saunders, Philadelphia 1967); J. L. Calais, *Arkiv Fysik* 28, 479, 511 539 (1965); *ibidem* 29, 255 (1965); *Intern. J. Quantum Chem.* 13, 661 (1967).
- (d) I. Mayer, *Adv. Quantum Chem.* 12, 189 (1980).
- [68] F. Martino and J. Ladik, *Phys. Rev.* A3, 862 (1971).
- [69] G. König and G. Stollhoff, *Phys. Rev. Lett.* 65, 1239 (1990).
- [70] S. Suhai, Habilitation Thesis, University Erlangen-Nürnberg, FRG (1983); see also Ref. [14]; J. Ladik, "Quantum Theory of Polymers as Solids", Plenum Publ. Corp. (New York, London, 1988) Chapter 5.
- [71] H. Kahlert, O. Leitner, and G. Leising, *Synth. Met.* 17, 467 (1987).
- [72] J. Avery, J. Packer, J. Ladik, and G., Biczó, *J. Mol. Spectr.* 29, 194 (1969); J. Ladik, *Int. J. Quant. Chem.* 4, 307 (1971).
- [73] J. Ladik, "Quantum Theory of Polymers as Solids", Plenum Publ. Corp. (New York, London, 1988) Chapter 5; S. Suhai and J. Ladik, *J. Phys.* C15, 4327 (1982).
- [74] D. Vanderbilt and E. J. Mele, *Phys. Rev.* B22, 3939 (1980).
- [75] G. Zannoni and G. Zerbi, *Chem. Phys. Lett.* 87, 55 (1982).
- [76] A. Peluso, M. Seel, and J. Ladik, *Can. J. Chem.* 63, 1553 (1985).

- [77] G. W. Hayden and Z. G. Soos, *Phys. Rev.* **B38**, 6075 (1988).
- [78] M. Takahashi and J. Paldus, *Can. J. Phys.* **62**, 1226 (1984).
- [79] D. S. Boudreaux, R. R. Chance, J. L. Bredas, and R. Silbey, *Phys. Rev.* **B28**, 6927 (1983).
- [80] M. Mehring, A. Grupp, P. Höfer, and H. Käss, *Synth. Met.* **28**, D399 (1989); H. Käss, P. Höfer, A. Grupp, P. K. Kahol, R. Weizenhöfer, G. Wegner, and M. Mehring, *Europhys. Lett.* **4**, 947 (1987).
- [81] S. Kuroda, H. Bando and H. Shirakawa, *Solid State Commun.* **52**, 893 (1984).
- [82] S. Kuroda, H. Bando and H. Shirakawa, *J. Phys. Soc. Jpn.* **54**, 3956 (1985); S. Kuroda and H. Shirakawa, *Synth. Met.* **17**, 423 (1987).
- [83] J. E. Hirsch and M. Grabowski, *Phys. Rev. Lett.* **52**, 1713 (1984).
- [84] D. K. Campbell, T. A. DeGrand and S. Mazumdar, *Phys. Rev. Lett.* **52**, 1717 (1984); *Mol. Cryst. Liq. Cryst.* **118**, 41 (1985).
- [85] Z. G. Soos and S. Ramasesha, *Phys. Rev. Lett.* **51**, 2374 (1983); S. Ramasesha and Z. G. Soos, *Synth. Met.* **9**, 283 (1984).
- [86] H. Fukutome and M. Sasai, *Prog. Theor. Phys.* **67**, 41 (1982); M. Sasai and H. Fukutome, *Prog. Theor. Phys.* **69**, 1, 373 (1983); **70**, 1471 (1983); and in "Dynamical Problems in Soliton Systems", ed. S. Takeno (Springer Verlag, Berlin, 1985) p. 205.
- [87] K. Tanaka, K. Yoshizawa and T. Yamabe, *Intern. J. Quantum Chem.* **40**, 305, 315 (1991).
- [88] C.-M. Liegener, *J. Phys.* **C18**, 6011 (1985); C.-M. Liegener and J. Ladik, *Phys. Rev.* **B35**, 6403 (1987).
- [89] Z. Vardeny, J. Orenstein and G. L. Baker, *Phys. Rev. Lett.* **50**, 2023 (1983); Z. Vardeny and J. Tauc, *Phys. Rev. Lett.* **54**, 1844 (1985); **56**, 1510 (1986).
- [90] R. H. Friend, D. D. C. Bradley and P. D. Townsend, *J. Phys.* **D20**, 1367 (1987); N. F. Colaneri, R. H. Friend, H. E. Schaffer, and A. J. Heeger, *Phys. Rev.* **B38**, 3960 (1988).
- [91] L. Robins, J. Orenstein and R. Superfine, *Phys. Rev. Lett.* **56**, 1850 (1986); C. G. Levy, D. V. Lang, S. Etamad, G. L. Baker, and J. Orenstein, *Synth. Met.* **17**, 569 (1987); M. Sinclair, D. Moses, R. H. Friend, and A. J. Heeger, *Phys. Rev.* **B36**, 4296 (1987); P. D. Townsend and R. H. Friend, *Phys. Rev.* **B40**, 3112 (1989).
- [92] S. Kivelson and W.-K. Wu, *Phys. Rev.* **B34**, 5540 (1986).
- [93] C.-M. Liegener, W. Förner, and J. Ladik, *Solid State Comm.* **61**, 203 (1987).
- [94] O. Rubner, W. Förner and J. Ladik, *Synth. Met.* (accepted).
- [95] W. Förner (unpublished)
- [96] A. S. Davydov, *Zh. Eksp. Teor. Fiz.* **78**, 789 (1980), *Sov. Phys. JETP* **51**, 397 (1980).

**MOLECULAR QUANTUM SIMILARITY:
THEORETICAL FRAMEWORK,
ORDERING PRINCIPLES, AND
VISUALIZATION TECHNIQUES^{1,2}**

*Ramon Carbó³
Blanca Calabuig
Leonel Vera⁴
Emili Besalú*

Institut de Química Computacional
Universitat de Girona
17071 Girona (Spain)

Keywords:: Quantum Molecular Similarity Measures.
Similarity Indices. Density Transforms. Molecular Point Clouds.
Visualization Techniques. Mendelev Postulates.

¹ *A contribution of the Grup de Química Quàntica de l'Institut d'Estudis Catalans.*

² *From the proceedings of the First Congress of the International Society for Theoretical Chemical Physics. June 28th - July 3rd, 1993, Girona, Spain.*

³ *To whom correspondence should be addressed.*

⁴ *Visiting Professor. Universidad Católica del Norte. Antofagasta. Chile.*

LIST OF CONTENTS

1. Introduction

2. Beyond the Expectation Value Concept: Expectation Functions

3. General Quantum Similarity Measures

3.1. General Definition

3.2. Particular Cases and Simplified Forms

3.3. Similarity Matrices

4. Approximate QSM

5. Manipulation of Similarity Measures: Similarity Indices

6. Ordering Molecular Sets: Mendeleev's Conjecture

7. Practical QSM Computation

8. Visualization Techniques

9. Application Examples

9.1. β -Diketones

9.2. Retinoid Compounds

9.3. Steroid Hormones

9.4. Odor molecules

10. Conclusions

Acknowledgments

Appendix A: Figures

Appendix B: Molecular Quantum Similarity Associated Integrals

A.1. Integrals over n S STO functions

A.2. General Integrals over GTO functions

References

1. INTRODUCTION

In previous papers [1,2] the authors have worked out the theoretical foundation of Quantum Similarity (QS). In this paper, several practical results will be listed. All of them came from some of the Quantum Similarity Measures (QSM) defined previously and are particular cases of the general QSM definition obtained in reference [1].

All the work developed in our Laboratory concerning QSM is focussed on defining and interpreting a QSM according to quantum mechanical principles and to accompany this by a construction of a QSM theoretical framework as general as possible. Also, our research is concerned with handling the obtained measures by means of results manipulation and visualization within pure geometrical ideas.

In order to construct a solid theoretical body concerning QSM, a set of rules which are referred to here as the **Mendeleev Postulates** are described. They are listed at the end of this paper in the Conclusions Section.

2. BEYOND THE EXPECTATION VALUE CONCEPT: EXPECTATION FUNCTIONS

The role played by the expectation value concept in Quantum Mechanics [3] is well known. Here it is shown how this concept can be generalized leading to the description of **Expectation functions**.

Quantum Similarity can be reformulated in terms of expectation functions as will be shown next.

Let us define the *transform* of a function $f(\mathbf{r}, \mathbf{u}, \mathbf{p})$ as the following integral [4]:

$$T(\mathbf{r}, \mathbf{s}, \mathbf{p}) = \int \Omega(\mathbf{r}, \mathbf{s}, \mathbf{u}, \mathbf{p}) f(\mathbf{r}, \mathbf{u}, \mathbf{p}) d\mathbf{u} , \quad (1)$$

where $\Omega(\mathbf{r}, \mathbf{s}, \mathbf{u}, \mathbf{p})$ is the *kernel* of the transform. The optional vector \mathbf{p} provides the dependence of the function, the operator or both with respect to a parameter set. In this manner, an ***n*-th order Density Integral Transform** can be defined as the transformation of a density matrix element [5] of the same order:

$$P^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) = \int \Omega(\mathbf{r}, \mathbf{s}, \mathbf{u}, \mathbf{p}) \rho^{(n)}(\mathbf{r}, \mathbf{u}, \mathbf{p}) d\mathbf{u} , \quad (2)$$

where \mathbf{p} is the optional parameter vector.

It can be said that the density transform (DT) in equation (2) constructs an ***n*-th order Expectation Function** of the operator Ω . From an expectation function of *n*-th order it can be obtained some **Expectation Value** by means of the evaluation of the expectation function at a fixed point $(\mathbf{r}_o, \mathbf{s}_o, \mathbf{p}_o)$ [1].

DT can have several forms depending on the structure adopted by the kernel which appears in equation (2), see reference [1] for more details. Here we will only deal with the following two cases:

- a) when the operator Ω appearing in equation (2) is defined as $\Omega(\mathbf{r}, \mathbf{s}, \mathbf{u}, \mathbf{p}) = \delta(\mathbf{u} - \mathbf{s})$, the DT becomes $P^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) = \rho^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p})$

and the transformation leaves invariant the density matrix element.

- b) if Ω is defined as $\Omega(\mathbf{r}, \mathbf{s}, \mathbf{u}, \mathbf{p}) = \delta(\mathbf{u} - \mathbf{r})$, then $P^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) = P^{(n)}(\mathbf{r}, \mathbf{p}) = \rho^{(n)}(\mathbf{r}, \mathbf{p})$ and the transformation returns a diagonal element of the density matrix: a density function [5].

3. GENERAL QUANTUM SIMILARITY MEASURES

We will now describe a general definition of QSM. Then, some special cases will be analyzed. The formulation tries to complete the development of various new computational techniques produced by other groups [6].

3.1. General Definition

The n -th order Quantum Similarity Measure between a set of m systems collected in the vector $\mathbf{A} = (A_1, A_2, \dots, A_m)$, with respect to an operator W , can be defined as an integral of the following kind [1], which involves a set of DT:

$$Z_A^{(n)}(W, \mathbf{R}^{(m+1,q)}, \mathbf{S}^{(m+1,q)}, \mathbf{p}) = \iint W(\mathbf{R}^{(1,q)}, \mathbf{S}^{(1,q)}, \mathbf{p}) \left(\prod_{i=1}^m P_{A_i}^{(n)}(\mathbf{r}_i, \mathbf{s}_i, \mathbf{p}) \right) d\mathbf{R}^{(1,m)} d\mathbf{S}^{(1,m)} . \quad (3)$$

Considering the previous equation (3) some particular aspects have to be taken into account, the most important are:

- a) The vector $\mathbf{n}=(n_1, n_2, \dots, n_m)$ collects the order of the DT of the set of systems which are active in the vector \mathbf{A} .
- b) Vector \mathbf{R} is defined as containing a set of vector coordinates: $\mathbf{R}^{(i,j)}=(\mathbf{r}_i, \mathbf{r}_{i+1}, \dots, \mathbf{r}_j)$. The corresponding differential vector has a related structure: $d\mathbf{R}^{(i,j)}=d\mathbf{r}_i d\mathbf{r}_{i+1} \dots d\mathbf{r}_j$. The same definitions can be considered to hold for the vector \mathbf{S} .
- c) If $m \geq q$ the measure in equation (3) does not depend on \mathbf{r} or \mathbf{s} coordinates.
- d) It is possible to define the operator \mathbf{W} in equation (3) with an embedded Dirac delta function, appearing as a multiplicative factor.
- e) The operator \mathbf{W} is usually taken as a positive definite operator.

3.2. Particular Cases and Simplified Forms

Despite the QSM general definition (3), only QSM involving two or three quantum systems will be treated here. Moreover, the \mathbf{W} operator will be taken in a simplified form as only depending on the $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{p})$ coordinates. A QSM involving two quantum systems \mathbf{A} and \mathbf{B} and their respective DT of order n and m is defined now as:

$$Z_{AB}^{(n,m)}(\mathbf{W}, \mathbf{p}) = \iiint \mathbf{W}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{p}) P_A^{(n)}(\mathbf{r}_1, \mathbf{s}_1, \mathbf{p}) P_B^{(m)}(\mathbf{r}_2, \mathbf{s}_2, \mathbf{p}) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{s}_1 d\mathbf{s}_2 . \quad (4)$$

Particular cases of previous equation (4) arise when the operator \mathbf{W} takes concrete forms and $n=m$.

- a) When W is defined as the following Dirac delta functions product:

$$W(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{p}) = \delta(\mathbf{r}_1 - \mathbf{r}_2) \delta(\mathbf{s}_1 - \mathbf{s}_2) \quad (5)$$

then, one is dealing with an **Overlap-like QSM**:

$$Z_{AB}^{(n)}(\delta(\mathbf{r}_1 - \mathbf{r}_2) \delta(\mathbf{s}_1 - \mathbf{s}_2), \mathbf{p}) = \iint P_A^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) P_B^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) d\mathbf{r} d\mathbf{s} . \quad (6)$$

- b) If W is a product of two Coulomb operators,

$$W(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{p}) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2| |\mathbf{s}_1 - \mathbf{s}_2|} , \quad (7)$$

one obtains the following QSM:

$$Z_{AB}^{(n)}\left(\frac{1}{r_{12} s_{12}}, \mathbf{p}\right) = \iiint \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \frac{1}{|\mathbf{s}_1 - \mathbf{s}_2|} \quad (8)$$

$$P_A^{(n)}(\mathbf{r}_1, \mathbf{s}_1, \mathbf{p}) P_B^{(n)}(\mathbf{r}_2, \mathbf{s}_2, \mathbf{p}) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{s}_1 d\mathbf{s}_2 ,$$

which can be considered an **Electrostatic Potential QSM**.

- c) When the involved DT are density functions, the counterparts of equations (6) and (8) read [2.i,2.j]:

$$Z_{AB}^{(n)}(\delta(\mathbf{r}_1 - \mathbf{r}_2), \mathbf{p}) = \int P_A^{(n)}(\mathbf{r}, \mathbf{p}) P_B^{(n)}(\mathbf{r}, \mathbf{p}) d\mathbf{r} . \quad (9)$$

and

$$Z_{AB}^{(n)}\left(\frac{1}{r_{12}}, \mathbf{p}\right) = \iint \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} P_A^{(n)}(\mathbf{r}_1, \mathbf{p}) P_B^{(n)}(\mathbf{r}_2, \mathbf{p}) d\mathbf{r}_1 d\mathbf{r}_2 . \quad (10)$$

respectively. The equation (10) may be referred as a **Coulomb-like QSM**.

- d) In a recent discussion on QSM [2.1] we have presented a practical computational framework derived from a previous idea [2.j], concerning the positive definite nature of the n -th order density functions. Due to this property, which can be attached to every density function, one can consider any other system's appropriate DT as the positive definite operator W in equation (4). A **Triple Density Transform Similarity Measure** (TDTSM) [2.1] is then constructed. In terms of n -th order density functions it reads:

$$T_{AB;C}^{(n)} = \int P_A^{(n)}(\mathbf{r}, \mathbf{p}) P_C^{(n)}(\mathbf{r}, \mathbf{p}) P_B^{(n)}(\mathbf{r}, \mathbf{p}) d\mathbf{r} . \quad (11)$$

- e) In this case, it is more general to consider the following TDTSM where the third system shares a non-diagonal element of its density matrix:

$$T_{AB;C}^{(n)} = \iint P_A^{(n)}(\mathbf{r}, \mathbf{p}) P_C^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) P_B^{(n)}(\mathbf{s}, \mathbf{p}) d\mathbf{r} d\mathbf{s} . \quad (12)$$

- f) This leads to a generalization of an idea of Cioslowski [6.d] concerning QSM involving first order density matrix elements: one can define an alternative form as the off-diagonal DT function matrix measure:

$$Z_{AB}^{(n)} = \iint P_A^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) P_B^{(n)}(\mathbf{r}, \mathbf{s}, \mathbf{p}) d\mathbf{r} d\mathbf{s} , \quad (13)$$

which can be easily extended inside our scheme, defining the counterpart of equation (13) as:

$$Z_{AB}^{(n)}(\mathbf{W}) = \iiint P_A^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}) W(\mathbf{r}_2, \mathbf{r}_3, \mathbf{p}) P_B^{(n)}(\mathbf{r}_3, \mathbf{r}_1, \mathbf{p}) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 . \quad (14)$$

g) This approach may be particularly useful in the TDTSM case, leading to the following integral:

$$T_{AB;C}^{(n)} = \iiint P_A^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}) P_C^{(n)}(\mathbf{r}_2, \mathbf{r}_3, \mathbf{p}) P_B^{(n)}(\mathbf{r}_3, \mathbf{r}_1, \mathbf{p}) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 . \quad (15)$$

3.3. Similarity Matrices

As it can be deduced from the discussion of Sections 3.1 and 3.2 above, by means of equation (3), and the subsequent simplifications, every quantum system A_i can be represented by a hypermatrix. In the framework of the particular cases which are developed here, the representation can be done by means of a matrix notation whose elements can be taken as:

$$\mathbf{Z}(\mathbf{W}) = \{Z_{IJ}(\mathbf{W})\} . \quad (16)$$

In the particular case of the TDTSM, the matrix notation can be achieved by means of the following reduction:

$$Z_{IJ}(W) = \sum_i \sum_j Z_{Iij}(W) Z_{Jij}(W) , \quad (17)$$

which possess a similar structure to a scalar product.

4. APPROXIMATE QSM.

When dealing with density matrix elements, it is interesting to have access to expressions describing these elements in the cheapest possible way. This is so because in the practical QSM computation there is involved an intermolecular position optimization. Obtaining good fittings for density matrix elements can be a time saving device in order to speed QSM computation. Several studies in this direction are performed in our Laboratory. An application to first order density matrix follows.

From a modified idea of Ruedenberg [7] concerning charge distributions, one can express a product of two orbital functions in the following manner:

$$\Omega_{ab}^{AB}(r, r') = |\phi_a^A(r) \rangle \langle \phi_b^B(r')| \approx \sum_K \sum_{\alpha, \beta \in K} g_{ab, \alpha \beta}^{AB, K} \omega_{\alpha \beta}^K(r, r') , \quad (18)$$

where the indices α, β collect all the pairs of function products centered at point K , $\{g_{ab, \alpha \beta}^{AB, K}\}$ are linear coefficients and $\{\omega_{\alpha \beta}^K\}$ one-center products of orbital functions:

$$\omega_{\alpha \beta}^K(r, r') = |\zeta_{\alpha}^K(r) \rangle \langle \zeta_{\beta}^K(r')| . \quad (19)$$

In general, centers K are arbitrary, covering the possibility to be the centers A or B among others. The nature of each one-center density

$\{\omega_{\alpha\beta}^K\}$ set must be chosen as to be rotationally invariant, this means that function $\{\zeta_{\alpha}^K\}$ sets, appearing in the definition (19) of one-center density $\{\omega_{\alpha\beta}^K\}$ set are complete.

A first order density matrix element is written as [5]:

$$D(\mathbf{r}, \mathbf{r}') = \sum_A \sum_B \sum_{a \in A} \sum_{b \in B} D_{ab}^{AB} \Omega_{ab}^{AB}(\mathbf{r}, \mathbf{r}') . \quad (20)$$

In this manner after substituting (18) into (20), one can collect the constant elements as:

$$G_{\alpha\beta}^K = \sum_A \sum_B \sum_{a \in A} \sum_{b \in B} D_{ab}^{AB} g_{ab, \alpha\beta}^{AB, K} , \quad (21)$$

so equation (20) now reads:

$$D(\mathbf{r}, \mathbf{r}') \approx \sum_K \sum_{\alpha, \beta \in K} G_{\alpha\beta}^K \omega_{\alpha\beta}^K(\mathbf{r}, \mathbf{r}') . \quad (22)$$

This general form can be transferred to density functions, taken as a particular case of the previous equation because they are the diagonal elements of the density matrix. In this context, the companion of equation (22) will be:

$$D(\mathbf{r}) \approx \sum_K \sum_{\alpha, \beta \in K} G_{\alpha\beta}^K \omega_{\alpha\beta}^K(\mathbf{r}) , \quad (23)$$

where here, the functions $\{\omega_{\alpha\beta}^K\}$ are one-center charge distributions defined as:

$$\omega_{\alpha\beta}^K(\mathbf{r}) = |\zeta_{\alpha}^K(\mathbf{r})\rangle \langle \zeta_{\beta}^K(\mathbf{r})| . \quad (24)$$

Under this approach, a QSM having an structure like the integral (4) becomes now:

$$Z_{AB}^{(1)}(\mathbf{W}, \mathbf{p}) = \sum_{K \in A} \sum_{L \in B} \sum_{\alpha, \beta \in K} \sum_{\lambda, \sigma \in L} G_{\alpha\beta}^K G_{\lambda\sigma}^L \iiint W(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{p}) \quad (25)$$

$$\omega_{\alpha\beta}^K(\mathbf{r}_1, \mathbf{s}_1) \omega_{\lambda\sigma}^L(\mathbf{r}_2, \mathbf{s}_2) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{s}_1 d\mathbf{s}_2 ,$$

or, in the charge distribution case, this will be:

$$Z_{AB}^{(1)}(\mathbf{W}, \mathbf{p}) = \sum_{K \in A} \sum_{L \in B} \sum_{\alpha, \beta \in K} \sum_{\lambda, \sigma \in L} G_{\alpha\beta}^K G_{\lambda\sigma}^L \quad (26)$$

$$\iint W(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}) \omega_{\alpha\beta}^K(\mathbf{r}_1) \omega_{\lambda\sigma}^L(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 .$$

This final form suggests a general expression for approximate QSM. Collecting the coefficients and the functions as follows:

$$R_{\alpha\beta}^K(\mathbf{r}, \mathbf{r}') = G_{\alpha\beta}^K \omega_{\alpha\beta}^K(\mathbf{r}, \mathbf{r}') , \quad (27)$$

the QSM in equation (25) becomes:

$$Z_{AB}^{(1)}(\mathbf{W}, \mathbf{p}) = \sum_{K \in A} \sum_{L \in B} \sum_{\alpha, \beta \in K} \sum_{\lambda, \sigma \in L} \iiint W(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{p}) \quad (28)$$

$$R_{\alpha\beta}^K(\mathbf{r}_1, \mathbf{s}_1) R_{\lambda\sigma}^L(\mathbf{r}_2, \mathbf{s}_2) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{s}_1 d\mathbf{s}_2 .$$

When one is dealing with a TDTQSM, expression (11) has the structure:

$$T_{AB;C}^{(1)} = \sum_{K \in A} \sum_{L \in B} \sum_{M \in C} \sum_{\alpha, \beta \in K} \sum_{\lambda, \sigma \in L} \sum_{\gamma, \delta \in M} \quad (29)$$

$$\iint R_{\alpha\beta}^K(\mathbf{r}, \mathbf{s}) R_{\lambda\sigma}^L(\mathbf{r}, \mathbf{s}) R_{\gamma\delta}^M(\mathbf{r}, \mathbf{s}) d\mathbf{r} d\mathbf{s} .$$

These kinds of equations have been used by the authors [2.j,2.k,2.l] in order to obtain practical QSM values for molecular systems.

The diagonal terms of functions (19) can be considered now. When rotational invariance is enforced the set of one-center functions becomes:

$$\{ \omega_{\alpha\beta}^K \delta(\alpha=\beta, \alpha, \beta \in \sigma) \} = \{ |S_{\alpha}^K\rangle \langle S_{\alpha}^K| \} , \quad (30)$$

where the delta symbol stands for a Logical Kronecker Delta [8], and gives the value of unity if all the logical expressions placed inside as arguments are true, otherwise it returns a value of zero. In this manner, in equation (30), one can take into account diagonal terms which share the same shell, σ .

In this context, within a CNDO-like approach, one can suppose that the distributions $\{R_{\alpha}^K\}$ of equation (27) are defined for each atomic center as the product of a Mulliken total gross shell population [9] and a nS distribution:

$$R_{\alpha}^K(\mathbf{r}) = Q_{\alpha\alpha}^K |S_{\alpha}^K(\mathbf{r})|^2 , \quad (31)$$

where $S_{\alpha}^K(\mathbf{r}) \equiv S_{\alpha}(\mathbf{r}-\mathbf{r}_K)$ is a nS orbital centered at \mathbf{r}_K .

Previous terms in equation (31) can be collected into a unique expression such as:

$$R^K(\mathbf{r}) = Q^K |S^K(\mathbf{r})|^2 , \quad (32)$$

where the term Q^K is a Mulliken gross atomic population:

$$Q^K = \sum_{\alpha \in K} Q_{\alpha}^K, \quad (33)$$

the orbital $S^K(r)$ can be taken as a unique nS function representing all the original nS functions $\{S_{\alpha}^K\}$ appearing in equation (31).

Under this point of view, it is also possible to consider other definitions for the distributions (27). For example using the diagonal population elements $\{N_{\alpha}^K\}$ described by Huzinaga et al. [10], the distribution (27), when applied to the charge distribution case, reads now:

$$R_{\alpha}^K(r) = N_{\alpha}^K |S_{\alpha}(r - \langle r \rangle_K)|^2, \quad (34)$$

where $\langle r \rangle_K$ is a centroid coordinate defined in such a manner as to reproduce the electronic part of the dipole moment.

5. MANIPULATION OF SIMILARITY MEASURES: SIMILARITY INDICES

As has been discussed in several previous papers [2.j,2.k], the set of QSM can be transformed or combined in order to obtain a new kind of auxiliary terms which can be named *Similarity Indices*. There are a great deal of possible QSM manipulations leading to a variety of QSI definitions. Some of them are:

- a) *Cosine-like Similarity Index* between two molecules I and J is constructed as:

$$C_{IJ}(W) = Z_{IJ}(W) [Z_{II}(W) Z_{JJ}(W)]^{-1/2}. \quad (35)$$

- b) *Distance Similarity Index* as is shown in reference [2.a], by using this, a Dissimilarity Index can be defined:

$$D_{IJ}(\mathbf{W}) = [Z_{II}(\mathbf{W}) + Z_{JJ}(\mathbf{W}) - (Z_{IJ}(\mathbf{W}) + Z_{JI}(\mathbf{W}))]^{1/2} , \quad (36)$$

with an euclidean distance-like form.

- c) The *Hodgkin-Richards* [6.l, 6.m] or *Tanimoto* [11] *Indices* which we cast into a unique formula:

$$V_{IJ}(\mathbf{W}) = (2-k)Z_{IJ}(\mathbf{W}) [Z_{II}(\mathbf{W})+Z_{JJ}(\mathbf{W})-kZ_{IJ}(\mathbf{W})]^{-1} , \quad k \in [0,2) , \quad (37)$$

where with $k=0$ the *Hodgkin-Richards Index* is reproduced and when $k=1$ the *Tanimoto Index* is obtained.

- d) In general, a *Distance-like Index* A_{IJ} can be defined from a *Cosine-like Index* of equation (35) in the following way [2.j]:

$$A_{IJ}(\mathbf{W}) = K \cos^{-1}(C_{IJ}(\mathbf{W})) , \quad (38)$$

where K is a scale factor.

- e) A *Cosine-like Index* M_{IJ} can be defined in general from the set of *Distance-like Indices* (36) [2.j]:

$$M_{IJ}(\mathbf{W}) = 1 - D_{IJ}(\mathbf{W}) [\max(\{D_{IJ}(\mathbf{W})\})]^{-1} . \quad (39)$$

- f) It is possible to transmute a *Distance-like Index* into a *Correlation-like* one, N_{IJ} , using the *Hyperbolic Transformation*:

$$N_{IJ}(\mathbf{W}) = 1 - \tanh(D_{IJ}(\mathbf{W})) , \quad (40)$$

- g) A *Correlation-like Index* transforms into a *Distance-like* one, B_{IJ} , as:

$$B_{IJ}(W) = K \coth(C_{IJ}(W)) , \quad (41)$$

where K is a scale factor.

QSI provide with a "fuzzy" relationship the connections between the elements of Quantum Sets. See references [2.c,2.f,2.i] for more details. This relationship can be extended to DT Measures involving three or more Quantum Objects.

6. ORDERING MOLECULAR SETS: MENDELEEV'S CONJECTURE

Ordering of molecular sets can be assisted by the *Mendeleev Conjecture*[1]. Once a set of molecular structures is defined, $S=\{S_i\}$, it can be associated with the corresponding set of density transforms, $D=\{D_i\}$. This allows the computation of the QSM or QSI table which can be used to establish an ordering into the density transforms set by means of a rule R . This ordering can be transferred to the molecular set S .

Mendeleev Conjecture states that if a new element s of S is known, whose density function transform d is also known, all QSM or QSI involving d can be computed and the rule R can act on the augmented set $DU\{d\}$. In this manner, a position into the ordered set D is found for d . This ordering can be transferred to the corresponding structure s in the set S .

Supposing known a new set $P=\{P_i\}$ built up with some values of some property attached in a one-to-one correspondence to the elements of set S , *Mendeleev Conjecture* also states that the ordering of these properties follow the same structure as the one established for the density transforms. So, an unknown value p of the property attached to the new structure s may be found somewhere within the order of the property set. Applying the rule R over the augmented property set $PU(p)$, an estimation of the property value p , can be made.

7. PRACTICAL QSM COMPUTATION

When performing a computation of a QSM between two molecules there appears implicit the idea that a particular QSM value is not relevant, but the relative position between the involved molecules giving the maximum QSM value is. So, in practice, when Born-Oppenheimer approximation is used, along the QSM computation there is embedded an optimization procedure with respect to the relative rotational and translational positions of the molecules. From the general QSM definition (3), the final aim is to achieve the condition:

$$\frac{\partial}{\partial \mathbf{p}} Z_A^{(n)}(\mathbf{W}, \mathbf{R}^{(m+1,q)}, \mathbf{S}^{(m+1,q)}, \mathbf{p}) = 0, \quad (42)$$

where the vector \mathbf{p} collects the atomic positions given as parameters of the QSM value and one is searching for a maximal value of the QSM.

Some experience has been obtained in our Laboratory in this direction. The optimization procedure becomes the bottleneck of the

whole computation, see references [2.a,2.c,2.e] for more details. This can be helped using fittings for density transforms as those described in Section 4.

In the present work, the relative orientation between molecules has been obtained referring all the molecular structures with respect to a common origin: the center of charges. The dipole moment is used to orient the molecules. For apolar structures, the quadrupole principal components have been used.

8. VISUALIZATION TECHNIQUES

The matrix representation of a molecular set can be associated to a set of finite dimensional vectors representing the molecules. This leads to the concept of **Point-Molecules** collected as a **Molecular Point Cloud**. If the set contains m molecules, one is dealing with a m -dimensional space.

It is postulated that there exists an *optimal* finite dimensional *projection* of a given object set. The term *optimal projection* can have several meanings and it is by no means unique. See references [1,2.h] for more details.

Under this context, several classification algorithms can be applied, for example clustering techniques or, in general, those related to pattern recognition [11,12]. Principal components of the similarity matrix \mathbf{Z} , collecting QSM or QSI, constitute a possible choice to represent the Molecular Point Cloud. This choice has been tried in the

present work to give a finite dimensional representation of the DT set D , hence obtaining a visual representation of the molecular structures set S as outlined in Section 6. This kind of projection ensures the minimal loss of information from the original ∞ -dimensional representation where the D set belongs. Successive dimension reductions can be made by various means, see [2.f] for example. Using spatial multidimensional rotations, the final visualization can take care of the multiple aspects present in an m -dimensional representation.

Internal relationships between elements of an Object Set or, in particular, the Molecular Point Clouds can also be obtained by means of the construction of Graphs and Trees. Besides using classical techniques [13] in our Laboratory we are interested in developing general drawing procedures. For example, from a set of m objects, it is possible to construct a d -dimensional polyhedron, usually a hypercube, a hypertetrahedron or a complex. All or some of the d -dimensional polyhedron vertices represent one of the object set elements. Several algorithms can generate connections between vertices, which manifest the internal relationship connections between objects. The present figures are obtained using a m -dimensional polyhedron directly to visualize the Object relationships of a Molecular Point Cloud.

The projection of these figures into 2- or 3-dimensional spaces is a complementary task which enables one to catch the information in an human-like manner.

9. APPLICATION EXAMPLES

Following, some examples will be given, belonging to several chemical fields, to illustrate some of the possibilities associated with Molecular Similarity calculations.

The practical computations of the present paper have been performed under four kinds of approximate QSM, all of them obtained under a CNDO-like approach framework. Thus, the used QSM are based on the considerations outlined in Section 4, with a definition of atomic density contribution as written in equation (32). With this approach in mind one needs to actualize equations (28) or (29) in order to obtain the desired practical computational forms, which can be written schematically as:

- a) *Cioslowski-like QSM*. Equation (28) with the unit operator can be used, then

$$\begin{aligned} Z_{AB} &= \sum_{K \in A} \sum_{L \in B} \iint R^K(\mathbf{r}, \mathbf{s}) R^L(\mathbf{r}, \mathbf{s}) d\mathbf{r} d\mathbf{s} \\ &= \sum_{K \in A} \sum_{L \in B} Q^K Q^L S_{KL}^2, \end{aligned} \quad (43)$$

where

$$S_{KL} = \int S^K(\mathbf{r}) S^L(\mathbf{r}) d\mathbf{r}. \quad (44)$$

- b) *Coulomb-like QSM*. Equation (28) with the Coulomb operator $|\mathbf{r}-\mathbf{s}|^{-1}$ can be used, thus

$$\begin{aligned}
 Z_{AB} &= \sum_{K \in A} \sum_{L \in B} \iint R^K(\mathbf{r}) |\mathbf{r}-\mathbf{s}|^{-1} R^L(\mathbf{s}) d\mathbf{r} d\mathbf{s} \\
 &= \sum_{K \in A} \sum_{L \in B} Q^K Q^L (KK|LL) ,
 \end{aligned} \tag{45}$$

where

$$(KK|LL) = \iint |S^K(\mathbf{r})|^2 |\mathbf{r}-\mathbf{s}|^{-1} |S^L(\mathbf{s})|^2 d\mathbf{r} d\mathbf{s} . \tag{46}$$

- c) *Overlap-like QSM*. Equation (28) with the unit operator but using the first order density function provides a QSM which can be related to the relativistic spin-spin coupling integrals [2.c], takes the practical form:

$$\begin{aligned}
 Z_{AB} &= \sum_{K \in A} \sum_{L \in B} \int R^K(\mathbf{r}) R^L(\mathbf{r}) d\mathbf{r} \\
 &= \sum_{K \in A} \sum_{L \in B} Q^K Q^L (KK||LL) ,
 \end{aligned} \tag{47}$$

where

$$(KK||LL) = \int |S^K(\mathbf{r})|^2 |S^L(\mathbf{r})|^2 d\mathbf{r} . \tag{48}$$

- d) *Triple Density QSM*. These measures come from the ideas developed in reference [2.1], but constitute a practical derivation of equation (29) in the present paper. In this case the following integral has been used:

$$\begin{aligned}
 Z_{AB;C} &= \sum_{K \in A} \sum_{L \in B} \sum_{M \in C} \int R^K(\mathbf{r}) R^L(\mathbf{r}) R^M(\mathbf{r}) d\mathbf{r} \\
 &= \sum_{K \in A} \sum_{L \in B} \sum_{M \in C} Q^K Q^L Q^M (KL||M) ,
 \end{aligned} \tag{49}$$

where

$$(KL||M) = \int |S^K(\mathbf{r})|^2 |S^L(\mathbf{r})|^2 |S^M(\mathbf{r})|^2 d\mathbf{r} . \quad (50)$$

These last integrals can be practically computed as shown by the authors [2.1]. The resultant hypermatrix can be considered as a set of matrix representations of the Molecular Set considered as for each molecule C.

There is a symmetric matrix constructed for each molecule C by means of the rule $T_C = \{T_{AB;C}\}$.

A similarity Matrix with the same dimension as those constructed with the previous QSM recipes can be constructed in particular by using the contracted products:

$$Z_{CD} = (T_C | T_D) = \sum_A \sum_B T_{AB;C} T_{AB;D} \quad (51)$$

as generally described in Section 3.3.

All Similarity calculations have been computed with an actualized version of MOLSIMIL-90 [2.e] program. The Triple Density Similarity has been implemented in the TRIDENT program. Visualization is performed by means of ND-CLOUD program. MOLSIMIL-90, TRIDENT and ND-CLOUD codes had been developed in our Laboratory.

From now on, the corresponding figures described in the following examples are shown in Appendix A.

9.1. β -Diketones

Some members of the β -diketones family act as acid quelating agents (dissolved in some organic solvent), in the hydrometalurgic extraction process of Cu(II) ions from ammoniacal aqueous solutions of leaching.

Table 1 shows extracting constants K_{extr} [14] for Cu(II) of ten different β -diketones, in the process symbolized by:



where HA stands for a generic β -diketone.

Table 1: Extracting constants for Cu(II) of β -diketones [14]

Molecule	$\log K_{\text{extr}}$
Acetylacetone (1,4-pentadione)	-3.47
Hexafluoroacetylacetone	-0.26
Trifluoroacetylacetone	-1.12
Dipropionylmethane (Heptan,3,5-dione)	-4.34
Pivaloylacetyl methane (2,2-dimethyl-hexan-3,5-dione)	-4.63
2-Theonylfluoroacetone	-1.08
Diisobutylmethane (2,6-dimethyl-heptan-3,5-dione)	-4.42
Benzoylacetone	-4.17
Dipivaloylmethane (2,2,6,6-tetramethyl-heptan-3,5-dione)	-5.19
Dibenzoylmethane	-3.80

Figures 1 to 4 show the projections of the point-molecule representation for the β -diketone compounds described in Table 1, on

the plane of two principal components coming from the Cioslowski-like (Figure 1), Coulomb-like (Figure 2), Overlap-like (Figure 3) and Triple Density (Figure 4) similarity measure matrices. The structure shown in all the four figures represents a 10-dimensional polyhedron. In all cases the elements of the set have been divided in two classes, depending on their ability for the Cu(II) ions extraction. We can see that elements with $\log K_{\text{extr}}$ larger than the average value (the most active), represented by circles, appear in a different place in the figure than the smaller ones, represented by squares.

Figures 5 to 8 show projections of the Point-Molecule representation for Cu(II)- β -diketones quelate compounds, on the plane of two principal components for the Cioslowski-like, Coulomb-like, overlap-like and triple density similarity measure matrices, respectively. As before, the figures represent two-dimensional projections of 10-dimensional polyhedrons and the elements of the set are divided in two classes, depending on their respective extraction constant (K_{extr}). We can see that the most active compounds (represented by circles) can be shown split from the less active ones (represented by squares).

In Figure 9 we present a Kruskal tree computed from euclidean distances for Cu(II)- β -diketone quelate compounds, on the plane of 3-6 principal components for the overlap-like similarity measure matrix. The quelate with acetylacetone as a ligand with constant value (-3.47) appears as a principal knot, acting as a bridge between more and less stable quelates.

In all cases, various similarity measures have provided figures whose elements with different values of the studied property appear at a separate region on the projection plane.

9.2. Retinoid Compounds [15]

Retinoids are compounds structurally related to retinol, more commonly known as vitamin A. The basic structure of vitamin A is that of a diterpene with a hydrophobic ring on one end and an isoprenic side chain containing a polar hydroxyl group on the other. It has long been known that vitamin A has at least three distinct functions in supporting mammalian life: systemic, optic and reproductive ones.

Although retinoic acid and other retinoids are not very useful as vitamin A substitutes in the diet, they have been found to possess therapeutic value in treating dermatologic conditions. Nevertheless, an important side effect has been found in these compounds, they have teratogenic effects [16]. In view of this, It is important to determine which structural features in retinoids are related to teratogenicity in order to design a useful retinoid without this side effect. Some studies in this direction have been seen performed in another theoretical context [15].

In Tables 2 and 3 the activity, expressed as concentration in mg equiv. of vitamin A/kg body wt. of 50% occurrence of teratogenesis, of the two sets of retinoid acid derived compounds studied, is given.

Table 2: Set of retinoid compounds

Common name/code	TD ₅₀ [*] [16]	Activity
β -C ₁₄ Aldehyde	980	-
β -Ionone	480	-
Abscisic acid	250	-
Retinylidene acetilacetone	190	-
13-cis-N-(4-Hydroxyphenyl)retinamide	139	-
Ro11-4768	87	+
trans-Aryltriene analog of retinoic acid ethyl ester	84	+
trans-5-[2,6-Dimethyl-8-(2,6,6-trimethylcyclohexen-1-yl)-1,3,5,7-octatetren-1-yl]tetrazole	81	+
9,10-Dihydroretinoic acid	76	+
2-Norbornenyl ethyl ester	75	+
15-Fluororetinone	54	++
trans-Retinylidene methyl nitron	39	++
7,8-Dehydroretinoic acid	37	++
9-C-Retinal	23.2	++
trans Retinol	22.9	++
R08-7699	10.9	+++
trans Retinoic acid/ro1-5488	10.5	+++
Ro13-4306	4.7	++++
Ro13-6307	0.66	++++
Ro13-9272	0.032	++++
TTNPB/Ro13-7410	0.019	++++

^{*}Concentration in mg equiv. of vitamin A/kg body wt. of 50% occurrence of teratogenesis.

Figures 10 to 12 show projections of the Point-Molecule representation for retinoid compounds of Table 2, on the plane constructed with two principal components of the similarity matrix. The elements of the set are divided in three classes in Figure 10 and Figure 11 and in two classes in Figure 12, depending on their teratogenetic activity given as log TD₅₀ (Figure 10), as TD₅₀ (Figure 11) or as relative strength of teratogenicity (Figure 12). In all cases one

can observe that most active compounds appear at different zones in the projection plane.

Table 3: Set of active retinoid compounds.

Common name/code	TD ₅₀ * [16]	Activity
Ro11-4768	87	+
trans-Aryltriene analog of retinoic acid ethyl ester	84	+
trans-5-[2,6-Dimethyl-8-(2,6,6-trimethylcyclohexen-1-yl)-1,3,5,7-octatetren-1-yl]tetrazole	81	+
9,10-Dihydroretinoic acid	76	+
2-Norbornenyl ethyl ester	75	+
15-Fluororetinone	54	++
trans-Retinylidene methyl nitron	39	++
7,8-Dihydroretinoic acid	38	++
7,8-Dehydroretinoic acid	37	++
9-C-Retinal	23.2	++
trans Retinol	22.9	++
13-cis-Retinoic acid/Ro4-3780	22.3	++
R08-7699	10.9	+++
trans Retinoic acid/ro1-5488	10.5	+++
Etretinate/Ro10-9359	5.7	++++
Ro13-4306	4.7	++++
Ro13-2389	0.70	++++
Ro13-6307	0.66	++++
Ro13-8320	0.036	++++
Ro13-9272	0.032	++++
Ro13-6298	0.023	++++
TTNPB/Ro13-7410	0.019	++++

*Concentration in mg equiv. of vitamin A/kg body wt. of 50% occurrence of teratogenesis.

Figures 13 and 14 are projections of the Point-Molecule representation for retinoid compounds of Table 3, in the plane of two principal components of the similarity matrix. The elements of the set

are divided into three or two classes, depending on their teratogenetic activity given as TD_{50} or as relative strength of teratogenicity, respectively. We can see that elements whose activities are larger than the average value appear at different parts of the figure than the rest of the Molecular Point-Cloud.

Figure 15 shows a graph of the elements described in Table 3, computed using the algorithm of nearest neighbors [2.f] whose connections are computed from euclidean distances.

9.3. Steroid Hormones

We have studied some natural or synthesized steroids. All of them bear a nucleus of cyclopentanophenanthrene in common. Although all of them have very similar chemical structures, small modifications in their frame can produce dramatic diverse effects in their biological function. We have grouped them depending on their activity in human body.

9.3.1. Androgens and anabolic agents

The compounds listed in Table 4 have the effect of male sex hormones, i.e., to stimulate sexual maturation. But all androgens have in some degree the ability to stimulate muscle development. In Table 4 there are included some synthetic substances which have a lessened androgenic, but a heightened anabolic effect.

Table 4: Melting point [17] of androgens and anabolic agents.

Compound	Mp°C	Compound	Mp°C
Androsterone	185	Fluoxomesterone	270
Aldosterone	164	Hydroxidione sodium	198
Spironolactone	190	Methandrosterone	167
Methylandrostenediol	206	Methyl testosterone	163
Norethandrolone	133	Oxandrolone	231
Oxymetholone	182	Prometholone	127
Testosterone	154	Testosterone cypionate	101
Testosterone enanthate	36	Testosterone phenylacetate	130
Testosterone propionate	120		

In Figure 16 one can see a projection of the Point-Molecule representation for compounds described in Table 4, on the plane defined for 5 and 7 principal components of the similarity matrix. The elements of the Molecular Point Cloud are divided in two classes depending on their melting point. We can see how points corresponding to substances which have a melting point larger than the average value of the group at the bottom left hand part of the figure. There is an exception: the point corresponding to testosterone propionate (Mp=120°C), with a melting point smaller than the average value, which appears in the region of large values.

Figure 17 is a graph showing the relationships between compounds of Table 4, computed using the minimal order graph algorithm [2.e] computed in turn from euclidean distances.

Figure 18 shows a Kruskal tree [2.c], computed from euclidean distances, showing the relationships between elements described in

Table 4. We can see how elements with melting points less than the average value are terminal branches of the tree (36, 130, 101) (120, 127, 133).

9.3.2. Estrogenic substances

Estrogenic agents, listed in Table 5, hasten sexual maturation in the female. The widest use of estrogens is in the treatment of the menopause. Estrogens are also used in the control of cancer of the prostate in the male.

Figure 19 shows a projection of the point-molecule representation for compounds on Table 5, on a plane defined by the (7,8) principal components of the similarity matrix. The elements of the set are split into two classes depending on their melting point. We can see how points corresponding to different classes appear in separated regions of the figure.

Table 5: Melting point [17] of estrogenic agents.

Compound	Mp°C	Compound	Mp°C
Equilenin	258.5	Equilin	238
Estradiol	176	Estradiol benzoate	194
Estradiol cypionate	152	Estradiol dipropionate	107
Estriol	282	Estrone	260
Estrone benzoate	220	Ethynyl estradiol	144
Mestranol	151		

In Figure 20 we have the Kruskal tree [2.c], computed from euclidean distances, showing the relationships between the elements on Table 5. In most cases we can see that elements with melting point values smaller than the average value are terminal branches of the tree.

Figure 21 represents a graph showing the relationships between compounds of Table 5, computed using a nearest neighbor graph algorithm [2.f] obtained from euclidean distances. We can see that compounds with melting point values larger than the average value are strongly correlated (are heavily linked) with the rest (282, 258.5, 238), whereas those with melting point values smaller than the average one are slightly correlated.

9.4. Odor Molecules

Although a relative agreement exists about what odors are, they probably constitute, one of the most difficult stimulus to classify. It is well known the hypothesis [18] on the possibility to distinguish, at least, seven types of odors (they can be considered as the primary odors for the human olfactory sense), these are: ethereal, camphoraceous, musky, floral, minty, pungent, and putrid odors. It is feasible that the most accepted assumption about the molecular basis of odors is the "stereochemical theory of odor" formulated by Amoore [18]. Amoore theory has the appeal of simplicity and elegance. This theory attributes substances' odors to certain molecular characteristics, which are not always related to their functional groups (except in two of the seven types). Odor character only depends on aspects such as size, shape and degree of molecular rigidity. Within this criterion,

molecules can be classified, depending on their degree of rigidity, as: invariant, determinate and articulate molecules. It is assumed that, for each primary odor, there exists an olfactory receptor site, geometrically different and complementary in its shape to the shape of molecules that exhibits this odor. However, flexibility of some molecules allows them to fit in different kinds of receptors. On the other hand, it is known that pungent and putrid odors are the only ones which are not exclusively governed by these factors, instead they depend strongly on their electronic structure. Thus, it is known that pungent odor molecules have to be of electrophilic type, whereas putrid odor ones have to exhibit nucleophilic character, and this is the reason for observed close relationship between these odor types and definite electro- or nucleophilic chemical functional groups.

In this work, we have chosen three substances with a high component of odor intensity for each one of primary odor types, and three of low intensity. In Table 6, selected molecules along with their odor intensities, are listed. As we can see, molecules with comparable odor intensity values need not have structural similarity. This has to be taken into account in any theoretical study about this field.

Similarity measures had been computed using overlap, Coulomb operators and triple density algorithms. Next, some of the obtained results for each one of the primary odor subsets are presented.

Table 6: Intensity odors of odorants* [18]

COMPOUND	CAM	ETH	FLO	MIN	MUS	PUN	PUT
1,2-Dichloroethane		6.25					
Chloroform	0.60	6.69			0.24		0.00
<i>cis</i> -1,2-Dichloroethylene		5.45					
Methylene chloride	0.33	6.27	0.22		0.42		
Cyclopentadecanone	0.42	0.35			5.46		
2,4-Dinitro-3-methyl-6- <i>tert</i> butylanisole		0.30					
2,4,6-Trinitro-1,3-dime- thyl-5- <i>tert</i> butylbenzene		0.37			5.06		
1,8-Cineole	5.23						
d,l-Camphor	5.42						
Bornyl chloride	3.39						
Cyclooctanone	3.28				0.29		
15-Hydroxypentadecanoic acid lactone					5.66		
Versalide					5.30		
d,l- β -Phenylethylmethyl ethyl carbinol			6.32				
Cyclopentyl acetate			4.20				0.00
Geraniol			4.38				
Isopulegol			4.18				
Adamantane			0.39				
<i>cis</i> -1,2-Dichloroethylene			0.69	0.60			
d,l-Menthone				6.60			0.00
l-Carvone				4.60			
Pulegone				5.43			
Δ^{16} Androsten-3- α -ol				0.66			
<i>tert</i> butylCarbinol				0.54			2.60
Menthyl acetate				3.41			
Formic acid						7.1	
Cyclooctane						1.13	
2-Pentyl acetate						1.16	
Phenyl acetic acid						1.08	
Δ^9 Hexadecenolactone						0.02	
α -Irone						0.02	
Vanilline						0.00	
Dimethyl disulfide							7.0
Hexachloroethane							1.92
Hexamethylethane							1.69

*CAM: Camphoraceous, ETH: Ethereal, FLO: Floral, MIN: Minty, MUS: Musky, PUN: Pungent, PUT: Putrid.

Figures 22 to 24 show results of similarity measures for camphoraceous odor molecules. In all cases the elements of the set have been divided into two classes depending on their intensity of odor, given in Table 6. Figure 22 shows a graph computed using the nearest neighbor algorithm [2.f] obtained from the overlap similarity measure matrix. We can see that all compounds possessing high camphoraceous odor intensity are linked to the highest one. In figure 23 we can see a Kruskal tree [2.c] computed from euclidean distances using the elements of the Coulomb similarity measure matrix. Here, the point corresponding to cyclooctanone (3.28) acts as junction between elements with high and low odor intensity. Figure 24 is computed using a Triple Density Measure. It shows a projection on the plane of (1,3) principal components of camphoraceous molecules Triple Density Similarity measure matrix. We can observe that elements split into two regions of the plane depending on the odor intensity.

Figures 25 to 27 correspond to the set of ethereal odor molecules. As before, the elements of the set are divided into two classes. In figure 25 a Kruskal tree [2.c], computed from euclidean distances of the overlap similarity measure, is drawn. We can see that elements with low odor intensity are terminal branches of the tree. Figures 26 and 27 represent graphs computed using a nearest neighbor algorithm [2.e] from the overlap similarity measure matrix and a minimal order algorithm [2.e] obtained from the Coulomb similarity measure matrix, respectively. In both cases we can observe that elements in the same class have preference to link.

Similarity results for the set of floral odor molecules are represented in figures 28 and 29. The first one shows a graph

computed from the overlap similarity measure using the nearest neighbor algorithm [2.e]. One can see how elements with high odor intensity are linked in a different way than the others. Figure 29 shows a projection on the plane of 4 and 5 principal components, for a Coulomb-like similarity measure matrix. It is easily observed that Point-Molecules with high odor intensity appear in the same zone of the plane.

In figure 30 we have a projection on the plane of 2 and 9 principal components from an overlap similarity measure matrix, for musky odor molecules. The Point-Molecules of the Molecular Point Cloud are divided into three classes which appear at different zones in the plane.

Figure 31 shows a nearest neighbor graph [2.e] computed using the overlap similarity measures, whereas figure 32 is a Kruskal tree [2.c] computed from the Coulomb measures using euclidean distances, for musky odor molecules. In both cases we can see a different trend to link related to different kind of elements.

Figures 33 and 34 refer to pungent odor molecules. The first one represents a projection on the plane of (4,6) principal components for an overlap-like similarity measure matrix. Elements in the set are divided into three classes, depending on their intensity odor, they appear collected at different sections of the figure. In Figure 34 we show a minimal order graph [2.e] for a Coulomb-like similarity measure. Here, we can observe how the standard Point-Molecule, with a 7.10 odor intensity value, is linked to the next highest odor intensity

elements in the set. Besides this, all Point-Molecules in the same class are linked together.

Finally, Figures 35 to 37 show results for putrid odor molecules. In all cases the elements of the set are divided into three classes. Figure 35 is a projection on the plane of (7,8) principal components of a Coulomb-like similarity measure matrix, whereas Figures 36 and 37 are nearest neighbor graphs [2.e] computed from euclidean distances for an overlap-like and a triple density similarity measures, respectively. In both cases we can observe that elements with 1.69 (hexamethylethane) and 2.60 (tertbutylcarbinol) odor intensities are strongly correlated.

10. CONCLUSIONS

After the discussion presented here, one can think about a set of rules which are at the origin of the QS theoretical scheme. One can use them to resume all the previous findings. We will call them *Mendeleev Postulates*:

P0 : Every Quantum System in a given state can be described by a Density Integral Transform. It has been shown in Sections 2 and 3 how the Density Transforms are defined and what role they play in the QSM Theory. This zeroth postulate is obviously borrowed from Quantum Mechanics.

P1 : Quantum Systems can be compared by means of a QSM or by a QSI. In Section 3 a general definition of a

Quantum Similarity Measure has been outlined and several related Quantum Similarity Indices have been defined in Section 5.

P2 : There always exists a Projection of a Quantum Object Set into some n -dimensional space. This possibility has been explored in Section 8.

P3 : A Quantum Object Set Ordering exists. In Section 6 it has been explained how an Object Set Ordering exists. This has been applied for several Quantum System examples in Section 9. It has been demonstrated how this ordering can be associated with the set by means of several techniques.

Results show how the general procedures described here allow us to obtain information directly related to any concrete chemical problem.

ACKNOWLEDGMENTS

This work has been financed by the CICYT-CIRIT, Fine Chemicals Programme of the "Generalitat de Catalunya" through a grant: #QFN91-4206. One of us (E.B.) benefits from a grant from the "Departament d'Ensenyament de la Generalitat de Catalunya". Another one of us (L.V.) benefits from a grant from the "Fundación Andes" (Chile). The authors want to thank PARAL-LEL 42 for its help in providing the program ND-CLOUD which has allowed us to construct all the figures and visualize in this way all the Molecular Object Sets studied in the present work.

APPENDIX A: FIGURES

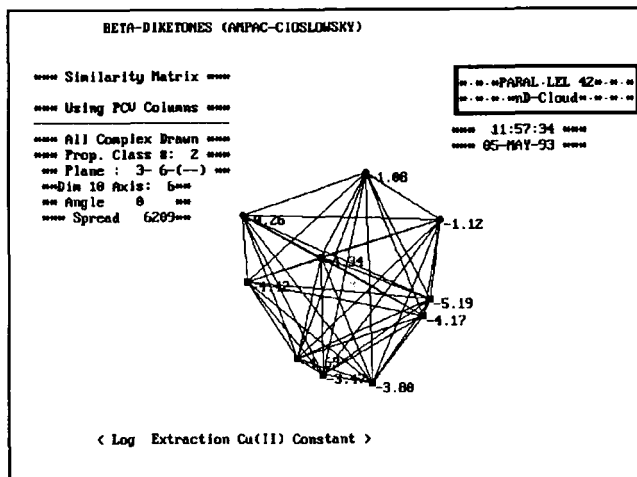


Figure 1.-Projection of the Point-Molecule representation, for beta-diketone compounds of Table 1, from a Cioslowski-like similarity measure.

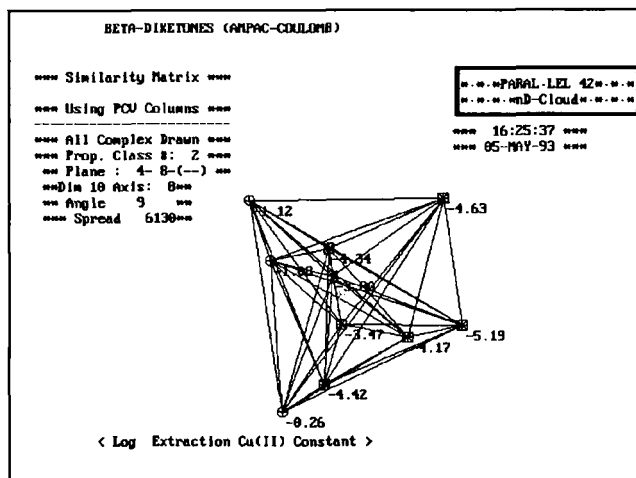


Figure 2.-Projection of the Point-Molecule representation, for beta-diketone compounds of Table 1, from a Coulomb-like similarity measure.

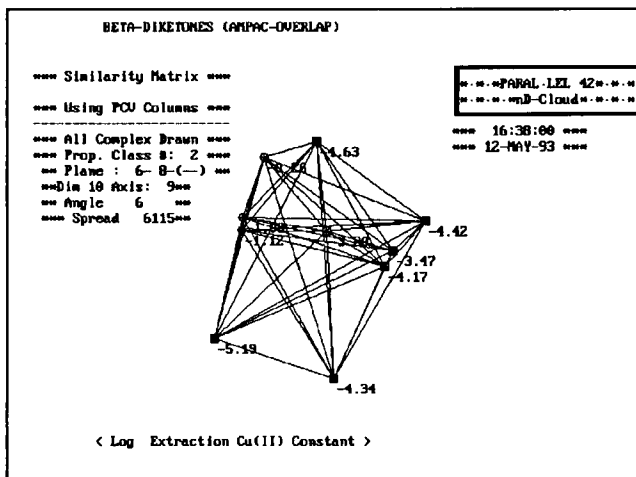


Figure 3.-Projection of the Point-Molecule representation, for beta-diketone compounds of Table 1, from an overlap-like similarity measure.

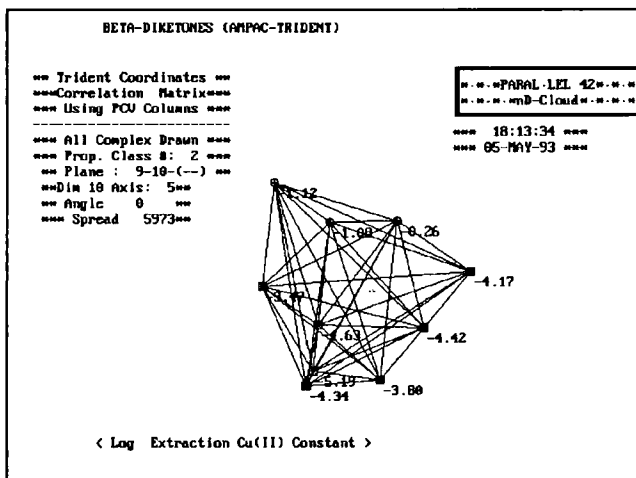


Figure 4.-Projection of the Point-Molecule representation, for beta-diketone compounds of Table 1, from a trident density similarity measure.

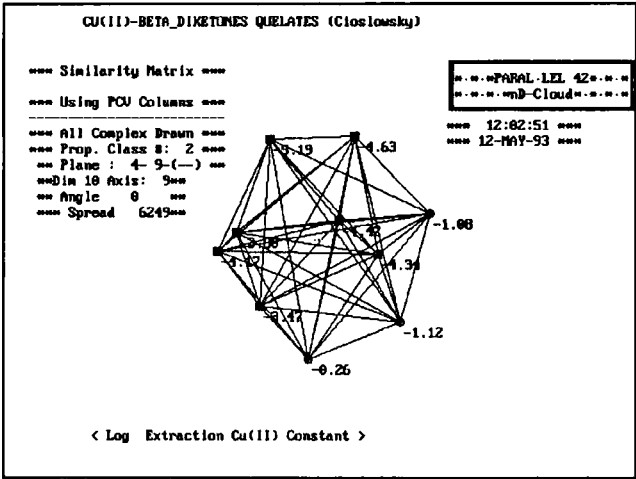


Figure 5.-Projection of the Point-Molecule representation, for Cu(II)-beta-diketones quelate compounds, from a Cioslowski-like similarity measure.

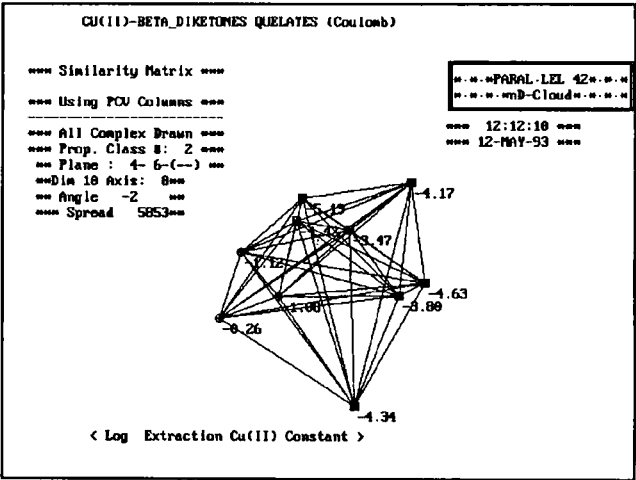


Figure 6.- Projection of the Point-Molecule representation for Cu(II)-beta-diketones quelate compounds, from a Coulomb-like similarity measure.

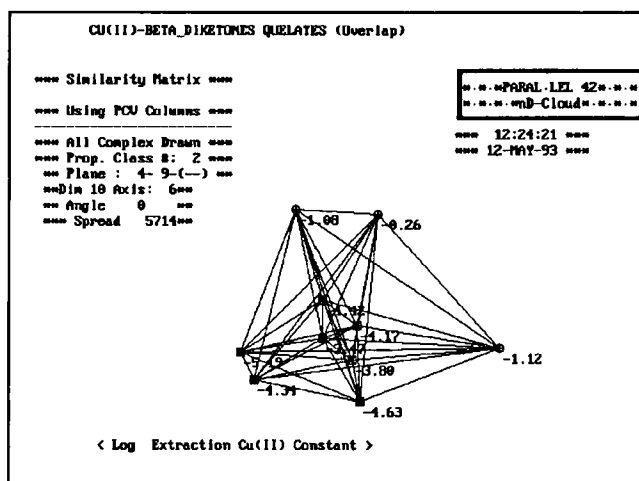


Figure 7.-Projection of the Point-Molecule representation for Cu(II)-beta-diketones quelate compounds, from an overlap-like similarity measure.

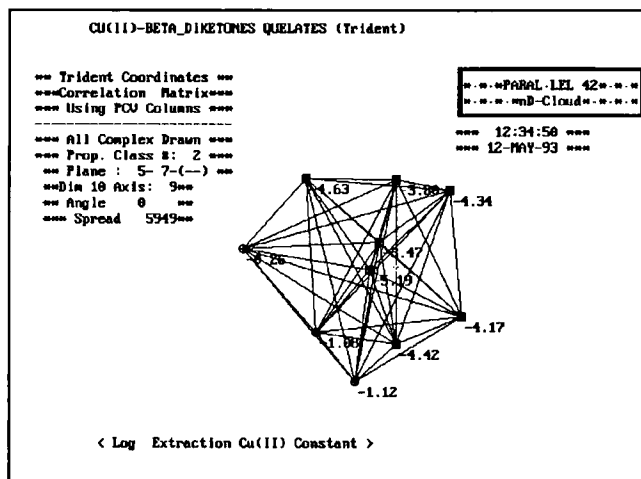


Figure 8: Projection of the Point-Molecule representation for Cu(II)-beta-diketones quelate compounds, from a triple density similarity measure.

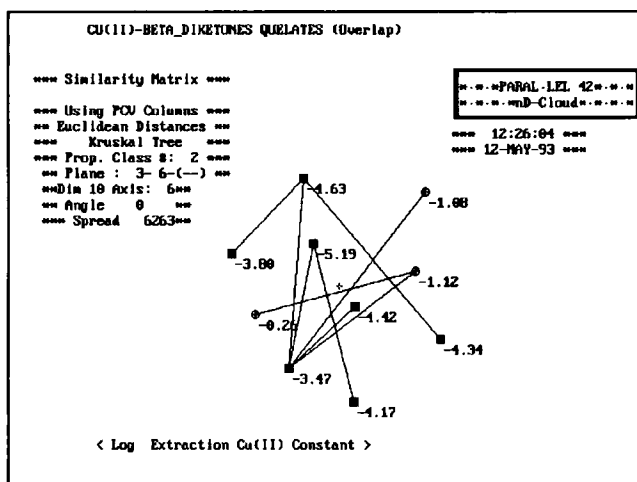


Figure 9.-Kruskal tree computed from euclidean distances for Cu(II)-beta-diketones quelate compounds, from overlap-like similarity measure.

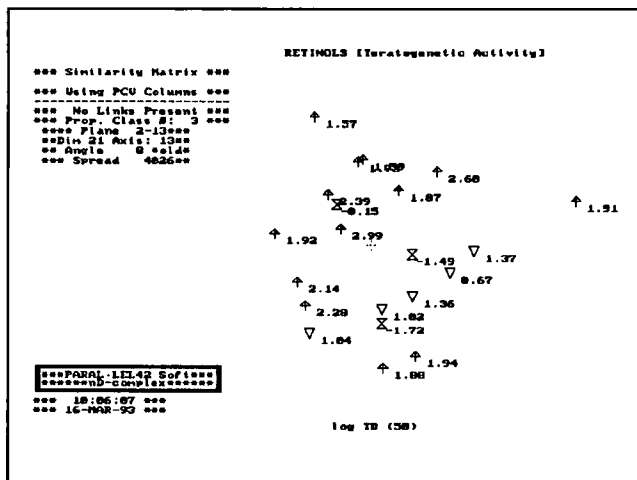


Figure 10.-Projection of the Point-Molecule representation for retinoid compounds of Table 2.

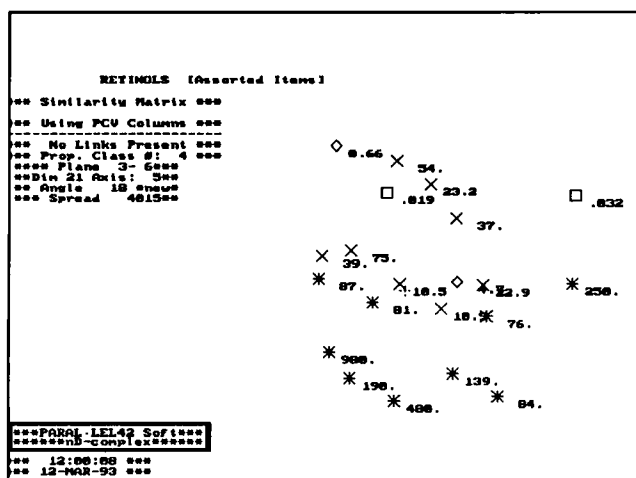


Figure 11.-Projection of the Point-Molecule representation for retinoid compounds of Table 2.

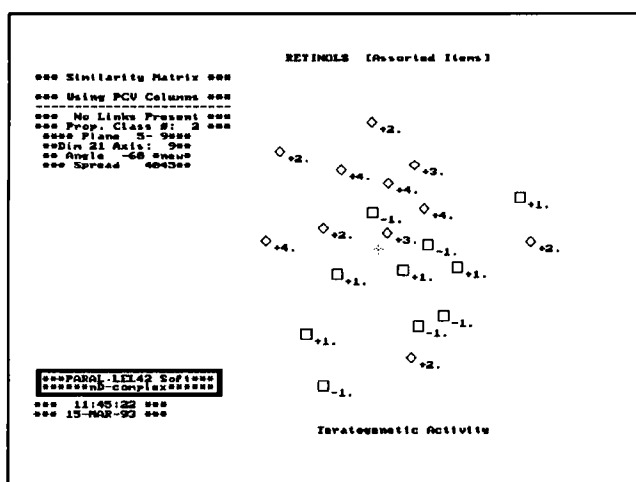


Figure 12.-Projection of the Point-Molecule representation for retinoid compounds of Table 2.

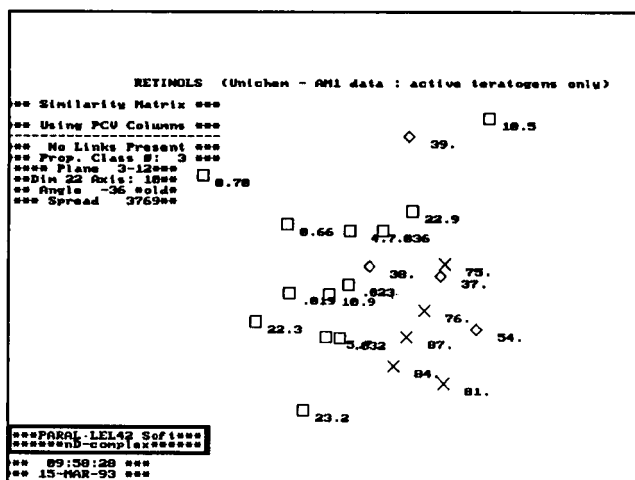


Figure 13.-Projection of the point-molecule representation for retinoid compounds of Table 3.

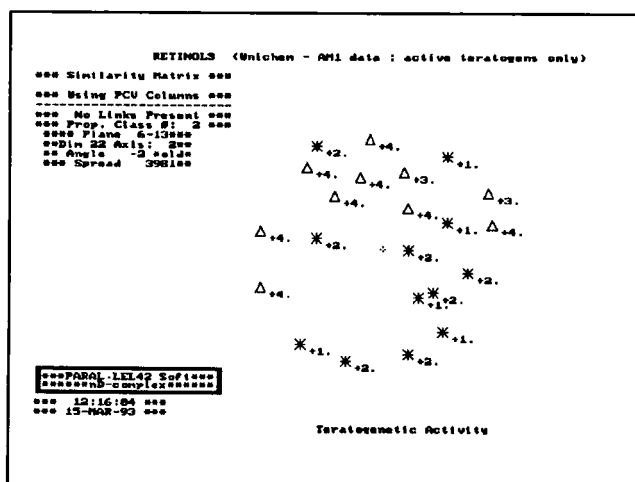


Figure 14.-Projection of the point-molecule representation for retinoid compounds of Table 3.

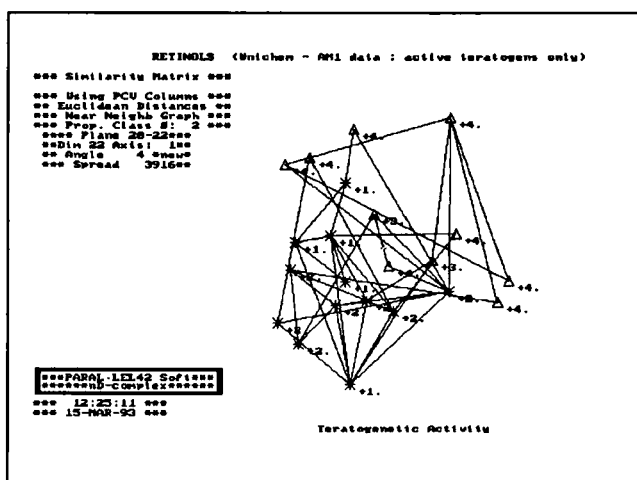


Figure 15.-Nearest neighbor graph, of elements on Table 3, from euclidean distances.

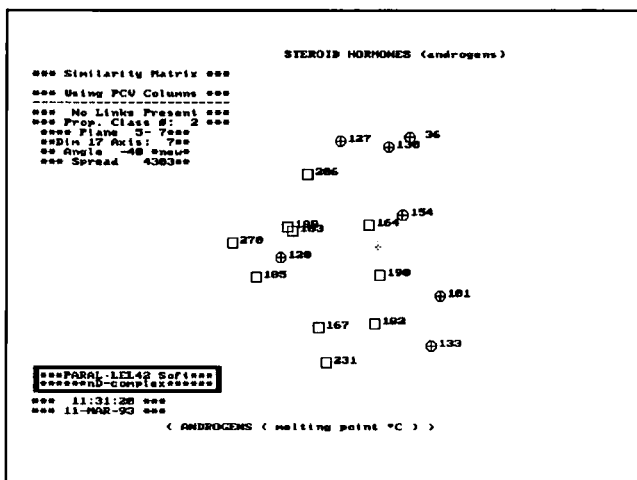


Figure 16.-Projection of the point-molecule representation for compounds on Table 4.

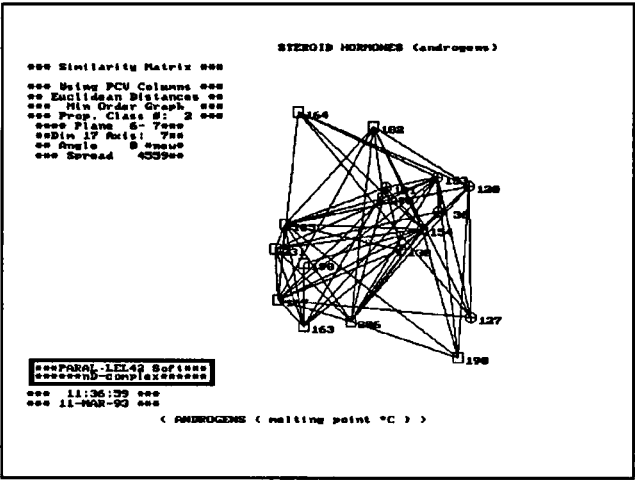


Figure 17.-Minimal order graph, of element of Table 4, from euclidean distances.

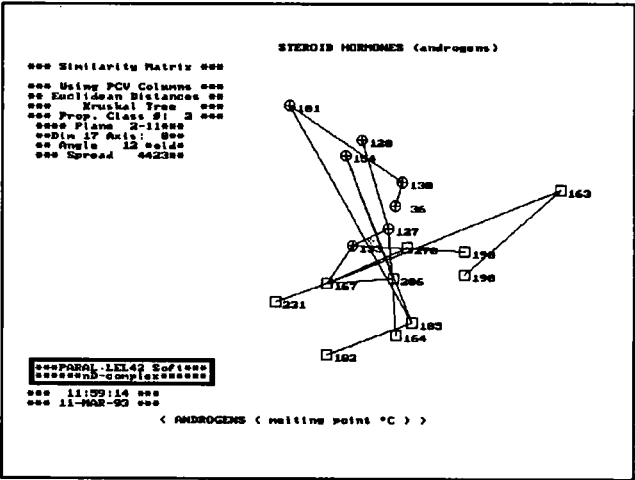


Figure 18.-Kruskal tree, computed from euclidean distances, of elements on Table 4.

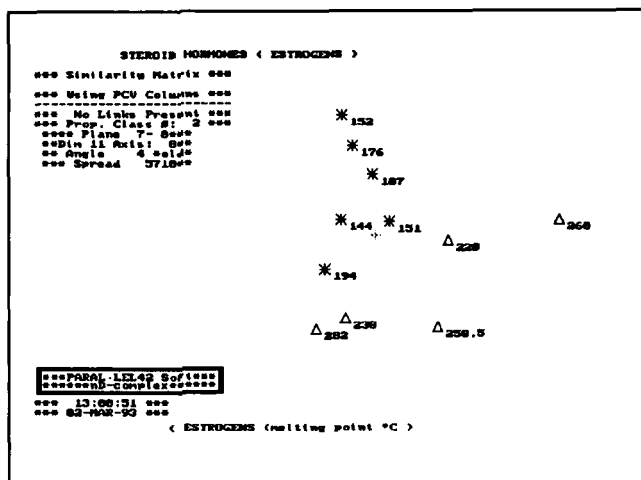


Figure 19.- Projection of the point-molecule representation for compounds on Table 5.

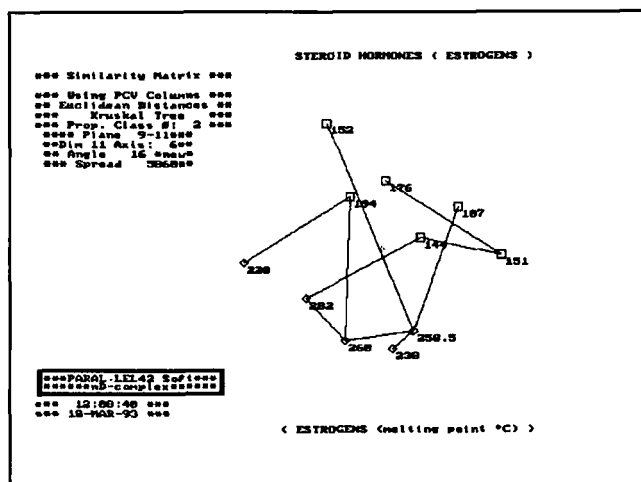


Figure 20.-Kruskal tree, computed from euclidean distances, of elements on Table 5.

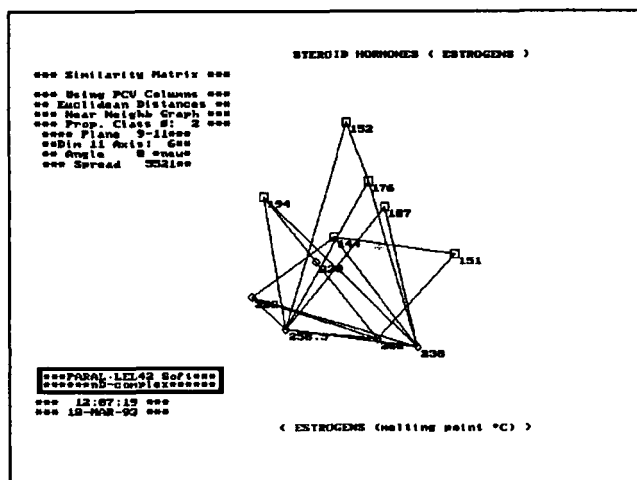


Figure 21.-Nearest neighbor graph, of compounds on Table 5, from euclidean distances.

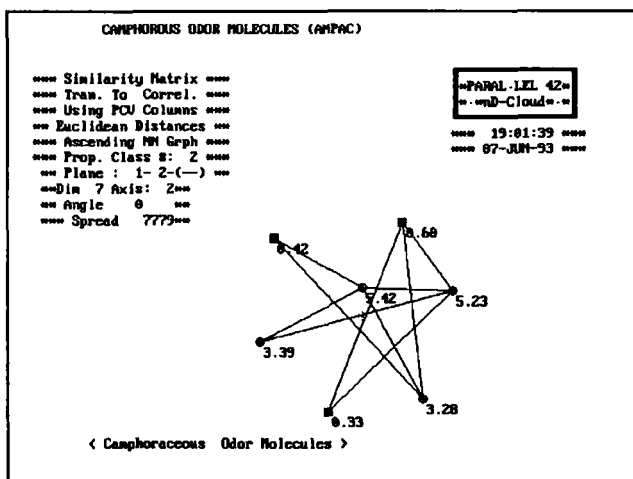


Figure 22.-Nearest neighbors graph from euclidean distances of the overlap-like similarity measure.

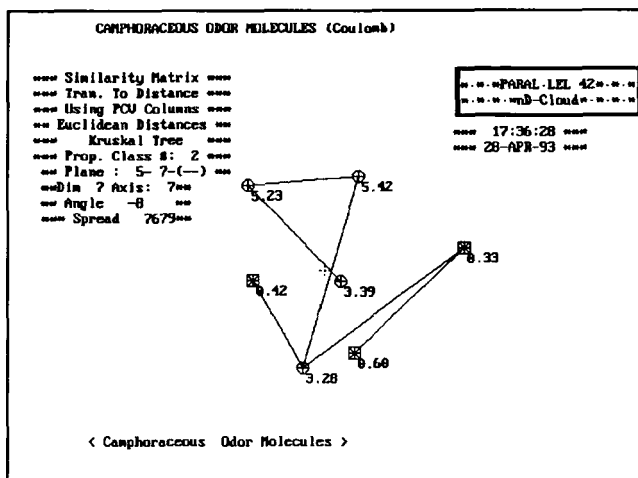


Figure 23.-Kruskal tree computed from euclidean distances of the Coulomb-like similarity measure.

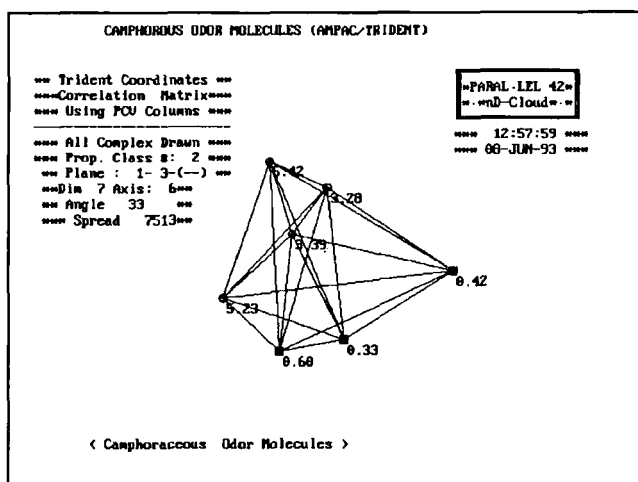


Figure 24.-Projection of the point-molecule representation of camphoraceous odor molecules using a triple density similarity measure.

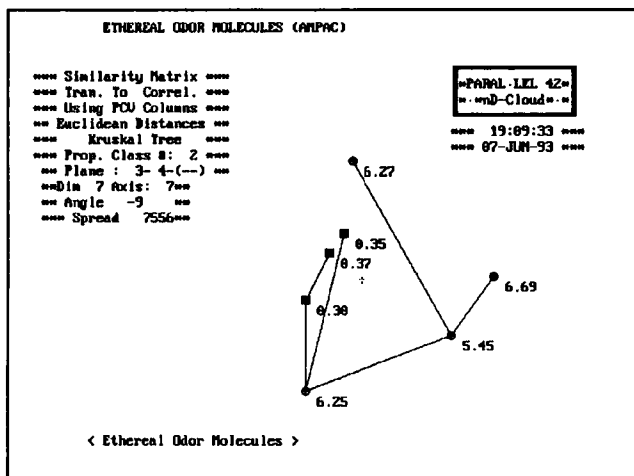


Figure 25.-Kruskal tree computed from euclidean distances of the overlap-like similarity measure.

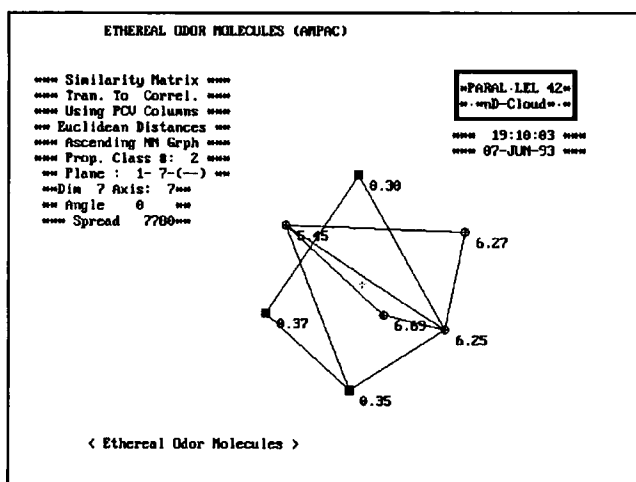


Figure 26.-Nearest neighbors graph from euclidean distances of the overlap-like similarity measure.

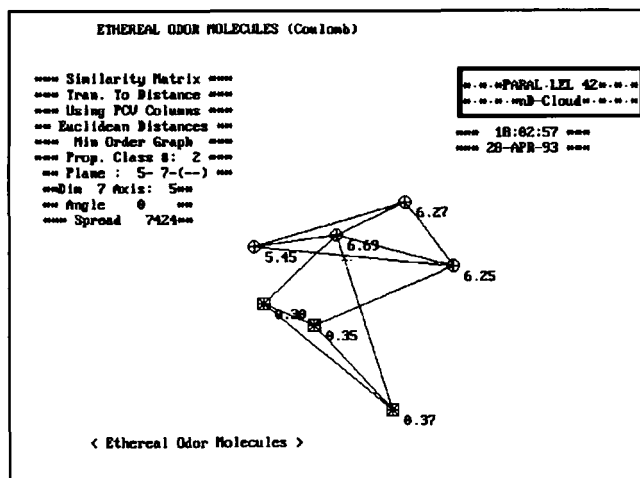


Figure 27.-Minimal order graph from euclidean distances of the Coulomb-like similarity measure.

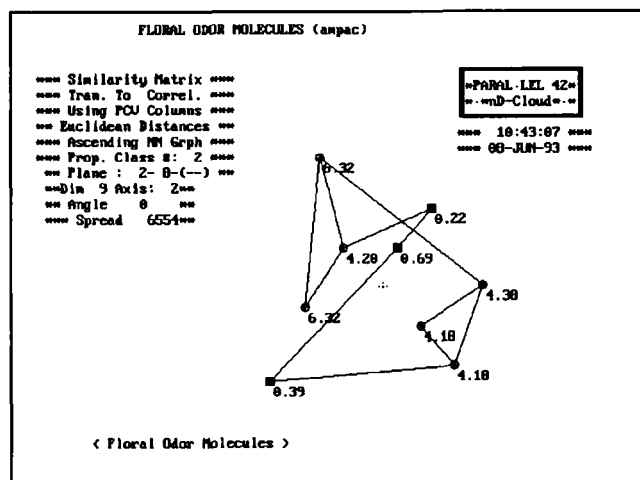


Figure 28.-Nearest neighbors graph from euclidean distances of the overlap-like similarity measure.

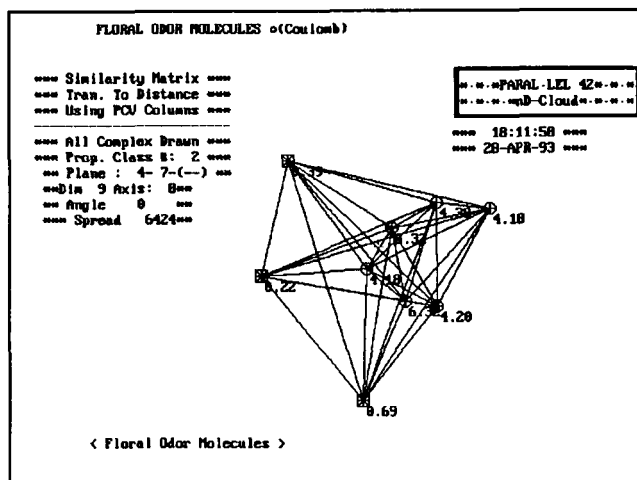


Figure 29.-Projection of the point-molecule representation of floral odor molecules using the Coulomb-like similarity measure.

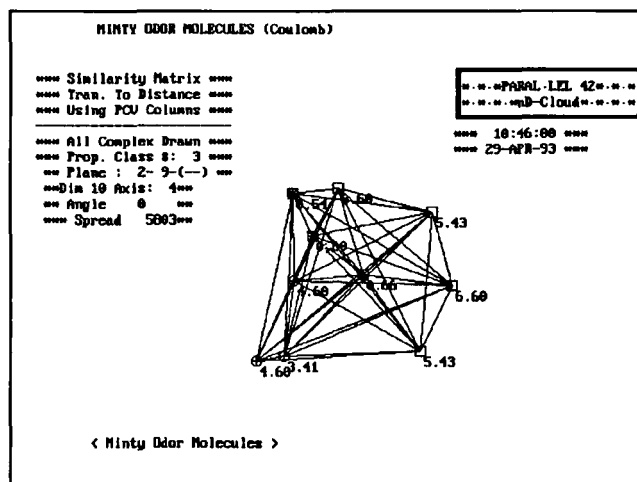


Figure 30.-Projection of the point-molecule representation of minty odor molecules using the Coulomb-like similarity measure.

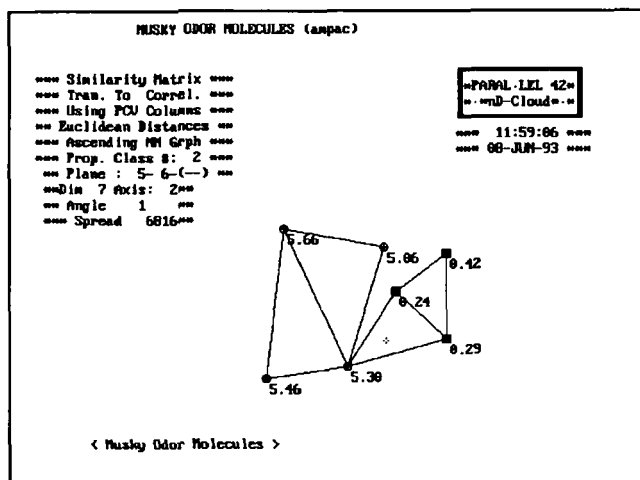


Figure 31.-Nearest neighbors graph from euclidean distances of the overlap-like similarity measure.

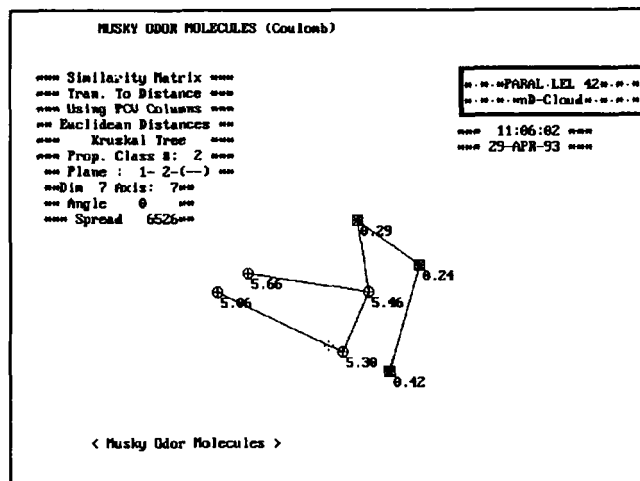


Figure 32.-Kruskal tree computed from euclidean distances of the Coulomb-like similarity measure.

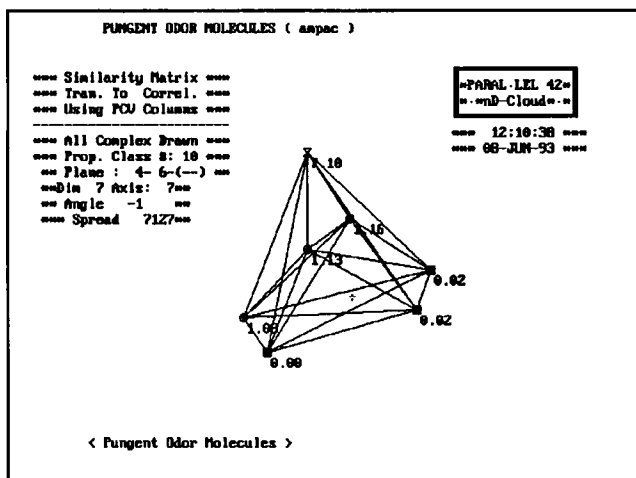


Figure 33.-Projection of the point-molecule representation of pungent odor molecules using the overlap-like similarity measure.

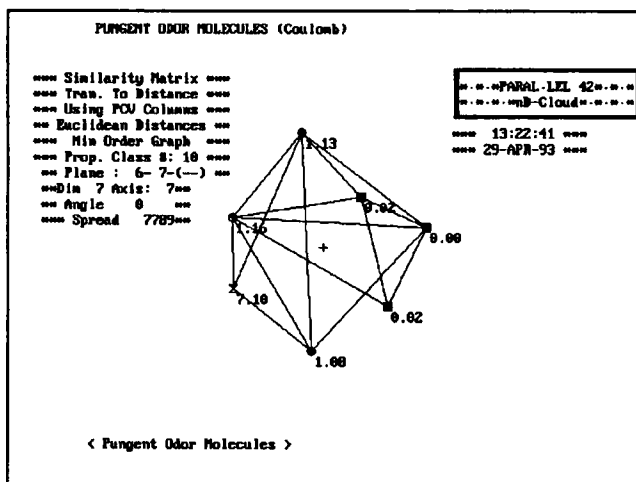


Figure 34.-Minimal order graph from euclidean distances of the Coulomb-like similarity measure.

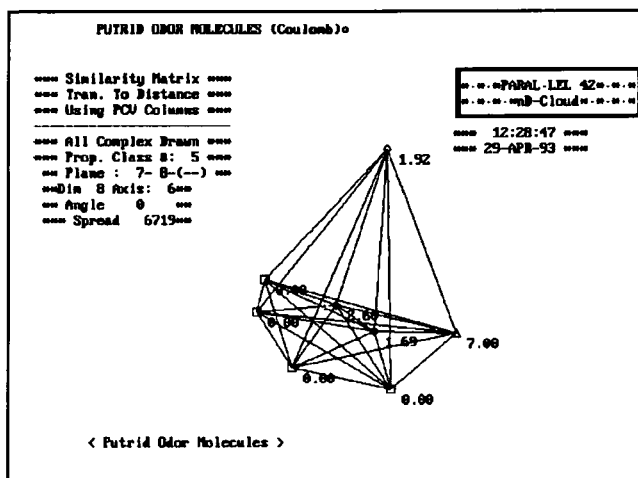


Figure 35.-Projection of the point-molecule representation of putrid odor molecules using the Coulomb-like similarity measure.

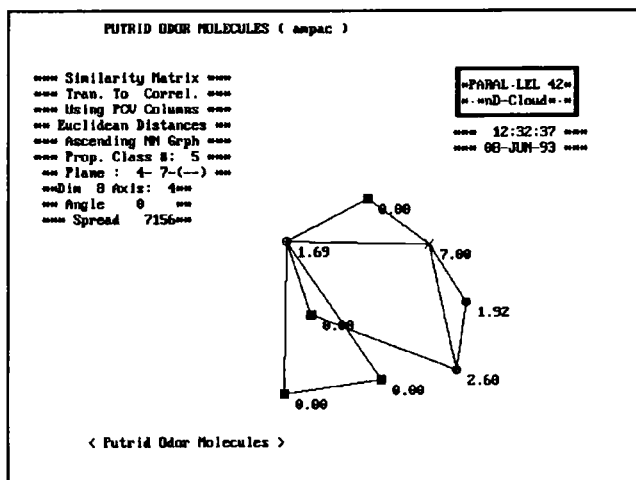


Figure 36.-Nearest neighbors graph from euclidean distances of the overlap-like similarity measure.

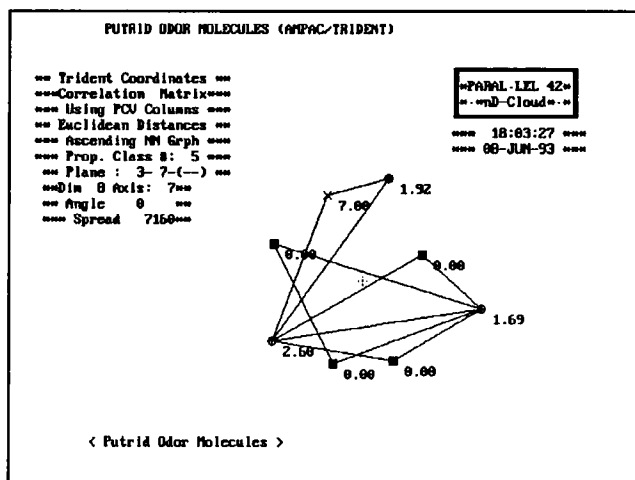


Figure 37.-Nearest neighbors graph from euclidean distances of a triple density similarity measure.

APPENDIX B: MOLECULAR QUANTUM SIMILARITY ASSOCIATED INTEGRALS.

B.1. Integrals over nS STO Functions.

The authors have some experience in computing QSM within the CNDO-Mulliken-like approach [2.1]. In this case the integrals of equation (26), for example, involve products between nS -STO functions. The following algorithm is proposed in order to compute these integrals:

- a) When in equation (26) there appears two or more n S STO functions sharing the same center, they can be multiplied and condensed into a unique S STO function of higher order.
- b) Two center product functions as the distributions of the type:

$$\Omega_{AB} = \chi_A \chi_B , \quad (53)$$

where χ_A and χ_B are S STO functions, can be expanded following the procedures outlined in references [2.1,8.a] or those described in Section 4. That is, a two-center charge distribution can be expressed as a linear combinations of one-center distributions. In this manner, the multicenter integral problem reduces to linear combinations of bicentric terms. As bicentric distributions are easily integrated, the entire problem can be considered solved.

B.2. General Integrals over GTO Functions.

The QSM integrals treated in the LCAO framework lead to the computation of linear combinations of multicenter integrals. We will describe here how to explicit this kind of integrals when dealing with spinorbital functions, whose space part is made of GTO functions or linear combinations of a set of them.

In reference [19] a General Gaussian Product Theorem is described, which allows us to obtain a convenient algorithm in order to contract the exponential part of each of the GTO functions involved

in the product. Also an explicit form for the resulting angular cartesian part products is provided. The final result gives the explicit form which adopts the product of GTO functions [19].

So, an integral arising from QSM computation, despite being able to involve several centers, can always be computed, in the worst case, as a linear combination of integrals involving two centers by electron.

REFERENCES

- [1] R.Carbó and E.Besalú. Proceedings of 1st Girona Seminar on Molecular Similarity. (1993)
- [2] a) R. Carbó, M. Arnau, L.Leyda, *Int.J.Quantum Chem.* **17** (1980) 1185.
- b) R. Carbó and C. Arnau, "Molecular Engineering: a general approach to QSAR", in: F.G. de las Heras and S. Vega (eds.) *Medicinal Chemistry Advances*, Pergamon Press, Oxford, 85 (1981).
- c) R. Carbó, Ll. Domingo, *Int.J.Quantum.Chem.* **23** (1987) 517.
- d) VII Escuela Latinoamericana de Química Teórica. Vol 1 (1989) 134. Centro Brasileiro de Pesquisas Fisicas. Coleção Galileo. Rio de Janeiro.
- e) R. Carbó, B. Calabuig, *Comp. Phys. Commun.* **55** (1989) 117.
- f) R. Carbó, B. Calabuig, "Molecular Similarity and Quantum Chemistry", chapter 6, p 147, in reference [6.k].
- g) R. Carbó, B. Calabuig, Proceedings del XIX Congresso Internazionale dei Chimici Teorici dei Paesi di Espressione

- Latina, Roma, Italy, September 10-14, 1990. *J.Mol.Struct. (Theochem)*, **254** (1992) 517.
- h) R.Carbó and B.Calabuig, *J.Chem.Inf.Comput.Sci.* **32** (1992) 600.
 - i) R.Carbó and B.Calabuig, "Quantum Similarity", in S.Fraga (ed.), *Structure, Interactions and Reactivity*, Elsevier Pub., Amsterdam (1992).
 - j) R.Carbó and B.Calabuig, *Int.J.Quantum Chem.* **42** (1992) 1681.
 - k) R.Carbó and B.Calabuig, *Int.J.Quantum Chem.* **42** (1992) 1695.
 - l) R.Carbó, B.Calabuig, E.Besalú and A.Martínez, *Molecular Engineering*, **2** (1992) 43.
- [3] A.Messiah, *Mechanique Quantique*, Vol I, E.Dunod. Paris (1959).
- [4] A.H.Zemanian, *Generalized Integral Transformations*. Dover Pub. Inc. New York (1987).
- [5] a) P.O.Löwdin, *Phys.Rev.* **97** (1955) 1474.
b) P.O.Löwdin, *Phys.Rev.* **97** (1955) 1490.
c) P.O.Löwdin, *Phys.Rev.* **97** (1955) 1509.
d) R.McWeeny, *Proc.Roy.Soc. A*, **232** (1955) 114.
e) R.McWeeny, *Proc.Roy.Soc. A*, **235** (1956) 496.
f) R.McWeeny, *Proc.Roy.Soc. A*, **253** (1959) 242.
- [6] a) D.L.Cooper and N.L.Allan, *J.Chem.Soc., Faraday Trans.* **83** (1987) 449.
b) D.L.Cooper and N.L.Allan, *J.Computer-Aided Mol.Design.* **3** (1989) 253.
c) D.L.Cooper and N.L.Allan, *J.Am.Chem.Soc.* **114** (1992) 4773.
d) J.Cioslowski and E.D.Fleischmann, *J.Am.Chem.Soc.*, **113** (1991) 64.

- e) J.Cioslowski and M.Challacombe, *Int.J.Quant.Chem.*, **25** (1991) 81.
- f) J.V.Ortiz and J.Cioslowski, *Chem.Phys.Lett.*, **185** (1991) 270.
- g) J.Cioslowski and P.R.Surján, *Jour.Mol.Struct.(Theochem)*, **255** (1992) 9.
- h) R.Ponec and M.Strnad; *Collect.Czech.Chem.Comm.* **55** (1990) 2583.
- i) R.Ponec and M.Strnad; *J.Phys.Org.Chem.* **4** (1991) 701.
- j) R.Ponec and M.Strnad; *Int.J.Quantum Chem.* **42** (1992) 501.
- k) M.A.Johnson and G.Maggiora (eds.), *Concepts and Applications of Molecular Similarity*, John Wiley & Sons Inc, New York (1990).
- l) E.E.Hodgkin and W.G.Richards, *Int.J.Quant.Chem.* **14** (1987) 105.
- m) A.C.Good, E.E.Hodgkin and W.G.Richards, *J.Chem.Inf.Comput.Sci.* **32** (1992) 188.
- n) M.Martín, F.Sanz, M.Campillo, L.Pardo, J.Pérez and J.Turmo, *Int.J.Quant.Chem.* **23** (1983) 1627.
- o) M.Martín, F.Sanz, M.Campillo, L.Pardo, J.Pérez, J.Turmo and J.M.Aulló, *Int.J.Quant.Chem.* **23** (1983) 1643.
- p) F.Sanz, M.Martín, J.Pérez, J.Turmo, A.Mitjana and V.Moreno, in: J.C.Dearden (ed.), *Quantitative Approaches to Drug Design*, Elsevier, Amsterdam (1983).
- q) F.Sanz, M.Martín, F.Lapeña and F.Manaut, *Quant.Struct.-Act.Relat.* **5** (1986) 54.
- r) F.Sanz, F.Manaut, J.José, J.Segura, M.Carbó and R. de la Torre, *Jour.Mol.Struct. (Theochem)* **170** (1988) 171.
- s) F.J.Luque, F.Sanz, F.Illas, R.Pouplana and Y.G.Smeyers, *Eur.J.Med.Chem* **23** (1988) 7.

- [7] K.Ruedenberg, *J.Chem.Phys.* **19** (1951) 1433
- [8] a) R.Carbó and E.Besalú, *Adv.Quant.Chem.* **24** (1992) 115.
b) R.Carbó and E.Besalú, *Can.J.Chem.* **70(2)** (1992) 353.
c) R.Carbó and E.Besalú, *J.Math.Chem.* **13** (1993) 331.
- [9] R.S.Mulliken, *J.Chem.Phys.* **23** (1955) 1833,1841,2338,2343.
- [10] S.Huzinaga, Y.Sakai, E.Miyoshi and E.Narita, *J.Chem.Phys.* **93** (1990) 3319.
- [11] J.T.Tou and R.C.González, *Pattern Recognition Principles*. Addison-Wesley Reading, MA (1974).
- [12] J.A.Hardigan, *Clustering Algorithms*. Wiley, New York (1975).
- [13] a) P.H.A.Sneath and R.R.Sokal, *Numerical Taxonomy*, W.H.Freeman, San Francisco (1973).
b) N.Jardine and R.Sibson, *Mathematical Taxonomy*, Wiley, New York (1977).
- [14] J. Stary ; H. Freiser, " Chelating Extractants", (Part IV of "Equilibrium Constants of Liquid-Liquid Distribution Reactions"), IUPAC Chemical Data Series N° 18, Pergamon Press, Oxford (1975).
- [15] G. Klopman, M.L. Dimayuga, *Journal of Computer-Aided Molecular Design*, **4** 117-130 (1990).
- [16] C.C. Willhite, In Dawson, M.I. and Okamura, W.H. (eds.)*Chemistry and Biology of Synthetic Retinoids*, CRC Press, Boca Raton, FL (1989).
- [17] D.L. Lide (ed.) *Handbook of Chemistry and Physics*, CRC Press, Boca Raton, FL (1991).
- [18] J.E. Amoore, "Molecular basis of odor", *American Lecture Series*, Charles C. Thomas, Publisher, Illinois, U.S.A. (1970).
- [19] R.Carbó and E.Besalú, *Int.J.Quant.Chem.* (submitted).

Index

A

Ab initio calculations, *trans*-butadiene, 210–217
 Algorithmic costs, as function of degree of contraction, 161
 Alkali metal ions, Group Ia, far infrared spectra, 59
 Anabolic agents, quantum similarity measure applications, 280–282, 297–298
 Androgens, quantum similarity measure applications, 280–282, 297–298
 Anharmonic potentials, 79–82
 Annihilated unrestricted Hartree–Fock approximation
 excited polyenes, 243
 singly charged polyenes, 237
 Annihilated unrestricted Hartree–Fock method, 209–210, 221–221, 227
 Axis-switch method, 149

B

Band states, nearly free ion, 70–71
 Basis functions, 144–147; *see also* Gaussian basis functions; Primitive basis functions
 constructing transfer state function, 117
 delta functions, 147
 Slater functions, 144
 Boys algorithm, 159
 Brackets, 163, 165
 Bras, 163–164
 Brillouin zone, first, 84–86, 88
trans-Butadiene, *ab initio* calculations, 210–217

C

Calcium
 role in membrane structure, 107–108
 as tunneling probe of membrane, 119

Carlson–Rushbrooke formula, 123–124
 Channel model, 52–58
 axial and transverse properties of potential, 55–58
 model potential energy functions, 53–54
 wall source location, 52–53
 Channels, 120
 bands, axial, 77–78
 conductors, band states, 65
 configuration, mechanical stability, 61–64
 as sequence of intersecting harmonic wells, 73
 Charge-dipole interaction, 128–130
 Conduction channel, mouth, 109
 Cotton–Mouton effect, 8, 35–38
 Coulomb interaction
 Gaussian screened, 121–122
 matrix elements, 125–128
 Coulomb operator, *ss*-matrix element, 98–99
 Coulomb potential
 Gaussian screened, matrix elements, 90–94
 Morse and Gaussian screened, 60
 Coupled Cluster Doubles theory, linear approximation, 210
 CPU-time, Gaussian integral algorithms, 152–153

D

dc-Kerr effect, 9
 dc-Pockels effect, 9
 Delocalization, 50
 Density Integral Transform, 256
 β -Diketones, quantum similarity measure application, 275–277, 290–294
 Dipole moment operator, matrix elements, 128–130
 Dirac delta function, 147
 three-dimensional, 154
 Dispersion coefficients, 32–33
 Dispersion formulas, 31–35

E

- Electric fields, 10–35
 - dispersion formulas, 31–35
 - electronic hyperpolarizabilities, 10–24
 - vibrational hyperpolarizabilities, 24–31
- Energy, change in, 5
- Energy function, potential
 - channel model, 53–54
 - three-body, 54
- Estrogenic substances, quantum similarity
 - measure applications, 282–283, 299–300
- Expectation functions, 255–257

F

- Flop-cost
 - Gaussian integral algorithms, 150–151
 - parameters for generating integral classes, 160–161
- Force constants, for vibration of ion along
 - channel, Cartesian forms, 122–124
- Fundamental Anti-Coulomb Integral, 158
- Fundamental Electron-Repulsion Integral, 157
- Fundamental Kinetic-Energy Integral, 156
- Fundamental Nuclear-Attraction Integral, 158
- Fundamental Overlap Integral, 155–156

G

- Gaussian basis functions, 145
 - contracted, 145–146
 - lobe functions, 146
 - three-dimensional, 82–90
- Gaussian charge distribution, Fourier
 - transform, 121
- Gaussian integral algorithms, 148–171; *see also* PRISM algorithm
 - axis-switch method, 149
 - Boys algorithm, 159
 - bras, kets, and brackets, 163–165
 - contraction problem, 160–161
 - CPU-time, 152–153
 - Flop-cost, 150–151
 - fundamental integrals, 153–158
 - anti-Coulomb integral, 158
 - electron-repulsion integral, 157
 - kinetic-energy integral, 156
 - nuclear-attraction integral, 158
 - overlap integral, 155–156
 - Head–Gordon–Pople algorithm, 169–171
 - McMurchie–Davidson algorithm, 166–167
 - Mop-cost, 151–152
 - Obara–Saika–Schlegel algorithm, 165
 - performance measures, 150–153
 - Pople–Hehre algorithm, 162
 - Second Generation methods, 149
 - two-electron methods, evolution, 149

H

- Hall interaction, three-body, 56
- Hall's potential, 55
- Hamiltonian
 - Pariser–Parr–Pople, 208–210, 219–220
 - Su–Schrieffer–Heeger, 208–210
- Hamiltonian operator
 - matrix elements, 74–75
 - nearly free ions, 66, 68
- Harmonic oscillator analysis, one-
 - dimensional, 72–79
- Harmonic potential, truncated, matrix
 - elements, 130–132
- Hartree–Fock approximation, *ab initio*, 211–212
- Hartree–Fock equations, 143
- Hartree–Fock method, extended, 220
- Hartree–Fock theory, time-dependent
 - coupled-perturbed, 19–22
- Head–Gordon–Pople algorithm, variations, 170–171
- Hodgkin–Richards index, 267
- Hückel matrix, 217–218
- Hypermagnetizability anisotropy, 8, 35–36
- Hyperpolarizabilities
 - electronic, 10–24
 - ab initio* calculations on medium-size systems, 18–24
 - frequency-dependent hyperpolarizabilities, 19–20
 - quasi-energy, 22–23
 - time-dependent coupled-perturbed Hartree–Fock theory, 19–22

- calibrational-quality calculations on
 - small systems, 11–18
 - electron-correlated functions of
 - James–Coolidge type, 15–16
 - helium basis functions, 13–14
 - perturbation-theoretic sum-over-states
 - formulas, 12
 - ratios, 17
 - wavefunctions, 13
 - vibrational, 24–31
 - anisotropy, 36
 - electric-property-function transition-
 - matrix elements, 27–28
 - non-zero terms in perturbation
 - expansion, 28–29
 - vibrational wavefunction, 24–25
- I**
- Interband transition, frequency dispersion
 - curve, 78
 - Ion-gating, into biological conduction
 - channels, 107–120
 - mechanism, 107–111
 - membrane model, 107–111
 - tunneling probes, 111–120
 - Ionic conduction channels, in lattice structure,
 - 108–109
 - Ionic migration, barrier height, 50–51
 - Ionic motion, 48–136, *see also* Ion-gating;
 - Mechanical harmonic oscillator analyses;
 - Tight binding analysis
 - channel model, 52–58
 - delocalization, 50
 - free ions, 66
 - multiple ion problem, 104–106
 - nearly free ions, 66–68
 - in non-biological systems, 49
 - small ions, 49
 - small polaron-like limit, 101–104
 - vibration along channel, force constants,
 - 122–124
 - vibration frequency, as function of linear
 - density, 105–106
 - Ions
 - free, 66
 - nearly free
 - band states, 70–71
 - general properties, 66–68
 - mobility, 71–72
 - model potential, direct use, 68–71
 - parameters, 70
 - Ion tunneling microscopy, 113–119
- K**
- Kets, 163, 165
 - Kleinman relations, 9–10
 - Kohn–Sham density functional theory, 143
 - Kruskal tree, 276, 281, 283, 286–287
- L**
- Lattice dilation waves, 108
 - Lobe functions, Gaussian, 146
 - Löwdin matrix, 86
 - Löwdin orthogonalization, 85
 - Löwdin orthonormalization, 118
- M**
- Magnetic fields, 35–40
 - Cotton–Mouton effect, 35–38
 - Magnetic flux density, perturbation-theoretic
 - expression, 8–9
 - McMurchie–Davidson algorithm, 166–167
 - Mechanical harmonic oscillator analyses,
 - 59–65
 - ion vibration at channel minima, 64–65
 - mechanical stability in channel
 - configurations, 61–64
 - model systems, 59
 - parameter choice, 59–61
 - Mechanical stability, in channel
 - configurations, 61–64
 - Membranes, biological
 - calcium role, 107–108
 - model, 107–111
 - surfaces and channels, tunneling probes,
 - 111–120
 - Mendelev conjecture, 268–269
 - Mobility, nearly free ion limit, 71–72
 - Molecular point cloud, 270–271, 287
 - Mop-cost, Gaussian integral algorithms,
 - 151–152
 - Morse potential, 55–57, 80, 118, 122–123
 - matrix elements, 125–128
 - Mulliken gross atomic population, 265–266

N

Non-linear-optics, 3–41, *see also* Electric fields
 applications, 3
 definitions, 5–10
 magnetic fields, 35–40
 materials, 3–4
 units, 10

O

Obara–Saika recursion relations, 91, 133
 Obara–Saika–Schlegel algorithm, 165
 Octatetraene, 215–216
 Odor molecules, quantum similarity measure
 applications, 283–288, 300–308

P

Parametrization, *trans*-polyacetylene, soliton
 dynamics, 223–224
 Pariser–Parr–Pople model, 217–223
 coupled eigenvalue equations, 219
 parametrizations, 224
 Perturbation-theoretic sum-over-states
 formulas, 12
 Phonon coupling, mobile ion-channel,
 97–101
 π - π correlations, 212–214
 π - π excitation energies, singlet and triplet,
 215–216
 Point-molecules, 270
 Polarizability
 dynamic dipole-dipole-quadrupole, 38
 Sellmeier dispersion formula, 31
 Polarization, 6
trans-Polyacetylene, soliton dynamics,
 207–247
ab initio calculations, 210–217
 bond alternation potential, 225
 charge density at odd numbered sites,
 235–236
 charged polymethine chains, dynamics,
 235–237
 different orbitals for different spins ansatz,
 220–221
 doubly charged polyenes, dynamics,
 241–242

energy level diagram, neutral and charged
 solitons, 229–230
 excited polyenes, dynamics, 243–246
 molecular orbital energies, 225
 neutral polymethine chains, dynamics,
 231–235
 normalized staggered coordinate, 238–240,
 242, 244–245
 parametrization, 223–224
 Pariser–Parr–Pople model, 217–223
 singly charged polyenes, dynamics,
 237–241
 soliton properties as functions of chain
 length, 227
 spin density, at odd numbered sites,
 232–234
 static properties, 225–231
 Su–Schrieffer–Heeger Hamiltonian,
 208–210
 total energy per CH unit, 225–226
 Polyenes
 doubly charged, dynamics, 241–242
 excited, dynamics, 243–246
 singly charged, dynamics, 237–241
 Polymethines
 charged chains, dynamics, 235–237
 neutral chains, dynamics, 231–235
 spin densities, 228–229
 Pople–Hehre algorithm, 162
 Potential
 exponential three-body, matrix elements,
 95–97
 model, direct use, nearly free ions, 68–71
 as sequence of intersecting harmonic wells,
 73
 Potential energy, profiles, 109–110
 Primitive basis function, 75, 83, 130–132
 Gaussian, 145
 matrix elements, 133–136
 Principal scaling, 179
 PRISM algorithm, 172–196
 contraction steps, 183
 HGP PRISM
 front face, 174
 full, 176
 one-electron transformations on, 193
 two-electron transformations on, 190
 loop structure, 194
 MD PRISM
 front face, 173
 full, 175

one-electron transformations on, 189
 two-electron transformations on,
 184–188
 [O]^(m) integral generation, 181–182
 performance in Gaussian 92, 195–196
 shell-pair data, 178–179
 shell-quartet selection, 180
 Proton transfer probe, 114–115

Q

Quantum similarity measures
 applications, 272–288
 androgens and anabolic agents,
 280–282, 297–298
 β -diketones, 275–277, 290–294
 estrogenic substances, 282–283,
 299–300
 odor molecules, 283–288, 300–308
 retinoid compounds, 277–280,
 294–297
 approximate, 262–266
 Cioslowski-like, 272
 computation, 269–270
 Coulomb-like, 260, 272–273
 definition, 257–258
 density integral transform, 288
 density matrix, 263
 electrostatic potential, 259
 expectation functions, 255–257
 integrals over GTO functions, 309–310
 integrals over STO functions, 308–309
 manipulation, 266–268
 ordering molecular sets, 268–269
 overlap-like, 259, 273
 particular cases, 258–261
 similarity indices, 266–268
 similarity matrices, 261–262
 simplified forms, 258–261
 triple density, 273–274
 visualization techniques, 270–271

R

Refractive index, electric-field-induced
 anisotropy, 8
 Retinoid compounds, quantum similarity
 measure applications, 277–280,
 294–297

S

Scanning tunneling microscopy, 112–113,
 120
 Schrödinger vibrational equation, 26
 Self-consistent-field procedure, 143
 Sellmeier dispersion formula, 31
 Shell-pairs, 178–179
 Shell-quartets, selection, PRISM algorithm,
 180
 σ bonds, *trans*-butadiene, 212–213
 Similarity index, 266–268
 cosine-like, 266
 distance, 267
 Slater basis functions, 144
 Sodium, parameters in potential energy,
 60–61
 Soliton, *see also trans*-Polyacetylene
 properties as functions of chain length,
 227
 Stereochemical theory of odor, 293
 Surface dilation wave, 110–111

T

Tanimoto index, 267
 Tight binding analysis, 72–104
 anharmonic potentials, 79–82
 exponential three-body potential, matrix
 elements, 95–97
 ionic mobility in small polarin-like limit,
 101–104
 mobile ion-channel phonon coupling,
 97–101
 one-dimensional harmonic oscillator
 analysis, 72–79
 screened potential, matrix elements,
 90–94
 three-dimensional Gaussian basis use of
 full potential, 82–90
 Transfer matrix element, 112, 118
 Transfer state function, constructing with
 basis functions, 117
 Transformations
 one-electron
 on HGP PRISM, 193
 on MD PRISM, 189
 two-electron
 on HGP PRISM, 190–192
 on MD PRISM, 184–188

Triple Density Transform Similarity Measure,
260–261

Tunnel current, 112

Tunneling probes, 111–120

U

Unrestricted Hartree–Fock approximation

π -electron energy, 218

singly charged polyenes, 237–240

spin density wave, 227–229

Unrestricted Hartree–Fock model,
207–208

V

Vibration, ions at channel minima, 64–65

W

Wall sources, location, 52–53

Z

Zero differential overlap approximation,
218–220