



# ADVANCES IN QUANTUM CHEMISTRY

Volume 20

Per-Olov Löwdin

ADVANCES IN  
QUANTUM CHEMISTRY

*VOLUME 20*

## 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)  
Masao Konani (Tokyo, Japan)  
Frederick A. Matsen (Austin, Texas)  
Roy McWeeney (Pisa, Italy)  
Joseph Paldus (Waterloo, Canada)  
Ruben Pauncz (Haifa, Israel)  
Siegrid Peyerimhoff (Bonn, Federal Republic of 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)  
Andrew Hurley (Clayton, Australia)  
Mu Shik Jhon (Seoul, Korea)  
Mel Levy (New Orleans, Louisiana)  
Jan Linderberg (Åarhus, 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**

ASSOCIATE EDITORS

**JOHN R. SABIN AND MICHAEL C. ZERNER**

QUANTUM THEORY PROJECT  
UNIVERSITY OF FLORIDA  
GAINESVILLE, FLORIDA

**VOLUME 20**



**ACADEMIC PRESS, INC.**

**Harcourt Brace Jovanovich, Publishers**

**San Diego New York Berkeley Boston**  
**London Sydney Tokyo Toronto**



COPYRIGHT © 1989 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.

San Diego, California 92101

*United Kingdom Edition published by*

ACADEMIC PRESS LIMITED

24-28 Oval Road, London NW1 7DX

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 64-8029

ISBN 0-12-034820-9 (alk. paper)

PRINTED IN THE UNITED STATES OF AMERICA

89 90 91 92 9 8 7 6 5 4 3 2 1

# CONTENTS

CONTRIBUTORS	ix
PREFACE	xi
ANNOUNCEMENT OF NEW FORMAT	xiii

## **The State of Surface Structural Chemistry: Theory, Experiment, and Results**

**M. A. VAN HOVE, S.-W. WANG, D. F. OGLETREE,  
AND G. A. SOMORJAI**

1. Introduction	2
2. Experimental Techniques for Surface Structure Determination	4
3. The Theory of Modern Techniques for Surface Structure Analysis	38
4. The Theory of Surface Chemistry and Bonding	81
5. Results of Surface Crystallography	112
6. Future Needs and Directions of Surface Studies	173
References	174

## **On the Change of Spectra Associated with Unbounded Similarity Transformations of a Many-Particle Hamiltonian and the Occurrence of Resonance States in the Method of Complex Scaling. Part II. Applications to the Hartree-Fock Scheme Based on the Bi-Variational Principle**

**PER-OLOV LÖWDIN, PIOTR FROELICH, AND MANOJ MISHRA**

1. Introduction	188
2. Brief Review of Some of the Results in the Previous Studies of the Stability Problem for a Pair of Adjoint Operators	190
3. Similarity Transformation of a Many-Particle Operator and Its Consequences in the Hartree-Fock Scheme	199
4. Properties of a Many-Particle Hamiltonian under Complex Scaling	217
5. Numerical Application of the Complex Scaling Method in the Hartree-Fock Approximation	222
6. Conclusions and Acknowledgments	231
Appendix A. Treatment of One-Particle Subspaces of Order N, Which Are Stable under Complex Conjugation	233
References	237

**Newton Based Optimization Methods for Obtaining Molecular Conformation**

JOHN D. HEAD AND MICHAEL C. ZERNER

1. Introduction	241
2. Characterization of Minima	243
3. Optimization Algorithms	245
4. Line Search Strategy	245
5. Search Directions	251
6. Analysis of Search Direction Methods	263
7. Computational Procedures	266
8. Results	267
9. Conclusions	285
Appendix	287
References	288

**Use of Cluster Expansion Methods in the Open-Shell Correlation Problem**

DEBASHIS MUKHERJEE AND SOURAV PAL

1. Introduction	292
2. "Theoretical Model Chemistry" and Its Relevance to the Open-Shell Formalisms	293
3. Cluster Expansion and Size-Extensivity	298
4. Categorization of the Open-Shell Many Body Formalisms	302
5. Semi-Cluster Expansion Theories for the Open-Shell States	310
6. Full Cluster Expansion Theories in Hilbert Space	324
7. Full Cluster Expansion Theories in Fock Space	332
8. Size-Extensive Formulations with Incomplete Model Space (IMS)	353
9. Concluding Remarks	364
References	365

**Jacobi Rotations: A General Procedure for Electronic Energy Optimization**RAMON CARBÓ, JOAN MIRÓ, LLORENÇ DOMINGO,  
AND JUAN J. NOVOA

1. Introduction	377
2. Elementary Jacobi Rotations	378
3. Variation of Quantum Mechanical Objects in the MO Framework via Elementary Jacobi Rotations	380
4. Some Application Examples	383
5. Monoconfigurational Energy Expressions	389

## Contents

vii

6. Multiconfigurational Energy Expression	395
7. Optimization of Rotation Sine	421
References	433
Appendix I	437
Appendix II	440
 INDEX	 443

This Page Intentionally Left Blank

## CONTRIBUTORS

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

RAMON CARBÓ (377), Unitat de Química, Col·legi Universitari de Girona, 17071-Girona, Spain

LLORENÇ DOMINGO (377), Secció de Química Quàntica, Departament de Quimiometria, Institut Químic de Sarrià, 08017-Barcelona, Spain

PIOTR FROELICH (188), Department of Quantum Chemistry, Uppsala University, S-75120, Uppsala, Sweden

JOHN D. HEAD (241), Department of Chemistry, University of Hawaii, Honolulu, Hawaii 96822

PER-OLOV LÖWDIN (188), Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611

JOAN MIRÓ (377), Unitat de Química, Col·legi Universitari de Girona, 17071-Girona, Spain

MANOJ MISHRA (188), Department of Chemistry, Indian Institute of Technology, Bombay 400076, India

DEBASHIS MUKHERJEE (292), Theory Group, Department of Physical Chemistry, Indian Association for the Cultivation of Science, Calcutta 700 032, India

JUAN J. NOVOA (377), Departament de Química Física, Facultat de Química, Universitat de Barcelona, 08028-Barcelona, Spain

D. F. OGLETREE (2), Materials and Chemical Sciences Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California

SOURAV PAL (292), Physical Chemistry Division, National Chemical Laboratory, Pune 411 008, India

G. A. SOMORJAI (2), Materials and Chemical Sciences Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720

**M. A. VAN HOVE (2), Materials and Chemical Sciences Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720**

**S.-W. WANG (2), Materials and Chemical Sciences Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720**

**MICHAEL C. ZERNER (241), Quantum Theory Project, University of Florida, Department of Chemistry, Gainesville, Florida 32611, and Guelph Waterloo Centre for Graduate Work in Chemistry, University of Guelph, Department of Chemistry, Guelph, Ontario, N1G 2W1 Canada**

## 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 not only shown that all the various kinds of matter are built up from a rather limited number of atoms, but also that these atoms are constituted 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 both boast of 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 try to present a survey of the current development of quantum chemistry as it is seen by a number of the 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 has seemed 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 is being used from Volume 20. A special announcement about the new format follows.

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 20, the Editors would like to thank the authors for their contributions, which give an interesting picture of part of the current state of art of the quantum theory of matter: from computational methods of optimizing the electronic energy and molecular conformations, over coupled-cluster expansion methods for the study of the open-shell correlation problem and the calculation of lifetimes of metastable states by means of the method of complex scaling, to a survey of the current state of surface structural chemistry.

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

## ANNOUNCEMENT OF NEW FORMAT

Starting with Volume 20, the *Advances in Quantum Chemistry* published by Academic Press assumes a new format. As before, all contributions will be *by invitation only*, but all authors will be asked to submit their manuscripts in *photoready form* in a format specified by the publisher. This procedure will eliminate copyediting, typesetting, and proofreading, and, in this way, will hopefully reduce the publication time to 6–8 months.

The text prepared by the author(s) should be single-spaced within a frame which measures 14 × 21 centimeters, and fonts of the type Courier 12, Times 12, Symbol 12, etc. should be used for the running text and the equations. The final text is conveniently prepared on 85% blue-line paper supplied by the publisher by means of a high-quality typewriter or printer. More detailed instructions are available from the publisher.

After going over to the new format, we further plan to publish two volumes of the *Advances* per year: one *regular volume* in the same style as before, and one *thematic volume* concerned with one specific subject, for example, computational methods in quantum chemistry, theoretical organic chemistry, quantum pharmacology and drug design, density functional theory, and relativistic quantum chemistry.

There has further been established an *Editorial Board* and an *Advisory Editorial Board* (the latter on a rotating basis) of eminent specialists in the field to suggest authors to be invited to make contributions both to the regular and thematic volumes and to help with questions such as publication policies. The names are listed in the front of this volume.

The Editors will be happy to answer inquiries about the *Advances* sent to: the Editorial Office of the *Advances in Quantum Chemistry*, Florida Quantum Theory Project, 362 Williamson Hall, University of Florida, Gainesville, Florida 32611. Telephone: (904) 392-1597, Telex: 510 601 6813 QTP UF UD; Bitnet: advances.orange.qtp.ufl.edu; FAX telephone: 904-392-8722.

This Page Intentionally Left Blank

THE STATE OF SURFACE STRUCTURAL CHEMISTRY  
THEORY, EXPERIMENT AND RESULTS

*M. A. Van Hove*

*S.-W. Wang*

*D. F. Ogletree*

*G. A. Somorjai*

Materials and Chemical Sciences Division  
Lawrence Berkeley Laboratory  
University of California at Berkeley  
Berkeley, California 94720

1. INTRODUCTION
2. EXPERIMENTAL TECHNIQUES FOR SURFACE  
STRUCTURE DETERMINATION
  - 2.1. Electron Diffraction Techniques
  - 2.2. Fine Structure Techniques
  - 2.3. Surface Topography Techniques
  - 2.4. Ion Scattering
  - 2.5. Complementary Techniques
3. THE THEORY OF MODERN TECHNIQUES FOR  
SURFACE STRUCTURE ANALYSIS

- 3.1. General Four-Step Description of Electron-Diffraction Techniques
- 3.2. Structural Information
- 3.3. Theoretical Aspects of Individual Techniques
- 3.4. Formalism for Individual Processes
- 3.5. Unifying and Distinguishing Features
- 4. THE THEORY OF SURFACE CHEMISTRY AND BONDING
  - 4.1. The Extended Hückel Theory and Applications
  - 4.2. The SCF  $X\alpha$  Scattered Wave Method
  - 4.3. SCF- $X\alpha$ -SW Calculations
  - 4.4. Results of the  $X\alpha$ -Method
  - 4.5. The Hartree-Fock-Slater Method and Its Applications
  - 4.6. The *ab initio* Quantum Chemical Calculation
  - 4.7. Conclusion
- 5. RESULTS OF SURFACE CRYSTALLOGRAPHY
  - 5.1. Ordering Principles
  - 5.2. Notation for surface lattices
  - 5.3. Numbers and types of solved surface structures
  - 5.4. Major trends among surface structures
- 6. FUTURE NEEDS AND DIRECTIONS OF SURFACE STUDIES

## 1. INTRODUCTION

The atomic and molecular structures of clean solid surfaces and of monolayers of adsorbed molecules are at the heart of most problems in surface science. The surface chemical bond can be investigated by determining the bond distances and bond angles of surface atoms with respect to their neighbors. Changes of surface structure with temperature, adsorbate coverage and surface composition are important ingredients of surface phenomena, including catalysis, corrosion, surface phase transformations (reconstruction, sintering, crystal growth and evaporation) and lubrication.

During the last ten years profound changes have occurred in surface science that resulted in the rapid improvement of our understanding of surface structure. Electron, atom and ion scattering techniques proved to be very sensitive to the surface monolayer structure and composition. High intensity photon fluxes can also be employed in ways to obtain high surface structure sensitivity.

Low energy electron diffraction (LEED) has produced most of the information on surface structure, while other techniques increasingly contribute to the data base. From these studies we learned that clean and flat solid surfaces often restructure; atoms often "relax" to shorten the first interlayer distance or seek new equilibrium positions that alter the long range order of the surface (reconstruction). Atoms of more open, rough surfaces relax more readily and often exhibit periodic steps of one atom in height. The surface composition of polyatomic solids (alloys, oxides, sulfides, etc.) can be very different from that in the bulk, resulting in altered atomic surface structures. Adsorbed atoms and molecules often form ordered two-dimensional layers. The structure and bonding in these layers changes with coverage and temperature.

The two-dimensional world of surface structural chemistry is rich, diverse, and full of surprises that reflect our incomplete knowledge of bonding and chemical interactions in the surface phase. This review attempts to summarize much of the knowledge that was gained in recent years.

We first discuss the various experimental methods of surface structure determination. Next we discuss the theoretical questions that arise when experimental data are analyzed to obtain surface chemical bond distances and angles, in particular the problem of electron propagation in solids, and the progress that has been made in answering these questions. A number of important techniques in addition to LEED rely on a proper understanding of electron propagation in solids to enable surface structure determination. This is followed by a discussion of the application of theoretical quantum chemistry to the calculation of the structure and bonding of chemisorbed atoms and molecules. Various calculational methods are reviewed, along with the results that have been obtained. The last section of this review surveys the surface structural results that have become available, ranging from clean surfaces to reconstructed

surfaces, multi-component systems and chemisorbed atoms and molecules. In conclusion we indicate the needs and possible directions for future studies of surface structural chemistry.

## 2. EXPERIMENTAL TECHNIQUES FOR SURFACE STRUCTURE DETERMINATION

Most quantitative information on surface structure and chemical bonding comes from studies of the solid-vacuum interface. In large part this is because the most powerful probes of surface structure rely on the propagation of electron, ion or atom beams. New developments such as the scanning tunneling microscope (STM), which can also investigate the solid-gas and solid-liquid interface, and new optical techniques, which can potentially investigate all types of surfaces and interfaces, may greatly extend our understanding of surface structural chemistry in the future. At this time, however, the solid-vacuum interface is the focus of most active investigations in surface structural chemistry. In this chapter we review the major experimental methods used in the study of surface structure and their application, along with recent experimental developments.

The majority of the known surface structures have been solved by comparing experimental data with theoretical predictions based on models of the surface. A theoretical description of the interaction of the probe with the surface is used to calculate spectra for a given model geometry (see part 3). Different geometries are tried, and structural parameters within a given model are varied, until a good fit is obtained between the experimental data and the theoretically calculated spectra. This basic approach has been used to interpret results from low-energy electron diffraction (LEED) and other electron diffraction techniques, as well as from helium diffraction and ion scattering.

Structural information has also been obtained by matching the observed frequencies of surface vibrational and electronic excitations with spectra calculated from structural models (see part 4). This approach has been applied

to the frequencies of electronic states as measured by angle-resolved ultra-violet photo-emission (ARUPS) and to the frequencies of vibrational excitations as measured by high-resolution electron energy-loss spectroscopy (HREELS).

There has been an ongoing search for a way to “image” surface structure directly from experimental results without fitting experimental data to theoretical calculations for model systems. So far this search has not been successful, although certain techniques yield more direct information for some special cases. Structure sensitive techniques such as scanning tunneling microscopy (STM) and field ionization microscopy (FIM) do give results that can be interpreted directly, but these results do not provide complete information on atomic coordinates in the near-surface region. Extended fine-structure techniques provide direct information on bond-lengths in many cases, but this is sometimes insufficient to fully determine the surface structure.

Most surface structural studies combine results from several different surface science techniques applied to the same system. For an approach based on model calculations to be successful, it is necessary to construct reasonable structural models of surfaces. The application of techniques which cannot give explicit information on the arrangement of atoms in the near-surface region is often necessary to develop good surface structural models.

A wide range of techniques have been developed to study surfaces in vacuum, and many techniques are commonly referred to by acronyms. Table I lists acronyms and brief descriptions of most common techniques used in surface science.

## 2.1. Electron Diffraction Techniques

The majority of known surface structures have been solved with electron diffraction techniques. Low energy electrons (below  $\sim 400$  eV) interact strongly with atoms through both elastic and inelastic processes. Inelastic scattering in solids limits electrons with energies below a few hundred eV to a mean free path of  $\sim 2\text{-}20$  Å. Elastic interactions are strong enough that multiple scattering is important in this energy range, so elastically scattered electrons are sensitive to



TABLE I. Surface Science Techniques

Acronym	Name	Description
AD	Atom or Helium Diffraction	Monoenergetic beams of thermal energy neutral atoms are elastically scattered off of ordered surfaces and detected as a function of scattering angle. This gives structural information on the outermost layer of the surface. Atom diffraction is extremely sensitive to surface ordering and defects.
AEAPS	Auger Electron Appearance Potential Spectroscopy	The EAPFS cross-section is monitored by Auger electron intensity. Also known as APAES.
AES	Auger Electron Spectroscopy	Core-hole excitations are created, usually by 1-10 KeV incident electrons, and Auger electrons of characteristic energies are emitted through a two-electron process as excited atoms decay to their ground state. AES gives information on the near-surface chemical composition.
AFM	Atomic Force Microscopy	Similar to STM. An extremely delicate mechanical probe is used to scan the topography of a surface, and a STM-type tunneling-current probe is used to measure the deflection of the mechanical surface probe. This is designed to provide STM-type images of insulating surfaces.

TABLE I (continued)

Acronym	Name	Description
APAES	Appearance Potential Auger Electron Spec- troscopy	See AEAPS.
APXPS	Appearance Potential X-ray Photoemission Spectroscopy	The EAPFS excitation cross-section is monitored by fluorescence from core-hole decay (also known as SXAPS).
ARAES	Angle-Resolved Auger Electron Spectroscopy	Auger electrons are detected as a function of angle to provide information on the spatial distribution or environment of the excited atoms (see AES).
ARPEFS	Angle-Resolved Photo-Emission Fine Structure	Electrons are detected at given angles after being photoemitted by polarized synchrotron radiation. The interference in the detected photoemission intensity as a function of electron energy $\sim 100$ -500 eV above the excitation threshold gives structural information.

(continued)

TABLE I (continued)

Acronym	Name	Description
ARPES	Angle-Resolved Photo-Emission Spectroscopy	A general term for structure sensitive photoemission techniques, including ARPEFS, ARXPS, ARUPS, and ARXPD.
ARUPS	Angle-Resolved Ultra-violet Photoemission Spectroscopy	Electrons photoemitted from the valence and conduction bands of a surface are detected as a function of angle. This gives information on the dispersion of these bands (which is related to surface structure), and also structural information from the diffraction of the emitted electrons.
ARXPD	Angle-Resolved X-ray Photoemission Diffraction	Similar to ARXPS and ARPEFS. The angular variation in the photoemission intensity is measured at a fixed energy above the excitation threshold to provide structural information.
ARXPS	Angle-Resolved X-ray Photoemission Spectroscopy	The diffraction of electrons photoemitted from core-levels gives structural information on the surface.

(continued)

TABLE I (continued)

Acronym	Name	Description
CEMS	Conversion-Electron Mössbauer Spectroscopy	A surface-sensitive version of Mössbauer spectroscopy. Like Mössbauer Spectroscopy, this technique is limited to some isotopes of certain metals. After a nucleus is excited by ray absorbtion, it can undergo inverse $\beta$ -decay, creating a core-hole. The decay of core-holes by Auger processes within an electron mean free path of the surface produces a signal. Detecting emitted electrons as a function of energy gives some depth-profile information, because of the changing electron mean free path.
DAPS	Dis-Appearance Potential Spectroscopy	The EAPFS cross-section is monitored by variations in the intensity of electrons elastically back-scattered from the surface.

(continued)

TABLE I (continued)

Acronym	Name	Description
EAPFS	Electron Appearance Potential Fine- Structure	A fine-structure technique (see EXAFS). Core-holes are excited by monoenergetic electrons at $\sim 1$ KeV. The modulation in the excitation cross section may be monitored through adsorption, fluorescence, or Auger emission.
ELNES	Electron energy-Loss Near-Edge Structure	Similar to NEXAFS, except monoenergetic high-energy electrons $\sim 60$ -300 KeV excite core-holes.
ELS	Electron Energy Loss Spectroscopy	Monoenergetic electrons $\sim 5$ -50 eV are scattered off a surface the energy losses are measured. This gives information on the electronic excitations of the surface and adsorbed molecules (see HREELS). Sometimes called EELS.
ESCA	Electron Spectroscopy for Chemical Analysis	Now generally called XPS.

(continued)

TABLE I (continued)

Acronym	Name	Description
ESDIAD	Electron Stimulated Desorption Ion Angular Distribution	Electrons break chemical bonds in adsorbed atoms or molecules, causing ionized atoms or radicals to be ejected from the surface along the axis of the broken bond by Coulomb repulsion. The angular distribution of these ions gives information on the bonding geometry of adsorbed molecules.
EXAFS	Extended X-ray Adsorption Fine-Structure	Monoenergetic photons excite a core-hole. The modulation of the adsorption cross-section with energy $\sim$ 100-500 eV above the excitation threshold yields information on the radial distances to neighboring atoms. The cross-section can be monitored by fluorescence as core-holes decay or by the attenuation of the transmitted photon beam. EXAFS is one of many "fine-structure" techniques. This is not intrinsically surface sensitive (see SEXAFS).

(continued)

TABLE I (continued)

Acronym	Name	Description
EXELFS	Extended X-ray Energy Loss Fine Structure	A fine-structure technique similar to EXAFS, except that 60-300 KeV electrons rather than photons excite core-holes. Like EXAFS, this techniques is not explicitly surface sensitive.
FIM	Field-Ionization Microscopy	A strong electric field $\sim$ volts/angstrom is created at the tip of a sharp, single crystal wire. Gas atoms, usually He, are polarized and attracted to the tip by the strong electrostatic field, and then ionized by electrons tunneling from the gas atoms into the tip. These ions, accelerated along radial trajectories by Coulomb repulsion, map out the variations in the electric-field strength across the surface with atomic resolution, showing the surface topography.

(continued)

TABLE I (continued)

Acronym	Name	Description
FTIR	Fourier-Transform Infra-Red spectroscopy	Broad-band IRAS experiments are performed, and the IR adsorption spectrum is deconvoluted by using a doppler-shifted source and Fourier analysis of the data. This technique is not restricted to surfaces.
HEIS	High-Energy Ion Scattering	High-energy ions, above $\sim 500$ KeV, are scattered off of a single crystal surface. The "channeling" and "blocking" of scattered ions within the crystal can be used to triangulate deviations from the bulk structure. HEIS has been used in particular to study surface reconstructions and the thermal vibrations of surface atoms (see also MEIS, ISS)
HREELS	High-Resolution Electron Energy Loss Spectroscopy	A monoenergetic electron beam, usually $\sim 2$ -10 eV, is scattered off a surface and energy losses below $\sim 0.5$ eV to bulk and surface phonons and vibrational excitations of adsorbates are measured as a function of angle and energy (also called EELS).

(continued)



TABLE I (continued)

Acronym	Name	Description
INS	Ion Neutralization Spectroscopy	Slow ionized atoms, typically $\text{He}^+$ , are incident on a surface where they are neutralized in a two-electron process which can eject a surface electron, a process similar to Auger emission from the valence band. The ejected electrons are detected as a function of energy, and the surface density of states can be determined from the energy distribution. The interpretation of the data is more complicated than for SPI or UPS.
IRAS	Infrared Reflection Adsorption Spectroscopy	Monoenergetic IR photons are reflected off a surface, and the attenuation of the IR intensity is measured as a function of frequency. This yields a spectrum of the vibrational excitations of adsorbed molecules. Recent improvements in the sensitivity of this technique allow IRAS measurements to be made on single crystal surfaces.

(continued)

TABLE I (continued)

Acronym	Name	Description
IRES	Infra-Red Emission Spectroscopy	The vibrational modes of adsorbed molecules on a surface are studied by detecting the spontaneous emission of infra-red radiation from thermally excited vibrational modes as a function of energy.
ISS	Ion-Scattering Spectroscopy	Ions are inelastically scattered from a surface, and the chemical composition of the surface is determined from the momentum transfer to surface atoms. The energy range is $\sim 1$ KeV to 10 MeV, and the lower energies are more surface sensitive. At higher energies this technique is also known as Rutherford Back-Scattering (RBS).
LEED	Low Energy Electron Diffraction	Monoenergetic electrons below $\sim 500$ eV are elastically back-scattered from a surface and detected as a function of energy and angle. This gives information on the structure of the near surface region.

(continued)

TABLE I (continued)

Acronym	Name	Description
LEIS	Low-Energy Ion Scattering	Low-energy ions, below $\sim 5$ KeV, are scattered from a surface, and the ion "shadowing" gives information on surface structure. At these low energies the surface atom ion scattering cross-section is very large, resulting in large surface sensitivity. Accuracy is limited because the low energy ion scattering cross-sections are not well known.
LEPD	Low Energy Positron Diffraction	Similar to LEED with positrons as the incident particle. The interaction potential for positrons is somewhat different than for electrons, so the form of the structural information is modified.
MEED	Medium Energy Electron Diffraction	Similar to LEED, except the energy range is higher, $\sim 300$ -1000 eV. LEED calculational methods break down in this energy range. New methods are being developed for glancing angle scattering, which emphasizes forward scattering.

(continued)

TABLE I (continued)

Acronym	Name	Description
MEIS	Medium-Energy Ion Scattering	Similar to HEIS, except that incident ion energies are $\sim 50$ -500 KeV.
	Neutron Diffraction	Neutron diffraction is not an explicitly surface-sensitive technique, but neutron diffraction experiments on large surface-area samples have provided important structural information on adsorbed molecules, and also on surface phase transitions.
NEXAFS	Near-Edge X-ray Adsorption Fine Structure	A core-hole is excited as in fine-structure techniques (see EXAFS), except the fine-structure within $\sim 30$ eV of the excitation threshold is measured. Multiple scattering is much stronger at low electron energies, so this technique is sensitive to the local 3-dimensional geometry, not just the radial separation between the source atom and its neighbors. The excitation cross-section may be monitored by detecting the photoemitted electrons or the Auger electrons emitted during core-hole decay.

(continued)

TABLE I (continued)

Acronym	Name	Description
NMR	Nuclear Magnetic Resonance	NMR is not an explicitly surface-sensitive technique, but NMR data on large surface-area samples has provided useful data on molecular adsorption geometries.
NPD	Normal Photoelectron Diffraction	Similar to ARPEFS with a somewhat lower energy range.
RBS	Rutherford Back-Scattering	Similar to ISS, except the main focus is on depth-profiling and composition. The momentum transfer in back-scattering collisions between nuclei is used to identify the nuclear masses in the sample, and the smaller, gradual momentum-loss of the incident nucleus through electron-nucleus interactions provides depth-profile information.

(continued)

TABLE I (continued)

Acronym	Name	Description
RHEED	Reflection High Energy Electron Diffraction	Monoenergetic electrons of $\sim 1$ -20 KeV are elastically scattered from a surface at glancing incidence, and detected as a function of angle and energy for small forward-scattering angles. Back-scattering is less important at high energies, and glancing incidence is used to enhance surface sensitivity.
SEELFS	Surface Electron Energy Loss Fine Structure	A fine structure technique similar to EXELFS, except the incident electron energies are $\sim 100$ -3000 eV. SEELFS is surface sensitive because of the lower excitation energy.
SERS	Surface Enhanced Raman Spectroscopy	Some surface geometries (rough surfaces) concentrate the electric fields of incident light sufficiently to enhance the Raman scattering cross-section so that it is surface sensitive. This gives information on surface vibrational modes, and some information on geometry via selection rules.

(continued)

TABLE I (continued)

Acronym	Name	Description
SEXAFS	Surface Extended X-ray Adsorption Fine-Structure	A surface-sensitive version of EXAFS, where the excitation cross-section fine-structure is monitored by detecting the photoemitted electrons (PE-SEXAFS), Auger electrons emitted during core-hole decay (Auger-SEXAFS), or ions excited by photoelectrons and desorbed from the surface (PSD-SEXAFS).
SHG	Second Harmonic Generation	A surface is illuminated with a high-intensity laser, and photons are generated at the second-harmonic frequency through non-linear optical process. For many materials only the surface region has the appropriate symmetry to produce a SHG signal. The non-linear polarizability tensor depends on the nature and geometry of adsorbed atoms and molecules.

(continued)

TABLE I (continued)

Acronym	Name	Description
SIMS	Secondary Ion Mass Spectroscopy	Ions and ionized clusters ejected from a surface during ion bombardment are detected with a mass spectrometer. Surface chemical composition and some information on bonding can be extracted from SIMS ion fragment distributions.
SPI	Surface Penning Ionization	Neutral atoms, usually He, in excited states are incident on a surface at thermal energies. A surface electron may tunnel into the unoccupied electronic level, causing the incident atom to become ionized and eject an electron, which is then detected. This technique measures the density of states near the Fermi-level, and is highly surface sensitive.
SPLEED	Spin-Polarized Low Energy Electron Diffraction	Similar to LEED, except the incident electron beam is spin-polarized. This is particularly useful for the study of surface magnetism and magnetic ordering.

(continued)



TABLE I (continued)

Acronym	Name	Description
STM	Scanning Tunneling Microscopy	The topography of a surface is measured by mechanically scanning a probe over a surface with Angstrom resolution. The distance from the probe to the surface is measured by the probe-surface tunneling current. Also known as Scanning Electron Tunneling Microscopy (SETM).
SXAPS	Soft X-ray Appearance Potential Spectroscopy	Another name for APXPS.
TEM	Transmission Electron Microscopy	TEM can provide surface information for carefully prepared and oriented bulk samples. Real images have been formed of the edges of crystals where surface planes and surface diffusions have been observed. Diffraction patterns of reconstructed surfaces, superimposed on the bulk diffraction pattern, have also provided surface structural information.

(continued)

TABLE I (continued)

Acronym	Name	Description
TDS	Thermal Desorption Spectroscopy	An adsorbate-covered surface is heated, usually at a linear rate, and the desorbing atoms or molecules are detected with a mass spectrometer. This gives information on the nature of adsorbate species and some information on adsorption energies.
TPD	Temperature Programmed Desorption	Similar to TDS, except the surface may be heated at a non-uniform rate to get more selective information on adsorption energies.
UPS	Ultra-violet Photoemission Spectroscopy	Electrons photoemitted from the valence and conduction bands are detected as a function of energy to measure the electronic density of states near the surface. This gives information on the bonding of adsorbates to the surface (see ARUPS).
WF	Work Function measurements	Changes in the work-function during the adsorption of atoms and molecules provide information on charge-transfer and chemical bonding.

(continued)

TABLE I (continued)

Acronym	Name	Description
XANES	X-ray Adsorption Near-Edge Structure	Another name for NEXAFS.
XPS	X-ray Photoemission Spectroscopy	Electrons photoemitted from atomic core levels are detected as a function of energy. The shifts of core level energies gives information on the chemical environment of the atoms (see ARXPS, ARXPD).
XRD	X-Ray Diffraction	X-ray diffraction has been carried out at extreme glancing angles of incidence where total reflection assures surface sensitivity. This provides structural information that can be interpreted by well-known methods. An extremely high x-ray flux is required to get useful data from single crystal surfaces. Bulk x-ray diffraction is used to determine the structure of organo-metallic clusters, which provide comparisons to molecules adsorbed on surfaces. X-ray diffraction has also given structural information on large surface-area samples.

the three-dimensional geometry of the near-surface region. Electron scattering also depends on the chemical identity of the scattering atoms, so electron diffraction techniques are sensitive to both chemical composition and structure throughout the near-surface region. Therefore low-energy electrons are an ideal probe of surface structural chemistry. The main difficulty with electron diffraction methods is that the strong electron-surface interaction makes data analysis difficult -- it is not possible in most cases to "invert" diffraction data to obtain the original surface structure.

### 2.1.1. Low energy electron diffraction

In a conventional LEED experiment a focused beam of monoenergetic electrons is reflected off an ordered single crystal surface and the elastically backscattered electrons are detected. Several types of information can be extracted from LEED patterns. First, the symmetry of the LEED pattern is related to the symmetry of the surface, so changes in surface symmetry because of reconstruction or chemisorption are immediately detected. Second, imperfections and deviations in long-range order change the shape of diffraction beams. This can be used to study the process of ordering in phase transitions or the growth of ordered domains within overlayers. Finally, the integrated intensity of diffraction beams depends on the detailed bonding structure of the surface. This information is most important for the study of surface chemical bonding. The intensity of LEED beam is a function of the incident electron beam energy and angles of incidence on the surface.

The intensity of a given diffraction beam can be measured as the incident electron energy or beam voltage is varied, producing an intensity-voltage or I-V curve; or as the angle of the incident electron beam relative to the surface is varied, producing an I- $\theta$  or rocking curve; or as the crystal is rotated around its own normal at a given angle of incidence, producing an I- $\phi$  curve. For structure determination LEED data are most often recorded in the form of I-V curves with the incident electron beam at normal incidence or lying in a mirror plane of the surface. This form of data are usually simpler to work with both experimentally and theoretically.

One advantage of LEED is that the diffraction process filters out effects due to local defects or deviations from long-range order. The contribution of defects to I-V curves is proportional to the first power of the number of defects, while the contribution of the part of the surface with long-range order is proportional to the square of the number of atoms involved, so the LEED beam integrated intensity reflects the equilibrium geometry of the ordered surface structure.

LEED has been used to determine the structure of a wide variety of surfaces, including clean and reconstructed surfaces of metals and semiconductors, and atomic and molecular physisorption and chemisorption on many different substrates (see part 5). As the theoretical and experimental tools of LEED have improved, the structure of systems with larger and more complex unit cells have been determined. Successful LEED structure determinations have been carried out for systems with several molecules adsorbed in unit cells up to 16 times larger than the substrate unit cell,<sup>/1/</sup> and for reconstructed surfaces where the structural rearrangement involves several surface layers.<sup>/2/</sup>

There are still a number of surface systems where the structure cannot be determined by LEED for theoretical and experimental reasons. High Miller-index surfaces, such as stepped or kinked surfaces, have layers separated by very small distances normal to the surface. The calculational tools normally used for LEED break down in this case, and no new approach has yet been developed to solve this problem. Experimental difficulties restrict the study of insulator surfaces, because of charging problems, and of molecular crystal surfaces, because of beam damage problems.

A reliable LEED structure determination requires a large data base, usually a number of different I-V curves collected at more than one angle of incidence. Electron beam damage to the surface must be prevented during data acquisition.

Over the last decade both the speed and accuracy of data acquisition and the sensitivity of the LEED experiment have been significantly improved. In most structure determination experiments the LEED patterns are displayed on a phosphor screen. LEED I-V curves are then measured by photometric or photographic means. Fast data acquisition systems have been developed that record LEED patterns with video cameras interfaced to computers and generate I-V curves during the experiment.<sup>/3,4/</sup> The sensitivity of LEED is also being greatly increased with the development of instruments that can detect single

diffracted electrons, which completely solves the problem of electron beam damage./5,6,7/ Such instruments should be able to extend the range of LEED structure determination to delicate systems such as molecular crystals and physisorption systems, and to systems without long-range order, where the diffracted intensity is much lower than for ordered systems.

There are variations of LEED that use positrons (low energy positron diffraction, or LEPD)/8/ or spin-polarized electrons as a probe./9/ The latter technique is especially useful as a probe of magnetic ordering in surfaces. Spin-polarized LEED, or SPLEED, has been used to study surface magnetism and phase transitions on nickel.

At higher energies the surface sensitivity and the importance of multiple scattering is reduced. Diffraction techniques in this energy range have primarily been applied to the study of defects and deviations from long range order, as in the use of reflection high-energy electron diffraction (RHEED) to monitor epitaxial growth on surfaces. The higher-energy diffraction techniques have only infrequently been used for surface structure determination/10/ due to theoretical complications, but new approaches are promising./11/

### 2.1.2. Diffuse low energy electron diffraction

Conventional LEED experiments are done on systems with long range order. This is not a fundamental restriction. The energy-dependent variation in diffracted intensity at a given point in the Brillouin zone is primarily determined by the local scattering geometry at the surface. Long range order gives rise to sharp diffraction beams which reflect this intensity dependence. If angle resolved intensity data are collected from a system with definite local geometry but without long range order, the local geometry can still be determined by LEED calculations./12/ This "diffuse" LEED experiment will be difficult to interpret unless there is one predominant local scattering geometry. Unlike the conventional LEED experiment, the diffraction process does not filter out the contributions of defects and impurities from the contribution of the equilibrium

structure in the diffuse LEED experiment. If the scattering from different kinds of sites is superimposed in the experimental data it will be difficult to construct a reasonable model of the surface.

There are a number of "lattice-gas" chemisorption systems, where atoms or molecules are adsorbed in well-defined sites on the surface, as determined by vibrational spectroscopy, but where there is no long range order. The adsorption sites and bond lengths for such systems can be determined by diffuse LEED calculations, as has been done for oxygen on the tungsten (100) surface.<sup>13/</sup>

### 2.1.3. Photoelectron diffraction

The interference and diffraction of photoemitted electrons can provide structural information, as in angle-resolved photo-electron fine structure (ARPEFS) spectroscopy, angle-resolved x-ray photoelectron spectroscopy (ARXPS), and angle-resolved x-ray photoelectron diffraction (ARXPD). The main advantage of photoemission over LEED is that the initial state of the photoemitted electron is a simple spherical wave, and can be controlled to emphasize particular structural features, especially in chemisorption systems. Because the initial state is simple, in some high-energy cases the data analysis can be more straight forward than a full LEED multiple-scattering calculation.

In photoelectron diffraction experiments monoenergetic photons excite electrons from a particular atomic core level. Angular momentum is conserved, so the emitted electron wave-function is a spherical wave centered on the source atom, with angular momentum components  $l \pm 1$ , where  $l$  is the angular momentum of the core level. If the incident photon beam is polarized, the orientation of the emitted electron wave-function can be controlled. These electrons then propagate through the surface and are detected and analyzed as in LEED experiments. A synchrotron x-ray source normally produces the intense beams of variable energy polarized photons needed for photoelectron diffraction.

In angle-resolved photoemission fine-structure (ARPEFS) experiments electrons are detected at a given angle as a function of energy. Structural information is obtained from the variation in intensity due to interference between different scattering paths for electrons over an energy range of  $\sim 100$ -500 eV. In this energy range the interference effects are dominated by single

scattering, so the scattering amplitude can be readily calculated for a model system. A Fourier transform of the interference pattern will show the distribution of different scattering path-lengths that contribute to the detected intensity. This can be a guide to the proper surface structure, but this result must be confirmed by a calculation of the interference for a model geometry./14/ ARPEFS has been used to solve several surface structures, including sulfur on nickel./15/ Long-range order is not required for ARPEFS experiments. In spite of the name, ARPEFS is not a normal "fine-structure" technique (see below). The variations in the detected intensity as a function of energy are orders of magnitude larger than the fine-structure modulation in core-level excitation cross-sections.

Photoelectron diffraction is most useful for systems where the photoexcited atoms all have the same local geometry, as in a chemisorption problem. If there are source atoms in different local geometries, there will be interference between multiple sets of scattering paths, and the resulting interference spectrum will be harder to interpret. For these cases LEED experiments are probably better, with diffuse LEED used for disordered systems.

## 2.2. Fine Structure Techniques

There are a large number of "fine structure" techniques, all based on the same physical principle. In all of these techniques an electron is ejected from an atomic core-level by incident photons or other particles. The different names refer to different experimental arrangements for the excitation and detection of core-holes.

As the excitation energy is varied the energy of the emitted electron varies. The emitted electron wave can be scattered back to the source atom from neighboring atoms, where it interferes with the source wave-function with a phase that depends on the electron energy, interatomic distance and the identity of the neighboring atom. This energy dependent interference changes the coupling of the incident excitation to the final state, producing a modulation or "fine-structure" in the excitation cross section.



Fine structure experiments are often carried out with synchrotron sources, since the initial electron state is better defined for photoemission than for electron excitation. When core-hole decay is detected by Auger or secondary electron emission, the technique is surface sensitive. Core-hole decay can also be detected by fluorescence, or by adsorption of the incident photon beam. These methods are not intrinsically surface sensitive, but they are useful when the source atoms are exclusively located at the surface.

If there are different local geometries for source atoms, the different fine structures will be superimposed in the experimental spectrum, which will then be more difficult to interpret. Fine structure techniques are particularly useful when a chemisorbed atom is used as the source atom.

### 2.2.1. Near-edge fine structure techniques

In the "near-edge fine structure" region emitted electrons have energies up to  $\sim 50$  eV and multiple scattering effects are predominant. An emitted electron wave can scatter several times and still return with a significant amplitude to the source atom. Therefore the variation in the observed cross-section depend not only on interatomic bond distances, but also on bond angles, so the full three-dimensional geometry around the source atom can be determined. The theoretical analysis of near-edge fine structure data is very similar to LEED data analysis.

A few structures have been solved by x-ray adsorption near-edge fine structure (XANES), including oxygen on Ni(100).<sup>/16/</sup> This technique is also known as near-edge x-ray adsorption fine structure (NEXAFS).

A variation of XANES or NEXAFS has been used to determine the structure of molecules chemisorbed on surfaces. In this approach photoemitted electrons excite molecular orbitals in the chemisorbed molecules. By varying the polarization of the incident photons, molecular orientation can be determined from selection rules for excitation. The bond lengths can be determined from a quasi-empirical correlation between bond-length and the shift in the molecular orbital excitation energy. This technique has been used to study the chemisorption of several hydrocarbon molecules on different metal surfaces.<sup>/17/</sup>

### 2.2.2. Extended fine-structure techniques

For emitted electrons above  $\sim 50$  eV, the “extended fine structure” region, the modulation of the excitation cross section is dominated by single backscattering between near-neighbor atoms and the source atom. A Fourier transform of the extended fine structure as a function of momentum transfer gives the distribution of radial distances between the source atom and neighboring atoms. With empirical or theoretical correction for scattering phase shifts, which have been calibrated using results from bulk structures known from x-ray diffraction, this gives chemical bond lengths with an accuracy of better than 0.05 Å.

There are a large number of extended fine-structure results available. These techniques, after LEED, have provided the largest amount of quantitative information on surface structure and chemical bonding. The extended fine structure techniques give quantitative information on surface structure without the need for complex model calculations. In simple systems and when combined with qualitative data from other experiments, knowledge of bond lengths may be sufficient to completely describe the surface geometrical structure. These techniques are most useful in multi-component or chemisorption systems, where atoms of a particular chemical species have only one local geometry. By selectively exciting an appropriate core level, near-neighbor bond lengths are determined. In systems with non-equivalent atomic sites, such as reconstructed surfaces, the extended fine structure will be complicated by a superposition of radial distribution functions for the different sites. Additional data will be needed to solve these structures. Data interpretation is generally simpler with photon excitation, since the polarization and orientation of final-state electrons can be controlled by using single crystal samples and polarized photons from a synchrotron source. One important advantage of extended fine structure techniques is that they do not require long range order or single crystal substrates, so they can be directly applied to many systems of technological importance.

A number of different experimental methods are available to observe fine structure. In extended x-ray adsorption fine structure (EXAFS) experiments x-ray photons excite core levels and the cross section is determined by x-ray absorption

or fluorescence from excited atoms. When Auger or secondary electrons emitted from the excited atoms are detected, the process, now called surface EXAFS (SEXAFS), is surface sensitive. SEXAFS experiments have also been done by detecting ions emitted from surfaces in photon stimulated desorption SEXAFS (PSD-SEXAFS).

There are analogous processes using electrons instead of photons to excite atomic core levels. Diffraction effects may perturb the cross section fine structure when electrons are used to excite core levels or monitor cross section. This is not a problem for systems without long range order. For cases involving long range order it is desirable to integrate over detector angles, and also electron beam incidence angles where this is practical, to average out multiple-scattering effects in the initial and final electron states. In electron appearance potential fine structure spectroscopy (EAPFS) incident electrons at  $\sim 1$  KeV excite shallow core holes, and the cross section is monitored by soft x-ray fluorescence in soft x-ray appearance potential spectroscopy (SXAPS-EAPFS), Auger electron appearance potential spectroscopy, (AEAPS-EAPFS), or by the variation in elastic backscattering in disappearance potential spectroscopy (DAPS-EAPFS), which is also referred to as surface electron energy loss fine structure (SEELFS). With excitation by high energy electrons (60-300 KeV) and detection by fluorescence the process is called extended x-ray energy loss fine structure (EXELFS), or electron energy loss near-edge spectroscopy (ELNES) in the near edge region. This last process, like EXAFS, is not inherently surface sensitive.

### 2.3. Surface Topography Techniques

Field ion microscopy (FIM), atom diffraction (AD) and scanning tunneling microscopy (STM) provide atomic scale information on surface topography. These techniques produce good qualitative images of surfaces. It is difficult to get atomic coordinates directly, but knowledge of the surface topography can lead directly to structural models of a surface.

### 2.3.1. Field ion microscopy

Field ion microscopy (FIM) is the oldest of these techniques, developed in the 50's./18/ A sharp, single crystal metal tip with a radius of  $\sim 1000 \text{ \AA}$  is maintained at  $\sim 10 \text{ kV}$  in a low pressure ( $\sim < 10^{-4} \text{ torr}$ ) gas, usually helium. Gas molecules are polarized and attracted to the tip by the strong, inhomogeneous electrostatic fields. The field strength near the tip is several  $\text{V/\AA}$ . Valence electrons in a gas atom can tunnel into the metal surface as the neutral atom approaches the tip, creating an ion which is repelled by the high field around the tip. The ions project an image of the high field regions of the tip onto a phosphor screen. If the tip is cooled, individual surface atoms can be imaged. Pairs of atoms separated by as little as  $1.5 \text{ \AA}$  have been resolved.

A field-ion microscopy image is a two-dimensional projection of the outermost surface layer. This image provides a qualitative image of the surface, but very little information on distances normal to the surface. Although the direct information on the surface chemical bond is limited, FIM has added greatly to the qualitative understanding of surface structure. Because of the high field strengths required, application of the FIM has been limited to refractory metals, although some chemisorption systems on these metals have been studied./19/

### 2.3.2. Atomic beam diffraction

A thermal energy atomic beam ( $20\text{-}200 \text{ meV}$ ) has a wavelength on the order of inter-atomic distances. The atomic beam diffracts from a contour of the surface potential corresponding to the beam energy. This contour is located  $3\text{-}4 \text{ \AA}$  above the ion cores in the outermost layer of the surface. Atomic beam diffraction patterns are normally interpreted using model surface scattering calculations, where the scattering is described as a Van der Waals interaction.

The low energy probe of atomic diffraction does not damage even delicate physisorbed overlayers, and it is sensitive to hydrogen, which is an important component of many surface systems of current interest. Electron scattering techniques are relatively insensitive to hydrogen because of its small scattering

cross section. The position of hydrogen chemisorbed on Pt(111) was determined using helium diffraction.<sup>/20/</sup> Finally, atom diffraction is extremely sensitive to surface order and defects, and it has been very useful in the study of disorder and kinetic processes on surfaces. Because neutral atoms scatter well outside the ion cores of the surface atoms, structural details can be lost and no information on subsurface structure can be obtained.

### 2.3.3. Scanning tunneling microscopy

The scanning tunneling microscope (STM) is one of the most recently developed surface sensitive techniques.<sup>/21/</sup> In STM a metal probe is brought close enough to the surface under study for the electron wave functions to overlap. A fixed potential difference between probe and surface is maintained and the probe approaches the surface until a given tunneling current is observed. The probe is mechanically scanned over the surface, and a feedback loop adjusts the vertical spacing to maintain a constant tunneling current. STM has achieved a horizontal resolution of  $\sim 2 \text{ \AA}$  and a vertical resolution of  $\sim 0.1 \text{ \AA}$  under optimum conditions. By changing the probe-surface potential difference it is possible to map out different surface profiles of wave function overlap. The STM is very flexible and can be applied to all kinds of surfaces. Unlike electron techniques it is not limited to the solid-vacuum interface, although the best resolution has been obtained in vacuum.

The STM can give a direct qualitative image of surface topography. An early STM experiment confirmed the LEED result that the "missing row" model correctly describes the  $2 \times 1$  reconstruction of the gold (111) surface.<sup>/22/</sup> Because the tunneling current is a complicated function of the electronic and geometric structure of the tip and surface, so it is difficult to relate a constant tunneling current contour to the geometrical structure of the surface, especially if more than one type of atom is involved. Information on the surface electronic structure is available, however. The lateral resolution of STM images is not good enough to provide accurate information on chemical bond lengths and angles. As the experimental and theoretical tools develop, scanning tunneling microscopy should provide increasingly useful information on surface structure.

## 2.4. Ion Scattering

At high energies ( $\sim 1$  MeV) the interaction of ions with surfaces can be described by classical Rutherford scattering. This simple interaction has been used to study surface structure by directing ion beams along bulk crystal axes at solid surfaces. The "channeling" and "blocking" of these beams is very sensitive to deviations from bulk structure. Ion scattering has been used in particular to study relaxation and reconstruction at crystal surfaces. Measurements of surface relaxation have provided accurate structure determinations, as in the case of reconstructed Au(110).<sup>23</sup> However, if a surface structure involves a large departure from the bulk crystal structure, it can be difficult to solve from ion scattering results.

In any case, because ion scattering is strongly affected by the thermal vibrations of surface atoms, experimental data must be compared to Monte-Carlo simulations for model surfaces to achieve quantitative results. The available data base for structure-fitting is rather small compared to electron spectroscopies, so the sensitivity to structural parameters is sometimes limited. But when the surface structure is close to the bulk structure, ion channeling data can be strongly sensitive to small variations in structural parameters.

## 2.5. Complementary Techniques

There are a number of other surface probes that are sensitive to the local geometry of the surface or give important information on surface composition, but which do not give direct information on atomic coordinates in the surface region. These techniques provide vital information needed to construct reasonable models of surface structure, which are needed to interpret data from quantitative techniques such as LEED and SEXAFS.

### 2.5.1. Chemical composition

Chemical composition is the most basic information needed to describe a surface, and a pre-requisite for structure determination. Auger electron spectroscopy is the most generally used techniques for measuring surface composition, and it is sensitive to all elements except hydrogen and helium. Thermal desorption spectroscopy (TDS) is a simple and particularly useful technique for chemisorption studies -- it gives a general idea of the chemical identity of adsorbates, and some information on the type of binding to the surface. Secondary-ion mass spectrometry (SIMS) is another way to get information on surface chemical composition -- its interpretation is generally more complicated, but SIMS is sensitive to hydrogen, and also gives some information on the molecules present on the surface.

### 2.5.2. Electronic spectroscopy

One group of techniques is sensitive to electronic structure at the surface, and can probe the electronic band structure and density of states near the Fermi level. This electronic information is useful for understanding the bonding mechanisms responsible for chemical process operating at the surface. Structural information can also be obtained by comparing experimentally observed electronic structure with theoretical calculations of electronic structure for model systems (see part 4).

Ultra-violet photoemission spectroscopy (UPS) probes the density of states, and ion neutralization spectroscopy (INS) and surface Penning ionization (SPI) provide similar information with probes of ions and metastable atoms, respectively. Angle-resolved UPS can determine the valence band structure. X-ray Photoelectron Spectroscopy (XPS) provides information on chemical shifts of the atomic core levels, and this can also help in understanding chemical bonding at the surface.

### 2.5.3. Vibrational spectroscopy

Another class of techniques monitors surface vibration frequencies. High-resolution electron energy loss spectroscopy (HREELS) measures the inelastic scattering of low energy ( $\sim 5\text{eV}$ ) electrons from surfaces. It is sensitive to the vibrational excitation of adsorbed atoms and molecules as well as surface phonons. This is particularly useful for chemisorption systems, allowing the identification of surface species. Application of normal mode analysis and selection rules can determine the point symmetry of the adsorption sites./24/ Infra-red reflectance-adsorption spectroscopy (IRRAS) is also used to study surface systems, although it is not intrinsically surface sensitive. IRRAS is less sensitive than HREELS but has much higher resolution.

### 2.5.4. Optical techniques

Non-linear optical techniques, such as second harmonic generation (SHG), have recently been used as surface probes. Bulk materials with inversion symmetry do not generate second harmonic signals, while surfaces and interfaces cannot have inversion symmetry, so the total SHG signal will come from the surface region for many systems. The components of the non-linear polarizability tensor have been used to determine the orientation of chemisorbed molecules. /25/ Optical techniques are not limited to solid-vacuum interfaces like charged particle techniques, so their further development can expand the range of surface structural studies to solid-solid, solid-liquid, solid-gas and liquid-gas interfaces.

### 2.5.5. Electron-stimulated desorption

Observation of the ion angular distribution after electron stimulated desorption of chemisorbed species (ESDIAD) can provide direct quantitative information on the orientation of adsorbed molecules on surfaces. Electrons incident on the surface can excite chemical bonds into non-bonding states, causing molecular decomposition. The excess energy can be converted into kinetic energy, which accelerates an ionic fragment of the molecule along the axis



of the broken bond. The angular distribution of desorbed ions can be related to the orientation of the bonds in the unperturbed adsorbed molecules. This technique gives direct information on the number and symmetry of sites for chemisorption, and approximate information on bond angles.

### 3. THE THEORY OF MODERN TECHNIQUES FOR SURFACE STRUCTURE ANALYSIS

Although the wide range of experimental techniques described in part 2 have made significant contributions to understanding surface structure and bonding, many of these techniques do not yield explicit information on atomic coordinates in the near surface region, the information required to characterize the surface chemical bond. Other techniques can provide some of this information in certain cases, but can only be applied to a limited range of surface systems, for example ion channeling studies have provided important information on the reconstructions of solid surfaces, but are much more difficult to apply to most chemisorption systems. Optical techniques are being developed which have the potential to contribute direct, quantitative information on surface structure, including surface x-ray diffraction and non-linear optical probes, but these have not yet reached the stage of routine use.

Almost all of the existing quantitative data on surface structure was obtained through experimental techniques that involve the propagation and scattering of electrons in solids. This class of surface probes includes the whole range of "fine-structure" techniques, in addition to LEED and angle-resolved photoemission experiments. This chapter will provide a theoretical description of these techniques and the methods of analyzing the experimental data to determine surface structure.

### 3.1. General Four-Step Description of Electron-Diffraction Techniques

The most promising techniques for obtaining detailed surface structural information about molecular adsorbates rely on electron diffraction in one way or another. These include LEED,/27,28,29/ IV-HREELS,/30/ EAPFS,/31/ SEELFS,/32,33/ EXELFS,/34,35/ ARUPS, ARXPS, ARPEFS,/36,37,38/ PE-SEXAFS, SEXAFS, EXAFS, and NEXAFS (XANES). These and other techniques have been discussed above in part 2, and were summarized in Table I. Among these techniques, LEED has been the most productive.

We shall in the following discuss the surface sensitive techniques involving electron diffraction from a common viewpoint. These techniques can all be fit within a four-step description, which will be very useful for comparison. Various techniques are illustrated in terms of this description in Figure 1, and the four steps are defined below.

#### 3.1.1. Step 1: incoming particle

Either a photon or an electron impinges on the surface. This photon or electron may be polarized to yield additional surface information. The electron, part of a well-collimated and monoenergetic electron beam, may be multiply scattered by surface atoms before reaching step 2. Multiple scattering is most marked at kinetic energies below about 200 eV and can make the process particularly sensitive to the surface structure at these lower energies.

#### 3.1.2. Step 2: primary event

A variety of processes may take place at this stage, distinguishing the different techniques -- photoelectron emission, Auger electron emission, ion emission, energy loss due to core-hole, plasmon or phonon excitation, as well as

elastic electron scattering. Thus, in the primary event, the incoming particle generates an outgoing particle. More than one type of particle may be emitted simultaneously, e.g. an electron and a photon in SEXAFS.

From the point of view of obtaining surface structure information, the nature of this primary process is important in two ways. First, the process may be atom-specific, giving chemical selectivity, as in photoelectron emission and Auger electron emission. Second, the process can affect the characteristics of the outgoing particles, especially in terms of angular distribution, and in terms of the exact location within the surface where this primary process takes place. The angular distribution can give direct evidence of the molecular orientation with respect to the surface; this happens, for instance, when an electron is emitted from a valence orbital.

In terms of location, the primary process may be concentrated at one atom, as in an electron-atom scattering in LEED or in electron emission from core levels. It may also be delocalized over many atoms, as in photoelectron emission from a delocalized valence band level. The latter case is again of value to obtain molecular orientations directly. The polarization of the incident electron may also be used to determine molecular orientations, through its effect on the primary process.

### 3.1.3. Step 3: secondary event

Part of the outgoing electron wave resulting from step 2 will back-scatter from neighboring atoms to the site of emission or initial scattering. If the kinetic energy of the outgoing electrons does not exceed a few hundred eV this results in detectable interference between the outgoing and back-scattered electron waves, which modulates the probability of the primary emission or scattering process. Such modulation is usually called "fine structure" and is the basis of techniques like SEXAFS (EXAFS) and EAPFS (EXELFS, SEELFS), as well as in NEXAFS (XANES). Note that the electrons responsible for this fine structure are rarely collected themselves -- other outgoing electrons or photons (or even ions) are normally detected experimentally.

Two energy regimes are of importance here. At electron kinetic energies above about 50 eV, single back-scattering from neighboring atoms is usually predominant. By contrast, at lower energies electrons can also be multiply scattered by several nearby atoms before returning to the location of the primary process. The single scattering regime modulates the outgoing particle current in a simple oscillatory manner. The near-neighbor interatomic distances can be extracted relatively easily by Fourier transformation given appropriate scattering phase shifts (cf. SEXAFS). When multiple scattering is significant, a multiple-scattering computer simulation is required. This makes the analysis more complex, but gives access to bond angles as well as bond lengths (cf. NEXAFS).

#### 3.1.4. Step 4: outgoing particle

Photons, ions or electrons are detected outside of the crystal. On their way out, electrons may be multiply scattered by surface atoms, especially at the lower energies, which again gives added sensitivities to the surface structure.

### 3.2. Structural Information

Two modes of electron detection are commonly used, which affect the kind of structural information that can be extracted.






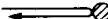



In the angle resolved mode, which selects electrons traveling in well-defined directions, emphasis is laid on the properties of plane-wave diffraction within the surface -- this provides access to all interplanar distances (not only between atomic planes parallel to the surface) and thus, to the complete three-dimensional crystallographic structure (as for example in LEED and in ARPES). Significant structural information can come from steps 1 and 4.

In the angle integrated mode, electrons emitted in many different directions are accumulated simultaneously. To an appreciable extent this averages out the structure-dependent effects of step 4, and therefore emphasizes the information arising from the primary and secondary events of steps 2 and 3, such as the radial distances determined in fine-structure techniques like SEXAFS.

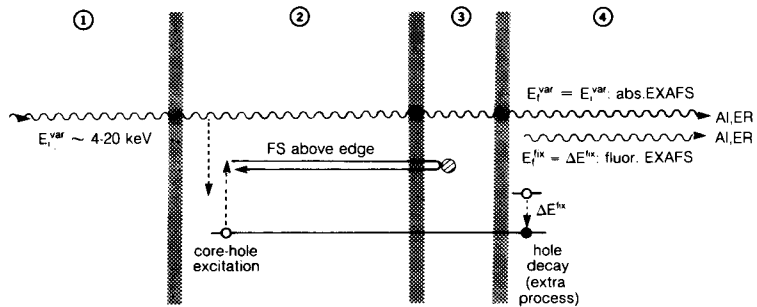
Energy vs. time diagrams

for the event sequences in surface crystallography techniques that use electron scattering/diffraction.

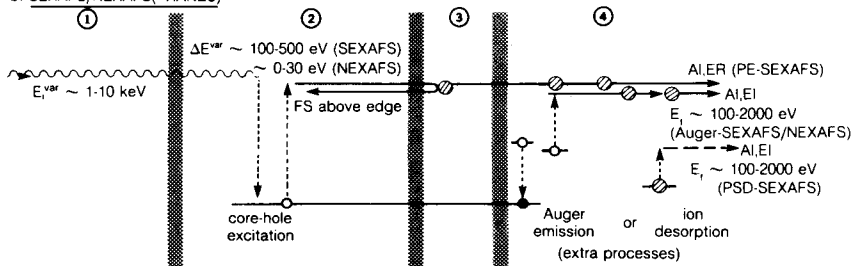
Legend:

	electron propagation
	photon propagation
	ion propagation
	multiple electron forward-scattering by atoms is important
	multiple electron backscattering by atoms is important
	electron backscattering by nearby atoms is important
	bound electron
	electron hole
	energy gain, loss
$E_i$	incident particle energy
$E_f$	final particle energy
superscript "fix":	energy is kept fixed in experiment
superscript "var":	energy is varied in experiment
$E_b$	electronic binding energy
$E_A$	atomic binding energy in surface lattice
AR	angle-resolved detection
AI	angle-integrated detection
ER	energy-resolved detection
EI	energy-integrated detection
FS	fine structure is important

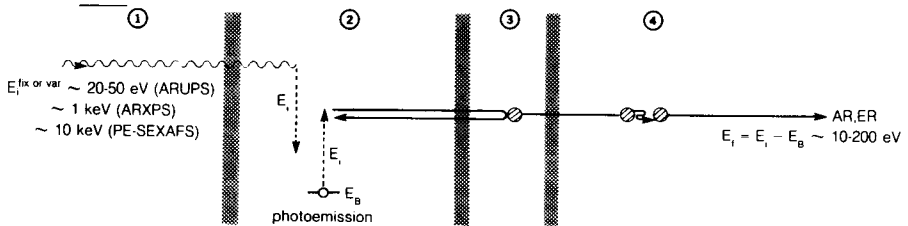
a. EXAFS



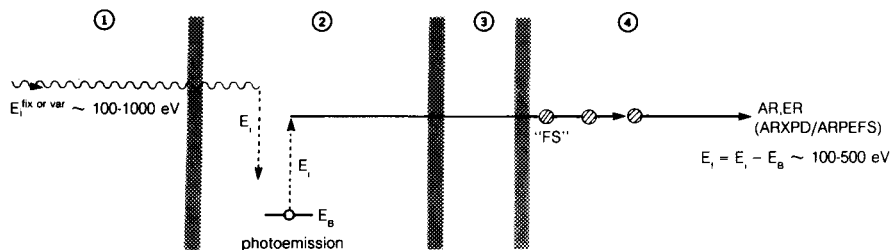
b. SEXAFS/NEXAFS(=XANES)



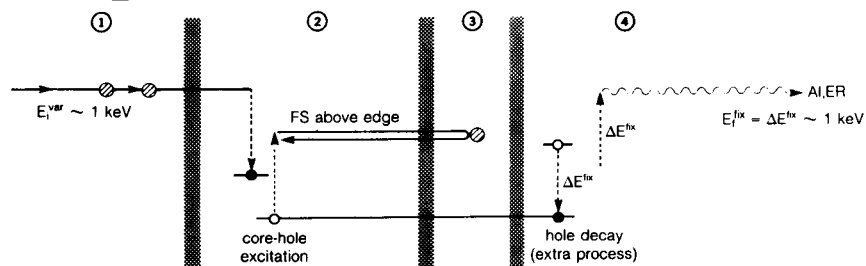
c. ARPES



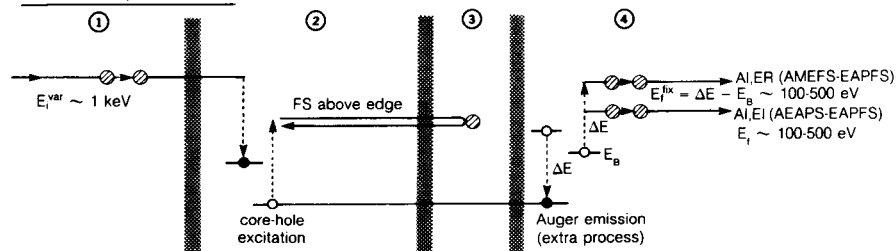
## d. ARXPD/ARPEFS



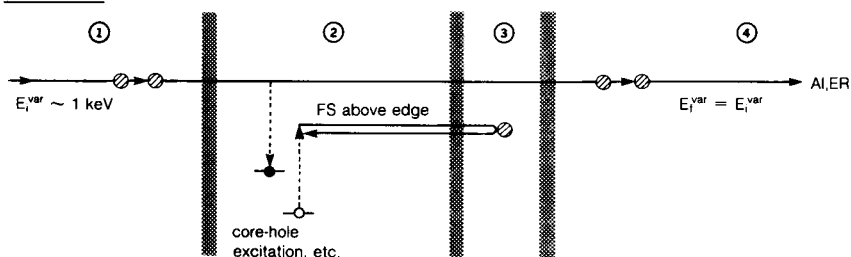
## e. SXAPS-EAPFS



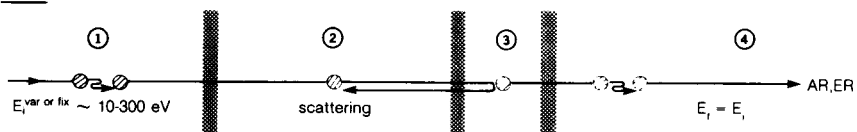
## f. AMEFS-EAPFS/AEAPS-EAPFS



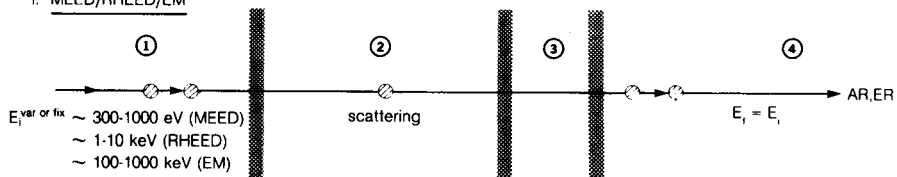
## g. DAPS-EAPFS



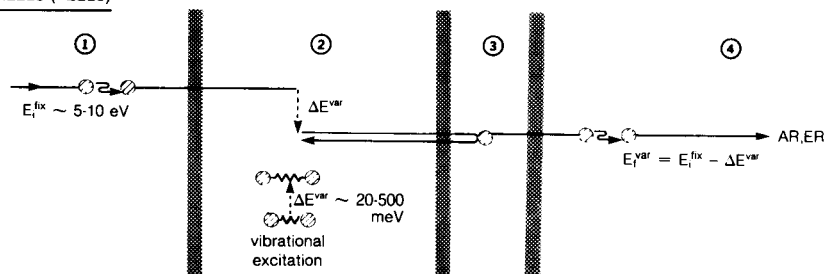
## h. LEED



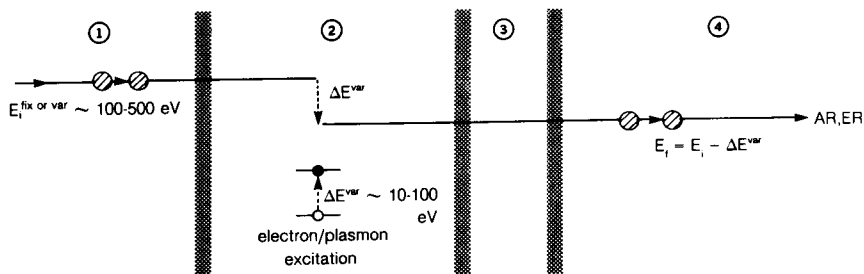
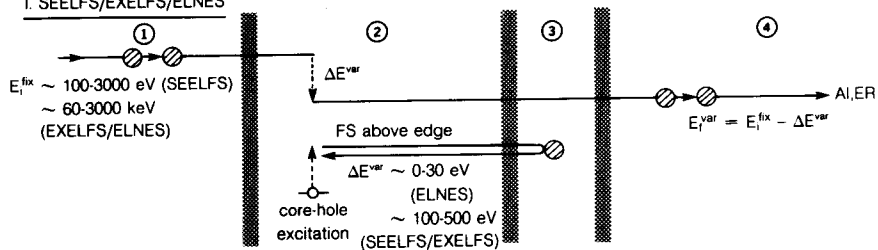
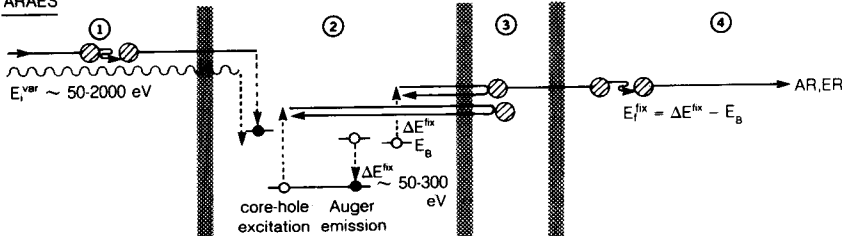
## i. MEED/RHEED/EM



## j. HREELS (=EELS)





k. ELSl. SEELFS/EXELFS/ELNESm. ARAES

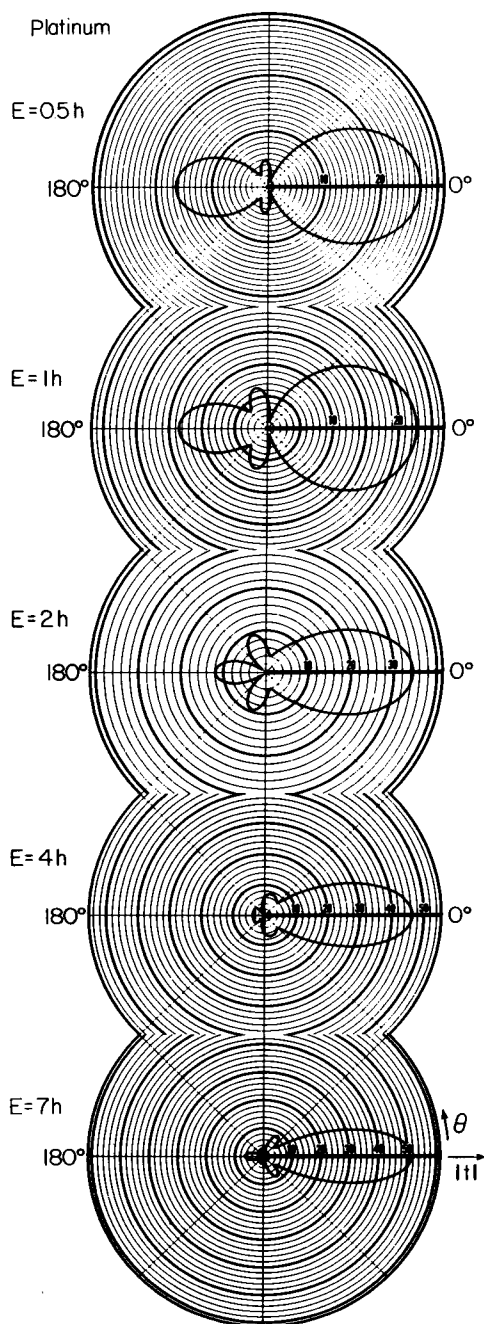
In several techniques an additional process associated with one of the four steps generates the detected particles, e.g. Auger emission in Auger-SEXAFS or ion emission in PSD-SEXAFS. This additional process does not normally provide structure sensitivity, but produces particles that are more conveniently detected. We shall include such processes in the discussion of step 4.

In any of the steps 1, 3 and 4, multiple scattering of electrons can play an important role in the techniques of interest here. On the one hand, multiple scattering provides a greater sensitivity to various aspects of structure; on the other hand, it complicates the theoretical treatment, i.e. it complicates the extraction of the structural information. For the following discussion it will be useful to first describe in more detail a couple of important features of electron scattering by individual atoms in a surface.

At low kinetic energies (10 eV), electrons are scattered strongly but somewhat isotropically by each atom (Figure 2), generating multiple scattering in all directions. As the kinetic energy rises, back-scattering weakens considerably, while forward-scattering remains strong, and multiple scattering remains important only for small scattering angles. This effect is already noticeable around 100 eV and becomes more and more pronounced towards KeV and MeV energies as scattering concentrates in an ever narrower forward cone. This gives the appearance of a more kinematic behavior, i.e. multiple scattering effects becomes less obvious. The reason is that only small momentum transfers occur, which also provides little structure sensitivity. Thus, with rising energy, little back-scattering and therefore, little back-and-forth scattering between any pair of atoms can occur, although multiple forward-scattering along a chain of atoms can be significant. As we shall describe more explicitly in section 3.5.1 one often disregards multiple forward-scattering and talks of a kinematic or single-scattering situation.

### 3.3. Theoretical Aspects of Individual Techniques

We shall now individually discuss the various techniques in terms of the above four-step description. In all these techniques the flux of outgoing particles is measured as a function of the energy (or sometimes direction) of either the



incident or outgoing particles. The flux (intensity) exhibits variations as a function of energy (or angle), from which the surface structure is extracted. In some cases, the outgoing particles are not energy-analyzed but measured collectively. This is done to enhance weak signal-to-noise ratios, as in SEXAFS.

### 3.3.1. EXAFS

Extended X-ray absorption fine structure (EXAFS) was the first technique in which "fine structure" was used to obtain structural information./39,35,40/ An incoming photon, penetrating the surface in step 1, undergoes an absorption process creating a core-hole excitation at an atomic site as the primary event (step 2), cf. Figure 1a. This is accompanied by the photo-emission of electrons which back-scatter from near-neighbor atoms as the secondary event (step 3). Interference by the back-scattered electrons modulates (i.e. produces fine structure in) the overall absorption cross-section as their kinetic energy varies in step with the incident photon energy. The fine structure can be observed in the loss of intensity from the incident photon beam ("absorption EXAFS") or instead in the intensity of fluorescence as the core-holes decay ("fluorescence EXAFS"). EXAFS is not intrinsically surface sensitive, and provides surface information only for high-surface-area materials, such as exfoliated graphite or highly dispersed particles, or when the signal comes from atoms which are only present at the surface,/39/ or when total reflection provides surface selectivity.

### 3.3.2. SEXAFS

To obtain surface sensitivity with EXAFS, thereby creating surface extended x-ray absorption fine structure (SEXAFS), one may monitor instead the outgoing Auger electrons (Auger-SEXAFS), outgoing ions ejected during core-hole decay (PSD-SEXAFS), or outgoing photoelectrons (PE-SEXAFS) instead of a photon signal. Surface sensitivity comes from the short escape depths of electrons and ions./31,41,42,43/ The electron kinetic energies involved in step 3 are sufficiently high ( $> 50$  eV) that single back-scattering prevails over multiple back-scattering. Coupled with angle-integrated detection, this allows a simple Fourier transform

approach to obtain near-neighbor distances (although theoretical or empirical atomic scattering phase-shifts need to be included). Polarization of the incoming photons can be used to obtain additional information about bond directions if valence levels are involved, since the excitation cross-sections include simple selection rules that depend on bond orientations. Chemical selectivity is provided by monitoring only Auger electrons or photoelectrons due to a particular chemical element. One may selectively investigate the structure in the immediate neighborhood of atoms of a particular element. The outgoing particles are normally energy-integrated to provide a measurable signal.

SEXAFS-type fine-structure can be obtained with photo-electrons in the low-energy region ( $\sim 100$ -300 eV). To that end, one must reduce the effect of multiple scattering in step 4 until the effects of step 3 dominate. This can be accomplished with polycrystalline surfaces and with angle-integrated detection, as has been done recently in PE-SEXAFS./44/

### 3.3.3. NEXAFS and XANES

If in step 3 of EXAFS or SEXAFS, the electron kinetic energy is held within about 50 eV of the excitation edge, one reaches the "near-edge" regime where multiple back-scattering becomes important in the secondary process cf. Figure 1b. This technique is referred to as near-edge x-ray absorption fine structure (NEXAFS) or x-ray absorption near-edge structure (XANES)./45,46,47,48/ Compared to EXAFS and SEXAFS, this multiple scattering gives more complete sensitivity to bond lengths and bond angles in the neighborhood of the atom where the primary process takes place. But model calculations are generally needed to extract this information. As with (S)EXAFS, polarization of the incoming photon can also give direct access to bond directions. Furthermore, combining an empirical correlation between carbon-carbon bond lengths and resonance energies, it appears to be possible to obtain these bond lengths./48/

### 3.3.4. ARPES, ARUPS and ARXPS

In angle-resolved photoelectron emission spectroscopy (ARPES), incoming photons give rise to the primary event of photoemission (step 2), cf. Figure 1c. This can involve excitation from delocalized electron orbitals (as in angle-resolved ultraviolet photoemission spectroscopy (ARUPS)/49,50/ from valence levels) or from deeper-lying localized orbitals (as in angle-resolved x-ray photoemission (ARXPS) from atomic core-levels)./51/ If the photoemitted electrons have kinetic energies in the range 10-200 eV, they undergo multiple scattering (similar to that encountered in LEED) in the third and fourth steps, of the four-step process. This yields sensitivity to the surface within a photo-electron mean free path length of the photoexcited atom, which can be chosen with chemical selectivity. Compared to ARUPS, ARXPS is a more convenient approach for the extraction of structural information, since core-level excitations are more easily described than valence-level excitations. Angle-resolved detection is normally used to obtain the most complete structural information, which is interpreted in a procedure very similar to that of LEED with full simulation of the electron emission intensity as a function of energy or angle.

### 3.3.5. ARXPS and ARXPEFS

If, starting with angle-resolved x-ray photoelectron spectroscopy (ARXPS),/51/ the emitted electron kinetic energy is increased to a few hundred electron volts, the back-scattering of photoemitted electrons to the source atom is considerably reduced, cf. Figure 1b and d. Then the secondary event of step 3 becomes negligible compared to the events of step 4. At the same time, step 4 simplifies to include predominantly single scattering by neighboring atoms, (for convenience, we neglect the multiple forward-scattering in this description, since it yields little structure sensitivity). In this case, electrons reach the detector either directly from the photoemitting atom or after single scattering by atoms that are near the photoemitting atom (again, x-rays, as opposed to ultraviolet radiation, provide the most convenient core-level emission). The interference

between these different outgoing electron paths modulates the detected intensity as a function of the angle of detection: this measurement mode is used in ARXPD.

Similarly, "fine-structure" arises as a function of the emitted-electron kinetic energy in the measurement of angle resolved photoemission extended fine structure (ARPEFS).<sup>/37,36,38/</sup> This "fine-structure" is different from that of SEXAFS, EXAFS, etc., since it does not originate from back-reflection to and interference at the electron source atom. It yields the surface structure information relatively easily through Fourier-transformation if angle-resolved detection is used (suitable scattering phase-shifts must be available, however). The full three-dimensional structural information can be so obtained. Sensitivity to a particular structural feature may be emphasized through appropriate selection of the detection angle and and the polarization of the incident photons.

### 3.3.6. Extended Appearance Potential Fine Structure (EAPFS)

In EAPFS,<sup>/31/</sup> incident electrons excite core-level electrons into states above the Fermi level  $E_F$  in the primary process cf. Figure 1e-g. This requires a minimum incident energy for any given core-level, namely  $E_F - E_1$ , where  $E_1$  is the core-level energy (hence the designation "Appearance Potential"). The choice of core-level provides chemical selectivity. The excited electron can then undergo back-reflection from neighboring atoms as the secondary step. This yields fine structure as a function of the incident energy that can be used to obtain structural information by Fourier transformation. Since the incident electrons can undergo multiple scattering prior to the core-level excitation, amorphous materials are preferred to minimize any disturbing LEED-like modulations due to such scattering.

There are three main detection modes for EAPFS within the appearance potential spectroscopy (APS) technique.<sup>/31/</sup> First, one may monitor soft-x-ray emission due to the decay of the core hole left by the primary process. This is called SXAPS-EAPFS (Figure 1e). Second, it is also possible to monitor Auger electrons due to the same core-hole decay, as in AEAPS-EAPFS and AMEFS-EAPFS, cf. Figure 1f. Third, one may measure the remaining total intensity of

the elastically scattered electrons in DAPS-EAPFS (DA = Dis-Appearance) cf. Figure 1g. Angle-integrated detection removes most of the structural sensitivity of step 4 (especially with amorphous materials). Normally one collects all emitted electrons in these experiments, i.e. energy-integration is used. However, in the Auger-Monitored Extended Fine Structure version (AMEFS-EAPES)/52/ energy resolution is used, which achieves chemical sensitivity at the cost of a lower signal/noise ratio, cf. Figure 1f.

### 3.3.7. LEED, MEED, RHEED and Electron Microscopy (EM)

In low-energy electron diffraction (LEED) only elastic scattering events are considered, i.e., no energy loss is allowed./27,28,29/ This situation is sometimes referred to as ELEED (Elastic LEED). In LEED, the different steps of the general four-step process are equivalent to each other, in the sense that each step describes possible elastic electron-scattering events in a multiple-scattering chain (which of these scattering events is termed the primary event becomes arbitrary). Figure 1h schematically shows possible scatterings in LEED. The four-step description makes vividly clear that multiple scattering plays a large role in LEED, especially when compared to other techniques described here. This multiple scattering, together with angle-resolved detection, gives LEED a high sensitivity to all details of the surface structure: layer spacings, bond lengths and bond angles throughout the surface to a depth determined by the energy-dependent electron mean free path. To obtain structural information from LEED, the multiple scattering must be simulated on a computer for each plausible geometry until the correct geometry is found by trial-and-error.

At increasingly higher energies, LEED becomes in turn medium-energy electron diffraction (MEED),/53,54,55/ reflection high-energy electron diffraction (RHEED)/56/ and electron microscopy (EM)/57,58/ cf. Figure 1i. Multiple scattering remains important, although the mean distance between successive scatterings increases gradually towards hundreds of Angstroms. On the other hand, to an increasing degree the only significant scattering events at these energies occur near the forward direction. A perfectly ordered surface yields sharp diffracted electron beams, unlike the other techniques which we describe in



this chapter. Any kind of disorder or defects generate diffuse intensities between the sharp diffraction beams. The diffuse intensity contains information not only about deviations from long-range order, but also about the short-range order or local geometry, including chemical bond lengths and bond angles.

### 3.3.8. ILEED, HREELS and ELS

If at least one of the scattering events in a LEED multiple-scattering chain involves an energy loss, one speaks of inelastic low-energy electron diffraction (ILEED),/59,60/ (multiple losses are possible but infrequent). The lost energy can be transferred to surface vibrations (e.g. molecular vibrations or phonons), as is the case in high-resolution electron loss spectroscopy (HREELS)/61,62,63/ or to electronic excitations (e.g. single-electron excitations or plasmons), as is the case in electron loss spectroscopy (ELS)./64,65/ In ILEED, the primary event of the four-step process is the inelastic electron scattering. It is not necessarily localized at an atomic site (as illustrated for convenience in Figure 1j,k) instead, it is quite often delocalized over many atoms, as with dipole scattering in HREELS or plasmon excitation ELS. In HREELS cf. Figure 1j, multiple scattering is at least as important as it is in LEED, and angle-resolved detection is also used. HREELS can therefore yield potentially the same structural information as LEED, e.g. when intensities are measured as a function of incident energy, at the cost of having to describe an additional inelastic process in the theory./63/ This approach may be called IV-HREELS, since LEED-like IV curves are used. HREELS is however mostly used in a mode where the multiple scattering effects are virtually constant, namely by scanning the energy loss over a small range at fixed incident energy./61,62/ Then all that matters is the occurrence of sharp energy losses at values corresponding to frequencies of surface vibration. This gives a loss spectrum similar to that familiar in infra-red adsorption spectroscopy (IRAS) in the gas phase and at surfaces. This is also the approach with ELS, cf. Figure 1k, although here the wide scan of energy loss allows multiple scattering to significantly modulate the heights of the peaks corresponding to individual excitations. Under these circumstances the structural information is obtained indirectly through the measured vibrational and excitational properties of the surface, rather than through the diffraction effects.

### 3.3.9. Surface Extended Energy Loss Fine Structure (SEELFS)

Another electron-based energy-loss technique is SEELFS./32,33/ Compared to HREELS and ELS, higher electron energies are used, so that a core-hole excitation-edge and the associated fine structure become available, cf. Figure 11. There is a great similarity between SEELFS and EAPFS. However, in SEELFS the energy-loss is kept small compared to the incident energy, thus, one has a single-electron final-state problem in SEELFS rather than a two-electron final-state problem as in EAPFS. By keeping the incident energy fixed in SEELFS and by angle-integrating the detection, variable diffraction effects in steps 1 and 4 are minimized. This yields fine structure that can be Fourier-analyzed to yield bond lengths around chemically-selected atoms. Complication can arise however, when several spherical waves with different angular momenta are needed to describe the electrons of step 3. This produces in effect several overlapping fine structures, which may not be easy to isolate. Nevertheless, experimental evidence suggests that normally one of the angular momentum components is sufficiently dominant to avoid these complications.

### 3.3.10. Extended X-Ray Energy Loss Fine Structure (EXELFS)

By substantially increasing the electron kinetic energy in ILEED and SEELFS to 60-300 KeV, one reaches a situation that can resemble EXAFS, with electrons replacing the EXAFS photons, cf. Figure 11. Chemical selectivity. is obtained by choosing an appropriate core excitation edge. And surface sensitivity on the atomic scale exists mainly for high-surface-area materials, as with EXAFS.

### 3.3.11. Electron Energy Loss Near-Edge Structure (ELNES)

ELNES/66/ is the near-edge equivalent of EXELFS, with the same high-energy (60-300 KeV) electrons as incident particles. It focuses on final energies within about 30 eV of an absorption edge, so that the fine structure arising in step 3 behaves very much as in NEXAFS, cf. Figure 11. Here multiple scattering

is important, giving full structural information in the immediate vicinity of a selected chemical element. However, there is no simplifying dipole selection rule in the primary process, since electrons replace the photons.

### 3.3.12. Angle-Resolved Auger Electron Spectroscopy (ARAES)

One may directly monitor Auger electrons emitted following core-hole excitation induced by an incident electron or photon beam, using angle resolved detection, cf. Figure 1m./67,68/ Thus the core-hole excitation together with the Auger process constitutes our step 2. At the lower energies used in ARAES ( $\sim 20$ -200 eV), multiple scattering is often important as in LEED, especially in step 4. Thus step 3, as in LEED, is not distinguished by convenient fine-structure, but is much more complex due to multiple scattering. In addition, to obtain structural information, the complicated atomic Auger process must be modeled accurately. This has only been done in a few special cases involving atomic orbitals (as opposed to solid state orbitals)./69,70/ As a result, the Auger transition probabilities into the various outgoing spherical waves in step 2 are generally unknown. So far these probabilities have been fit to experiment, with limited success in terms of structural determination. The orientation of adsorbed molecules appears to be accessible with the photon-excited version of this technique, by using photon polarization and symmetry selection rules./71/ Otherwise the angle-resolved detection of ARAES can yield the same relatively complete three-dimensional structural information as does LEED.

### 3.4. Formalism for Individual Processes

In section 2.3 we have decomposed various surface-sensitive techniques into a small set of more elementary processes. In this section we shall present appropriate formalisms that describe these individual processes. Namely, we shall discuss the theoretical treatment of the propagation of electrons in the surface region, as well as the treatment of elastic and inelastic electron-atom scattering, photoelectron emission and of Auger electron emission. These parts

can then be easily recombined to provide a complete theoretical description of each surface-sensitive technique. Some general features of particular interest will be highlighted in section 3.5.

#### 3.4.1. Overview of process formalisms

We shall deal mainly with the propagation and scattering of electrons through solid surfaces. The most appropriate formalism to describe these processes will depend very much on the situation, as computational efficiency is at a premium. Spherical-wave and plane-wave expansions are frequently used, often side by side within the same theory. Spherical waves are convenient to describe the scattering or emission of electrons by an individual atom. An important quantity will then be the complex amplitude of each spherical wave leaving that atom.

Plane waves are often used when two conditions are satisfied: 1) many (but not necessarily all) atomic layers of the surface have a two-dimensional periodicity, and 2) either a plane-wave incident electron beam is present or angle-resolved electron detection is applied. Computation based on the plane-wave expansion are often much more efficient than those based on the spherical-wave expansion. This explains their frequent use even in problems that do not involve strict two-dimensional periodicity, as with disordered overlayers on an otherwise periodic substrate.

The fate of scattered or emitted electron waves depends on the occurrence of multiple scattering from atom to atom. With strong atomic scattering, a large number of scattering paths must be considered: this applies especially at low energies and with angle-resolved detection, e.g. in LEED and ARPES. At higher energies fewer scattering paths are important, because the atomic scattering is relatively weaker (except for forward-scattering). Paths involving mostly forward-scattering dominate in angle-resolved techniques such as ARPEFS, MEED, RHEED, and EM. By contrast, single back-scattering to the emitting atom dominates in angle-integrated fine structure techniques such as (S)EXAFS, EAPFS and SEELFS. The individual formalisms strongly reflect these different circumstances.

### 3.4.2. Electron propagation

The interference effects resulting from electron propagation are responsible for providing the desired structural information in all the techniques discussed here. The extraction of structural information, therefore, requires a knowledge of the electron wavelengths and of any phase shifts that may occur in electron emission and electron-atom scattering. Failure to understand these processes can result in quite erroneous results.

The well-known muffin-tin model of the electron atom interaction potential has in most cases proved to be an adequate compromise between accuracy and computational efficiency, at least for electron kinetic energies exceeding  $\sim 20$  eV./27,28,29/ The muffin-tin potential is spherical inside the muffin-tin spheres. This "ion core" usually provides the dominant contribution to electron scattering and emission, to be discussed shortly.

The interstitial region between these spheres is represented by a constant potential value (the muffin-tin constant or muffin-tin zero). This is the region of unscattered electron propagation, in which simple plane-wave or spherical-wave behavior describes the electron wave-field. The plane-wave description is computationally advantageous when dealing with multiple scattering by periodic lattices, since such lattices diffract any plane wave into other well-defined plane waves. This approach is often useful even with non-periodic overlayers adsorbed on an otherwise periodic substrate.

The spherical-wave description has its own great advantages. It is best adapted to the scattering and emission by the spherical ion cores and it does not require the presence of any structural periodicity. In particular, it is well suited to the treatment of multiple scattering between different atoms within any cluster of atoms, in particular within a periodic unit cell as in LEED. It is also convenient for the treatment of fine structure arising from back-scattering by nearby atoms, as in (S)EXAFS, NEXAFS, EAPFS, etc. (i.e. in step 3 in our four-step description).

For energies lower than  $\sim 20$  eV, the muffin-tin model is often inadequate outside the muffin-tin spheres. In the case of surfaces, this applies especially to low-energy studies of adsorbed molecules. For instance, HREELS is normally performed at kinetic energies around 5-10 eV (to benefit from a strong reflection

coefficient from the substrate). HREELS is therefore highly sensitive to the potential in the molecular bonding regions and couples strongly with weakly bound and relatively delocalized electron states. Also, NEXAFS is often measured down to energies of about 10 eV above an emission edge, giving similar sensitivity to bonding details. No satisfactory alternative treatment to the muffin-tin model has yet been applied in the surface context, mainly because excessive computational efforts would be involved (however this problem has been dealt with in the case of low-energy electron scattering by gas-phase molecules/72/ ).

### 3.4.3. Single electron-atom scattering

Electron-atom scattering is central to all techniques under discussion here. As we have mentioned, electron-atom scattering at surfaces has been treated almost exclusively by means of the muffin-tin model.

The formalism of electron-atom scattering has been extensively dealt with elsewhere./27,28,29/ We shall only recall its main features here. Because of the assumed spherical symmetry, the partial-wave scattering approach is convenient. Namely, an incoming spherical wave  $h_l^{(2)}(kr) Y_l^m(\mathbf{r})$ , ( $-l \leq m \leq l$ ) can scatter only into the outgoing spherical wave  $h_l^{(1)}(kr) Y_l^m(\mathbf{r})$  (here  $h_l^{(1)}$  and  $h_l^{(2)}$  are Hankel functions of the first and second kinds,  $k = 2\pi/\hbar(2mE)^{1/2}$ ,  $E$  is the kinetic energy and  $\mathbf{r} = |\mathbf{r}|$ ). This occurs with amplitude  $t_l$  ( $t_l$  is an element of the diagonal atomic  $t$ -matrix), which is related to the phase shifts  $\delta_l$  through

$$t_l = -\frac{\hbar^2}{8\pi m} \frac{1}{k} \sin \delta_l e^{i\delta_l} \quad (1)$$

The phase shifts  $\delta_l$  are calculated by standard partial-wave scattering theory. It involves the electron-atom interaction potential of the muffin-tin model. There are a variety of ways to obtain this potential, which consists of electrostatic and exchange parts (spin dependence may be included, especially when the spin polarization of the outgoing electrons is of interest). One usually starts from known atomic wave functions within one muffin-tin sphere and spherically averages contributions to the total charge density or potential from nearby

atoms. The exchange term may be treated as non-local (e.g. Hartree-Fock) or local (e.g.  $X_\alpha$ ). The muffin-tin constant with respect to the vacuum zero level is a by product of this approach, however in practice this constant is normally adjusted *a posteriori* to take into account surface-related work function changes.

The  $t$ -matrix of Eq. (1) will be useful in describing the scattering of a *plane* wave by an atom. Expansion of the plane wave  $e^{i\mathbf{k}\cdot\mathbf{r}}$ , with  $|\mathbf{k}| = k$ , into spherical waves yields the total scattered wave

$$e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta) \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{r} \quad (2)$$

at large values of  $kr$ . The atomic scattering amplitude  $f(\theta)$  depends on the scattering angle  $\theta$  through the well-known relation

$$f(\theta) = -4\pi \sum_l (2l+1) t_l P_l(\cos\theta) \quad (3)$$

where the functions  $P_l(\cos\theta)$  are Legendre polynomials. The scattering amplitude  $f(\theta)$  can be used to illustrate the nature of electron scattering in solids. A large scattering amplitude implies strong multiple scattering. To discuss this issue, we need to first include the effect of atomic thermal vibrations. It has been found adequate for this purpose to replace the scattering amplitude  $f(\theta)$  by a statistical average over certain atoms, even in the presence of multiple scattering. In the case of diffraction by a periodic lattice, the average is performed over atoms that are equivalent by two-dimensional periodicity, and in the case of back-scattering by shells of vibrating neighbors, around equivalent source atoms. In either case, one ends up multiplying  $f(\theta)$  by a Debye-Waller attenuation factor  $e^{-M}$ , with

$$M = \frac{1}{2} |\mathbf{s}|^2 \langle (\Delta\mathbf{r})^2 \rangle \quad (4)$$

where  $\mathbf{s}$  is the linear momentum transfer and  $\langle (\Delta\mathbf{r})^2 \rangle$  is the mean-square atomic displacement.

Note that the product  $e^{-M} f(\theta)$  can be expressed in terms of effective phase shifts in direct analogy with Eqs. (1) and (3), then  $\delta_l$  is replaced by an effective

$\delta_i(T)$  in those relations. These so-called “temperature-dependent” phase shifts  $\delta_i(T)$  are commonly used in multiple scattering theory, because the effect of thermal vibrations is felt again and again at each scattering in a chain of scatterings.

The product  $e^{-M} f(\theta)$ , shown as  $|f(\theta)|$  in Figure 2, has a marked lobe structure in its angular dependence. There is always a pronounced lobe of strong scattering amplitude in the forward-scattering direction ( $\theta = 0$ ). This lobe becomes narrower (more forward-focused) with increasing kinetic energy, becoming quite narrow (a few degrees) at energies above 1000 eV. High temperatures also favor a narrow forward scattering lobe since the Debye-Waller factor attenuates forward-scattering less than large angle scattering. In addition, there is usually a lobe centered on the back-scattering direction ( $\theta = \pi$ ). At low energies ( $\sim 10$  eV) this lobe is wide and comparable in strength to the forward lobe, yielding nearly isotropic scattering. But at higher energies the back-scattering lobe rapidly loses strength, while also becoming narrower. Finally there are one or more side-scattering conical lobes which are usually of a strength comparable to or smaller than the back-scattering lobe. These become more numerous and narrower with increasing energy. Between lobes there may be directions with vanishing scattering amplitude, when accidental cancellation of partial scattered waves occurs.

#### 3.4.4. Multiple electron scattering

We have just described electron scattering by a single atom. The main feature to remember is the pronounced forward scattering within a more or less narrow forward cone, which occurs especially at the higher energies. The implications for multiple scattering are as follows.

At low energies ( $\sim 10$  eV) the scattering is relatively isotropic and relatively strong, so that scattering in any direction is likely to be followed by scattering from another atom that finds itself in the electron path. In fact, several successive scatterings are quite possible, especially with heavy (high- $Z$ ) atoms, which scatter most strongly. This includes the possibility of repeated back-and-forth scattering between two or more atoms, which is, when the interference is



constructive, similar to a negative-ion resonance in a molecule. With such strong multiple scattering, a perturbation expansion in terms of the number of scatterings will often fail to converge and a self-consistent solution is then necessary.

At higher energies, back-scattering and side-scattering become increasingly less probable, leaving mainly forward-scattering, which itself remain quite strong. Even for very high energies, e.g. in HEED or EM at  $E \sim 10\text{-}1000$  KeV, multiple scattering is still important. But now all scattering paths remain bundled within a few degrees of the incidence direction.

These facts explain many differences between surface-sensitive techniques. Those techniques which use low kinetic energies, such as LEED, ELS, HREELS, ARAES, ARPES and NEXAFS, require a rather complete multiple scattering treatment. Some perturbation theories do apply to them but close attention to multiple forward-scattering as well as to several back-scattering events remain essential. Techniques using higher energies split up in two groups. The fine-structure techniques, such as (S)EXAFS, EAPFS, SEELFS and EXELFS, rely almost exclusively on the (weak) back-scattering from neighboring atoms to the emitting atom. Forward-scattering, though strong, is largely irrelevant in this situation (this is enhanced by angle-integrated detection, which averages over many exit directions). Other high-energy techniques are much more sensitive to multiple forward scattering events, e.g. MEED, HEED, EM and ARPEFS (partly because of their angle-resolved detection mode in the forward direction).

We shall next describe several formalisms used to calculate the effects of multiple scattering. We start with the most accurate formalism in common usage and then introduce simplifications and specializations.

Within the muffin-tin model of spherical scatterers, all multiple scattering can be properly included with the Green function formalism. We already have given the  $t$ -matrix for scattering by a single atom. It is now necessary to describe the propagation of the scattered spherical waves in the "near-field" region (without asymptotic forms). Consider the propagation of a wave of angular momentum quantum number  $L \equiv (l, m)$  from an atom at location  $\mathbf{r}_i$  to an atom at location  $\mathbf{r}_j$ . This wave can be expressed in terms of spherical waves  $L' \equiv (l', m')$  incident on atom  $j$  with amplitudes given by the Green function

$$G_{LL'}^{ij} = -4\pi i \frac{8m}{h^2} k \sum_{L_1} i^{l'_1} a(L, L', L_1) h_{l_1}^{(1)}(k | \mathbf{r}_i - \mathbf{r}_j |) Y_{L_1}(\mathbf{r}_i - \mathbf{r}_j) \quad (5)$$

Here  $k = |\mathbf{k}|$  and  $L_1 = (l_1, m_1)$  extends over all values of  $l_1$  and  $m_1$  compatible with  $L$  and  $L'$ , namely such that  $|l - l'| \leq l_1 \leq l + l'$  and  $m + m' = M_1$ . This compatibility is manifest as non-zero values of the Clebsch-Gordon (or Gaunt) coefficients

$$a(L, L', L_1) = \int_{\Omega} Y_L^*(\Omega) Y_{L'}(\Omega) Y_{L_1}^*(\Omega) d\Omega \quad (6)$$

With the  $t$ -matrix and the Green function we can represent any multiple-scattering path explicitly. For instance, scattering by atom 1 to atom 2, then back to atom 1 and on to final scattering by atom 3 yields a scattering amplitude of

$$t^3 G^{31} t^1 G^{12} t^2 G^{21} t^1 \quad (7)$$

where one should read from right to left (in the fashion of matrix multiplication) and where the superscripts on the  $t$ -matrices allow the different atoms to have different scattering properties. When strong multiple scattering occurs, the contributions from all such scattering paths should be included. This can be done self-consistently by replacing the infinite series due to longer and longer scattering paths by a matrix inversion.<sup>/73/</sup> The result is the following expression, for diffraction by a set of  $N$  atoms labeled  $i = 1, 2, \dots, N$ :

$$\begin{pmatrix} T^1 \\ T^2 \\ \vdots \\ T^N \end{pmatrix} = \begin{pmatrix} I & -t^1 G^{12} & \dots & -t^1 G^{1N} \\ -t^2 G^{21} & I & \dots & -t^2 G^{2N} \\ \vdots & \vdots & \ddots & \vdots \\ -t^N G^{N1} & -t^N G^{N2} & \dots & I \end{pmatrix}^{-1} \begin{pmatrix} t^1 \\ t^2 \\ \vdots \\ t^N \end{pmatrix} \quad (8)$$

Here  $t^i$ ,  $T^i$ ,  $G^{ij}$  and  $I$  are matrices indexed by  $L = (l, m)$  and  $L' = (l', m')$ ;  $I$  is the unit matrix. A multicenter expansion is used, i.e. the spherical waves are centered on the respective atoms. The quantity  $T_{LL'}^i$  is the total amplitude of the

spherical wave  $Y_L$  that leaves atom  $i$  after spherical waves  $Y_{L'}$  have initially impinged on every atom of the cluster and multiple scattering within the cluster has been taken into account. The scattered wave accumulated over all terminal atoms  $i$  is obtained by simple addition of the  $T^i$  values with proper consideration of the relative amplitudes and phases of the initial wave incident on the cluster.

The dimension of the large matrix to be inverted in Eq. (8) is  $N(l_{\max}+1)^2$ , where  $l_{\max}$  is the largest value of  $l$  included. This dimension critically affects the required computing effort. The number of spherical waves used,  $(l_{\max}+1)^2$  is typically 16 to 64 at LEED energies ( $l_{\max}$  is  $\sim$  inversely proportional to the electron wavelength). The value of  $N$  is in general the number of atoms in the cluster needed to adequately account for all multiple paths involving any given surface atom. At 20 eV this might be several hundred atoms, as defined by a sphere around the given atom and whose radius is a few times the electron mean free path. Thus, matrix dimensions in Eq. (8) of the order of  $10^4$  would be required within this formalism. This is the situation for NEXAFS, ELNES and diffuse LEED. A similar single-center expansion approach has also been formulated for this problem./74,75/ It involves a new "wallpapering" scheme that bypasses the larger matrix inversion of Eq. (8); instead it builds up the cluster scattering properties as seen from a central atom in a shell-by-shell iteration, each step of which involves the inversion of a smaller matrix. However, the single-center expansion gives rise to much larger values of  $l_{\max}$  (proportional now to the radius of the cluster instead of the muffin-tin radius) which still results in large computation times.

Simplifications can be brought about whenever the surface structure has symmetries. Point-group symmetries help moderately to reduce the matrix dimensions. On the other hand, two-dimensional periodicity can help drastically by reducing the number  $N$  to the number of atoms within a single two-dimensional unit cell with a depth perpendicular to the surface of a few times the electron mean free path. For surface crystallography this is, however, not yet sufficient, because surface structural determination requires repeating such calculations for hundreds of different geometrical models of the surface structure.

Great computational advantage can ensue from using the plane-wave expansion between atomic layers. This occurs when such a representation converges, which requires sufficiently large spacings between the layers. Various

calculational methods based on plane waves have been developed over the years, such as the Bloch-wave, Layer Doubling and Renormalized Forward Scattering methods. These have been extensively described previously in the context of LEED,<sup>/27,28,29/</sup> but their use in other techniques merits further discussion here. In conventional LEED, the two-dimensional surface periodicity clearly defines a finite set of outgoing plane waves (corresponding to the discrete diffracted beams defined by the two-dimensional reciprocal lattice). Most other techniques do not share this feature -- no sharp electron beams emerge from the surface in photoemission, SEXAFS, diffuse LEED, etc. However, the use of angle-resolved detection is equivalent to defining a finite set of plane waves. This situation has been described as inverse LEED or time-reversed LEED.<sup>/76/</sup> The detector essentially singles out an emerging plane wave, which, when time is reversed, acts like a plane wave incident on the surface. This plane wave scatters through the periodic layers as does normal LEED, via the set of plane waves familiar in LEED, until they reach the original source of electrons. The source can be a photoemitting atom, for instance. The calculation need not proceed in this time-reversed manner, however, given the source of emission in the form of amplitudes of outgoing partial waves (which can be corrected for intra-layer multiple scattering), one can deduce the resulting amplitudes of the above-mentioned set of plane waves as they leave the source. From then on multiple scattering through the lattice proceeds as in conventional LEED using the finite set of outgoing plane waves defined by the direction of detection.

Some techniques also involve a well-defined *incident* electron beam, even though the primary process at some point imparts an arbitrary parallel momentum to the electrons. This happens, for example, with energy loss in HREELS and ILEED, with diffuse scattering in LEED and with Auger emission in ARAES. In these cases the direction of the electrons leaving the surface has an arbitrary relationship to the incident beam direction. Up to the primary process, however, conventional LEED can be applied in the plane-wave representation, at least in the ordered part of the surface, using the finite set of plane waves defined by the direction of incidence.

### 3.4.5. Approximations in multiple scattering

We shall now describe a series of approximations of increasing simplicity which have been developed to reduce computational effort required to describe multiple scattering. These approximations are usually tailored to specific circumstances and are used only in those particular circumstances in order to maintain the highest possible accuracy.

We have mentioned the preponderance of forward scattering of electrons by atoms. Perhaps the earliest method used to exploit this feature was Renormalized Forward Scattering in LEED.<sup>/27/</sup> This plane-wave method is a perturbation expansion in terms of the number of back-scatterings between atomic layers. It expressly does not count forward scatterings in the expansion. Instead, all forward scattering events that may occur between successive back-scattering events are explicitly included. This approach is very economical and converges well except at the lowest energies ( $\sim 10$  eV) where multiple scattering is too strong (as with all plane-wave methods, a sufficiently large interlayer spacing is also required). More recently this idea has been applied to conventional LEED within the spherical-wave representation under the name Reverse Scattering Perturbation.<sup>/77/</sup> Most recently, this idea has been adapted to a shell-by-shell treatment (as opposed to a layer-by-layer treatment),<sup>/78/</sup> similar to the “wallpapering” scheme mentioned above. One then considers spherical waves leaving an atom or cluster and reflecting back from shells of neighboring atoms, including multiple scattering. This approach is particularly useful when long-range order is absent or unimportant, as in NEXAFS<sup>/78/</sup> and in one part of a diffuse LEED theory.<sup>/79/</sup>

This diffuse LEED theory consists of three parts:

- 1) multiple scattering of the incident electrons through the ordered region of the surface, yielding to a conventional LEED treatment;
- 2) scattering by a defect site, e.g. an isolated disordered adsorbate, including multiple scattering from shells of neighboring atoms back to the defect site;
- 3) multiple scattering on the way out from the defect site to the detector through the ordered region of the surface, which corresponds to the inverse LEED problem.

Another method based on the number of back-scatterings by atomic layers is Beam Set Neglect in LEED,/80/ which is particularly appropriate for adsorbates on simple substrates. This method takes advantage of the fact that some sets of diffraction beams only affect the intensities of other beams through multiple scattering of third- or higher-order, therefore such sets of beams may be neglected. The method is very efficient both large superlattice unit cells (which create large number of beams, most of which may be neglected) and also yields an efficient solution to the diffuse LEED problem./79/

As the electron energy increases, back-scattering rapidly becomes weaker and multiple scattering occurs primarily within a narrow cone in the forward direction. This is exploited in the chain method developed for MEED and RHEED./53,54/ In these techniques, grazing incidence is normally used to obtain surface sensitivity and sufficient reflection. These factors favor scattering along chains of atoms oriented parallel to the surface in the direction defined by the incident beam. The chain method uses this as a basis to calculate MEED and RHEED intensities exactly (allowing application to LEED as well). But when back-scattering and side-scattering are very weak, they can be selectively and efficiently neglected in favor of the multiple forward-scattering that occurs within the chains of atoms.

We next turn to the high-energy fine-structure techniques: EXAFS/SEXAFS, EAPFS, SEELFS, EXELFS and ARPEFS. All but one of these involve single scattering back to a source atom, resulting in interference modulation of the primary process (Auger emission, etc.). The exception is ARPEFS, where the source atom emits photoelectrons which are scattered only once by neighboring atoms directly towards the detector (not counting multiple scattering which is insensitive to the structure). The structural information stems from the interference at the detector between the waves directly traveling from the source atom to the detector and the waves scattered emitted from the source atom and scattered into the detector by different neighboring atoms. Furthermore, angle-resolved detection is used in ARPEFS unlike the other fine-structure techniques. The common feature of all these techniques is high electron energies and a diffraction geometry which de-emphasizes forward scattering (although forward scattering effects have been observed)./47/ Under these circumstances substantial simplifications can be considered. First, single

scattering by neighboring atoms may be assumed. Second, one may ignore the finite radius of curvature of the spherical waves reaching the neighboring atoms, yielding the "plane-wave" or "small-atom" approximation. Third, the radial spherical functions may be given their asymptotic form  $\frac{e^{ik \cdot r}}{r}$ .

These simplifications lead to the following relative intensity modulations for ARPEFS:

$$\begin{aligned} \frac{(I - I_0)}{I_0} &\equiv \chi(k) \\ &= \sum_j \frac{\cos \beta_j}{\cos \gamma} \frac{|f(\pi) + f(0)|}{r_j} \cos[kr_j (1 - \cos \alpha_j) + \phi_j] \end{aligned} \quad (9)$$

Here  $k$  is the electron momentum,  $\alpha_i$  is the scattering angle at the neighboring atom situated at position  $\mathbf{r}_j$ ,  $|f(\alpha_i)|$  is the corresponding scattering amplitude and  $\phi_j$  the corresponding phase shift (i.e. the phase of  $f(\alpha_i)$ ), while  $\gamma$  is the angle between the polarization direction and the direct emission path to the detector, and  $\beta_j$  is the angle between the polarization direction and the initial path of an electron scattered from site  $j$ . (A Debye-Waller factor and an inelastic damping factor may be added to Eq. (9)). This expression may involve many terms  $j$ , one for each neighboring atom which contributes. Nevertheless the expression remains simple to analyze in many circumstances. Namely, by suitable choice of the energy, which affects  $f(\alpha_i)$ , and polarization direction, most terms can be made negligible. Then it becomes relatively easy to extract the structural information in the form of the path length differences  $r_j (1 - \cos \alpha_j)$ . With the energy scans of ARPEFS, Fourier transformation of the  $k$ -dependent data is sufficient, although some care with data manipulation is required. In particular, the phase shifts  $\phi_j$  have to be known from theory. There is also controversy about whether the plane-wave approximation is adequate in ARPEFS.

With the other fine-structure techniques (EXAFS, SEXAFS, EAPFS, etc.), the above mentioned approximations lead to the following general relationship for the intensity modulation at the detector: /41,42,43/

$$\chi(k) = \sum_j \frac{N_j}{kr_j^2} \sin(2kr_j + 2\delta_l + \phi_j) |f_j(\pi)| \quad (10)$$

Here  $\delta_l$  is the phase shift of the central atom for a partial wave of angular momentum  $l$ , while  $N_j$  is the number of identical atoms at a distance  $r_j$ . (Here again a Debye-Waller and an inelastic damping factor may be added; also a polarization dependence may be included for EXAFS and SEXAFS.) Fourier transformation can yield the distances  $r_j$  with proper care. In this case the phase shifts are usually obtained from EXAFS results for bulk materials where the phase shifts are determined from the interatomic distances known from x-ray diffraction. Alternately, the phase shifts may be computed as is done in LEED. Note that multiple forward scattering is not included in Eq. (10).

#### 3.4.6. Photoelectron emission

The first step in photoemission (and some other techniques) involves photon propagation through the surface region to the location of the electron excitation.

#### 3.4.7. Photon propagation

The propagation of electromagnetic waves in bulk solid materials and in molecules is well understood. This understanding is appropriate for bulk EXAFS, for instance, but not necessarily for surface studies. The situation of a solid surface has only been studied recently in terms of atomic-scale phenomena, as is required by techniques such as photoemission. Although photons typically penetrate to a depth of about  $1 \mu\text{m}$  into the metal surfaces of interest, the only important photon-induced excitations occur within the electron escape depth of about 5 to  $10 \text{ \AA}$ . Therefore, the shape of the electromagnetic fields should be known in this near-surface region of a few atomic layers.

It has been found that for photon energies near and below plasmon energies ( $\sim 20 \text{ eV}$ ) and near other excitations, the electromagnetic fields are quite structured, both as a function of depth from the surface and as a function of energy.<sup>81</sup> For any quantitative structural studies, this situation is to be



avoided. However, at all other photon energies, the fields are quite uniform and well described by the macroscopic bulk dielectric theory./82,83/ There remains, nevertheless, the unanswered question of where exactly to locate the boundary between solid and vacuum, especially in the case of adsorbed layers: at the substrate-overlayer interface, or at the overlayer-vacuum interface or elsewhere. It has been estimated that this uncertainty leads to errors which do not exceed approximately 20% in absolute emission cross-sections and which do not vary when the polar incidence angle with respect to the surface normal is kept constant./82/

To summarize, for photon energies above the plasmon energies (and away from other excitation energies), the electromagnetic fields can be described adequately with macroscopic dielectric theory.

### 3.4.8. Photoemission matrix elements

The cross-section for photoexcitation from an initial state  $\psi_i$  to a final state  $\psi_f$  is: /81/

$$\frac{d\sigma}{d\Omega d\omega} = \frac{2\pi e^2}{hc} \frac{2\pi^3 k}{|\mathbf{A}_0|^2 m\hbar\omega} \left| \int \mathbf{j}_f(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r}) d^3r \right|^2 \quad (11)$$

$$\text{where } \mathbf{j}_f = \psi_f^*(\mathbf{r}) \nabla \psi_i(\mathbf{r}) - \psi_i(\mathbf{r}) \nabla \psi_f^*(\mathbf{r})$$

is the transition current density and  $\mathbf{A}(\mathbf{r})$  is the vector potential of the photon field. Assuming that  $\mathbf{A}(\mathbf{r})$  is spatially constant within the surface, one obtains the following dipole matrix element, written in the single particle model using the electron potential  $V(\mathbf{r})$ :

$$\frac{d\sigma}{d\Omega d\omega} = \frac{2\pi e^2}{hc} \frac{2\pi\hbar^2 k}{m\hbar\omega(E_f - E_i)^2} \left| \langle \psi_f | \mathbf{e}_A \cdot \nabla V(\mathbf{r}) | \psi_i \rangle \right|^2 \quad (12)$$

$E_f$  being the kinetic energy of the emitted electron.

In the case of photoexcitation from a bound state  $\psi_i$  of a single atom  $\alpha$  with a muffin-tin potential, one obtains/82/ the following amplitudes for out going spherical waves  $L = (l, m)$ :

$$M_L^\alpha = -(-i)^l \int d^3r [e^{i\delta_l^\alpha} R_l(|\mathbf{r} - \mathbf{R}_\alpha|) - j_l(|\mathbf{r} - \mathbf{R}_\alpha|)] \\ \times Y_L^*(\mathbf{r} - \mathbf{R}_\alpha) - \frac{i\hbar}{2\pi} \frac{e}{mc} \frac{1}{(E_f - E_i)} \mathbf{A} \cdot \nabla V(\mathbf{r}) \psi_i(\mathbf{r}) \quad (13)$$

where the  $\delta_l^\alpha$  are the phase shifts of atom  $\alpha$ , the  $j_l$  are Bessel functions and the  $R_l$  are the normalized radial parts of the solutions to Schrödinger's equation inside atom  $\alpha$  at energy  $E_f$ . If  $\psi_i$  is an atomic core state with quantum number  $l'$ , then the photoemission dipole selection rule (implicit in Eq. (13)) implies that  $l = l' \pm 1$ . Usually, one of the two outgoing values of  $l$  is by far dominant, as is the case in the gas phase.

Once  $M_L^\alpha$  has been computed, one knows the electron waves leaving the source atom and one can proceed to calculate the electron propagation through the surrounding medium, as described in preceding sections (this is the inverse LEED problem). It should be recalled that in this subsequent propagation, scattering by the source atom itself may occur. This is usually treated as if the source atom had already completely relaxed to its initial condition, i.e. as if the hole left by the emitted electron had already been filled. This assumption does not appear to impair the ability to perform structural determination with photoelectrons.

#### 3.4.9. Auger electron emission

The transition matrix element needed to describe Auger electron emission corresponds to the process in which an electron is ejected as a result of the filling of a deeper lying empty level by another electron. This process can be localized on one atom or involve neighboring atoms, especially if valence and conduction levels are involved. The Auger matrix element is/84/

$$M = \langle \psi_c \psi_e | \frac{e^2}{r} | \psi_1 \psi_2 \rangle \quad (14)$$

where  $\psi_c$ ,  $\psi_e$ ,  $\psi_1$  and  $\psi_2$  are respectively the wave functions of the core hole being filled, the emitted Auger electron and the initial state wave-functions of the two participating electrons. It is rather difficult to evaluate this matrix element, especially in the solid state. In the gas-phase atomic state, some results have been tabulated./69/ However, the difficulty of obtaining reliable values for the solid state is such that few attempts have been made./67,68/ In addition, there are no selection rules that limit the number of excited partial waves. Therefore, in ARAES the amplitudes of the outgoing partial waves have usually been left as adjustable parameters; and their relative phases have been neglected by assuming incoherent emission into the different partial waves. This inconvenient situation is a major cause for the infrequent use of ARAES in surface structural determination. While some success has been achieved in understanding the angular dependence of AES (and to a limited extent some structural determination has resulted),/67,68,71/ rather more accuracy is needed to achieve structural results that can compete with other techniques.

Note that in AEAPS-EAPFS, the Auger process is not the primary process which is modulated by fine structure since the energies involved in the Auger process are constant. Given the angle-integrated deflection mode in EAPFS, the Auger step, therefore, need not be described in any detail.

### 3.4.10. Inelastic electron-electron scattering

The matrix element for inelastic electron-electron scattering is given by the following golden-rule expression (spin-averaged): /85/

$$\begin{aligned} \langle \rho_I \rangle = & \frac{4\pi^2 N_c}{h} \sum_{\mathbf{k}_1, \mathbf{k}_2 > k_F} \delta(E_i + E_c - E_1 - E_2) \\ & \times \left\{ \frac{3}{4} | -\langle \mathbf{k}_1 \mathbf{k}_2 | V_c | \mathbf{k}_i \cdot \boldsymbol{\gamma}_l \rangle - |^2 \right. \\ & \left. + \frac{1}{4} | +\langle \mathbf{k}_1 \mathbf{k}_2 | V_c | \mathbf{K}_i \cdot \boldsymbol{\gamma}_c \rangle - |^2 \right\} \end{aligned} \quad (15)$$

Here  $\langle \dots \rangle$  indicates an average over the possible magnetic quantum numbers of the core-hole;  $E_c$  is the core-hole binding energy;  $\gamma_c$  represents the quantum numbers of the core electron, and  $\mathbf{k}_1$ ,  $\mathbf{k}_2$  and  $E_1$ ,  $E_2$  correspond to the momentum and energies of the electrons in the final state. The + and - subscripts indicate the symmetry of the wave-functions under exchange of the spatial coordinates of the two electrons;  $N_c$  is the number of scattering sites, and  $k_F$  is the Fermi momentum. The above expression is applicable in particular to the primary process of EAPFS. No strict selection rules emerge in this situation, but a "pseudo-selection rule" does exist:/85/ for nodeless core wave-functions, both final state electrons in the dominant channel have angular momentum increased by one over that of the core state, at least for incident energies which do not exceed the ionization threshold by 500 eV. In addition, it is found that the dipole component in a partial-wave expansion of the Coulomb interaction dominates the matrix elements.

Phase shifts calculated following the above formalism are incorporated into the Fourier-transformation of the experimental EAPFS oscillations, together with phase shifts for back-scattering from nearby atoms./31/ This allows the bond length to be determined with an accuracy on the order of 0.02 Å.

At the high energies  $E$  of SEELFS/32,33/ and ELNES,/66/ the final state has only a single "active" electron (the incident electron). One obtains then, for momentum transfer  $q$ , a differential cross-section/33/

$$\frac{d^2N(E,q)}{dEdq} \approx -\frac{1}{q^3} \text{Im} \frac{1}{\tilde{\epsilon}_L(E,q)} \quad (16)$$

where  $\tilde{\epsilon}_L = \epsilon_1 + i\epsilon_2$  is the longitudinal dielectric function. Moreover, a dipole approximation is valid when the momentum transfer  $q$  is sufficiently small. As a result, one obtains a situation very similar to that of (S)EXAFS. By utilizing phase shifts calculated for (S)EXAFS, one may then determine/32/ near-neighbor bond distances with a quoted accuracy of 0.05 Å.

## 3.4.11. Electron-phonon scattering

We are concerned here with HREELS,/61,62/ where the probe electrons lose a small amount of energy ( $\leq 0.5$  eV) to atomic vibrations in the surface. Two regimes are commonly discussed. In the case of "small-angle inelastic scattering", also called "dipole scattering", the energy is lost to a long-wavelength phonon with small momentum transfer. The scattered electrons are made to reappear outside the surface by a LEED-type scattering, which supplies the large momentum transfer required for reflection, thus, these electrons emerge within a few degrees of LEED beam directions.

The second case is that of "impact scattering", where the lost energy is transferred to a vibrational mode by a localized impact (distances in the Angstrom range). Impact scattering can occur with small as well as with large scattering angles, but is often overshadowed by dipole scattering at the smaller angles. LEED-type elastic scatterings can occur in addition to impact scattering, but are not essential to obtain a measurable effect. However, they are unavoidable and must therefore be understood./30/

Dipole scattering does not require an atomistic theory. A phenomenological theory suffices, which includes a response function dependent on dielectric constants. The cross-section for dipole scattering based on these assumptions is given in Eqs. 3.7 and 3.9 of Ibach and Mills./61/ These formulae include plane-wave reflection coefficients from the surface, which are solutions of the standard LEED problem. Since dipole scattering involves essentially only forward scattering, it is not necessary in practice to adopt the spherical-wave picture of our step 2 (cf. section 3.4.3), the plane-wave approach is adequate in this situation.

In the case of impact scattering, an atomistic description is essential, since a short-range interaction is operative. The cross-section is then (Eqs. 3.77-79 of Ibach and Mills/61/ )

$$\frac{dS}{d\Omega} = \frac{2mE_I}{h} \frac{\cos^2\theta_{out}}{\cos\theta_{in}} A |M|^2 \quad (17)$$

where

$$M = (n_s + 1)^{\frac{1}{2}} \left( \frac{h}{2N\omega_s} \right)^{\frac{1}{2}} \frac{\partial f}{\partial Q_s} \quad (18)$$

Here  $m$  and  $h$  are the electron mass and Planck's constant,  $\theta_{in}$  and  $\theta_{out}$  the polar angles of incidence and emergence,  $A$  the surface area and  $n_s$  the number of vibrational quanta present (following the Bose-Einstein statistics);  $N$  is the number of unit cells of the surface,  $\omega_s$  the frequency of the vibrational mode and  $\frac{\partial f}{\partial \theta_s}$  the partial derivative of the atomic scattering amplitude with respect to all the nuclear displacements occurring in the vibrational mode. The latter quantity can be related to the corresponding partial derivative of the atomic scattering potential  $V(\mathbf{r})$ . If  $R_{i\alpha}$  is the nuclear displacement of atom  $i$  in vibration mode  $\alpha$ , one obtains in the Green function language (Eqs. 3.80-81 of Ibach and Mills/61/ )

$$\begin{aligned} \frac{\partial f}{\partial R_{i\alpha}} &= \langle \mathbf{k}_{out} | (G + GT_o G) \frac{\partial V(\{\mathbf{R}\})}{\partial R_{i\alpha}} (1 + GT_o) | \mathbf{k}_{in} \rangle \\ &= \langle \mathbf{k}_{out} | g_{PE} \frac{\partial V(\{\mathbf{R}\})}{\partial R_{i\alpha}} | \psi_{LEED}^{(in)} \rangle \end{aligned} \quad (19)$$

Here  $V(\{\mathbf{R}\})$  is the scattering potential for the ensemble of scatters at locations  $\mathbf{R}_i$ ,  $G$  is the free-space electron propagator and  $T_o$  the  $t$ -matrix for multiple scattering of the electron by the surface.

Eq. (17) clearly exhibits the different steps of the HREELS process:  $\psi_{LEED}^{(in)}$  is the LEED wave-function of the electron before the inelastic event;  $g_{PE}$  is the photo-electron propagator of the electron following the inelastic event (this includes all multiple scattering paths ending at the detector). Both  $\psi_{LEED}^{(in)}$  and  $g_{PE}$  provide the sensitivity to surface structure, while  $\frac{\partial V}{\partial R_{i\alpha}}$  contains the vibrational properties of the surface. Given the low energies in most HREELS experiments ( $\sim 5$ -10 eV), the multiple scattering is extremely important and complex. Only a full multiple scattering calculation can properly simulate the diffraction effects.

Nevertheless, as a function of loss energy, the widely used loss peak structure appears.<sup>61,62</sup> This often leads to a relatively simple vibrational mode assignment in the case of adsorbates.

### 3.5. Unifying and Distinguishing Features

In this section, we shall emphasize some aspects of electron diffraction that are common to or differ among the various experimental techniques mentioned so far.

#### 3.5.1. Single vs. multiple scattering

A clarification of terminology is necessary here. In essence, we shall describe why multiple-forward scattering can often be neglected, leaving only single (back-) scattering. We recall from section 3.2 the relative predominance of forward scattering of electrons by atoms over scattering into other directions, especially at higher energies (this is reinforced by the Debye-Waller factor). It is important to realize that, as a result, multiple forward scattering is always present. For instance, in SEXAFS, where one traditionally speaks of single scattering from shells surrounding the emitting atom, there is in reality also multiple forward scattering from one atom to the next in a radial direction, whether outbound or inbound relative to the source atom. This forward scattering often takes the form of focusing of electrons beyond the scattering atom. In this case, the main effect of forward scattering, as far as structure information is concerned, is to change the effective wavelength of the electron, which can be modeled as a changed inner potential. Otherwise, forward scattering does not produce a position-dependent phase change (since  $e^{i\mathbf{s}\cdot\mathbf{r}} = 1$  for forward-scattering where the momentum transfer  $\mathbf{s}$  vanishes, whatever the atomic position  $\mathbf{r}$ ).

This shows how SEXAFS can be described kinematically with a suitable inner potential. Since in SEXAFS the phase shifts are usually empirically determined from other experiments, this corrected inner potential is already

implicitly built into the phase shifts. Consequently, the multiple forward scattering need not be explicitly accounted for and one can use a single-scattering theory in an effective sense.

For the same reason, LEED I-V curves for clean unreconstructed surfaces often have recognizable Bragg peaks, despite strong multiple scattering. These Bragg peaks occur at energies that may be explained kinematically with a suitably adjusted inner potential, which can differ considerably from the muffin-tin constant of a multiple-scattering theory: the difference is due again to the effects of forward scattering./28/

### 3.5.2. Range of multiple scattering

Multiple scattering samples the environment of any surface atom and gives sensitivity to bond angles in addition to bond lengths within that environment. An important question is therefore how large that environment is. The determining factors are, primarily, the mean free path and, to a lesser degree, the atomic scattering amplitude. At kinetic energies of a few eV the mean free path can be large ( $> 10 \text{ \AA}$ ) which results in a large sampled environment. The sampling radius may then be on the order of  $20 \text{ \AA}$  in metals, for instance, including many tens of neighboring atoms.

As the energy is raised, the mean free path goes through a minimum near 50-100 eV and slowly increases again, while the scattering amplitude steadily but slowly decreases, especially in non-forward directions and when the Debye-Waller factor is counted in. Then the sampled environment becomes generally smaller.

We must now distinguish between different techniques. In fine-structure techniques (such as SEXAFS, EAPFS, etc.), weak back-scattering combined with the usual  $\frac{1}{r^2}$  radial decay of spherical waves limits the range of scattering to a few shells of nearby atoms. In techniques like LEED and ARPES, including ARPEFS and ARXPD, multiple forward scattering can occur along electron paths that leave any atom toward the detector or, in the case of LEED, that arrive toward any atom from the source. Such nearly straight paths extend the



sampling range relative to the fine-structure techniques. Therefore, the range of multiple scattering which matters in these techniques is mainly determined by the mean free path.

It is clear that the larger the range of multiple scattering, the more complex the corresponding calculations become. But it has been found, in the context of LEED calculations, that one may markedly cut the effective range of multiple scattering in the *less dense* materials, especially in molecular overlayers.<sup>/86/</sup> This approximation is valid because relatively few neighboring atoms are available to interfere with the primary kinematic process near a given atom. For example, in a typical organic molecule, each atom has at most four near-neighbors, compared with a maximum of 12 within a metal lattice. One may approximate the calculation by including only nearest neighbors in the multiple scattering process. This approach has been used successfully in LEED structural determination of adsorbed molecules in the method called "near-neighbor multiple scattering".<sup>/86/</sup>

### 3.5.3. Long- vs. short-range order

We address here the question whether LEED samples only long-range order, as is often stated, rather than short range order, as do most other techniques described here. This is an important question, because many surfaces will not take on long-range order, and this might prevent the application of LEED. We shall argue that LEED remains applicable. In the case of LEED we should make the following distinction: on the one hand, there is sensitivity to the long-range order, determined by the periodic lattice which produces the diffraction beams, and on the other hand, there is short-range sensitivity to local geometry, determined by the range of multiple scattering as discussed above. This distinction becomes clearest in the case of diffuse LEED from disordered surfaces. Then, the diffracted intensity can be expressed approximately<sup>/79/</sup> as the product of a kinematic structure factor representing the long-range arrangement of the surface and an effective form factor representing the short-range structure, including the sampling by multiple scattering in the neighborhood of any atom.

This factorization into structure and form factors is exact in the limit of dilute identical defects in an otherwise periodic lattice, and only approximate for

dense defects since multiple scattering between defects is neglected. Such a "defect" may be a vacancy or a step in an otherwise periodic surface, or a molecule in an adsorbed layer that has no long range periodicity despite long-range periodicity in the substrate (a "lattice gas"). The error made in the approximation mentioned above can be shown to be small in most cases of interest, since it involves higher-order multiple scattering events, which are relatively weak, as well as long range multiple-scattering paths, which are also relatively weak. Therefore, the main contributions to LEED intensities are due to short-range structure, through the effective form-factor, and to long-range structure, through the kinematic structure-factor, and these terms can be separated out.

This factorization into structure- and form-factors can be applied to ordered surfaces as well (the traditional application of LEED), it then decouples long-range order from short-range order. It is particularly valuable for large unit cells, because one may then simply ignore the long-range order, and also the large number of beams that give rise to very large computational efforts in the more exact traditional formalisms (where all beams are considered coupled). This approach has been implemented and tested in the Beam-Set Neglect method mentioned previously,<sup>/80/</sup> in which most beams are ignored and only selected sub-sets of beams are included in the calculation. This BSN method has enabled the study of large adsorbed molecules, e.g. benzene, including cases of several molecules per unit cell, with a much reduced computational effort compared to previous methods. In fact, the computational effort in this approach may even be chosen to be independent of the unit cell size, instead of being proportional to its second or third power. The same BSN method has also been programmed and tested for the case of disordered overlayers.<sup>/79/</sup>

#### 3.5.4. Various forms of disorder

Many kinds of disorder are known at surfaces. "Clean, well-ordered" surfaces present, among others, point defects (impurities, dislocations, etc.) and line defects (steps, crystallite boundaries, etc.). Surfaces with adsorbates or reconstruction-induced superlattices can have a variety of additional defects, e.g.

out-of-phase, rotated or mirrored domains, disordered regions, isolated ad-atoms, etc. Also, different adsorption geometries can occur at substrate defects and at domain boundaries. With adsorbed molecules (including adsorbed clusters), internal degrees of freedom of the molecules can lead to further disorder, e.g. rotational disorder, while chemical reactions can produce a variety of coadsorbed species.

Different surface-sensitive techniques respond differently to the various kinds of disorder. The measurement of LEED beams is unique in largely filtering out all defects that are unrelated to the superlattice periodicity that defines the beams. Other techniques (including diffuse LEED) generally include contributions from all defects, for instance from adsorbates located at undesired steps and crystallite boundaries. Then only a reduction of the defect concentration can remove defect contributions from the experimental data. Rotational disorder of adsorbed molecules does not matter for techniques which measure only bond lengths, and not bond orientations. Thus, NEXAFS is more sensitive to such disorder than SEXAFS.

In the presence of a variety of chemical species, one may distinguish between the species by various means, e.g. by selection of Auger lines in SEXAFS, or core-levels in ARPEFS, or vibrational frequencies in HREELS.

### 3.5.5. Layer spacings vs. bond lengths

As mentioned earlier, the electron diffraction techniques under discussion in this paper divide into two categories, depending on whether they are primarily sensitive to layer spacings and bond orientations or to bond lengths. The cause of this difference is the source of the structural information. In SEXAFS and EAPFS, for example, the structural information comes primarily from step 3 of the 4-step description, as electrons emitted from the source atom back-scatter from neighboring atoms. Angle-integrated detection is used, which tends to remove all directional information from propagation effects (step 4) and, therefore, leaves only radial information, mainly bond lengths. By contrast, angle-resolved techniques, such as LEED and ARPES, emphasize the behavior of plane waves. In these techniques a lot of the information comes from

propagation effects (steps 1 and 4 of the 4-step description). These plane waves carry structural information through the phase of the interference function  $e^{i\mathbf{s}\cdot\mathbf{r}}$  where  $\mathbf{s}$  is the momentum transfer between two plane waves. Therefore, only the projection of the atomic position vector  $\mathbf{r}$  onto the direction of  $\mathbf{s}$  is relevant, and this projection can be viewed as a layer spacing. When many beams are involved and multiple scattering occurs, many values and directions of  $\mathbf{s}$  are used simultaneously, such that layer spacings in many orientations are accessible. This gives sufficient information to completely determine a surface structure.

### 3.5.6. Bond angles vs. bond directions

We have pointed out two methods to obtain bond angles or bond directions. One uses the sensitivity of electron multiple scattering, as in LEED and NEXAFS, and the other uses variable polarization of incident photons, together with selection rules, as in SEXAFS. The first method gives access to bond angles, i.e. the angle between two inter-atomic bonds. The second method gives access to bond directions, i.e. the absolute orientation of a bond with respect to a laboratory coordinate system. The bond angle is the chemically more relevant quantity, which can only be obtained from the bond direction with additional information or assumptions. For example, the direction of the bond between atoms A and B only provides the bond angle between atoms A, B and C if the bond direction BC is also measured or is known from previous work.

## 4. THE THEORY OF SURFACE CHEMISTRY AND BONDING

The goal of theoretical surface chemistry is to understand the surface chemical bond, and from this to be able to describe and predict the properties of atoms and molecules adsorbed on surfaces. The primary properties of interest include adsorption sites and geometries, bond lengths and angles, the electronic structure of adsorbed species, adsorption energies, diffusion energies, and the

frequencies of vibrational excitations. Other properties that may be derived relate to reaction pathways, kinetics, rates and selectivities as a function of the nature of surface on which the reaction proceeds. Significant progress has been made toward this goal in the last decade using both semiempirical and *ab initio* calculational methods.

The basic approach of chemical theory to surface science is to model a surface with a cluster of a finite number of atoms, with one or more "adsorbate" atoms or molecules bonded to various sites on the cluster. In parallel with the chemical theory there is also the solid state physics approach. This starts from an "extended" surface model, where an array of atoms perfectly periodic in two dimensions represents both the substrate and any adsorbates. Many theoretical techniques have been developed for the extended-surface model. We can only refer the interested reader to the literature/87,88,89,90,91,92,93,94/ and remark that the relative merits of the cluster and extended-surface approaches are still very much under active debate. It is clear that certain properties, such as bonding, are very localized in character and are well represented in a cluster. On the other hand, there are properties that have a delocalized nature, such as adsorbate-adsorbate interactions and electrostatic effects, for which an extended surface model is more appropriate.

The cluster approach is, like the extended-surface approach, characterized by many different calculational schemes. A recent review stresses electronic aspects of bonding. In this review we have chosen to concentrate on geometric aspects. We shall discuss a number of major techniques in order of increasing computational complexity: the extended Hückel theory, self-consistent  $X\alpha$  scattered wave calculations, and self-consistent *ab initio* Hartree-Fock and valence bond methods. In that order these techniques allow increasing accuracy. However, the cluster size must decrease simultaneously due to calculational complexity, ultimately reducing the degree of analogy with surfaces.

#### 4.1. The Extended Hückel Theory and Applications

The most widely used semiempirical quantum chemistry technique for theoretical chemisorption studies is the Extended Hückel Theory (EHT). The method was first proposed by Hoffmann/95/ in its nonrelativistic form, and by Lohr and Pyykkö/96/ and also Messmer/97/ in its relativistic form, based on the molecular orbital theory for calculating molecular electronic and geometric properties. For a cluster the molecular orbitals are expanded as linear combinations of atomic orbitals

$$\psi_i = \sum_{j=1}^n C_{ij} \phi_j. \quad (20)$$

On minimizing the total energy, one obtains a set of secular equations

$$\sum_{i=1}^n [H_{ij} - \epsilon S_{ij}] C_{ij} \phi_j = 0 \quad \text{for } j = 1, 2, \dots, n. \quad (21)$$

Here  $i$  and  $j$  span the  $n$  atomic orbitals in the basis set of valence orbitals (core electrons are neglected) and the coefficients  $C_{ij}$  are chosen to diagonalize the Hamiltonian matrix. The quantity  $S_{ij}$  is the overlap integral between orbital  $i$  and orbital  $j$ . EHT approximates the off-diagonal matrix elements  $H_{ij}$  by the Wolfsberg-Helmholz formula/98/

$$H_{ij} = \frac{1}{2} K (H_{ii} + H_{jj}) S_{ij}, \quad (22)$$

where the diagonal matrix elements  $H_{ii}$  and  $H_{jj}$  are approximated as valence state ionization potentials (IP) for atomic orbitals on atom  $i$  and atom  $j$ . The constant  $K$  is commonly taken as 1.75. The total energy is computed as a sum of one-electron orbital energies

$$E_{EHT} = 2 \sum_i \epsilon_i. \quad (23)$$

The EHT total energy differs from the Hartree-Fock total energy  $E_{HF}$  by the neglect of the nuclear-nuclear repulsion energy  $E_{NN}$  and by the overcounting of electron-electron repulsion and exchange. The energy difference is

$$E_{HF} - E_{EHT} = E_{NN} - \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij}). \quad (24)$$

Here  $J_{ij}$  and  $K_{ij}$  are the Coulomb and exchange integrals of the  $i$  and  $j$  orbital pairs. The sum in Eq. (24) is usually positive.

EHT has the advantage of being computationally an extremely simple method and can be applied to large clusters, up to a few hundred atoms. However, due to the neglect of repulsion terms in Eq. (24), the EHT binding energies for molecules are over-estimated and bond distances are too short./95/ Unrealistic electron charge transfers have also been found for diatomic molecules when the difference between the atomic electro-negativities is large./95/ Improvements can be made by calculating the ionization potential self-consistently/99,100/ or by evaluating the matrix elements more accurately, as is done in the CNDO method./101/

Here we shall discuss a simple method of incorporating the repulsive energy terms as suggested by Anderson and Hoffmann./102/ The method is called Atom-Superposition Electron-Delocalization Molecular Orbital (ASED-MO) theory. It uses the fact that the cluster binding energy can be expressed exactly as the sum of pairwise repulsive energies  $E_R$ , due to rigid atom superposition, and attractive electron delocalization energies  $E_D$ ./103/

$$E_{ASED-MO} = E_R + E_D \quad (25)$$

It has been shown/103/ that  $E_D$  can be approximated by a one-electron molecular orbital energy  $E_{MO}$ :

$$E_{ASED-MO} = E_R + E_{MO} \quad (26)$$

$E_{MO}$  is obtained using Eq. (21) with matrix elements

$$H_{ii} = -IP_i \quad (27)$$

$$H_{ij} = 1.125 (H_{ii} + H_{jj}) e^{-0.13R_{ij}} S_{ij} \quad (28)$$

Here  $K = 2.5$  has been used in obtaining  $H_{ij}$ . An exponential damping factor depending on the inter-atomic distance  $R_{ij}$  is introduced which often improves the results.  $E_R$  is given as the Coulomb interaction between the nucleus of one atom (usually the more electronegative atom) and the nuclear and electronic charges of the other. It has been shown that the simulation of two-body interactions in this way appreciably improves the calculated molecular binding energies and bond lengths./103/

There is a vast literature on the application of EHT to surface chemistry using model clusters. There are also recent reviews on this subject./104,105/ We shall discuss next specific examples relevant to chemisorption of molecules on metal surfaces for illustrative purposes.

Our first example is the widely discussed problem of CO chemisorption on transition metal surfaces. Ray and Anderson/106/ investigated CO chemisorption on a Pt (111) surface using the ASED-MO theory. In their work,  $Pt_4$  and  $Pt_{10}$  clusters were used to model the Pt (111) surface. The bulk Pt-Pt bond distance was used and other theoretical parameters are reproduced in Table II. The calculated Pt-Pt and C-O bond distances and harmonic force constants are within 8% of experimental values for isolated  $Pt_2$  and CO. The results for dissociation energies are less satisfactory: 94.08 and 116 kcal/mole in comparison with experimental values of 84.5 and 256 kcal/mole, respectively.

The calculated binding energy for CO adsorbed on  $Pt_4$  and  $Pt_{10}$  is largest for the 1-fold coordinated site ("top-site"), indicating a definite preference for the 1-fold site. With the 3-fold coordinated site ("hollow-site"), there is a small preference for the 3-fold site which has a vacancy in the second metal layer, corresponding to the fcc hollow site of the (111) surface. The binding energies for 2-fold (bridging) and 3-fold sites are similar and the former is favored on the larger cluster. These results are supported by experiment. High-resolution



electron energy loss spectroscopy (HREELS) obtained by Froitzheim et al./107/ for CO adsorbed on the Pt (111) surface indicates sequential occupation of two binding sites with increasing coverage of CO: the 1-fold site is occupied first at lower coverages, and at higher coverages the two-fold sites are also occupied by CO./108/ Infrared reflectance-absorption spectroscopy (IRAS),/109/ XPS and work function measurements/110/ also support these results. In addition the experimental adsorption energy derived from molecular beam experiments (1.52 eV)/111/ is in good agreement with the theoretical value of 1.66 eV. Both the calculated and measured force constants/111/ indicate a weakening of the C-O bond relative to gas-phase CO when adsorbed in the 1-fold site. This can be explained by charge transfer from the *sd* band of Pt to the two antibonding  $\pi^*$

Table II. Extended Hückel Parameters for CO-Pt

orbital	atom	<i>n</i>	IP(ev)	$\zeta$	C
<i>s</i>	Pt	6	9.0	2.55	-
	O	2	28.48	2.246	-
	C	2	20.00	1.658	-
<i>p</i>	Pt	6	4.96	2.25	-
	O	2	13.62	2.227	-
	C	2	11.26	1.618	-
<i>d</i>	Pt	5	9.6	2.39	6.013
					0.5715
					0.6567

Here *n* = principal quantum number, IP = ionization potential,  $\zeta$  = orbital exponent and C = *d*-orbital coefficient. To mimic self-consistency in adsorption studies Pt ionization potentials are increased by 1.5 eV, O and C ionization potentials are decreased by 1.5 eV, and O exponents are decreased by 0.1 eV.

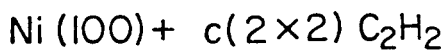
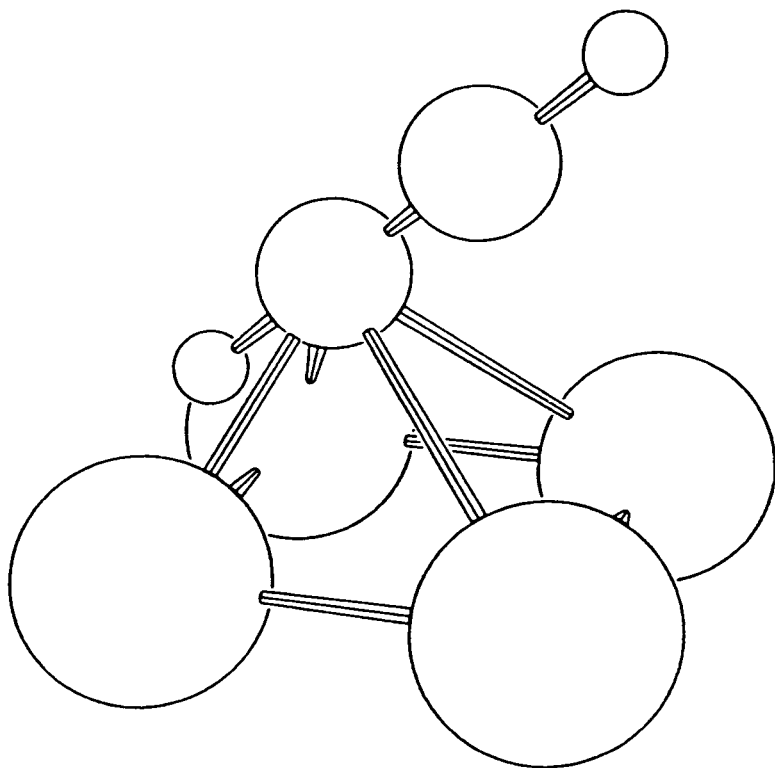
CO orbitals. The calculated bond stretch is 0.07 Å. On going to the 2-fold and 3-fold adsorption sites, the C-O force constant is progressively weakened and the C-O bond lengthened.

The interaction of adsorbed CO and adsorbed oxygen has also been studied to estimate the energy of activation for the oxidation of CO on Pt (111) surface. Two mechanisms have been proposed: the Langmuir-Hinshelwood mechanism and the Eley-Rideal mechanism. In the Langmuir-Hinshelwood mechanism, the associatively adsorbed CO interacts with adsorbed oxygen. The activation energy for oxidation is calculated as 1.6 eV for the transition state geometry. This is in reasonable agreement with experimental results of 1 eV./111/ In the Eley-Rideal mechanism, the CO molecule dissociates on the surface and only then interacts with the adsorbed oxygen atoms. The calculated barrier to CO<sub>2</sub> formation is 1 eV greater than that in the Langmuir-Hinshelwood case. Both the experimental evidence and the theoretical results favor the Langmuir-Hinshelwood mechanism.

The chemisorption of hydrocarbon molecules on surfaces presents another class of important and interesting systems for study. We shall discuss the case of acetylene chemisorption on the Ni (111), Rh (111) and Pt (111) surfaces, as they incorporate many features relevant to all hydrocarbon chemisorption systems.

Gavezzotti et al./112,113/ have applied EHT to predict the chemisorption binding energies of acetylene and a number of organic fragments such as CH, CCH<sub>2</sub>, CCH<sub>3</sub> and HCCH on the Ni (111), Rh (111), and Pt (111) surfaces. The metal surface was modeled by a four-atom cluster, and the acetylene bond lengths were held constant, although the C-C-H bond angle was allowed to vary. The bond length was 1.2 Å for C-C and 1.06 Å for C-H. They found the behavior of acetylene chemisorbed on these different metals to be similar. Ni, Rh and Pt catalyze the rearrangement of adsorbed acetylene to vinylidene ( $>C=C<\overset{H}{H}$ ), and Ni was also found to catalyze the reverse reaction. However, Ni was found to be the best catalyst for carbon-carbon bond breaking in acetylene. Moreover, adsorbate-adsorbate interactions force the acetylene molecule to tilt out of the plane parallel to the surface on Ni (100), as shown in Figure 3.

Anderson and Hubbard/114/ also performed an ASED-MO calculation for acetylene chemisorbed on the Pt (111) surface. Various adsorption sites were considered in this calculation. The hollow site with the C-C axis parallel to the



*Fig. 3. Perspective view of acetylene over a hollow site of Ni(100), represented by four Ni atoms. The C-Ni bonds are drawn to indicate geometrical relationships only, not the bonding character.*

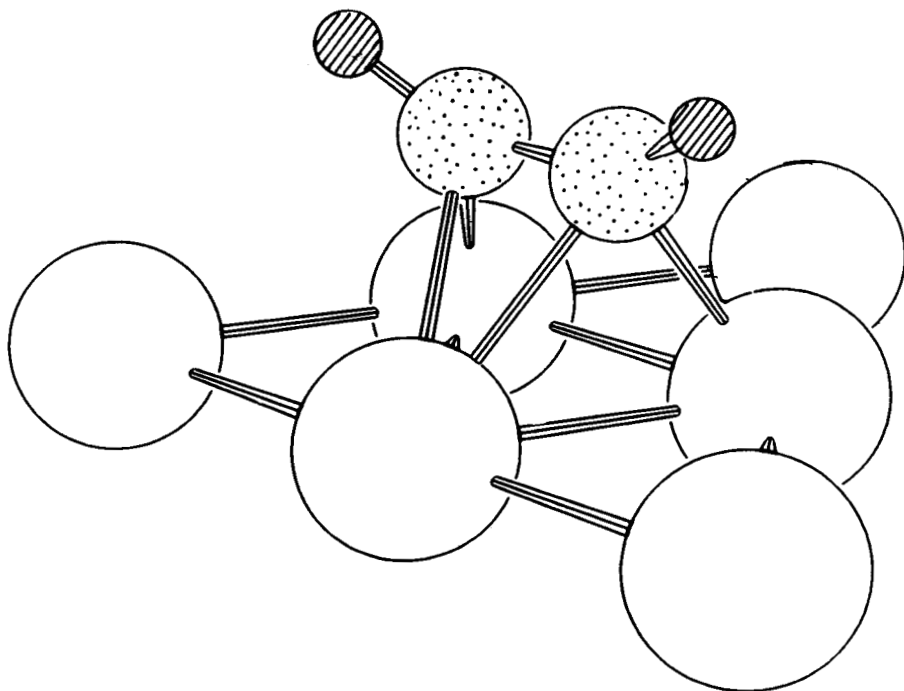
sides of the  $\text{Pt}_3$  triangle was definitely preferred, in agreement with the HREELS analysis of Ibach and Lehwald,/115/ (see Figure 4). In most sites the C-H bond axes rotated away from the surface by about  $55^\circ$  and the C-C bond lengthened by about  $0.3 \text{ \AA}$  relative to the gas-phase. This is in agreement with the bending results of Gavezzotti et al./112/ as well as the HREELS estimates of  $60^\circ$  and  $0.24 \text{ \AA}$  for these quantities.

A very stable species found on several metal surfaces, including  $\text{Pt}(111)$ , is ethylidyne,  $\text{CCH}_3$ . ASED-MO calculations agree that the  $\text{CCH}_3$  species does take on the structure deduced from LEED analysis, namely the C-C axis is perpendicular to the surface, with one carbon bonded equally to three metal atoms. The LEED estimates of a  $2.0 \text{ \AA}$  Pt-C distance and a  $1.50 \text{ \AA}$  C-C distance agree well with the calculated values of  $2.0 \text{ \AA}$  and  $1.55 \text{ \AA}$  (see Figure 5). Ethylidyne is produced either with acetylene co-adsorbed with  $\text{H}_2$  or with ethylene adsorption, both of which are unstable relative to ethylidyne. Structures of  $\text{CCH}_2$  and  $\text{CHCH}_2$  have also been calculated and were found to be less stable than acetylene and  $\text{CCH}_3$  when chemisorbed on Pt (111).

We conclude this section with the observation that the extended Hückel method can produce results in good agreement with experiments when used with caution and appropriate parameter searching procedures.

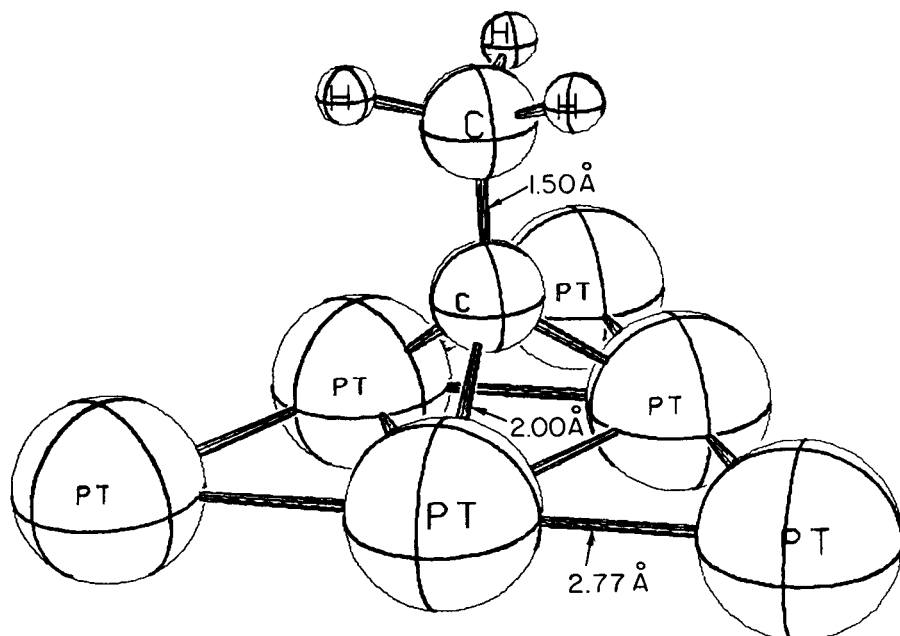
#### 4.2. The SCF $X\alpha$ Scattered Wave Method

The self-consistent-field- $X\alpha$ -Scattered Wave (SCF- $X\alpha$ -SW) method was developed by Johnson/116/ to calculate electronic energy levels and eigenfunctions for polyatomic molecules and solids. In this scheme, the electronic wave functions are solved by a multiple-scattering method, equivalent to the Korringa-Kohn-Rostoker/117, 118/ (KKR) method used for energy band calculations. The method has also been adapted to the calculation of magnetic problems, by use of a spin-polarized formalism,/119/ and to the case of clusters containing heavy elements, by the use of a relativistic formalism./120/ This method has been used quite successfully in correlating the electronic structure of chemisorption systems with the observed experimental photoemission spectrum



### Metastable acetylene on Pt(111)

*Fig. 4. The structure of acetylene on Pt(111) as determined from HREELS data. The lines between atoms are drawn to indicate geometrical relationships only, not the bonding character.*



Pt (111) + ethynidyne

Fig. 5. The adsorption geometry of ethynidyne on Pt(111) as determined by LEED intensity analysis.

assuming known adsorbate binding geometries. It has been less successful in obtaining the adsorption binding energies via total energy calculations.

#### 4.3. SCF- $X\alpha$ -SW Calculations

The SCF- $X\alpha$ -SW method makes the following major assumptions. First one replaces the nonlocal Hartree-Fock exchange potential with the  $X\alpha$  local exchange potential that corresponds to the average exchange potential of a free electron gas. The exchange potential is related to the local electronic density by

$$V_{X\alpha}(\mathbf{r}) = -6\alpha \frac{3}{8\pi} \rho^{1/3}(\mathbf{r}) \quad (29)$$

The value of  $\alpha$  varies between  $2/3$  ( the Kohn-Sham potential)/121/ and  $1$  (the Slater potential)./122/ Different methods for determining the value of  $\alpha$  have been proposed. Usually the value of  $\alpha$  for an isolated atom is determined by requiring that the total energy, using the statistical approximation, should equal the precise Hartree-Fock energy/123/ or the experimental value. The latter takes into account the atomic correlation energy in a phenomenological way. In a molecule or cluster  $\alpha$  values are usually chosen in one of two ways. The values of  $\alpha$  characteristic of the various atoms may be used within each atomic core, and an average may be used in the region between atoms; alternatively, one may set the parameter  $\alpha$  to a universal value in all regions. /124/ Using the statistical approximation, no Slater determinant is required for the electronic wave functions. Instead, we have a one electron equation

$$[-\nabla^2 + V(\mathbf{r})] \psi(\mathbf{r}) = E \psi(\mathbf{r}) \quad (30)$$

where  $V(\mathbf{r}) = V_c(\mathbf{r}) + V_{X\alpha}(\mathbf{r})$  and  $V_c$  is the Coulomb potential. Once the eigenvalues and eigenfunctions for this problem are determined, one assumes that the orbitals of the lowest energy are occupied up to a Fermi level. From the resulting charge densities, one may compute the total energy of the cluster, using the statistical approximation for the exchange-correlation energy.

The second assumption of the SCF- $X\alpha$ -SW method is the introduction of "muffin-tin" approximation to the potential. Each atom  $j$  in the cluster is represented by a sphere of radius  $b_j$  where the inside of these spheres constitutes the so-called region I. The potential  $V^j(\mathbf{r})$  in each sphere is spherically symmetric and nonoverlapping in the original SCF- $X\alpha$ -SW method. Later, overlapping spheres were introduced to improve the results of total energy calculations.<sup>/124/</sup> The entire cluster is then surrounded by a sphere of radius  $b_o$ , which is the smallest sphere containing all the atomic spheres of region I. The outside of this large sphere defines the region III, in which the potential  $V(\mathbf{r})$  is also taken to be spherically symmetric. The space between the small spheres of region I and within the large sphere of region III defines the so-called region II. Within region II, the potential is taken as a constant  $\bar{V}$ , which is the volume average of  $V^j(\mathbf{r})$  in region II. Assuming that the position of the nucleus of the  $j$ th atom is at  $\mathbf{R}_j$ , ( $j = 1, 2, \dots, N$ ) and the center of the outer sphere is at  $\mathbf{R}_o$ , then:

$$V(\mathbf{r}) = \begin{cases} V^j(r_j) & \text{region I; } r_j = |\mathbf{r} - \mathbf{R}_j| \leq b_j \\ \bar{V} & \text{region II;} \\ V^o(r_o) & \text{region III; } r_o = |\mathbf{r} - \mathbf{R}_o| > b_o \end{cases} \quad (31)$$

The self-consistent field calculation requires that as a first approximation  $V_j$  be calculated for each atom  $j$  with its associated atomic electron density. The total potential  $V$  is then obtained as a superposition of the atomic potentials:

$$V(\mathbf{r}) = \sum_{j=1}^N V^j(r_j) \quad (32)$$

Next the total potential is made spherically symmetric within each atomic sphere and in region III. The constant potential  $\bar{V}$  in region II is obtained by

$$\bar{V} = (\Omega_{II})^{-1} \int_{\Omega_{II}} V(\mathbf{r}) d\mathbf{r} \quad (33)$$

where  $\Omega_{II}$  is the volume of the interatomic region. It should be noticed that the



potential in each atomic region includes not only the contribution of the atom located at its center but also the spherically averaged contribution from all other atoms.

After constructing the  $X\alpha$  potentials in each region, the one-electron Schrödinger equation can be solved by matching together separate solutions in regions I, II and III, just as in the KKR method for crystals. The wave matching conditions lead to a determinantal equation, which can be solved only for certain values of  $E$ , the eigenvalues. The corresponding wave-functions yield a new electron density, from which the procedure can be iterated.

We will now discuss the advantages and disadvantages of the SCF- $X\alpha$ -Scattered Wave method. The *ab initio* approach is the major advantage of this method, while it has the flexibility of semi-empirical methods through the incorporation of adjustable parameters such as  $\alpha$  and the muffin tin radii. It also need not treat the core and valence electrons separately: both are handled as parts of the same calculation. On the other hand, the unrealistic "muffin-tin" approximation is disadvantageous in that the bonding between atoms in the constant potential region is overestimated./125,126/ This results in unreliable total energy values. Although overlapping atomic spheres/127/ and non muffin-tin corrections/128/ have been introduced to compensate for this problem, such corrections increase the computational work considerably.

This shortcoming is avoided in the Hartree-Fock-Slater method by removing the muffin-tin approximation./128/ This will be discussed in section 4.5. The SCF- $X\alpha$ -SW method has nevertheless been able to score successes in the determination of surface structures. This success comes from the sensitivity of the calculated surface density of states (DOS) to the surface geometry. It is true that the orbital energies obtained by the SCF- $X\alpha$ -SW method contain unwanted self-interaction energies and are different from the HF orbital energies./129/ However, by using Slater's transition-state theory,/128/ which assumes that  $\frac{1}{2}$  electron is in the initial-state orbital and  $\frac{1}{2}$  electron is in the final-state orbital, one obtains good  $X\alpha$  orbital energies, since the self-interaction terms partially cancel, and most of the relaxation energy is included. In this way good agreement between the calculated DOS with a given adsorbate geometry and the experimental photoemission spectra is often obtained, so that the most favorable adsorbate geometry may be determined. However, the interpretation of

experimental photoemission spectra presents its own complications, which can induce uncertainties in structural determinations. We shall discuss recent applications of this method in the next section.

#### 4.4. Results of the $X\alpha$ -Method

Determination of the adsorbate bonding geometries for CO chemisorbed on the Cu (100) surface provides an interesting example of the  $X\alpha$ -SW method. This example illustrates the problems of interpreting photoemission spectra, since the location of CO on Cu(100) is found to be on a top-site by LEED, while  $X\alpha$ -SW calculations either favor a 4-fold hollow site or are inconclusive. Yu/130/ has carried out calculations for CuCO and Cu<sub>5</sub>CO clusters with CO bonded to the Cu surface through the carbon atom. In CuCO the CO molecule was put directly on top of a Cu atom and in Cu<sub>5</sub>CO it was located in the four-fold coordinated hollow site, simulating CO adsorbed perpendicular to the (100) surface. The calculated valence density of states for the four-fold adsorption site gives a satisfactory interpretation of the two main peaks observed below the copper *d*-band in the UV photoemission spectrum (UPS)./131/ An observed weak third peak can be correlated with a minority of CO molecules adsorbed directly on top of a copper atom.

Messmer and Lamson later calculated core and valence photoemission spectra using the  $X\alpha$ -SW method and offered another explanation for the CO-Cu (100) UV photoemission spectra./132/ They pointed out that, in addition to the main photoionization transitions, the outgoing electron can also excite valence electrons. Assuming that a Cu<sub>5</sub> cluster with CO bonded in a four-fold coordinated hollow-site models the Cu (100) surface, they obtained theoretical DOS to compare with the experimental UPS of Norton et al./133

The experimental spectrum was resolved into four peaks. Two of these are assigned to three shake-up transitions labeled  $1\pi_+$ ,  $4\sigma_+$  and  $5\sigma_+$  arising from a hole in each of the  $1\pi$ ,  $5\sigma$ , and  $4\sigma$  orbitals. A third peak is assigned to the  $1\pi$  and

5 $\sigma$  ionizations and the fourth peak to the 4 $\sigma$  transition. Messmer and Lamson conclude that CO adsorption in the four-fold symmetric hollow site is sufficient to explain the observed UPS data.

A theoretical interpretation of the *core-level* satellite structure in the x-ray photoemission spectrum was also offered by Messmer et al./134/ using a larger Cu<sub>9</sub>CO cluster and assuming the chemisorption of CO in the one-fold and four-fold sites of the Cu (100) surface. The qualitative differences in the intensity distributions suggest that the Cu<sub>9</sub>CO cluster is a better model than the smaller Cu<sub>5</sub>CO cluster for the Cu (100) surface.

A three-peak structure of the carbon 1s x-ray photoemission spectrum for CO chemisorbed on Cu (100) is obtained experimentally. The structure is from the photoionization of a carbon 1s electron, where three final ion states with roughly the same energies but different degrees of hole screening are excited. These differences in hole-screening lead to the observed differences in the carbon 1s binding energies.

A molecular orbital study using the X $\alpha$ -SW method suggests that the first peak in the experimental XPS data arises from a transition between the ground state of the neutral chemisorption system and a final state in which a core hole in the CO molecule is created, together with the transfer of an electron from Cu to the 2 $\tilde{\pi}_g$ -orbital of the chemisorption system. The other two peaks are caused by shake-up excitations from this core-hole-ion ground state. One of these two peaks can be viewed as a transition from this ground state ion to an excited-state ion via an excitation of electron from the 2 $\tilde{\pi}_g$ -orbital to the higher energy 2 $\tilde{\pi}_a$ -orbital. The other peak is described as a one-electron excitation from the 1 $\tilde{\pi}'$  level to the 2 $\tilde{\pi}'$  level. Unfortunately the core level shake-up spectra calculated for the 1-fold and 4-fold sites are qualitatively very similar, and this rules out a definite conclusion regarding the CO binding site on the Cu (100) surfaces.

Another way of making geometric structural determinations with X $\alpha$ -SW calculations is to compare the adsorbate photoemission energy level shifts with theoretical level shifts obtained with the assumed binding geometries. An example is the adsorption of ethylene (C<sub>2</sub>H<sub>4</sub>) on Ni (111).

In the so-called  $\pi$ -bonded model, the C<sub>2</sub>H<sub>4</sub> molecule lies flat on the surface and a  $\sigma$ -bond is formed by the overlap of the Ni  $d_{z^2}$  orbital and the C<sub>2</sub>H<sub>4</sub>  $b_{3u}(\pi)$ -

orbital, which donates charge to the nickel atoms. The “ $\pi$ -back-bonding” refers to charge transfer from the Ni  $d_{zz,yz}$ -orbitals into the unoccupied  $C_2H_4$   $b_{2g}(\pi)^*$ -orbital. UPS spectra show an ethylene derived  $\pi$ -level with a bonding shift of 0.9 eV for adsorption on Ni (111), in addition to a relatively uniform relaxation shift for both the  $\pi$ - and  $\sigma$ -type orbitals./135/ The  $X\alpha$ -SW calculation of Rösch and Rhodin/136/ for the  $Ni_2(C_2H_4)$  cluster gave a 0.72 eV bonding shift for the  $C_2H_4$   $b_{3u}(\pi)$  orbital of the  $\pi$ -bonded complex and a 0.30 eV shift in the di- $\sigma$ -bonded complex.

Howard and Dresselhaus later did calculations for the  $\pi$ -bonded model using a  $Ni_{10}$  cluster and found that the level structure as a whole exhibits slow oscillations in energy from iteration to iteration./137/ An accurate level shift cannot be determined without artificially shifting and aligning the calculated Fermi energy with the experimental one for the chemisorption system. The resulting level structure can explain the reported activation energy difference for ethylene hydrogenation on paramagnetic and ferromagnetic nickel. It is due to shifts in the Ni  $spd$  levels relative to the ethylene  $b_{3u}(\pi)$  bonding orbital level, via an exchange splitting of  $\sim 0.1$  eV.

Finally we discuss the case of CO bonding to a platinum (100) surface using the relativistic version of the  $X\alpha$ -SW method./138/ Adsorption in the top, bridge and four-fold hollow sites with the carbon atom bonded to the metal were examined for  $PtCO$ ,  $Pt_2CO$  and  $Pt_5CO$  clusters, respectively. In general, the  $s$ -band in the transition metals hybridizes with the  $d$ -band and at the same time participates in bonding with an adsorbate. However, its hybridization with the  $d$ -band is sufficiently weak so the general appearance of the latter remains intact. When relativistic spin-orbit coupling is included the  $d$ -band splits into  $d_{3/2}$  and  $d_{5/2}$  components. Strong  $d_{3/2}$ - $d_{5/2}$  hybridization is observed only for the  $sp$  or  $spd$  orbitals, which are largely responsible for bonding among the Pt atoms. Upon addition of the CO molecule, the metal bands undergo different changes in each model cluster. In all cases, the  $s$ -band which overlaps the  $d$ -band, and which participates in bonding as well as back-bonding with CO, is partially depleted in the  $d$ -band region. The orbitals responsible for bonding between CO and Pt clusters are the  $5\sigma$  and  $2\pi^*$  in the generally accepted bonding picture of  $5\sigma$  donation and  $2\pi^*$  back-donation. The energy position of the  $5\sigma$ -orbital with respect to the centroid of the  $d$ -band remains fixed for all three

clusters. In all cases the C-Pt bond involves at least one antibonding orbital. In the  $\text{Pt}_5\text{CO}$  cluster, but not the other two, one of these antibonding orbitals is occupied. Thus, qualitatively, we expect the four-fold site to have the lowest binding energy among the three sites considered for CO-Pt (100). This prediction is confirmed by HREELS observations.

In the two previous sections, we have discussed the use of the basic SCF- $X\alpha$ -SW method. The advantage of this method is its computational efficiency relative to other *ab initio* methods which allows larger clusters to be handled. The disadvantage is the inadequacy of the "muffin-tin" approximation for systems with low symmetry or open structure, which renders the determination of orbital energies and total binding energies unreliable. In the following section we shall describe an improved method which relaxes the "muffin-tin" approximation and is potentially valuable in adsorbate structure determination.

#### 4.5. The Hartree-Fock-Slater Method and its Applications

In the Hartree-Fock-Slater (HFS) method, one solves the  $X\alpha$  one-electron equation (Eq. (21)) directly without using the muffin-tin approximation. Various approaches can be taken. The scheme called discrete variational linear combinations of atomic orbitals (DV-LCAO) uses Slater type orbitals (STO) as variational basis functions./139/ The charge density required to calculate  $V_c$  and  $V_{X\alpha}$  is cast into a STO-based multipolar form by applying a fitting procedure to each product of orbital pairs. Highly accurate results can be obtained, but the large basis sets are only treated with difficulty. Alternatively, the discrete variation self-consistent charge (DV-SCC) scheme uses numerical atomic-like wave functions as variational basis functions./140/ Since these numerical basis functions can be chosen to form a computationally highly efficient contracted basis set, larger clusters may be treated. To facilitate the calculations, a shape approximation to the charge density is obtained through a Mulliken population analysis/141,142/ of the cluster eigenvectors which leads to

spherically overlapping charge densities and Coulomb potentials. Generally DV-SCC results are found to be of intermediate quality, between  $X\alpha$ -SW and full potential DV-LCAO results.

An example of such an application is the calculation of molecular ionization energies for oxygen chemisorbed on Ni (100) surfaces using the DV-SCC method. A  $\text{Ni}_5\text{O}$  cluster was used to model chemisorbed oxygen in the  $c(2\times 2)$  configuration in four-fold symmetric hollow sites on the Ni (100) surface. The vertical distance of the oxygen atom to surface plane is taken as 1.7 a.u., a height near that inferred from LEED data. All 148 electrons of  $\text{Ni}_5\text{O}$  were explicitly treated with  $\alpha = 0.7$  and basis sets including the 4s and 4p nickel states. At this height the oxygen valence levels ( $\sigma$  and  $\pi$  type) are strongly perturbed by mixing with the metal 3d, 4s, and 4p orbitals of  $a_1$  and  $e$  symmetry. The 12e orbital of the  $\text{Ni}_5\text{O}$  cluster, which has a dominant oxygen 2p character, was found to have a binding energy of 10.8 eV. The UPS spectra of Eastman and Cashion/143/ and the ion neutralization spectra of Hagstrum and Becker/144/ for oxygen on Ni (100) show a broad peak ( $\sim 2$  eV wide) centered 5.5 - 6.0 eV below the Fermi energy. This corresponds well to the calculated oxygen 2p level. In addition, some high lying (4.7 eV above the Fermi energy) antibonding levels can be correlated directly with the HREEL spectrum obtained by Andersson./145,146/

Rosén et al./147/ and Ellis et al./148/ have carried out calculations using the DV-LCAO method for CO chemisorbed on the Ni (100) surface at various binding sites. They used  $\text{NiCO}$  and  $\text{Ni}_5\text{CO}$  clusters for the top site configuration, a  $\text{Ni}_2\text{CO}$  cluster for the bridge bonding configuration and a second  $\text{Ni}_5\text{CO}$  cluster for the four-fold hollow site configuration. The experimental HREELS measurements of Andersson for  $c(2\times 2)$  CO on Ni (100) show two vibrational losses, which are assigned to the Ni-C and C-O stretch vibrations of CO linearly bonded to nickel atoms (top-sites)./145,146/ At low coverage, Andersson also found vibrational excitations assigned to CO bridge bonded to two surface nickel atoms.

Total energy calculations of sufficient precision would be able to determine the energetically favorable adsorption site. Such calculations are still much more difficult than the calculation of orbital energy levels and vibrational frequencies. Values of the Ni-Ni bond distance were chosen to correspond to the bulk crystal structure. In the DV-LCAO calculations the C-O and C-Ni bond distances were

taken from known nickel-carbonyl bond distances./149/ The results showed that one-electron binding energies in these clusters are relatively insensitive to the site geometry. The comparison between the theoretical and experimental stretching frequencies for chemisorbed CO did not provide firm evidence for one specific chemisorption site. However, the theoretical results are sensitive to geometrical factors such as the Ni-C and C-O bond distances, and the possible variation of the CO molecular axis from the surface normal. Several LEED results confirm the HREELS observation of top sites without tilting, and support the bond lengths assumed in the DV-LCAO calculations. In addition, a comparison of experimental angle-resolved UPS measurements /150/ with theoretical calculations/151/ showed that CO is chemisorbed with its molecular axis normal to the surface within  $5^\circ$ . A tilted molecule would give a split  $\pi$  level. However, this splitting is estimated to be quite small./151/

The chemisorption of chalcogen atoms on Ni surfaces has been investigated by many techniques (see Table VI). As a result of LEED and other studies the geometrical structure of chalcogens chemisorbed on nickel surfaces is well established. Chalcogen atoms generally adsorb in the site with maximum coordination on close packed nickel surfaces; 4-fold coordinated on the (100) and (110) surfaces, and 3-fold coordinated on the (111) surface. Recently Cao et al./152/ carried out DV-LCAO calculations investigating the spectroscopic and bonding properties of S, Se and Te atoms chemisorbed on Ni (100), (110) and (111) surfaces. The model clusters they used were  $\text{Ni}_5\text{S}$  and  $\text{Ni}_9\text{S}_2$  for Ni (100),  $\text{Ni}_5\text{S}$  for Ni (110) and  $\text{Ni}_7\text{S}$  for Ni (111) corresponding to the binding sites found by LEED. Their calculations of the variation of adsorbate energy levels with adatom spacing above the metal surface are in good agreement with the results of the LEED structural analyses.

Finally, we conclude this section by drawing attention to the recent development of total energy calculations within the  $X\alpha$  approximation. A total energy algorithm specially implemented for calculating small energy differences in large systems has been developed by Ziegler for chemisorption studies./153/ Ellis et al. also developed a method within the HFS approximation to handle this problem and the results for a series of clean copper clusters, namely  $\text{Cu}_2$ ,  $\text{Cu}_4$ ,

$\text{Cu}_{13}$  and  $\text{Cu}_{79}$  are very encouraging. Direct binding site determination via accurate total energy calculations for clusters of sufficient sizes is likely to play a more important role in the near future.

#### 4.6. The *ab initio* Quantum Chemical Calculation

In the previous two sections we have discussed the semiempirical extended Hückel theory and the SCF- $X\alpha$ -SW method. We have detailed the advantages and disadvantages of these methods for surface structure determination. However, there exists a vast literature of *ab initio* quantum chemical methods which are described in terms of either the molecular orbital (MO) or valence bond (VB) schemes for determining the electronic and geometric structure of molecules. The application of these methods to surface problems has advanced rapidly in recent years, as we shall discuss in this section.

The most commonly used *ab initio* method is the self-consistent field Hartree-Fock (SCF-HF) scheme with or without configuration interaction (CI). In the Hartree-Fock method, the wave function for a closed shell singlet state has the antisymmetrized form

$$\psi_i = A(\phi_1\alpha\phi_1\beta\phi_2\alpha\phi_2\beta \cdots \phi_n\alpha\phi_n\beta) \quad (34)$$

with each orbital appearing twice for the two spin states  $\alpha$  and  $\beta$ . The equation satisfied by the one-electron orbital  $\phi_i$  is written as

$$[h_{\text{eff}} + \sum_j (2J_j - K_j)] \phi_i = \epsilon_i \phi_i \quad (35)$$

Here  $h_{\text{eff}}$  contains the kinetic energy operator and the electron-nucleus attraction interaction.  $J_j$  and  $K_j$  are the usual Coulomb and exchange operators. Eq. (35) can be solved by a self-consistent iteration procedure within the LCAO approximation. Filling these molecular orbitals up to the Fermi energy gives the ground state electronic configuration for the system. For open-shell systems when a few determinants of the form of Eq. (34) are combined, the resulting one-



electron equations can be solved using the multiconfiguration SCF (MCSCF) method. The SCF-HF results neglect electron-electron correlation effects. The configuration interaction includes these effects by formulating a multiconfiguration electron state as

$$\Psi = \sum_i c_i \psi_i \quad (36)$$

where the  $c_i$ 's are coefficients to be determined variationally.

Numerous studies have applied SCF-HF, MCSCF and SCF-HF-CI methods to surface problems./154,155,156,157,158,159,160,161,162,163/ One particularly interesting example is an explanation of the origin of the coverage-dependent vibrational frequency shift for oxygen on Ni (100). Bauschlicher and Bagus have carried out SCF studies for the  $\text{Ni}_{25}\text{O}$  and  $\text{Ni}_{25}\text{O}_5$  clusters to model the p(2x2) and c(2x2) structures of O on Ni (100), respectively./164/ The cluster contains a four-atom by four-atom square top layer and a three-atom by three-atom second layer. Four oxygen atoms are located in hollow sites for the (2x2) arrangement and a fifth centered atom is added for the c(2x2) arrangement. The spacing between the oxygen atoms and the first Ni layer is varied. For the  $\text{Ni}_{25}\text{O}_5$  cluster, the four atoms placed at the outer four-fold sites are given an equilibrium spacing determined from  $\text{Ni}_{25}\text{O}$  calculations where only one oxygen is present, centered on the square top layer. The spacing of the central atom is then allowed to vary.

The Ni atoms are treated as one-electron systems in which the effects of the Ar-like core and the nine 3d electrons are replaced by a modified effective potential (MEP) as suggested by Melius et al./165/ A contracted gaussian basis set is used for Ni, which includes two functions to describe the 4s and one to describe the 4p atomic orbitals. Since a previous study/159/ found the O-Ni spacing and the vibrational frequency  $\omega_e$  insensitive to correlations in this open-shell system, the authors have adopted the SCF calculation scheme. To check the approximation of treating the Ni atoms as a one-electron system, they performed both the MEP and an all electron SCF calculation for the  $\text{Ni}_5\text{O}$  cluster. They found that the MEP spacing is 0.37 Å or 35% smaller than the all-electron value; the MEP  $\omega_e$  value is 90  $\text{cm}^{-1}$  or 24% smaller. Since the 3d

orbital extends into the valence region where the 4s electron density is substantial, a correction needs to be added to the MEP calculation to take the 3d penetration into account. This was done phenomenologically by replacing the core charge of +1 by a distance-dependent  $Z_{eff}(R)$ . When this correction is included, the resulting spacing and  $\omega_e$  values are in good agreement with the all-electron values. They found that for  $\text{Ni}_{25}\text{O}$ , the spacing is 0.83 Å and for  $\text{Ni}_{25}\text{O}_5$  it is 0.72 Å. The  $\omega_e$  values are 400 and 315  $\text{cm}^{-1}$ , respectively./166/ The  $\omega_e$  values measured in HREELS experiments are 430 and 310  $\text{cm}^{-1}$  for p(2x2)-Ni(100)-O and c(2x2)-Ni(100)-O. The spacings obtained from EXAFS experiments are  $0.86 \pm 0.07$  Å for both p(2x2)-Ni(100)-O and c(2x2)-Ni(100)-O./167/ LEED analyses have yielded a spacing of  $0.9 \pm 0.1$  Å for both coverages./168,169/ These calculations do not support the possibility of having two states, i.e., an O radical state for p(2x2) and an oxide state for c(2x2) coverages. The two states were previously suggested by Upton and Goddard/170,171/ as an explanation for the two different Ni-O vibration frequencies.

Another interesting example of MCSCF calculation is the determination of the oxygen binding site on Si(100). Batra et al performed this work for a few model clusters for oxygen chemisorbed at one, two and four-fold coordinated sites on  $\text{Si}_3\text{H}_6$ ,  $\text{Si}_6\text{H}_{12}$ , and  $\text{Si}_7\text{H}_8$  clusters./172/

They attempted to resolve the controversy regarding the dissociative chemisorption of oxygen on semiconductor surfaces. The energy-minimized geometries and absolute binding energies are listed in Table III. Since the binding energy of  $\text{O}_2$  is 2.6 eV per oxygen, dissociative chemisorption is found to be energetically favorable and is exothermic by 3 eV. Moreover, the two-fold bridge site is calculated to be more stable than the one-fold top and four-fold hollow sites. The calculated interatomic distance for the on-top site, 1.64 Å, is in good agreement with the value obtained by Goddard et al., of 1.69 Å./173,174/ Since the treatment of correlation effects in MCSCF is limited, the dissociation energies are expected to increase when a more accurate configuration interaction calculation is performed.

Comparing the calculated vibrational frequencies with the HREELS data/175/ also yields a reasonable interpretation. For Si(100) at low  $\text{O}_2$  exposures at 700 K, two major EELS peaks were reported, one at 1060  $\text{cm}^{-1}$  and

Table III. Oxygen on Si(100)

coordination	1-fold	2-fold	4-fold
model cluster	Si <sub>3</sub> H <sub>6</sub> O	Si <sub>6</sub> H <sub>12</sub> O	Si <sub>7</sub> H <sub>3</sub> O
spacing (Å)	1.64	0.06	0.96
bond length (Å)	1.64	1.92	2.88
energy (eV)	3.81	4.28	0.27
$\omega_e$ (cm <sup>-1</sup> )	866	288	106

All silicon atoms which are not involved in Si-O bonding are 4-fold coordinated, with all dangling (non Si-Si) bonds taken up by Si-H bonds.

the second at 866 cm<sup>-1</sup>. A third low intensity peak is at 370 cm<sup>-1</sup>. The authors assigned the high-frequency peak to the top-site and the lower frequency 866 cm<sup>-1</sup> peak to the bridge site. The third low intensity peak can be associated with adsorbed O<sub>2</sub> or O atoms in an initial or intermediate stage of oxidation of Si to form SiO<sub>2</sub>.

While the above mentioned *ab initio* calculations produced good results for surface structure determinations, their computational cost is often high. In most formulations, the cost arises from the large number of two-electron integrals to be evaluated. The number of integrals increases as  $n^4$  for  $n$  basis functions, so that computations become lengthy when large basis sets are required to achieve accuracy. This places practical restrictions on the applicability of *ab initio* techniques. One possible simplifying approach is to consider exact evaluations of certain important integrals in the  $n^4$  list and to approximate the remainder. This idea was first applied to molecular problems by Whitten.<sup>176</sup> He introduced the notation

$$\langle g_i(1)g_j(1) | \frac{1}{r_{12}} | g_k(2)g_l(2) \rangle = [\phi_{ij} | \phi_{kl}] \quad (37)$$

for electron repulsion integrals where  $g$  denotes an arbitrary basis function. He then showed that

$$| [\phi_{ij} | \phi_{kl}] - [\phi_{ij}' | \phi_{kl}'] | \leq \delta \quad (38)$$

where

$$\delta = \epsilon_a^{1/2} \epsilon_b^{1/2} + \epsilon_a^{1/2} K_b^{1/2} + \epsilon_b^{1/2} K_a^{1/2} \quad (39)$$

$$\epsilon_m = [\phi_m - \phi_m' | \phi_m - \phi_m'] \text{ and } K_m = [\phi_m | \phi_m'], \quad m = a, b \quad (40)$$

Here  $\phi_{ij} = g_i g_j$  and  $\phi_{kl} = g_k g_l$  are used to approximate the electron densities  $\phi_{ij}'$  and  $\phi_{kl}'$ . The attractiveness of this approach is due to the simple structure of the error bound  $\delta$  which can be calculated separately for the pair of densities  $\phi_{ij}$  and  $\phi_{kl}$ , yielding a number of integrals proportional to  $n^2$ . One can then obtain the approximated density by minimizing  $\delta$  with respect to all parameters in  $\phi_{ij}$  and  $\phi_{ij}'$ . This method is called the error bound method. Its results have been shown to be quite successful in many cases.

In addition to the error bound method to approximate two-electron integrals, Whitten and Pakkanen also proposed a new embedding theory to increase the cluster size that could be used for chemisorption studies.<sup>177</sup> Let us assume that molecular adsorption is taking place at a chosen site of well-defined geometry. Let the adsorbate be characterized by a set of orbitals  $\{\phi_k^A\}$  with energies  $\{\epsilon_k^A\}$  and let the metal substrate be characterized by another set of orbitals  $\{\phi_l\}$  with eigenvalues  $\{\epsilon_l\}$ . For large metal clusters, the number of orbitals in  $\{\phi_l\}$  is very large and the energy levels  $\{\epsilon_l\}$  are densely spaced. Whitten and Pakkanen proposed a so-called "localization" scheme in which the interaction of the adsorbate levels with the large number of metal levels can be described by those orbitals for which the interaction  $| \langle \phi_k^A | H_{eff} | \phi_l \rangle |$  is large. A unitary transformation can be introduced such that  $| \langle \phi_k^A | H_{eff} | \phi_l \rangle |$  is maximized. This procedure yields one set of functions which physically represent orbitals localized on the designated surface atoms, bonds between these atoms, as well as bonds linking the designated atoms with the remainder of the cluster. A second set of functions consists of so-called interior orbitals. They are treated as an invariant core in each configuration for the step of the configuration

interaction. Thus, for an  $n$ -electron chemisorption system the interior orbitals contribute a fixed Coulomb and exchange field acting on the  $n$  electrons included explicitly.

The application of the error bound method and the embedding theory may be illustrated by a SCF-HF-CI calculation for  $H_2$  on Ti (0001)./178/ It was found by this method that dissociative adsorption of  $H_2$  occurs for the  $Ti_7H_2$ ,  $Ti_{10}H_2$  and  $Ti_{38}H_2$  clusters. The top, bridge and three-fold hollow sites were considered.  $H_2$  was allowed to approach the surface with its axis parallel and perpendicular to the surface plane and the H-H bond was also allowed to stretch. The authors found that dissociative chemisorption occurs and the most stable adsorption geometry is in the 3-fold symmetric site with H spaced 1.3 Å above the surface. Adsorption of two hydrogens in adjacent 3-fold sites (about 1.5 Å apart) is less stable than in next-nearest 3-fold sites or more distant sites. The calculated adsorption energy of 45 kcal/mol for  $Ti_{38}H_2$  compares favorably with experiment (the adsorption geometry is otherwise not known, but by analogy with Ni (111), 3-fold hollow sites are likely). The chemical bond involves mainly the 4s electrons of the metal interacting with the hydrogen and a polarized substrate electron distribution, but  $d$ -electron bonding and correlation significantly increase the binding energy. The binding energies calculated for the smaller clusters are in most cases a few tenths of an eV larger than those calculated for the larger clusters. Similar calculations have also been carried out by Madhavan and Whitten for the chemisorption of  $H_2$  on Cu (100) using a  $Cu_{38}H_2$  cluster./179/ They found that  $H_2$  dissociative chemisorption is activated in that case. The  $H_2$  binding energies are in the range of 13-22 kcal/mol. The activation barrier is about 35-40 kcal/mol due to the repulsion of molecular  $H_2$  by the surface and the difficulty of stretching  $H_2$  at a significant distance above the surface. It was also concluded from these calculations that the effect of the 3d-electrons, either in terms of direct 3d-4d mixing or in terms of 3d-4s promotion is unlikely to assist the dissociation of  $H_2$  on the Cu surface.

Finally we discuss the Generalized Valence Bond (GVB) method as proposed by Hunt, Hay and Goddard./180,181,182/ This is a method aimed at improving some of the HF procedures, for instance to obtain the correct dissociation limit, and to partially include correlation effects. The basic idea is that the  $\phi_i\alpha\phi_i\beta$  type of orbitals in Eq. (34) be replaced by

$$\phi_{ia}\alpha\phi_{ib}\beta, \quad (41)$$

such that each spin can have a different spatial orbital  $\phi_{ia}$  or  $\phi_{ib}$ , instead of having the same spatial orbital  $\phi_i$  for both spins. This yields the wave function

$$\begin{aligned} \psi_i &= A(\phi_{1a}\alpha\phi_{1b}\beta\phi_{2a}\alpha\phi_{2b}\beta \cdots \phi_{Na}\alpha\phi_{Nb}\beta) \\ &= (\phi_{1a}\phi_{1b}\phi_{2a}\phi_{2b} \cdots \phi_{Na}\phi_{Nb}\chi) \end{aligned} \quad (42)$$

where  $\chi$  is a general  $N$ -electron spin function and the orbitals  $\phi_{ia}$  and  $\phi_{ib}$  are solved for self-consistently. The dependence of the energy upon the orbital pair  $i$  has the form

$$\begin{aligned} E &= E_{(i)} + f_{1i} \langle \phi_{1i} | 2h_{eff} + J_{1i} | \phi_{1i} \rangle + f_{2i} \langle \phi_{2i} | 2h_{eff} + J_{2i} | \phi_{2i} \rangle \\ &+ C_{1i}C_{2i} \langle \phi_{1i} | K_{2i} | \phi_{1i} \rangle \end{aligned} \quad (43)$$

where  $i = a, b$ ;  $h_{eff}$  is the usual HF Hamiltonian excluding the orbital pair  $i$ ,  $E_{(i)}$  is independent of the orbitals in pair  $i$  and  $f_k$  is a constant which depends on the orbital occupancy. Separating from  $E_{(i)}$  the terms involving the other pairs, the authors obtained the general expression

$$E = \sum_k f_k h_k + \sum_{k,l} (a_{kl} J_{kl} + b_{kl} K_{kl}) \quad (44)$$

which gives, upon using the variation principle

$$\bar{H}_k \phi_k = [H_k - \sum_{j \neq k} |j\rangle \langle j| H_j] \phi_k = \epsilon_k \phi_k, \quad k = 1, 2, \dots, n \quad (45)$$

where

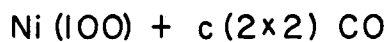
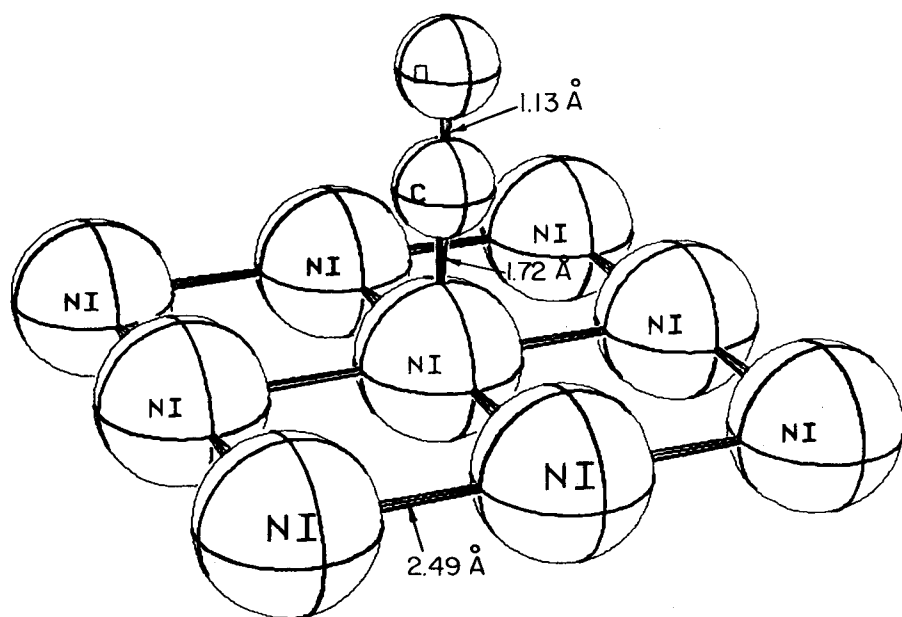
$$H_k = f_k h + \sum_l (a_{kl} J_l + b_{kl} K_l) \quad (46)$$

and  $n$  is the number of distinct orbitals.  $J$  and  $K$  are the usual Coulomb and exchange operators from HF theory. Eq. (46) can be solved self-consistently. Since the GVB method allows paired electrons which have different spins to occupy different spatial orbitals, some correlation is included in solving the GVB equation, which removes many difficulties and inconsistencies of the HF method. Using the GVB wave functions one can include further correlation via the configuration interaction in the so-called GVB-CI method.

To illustrate applications of the GVB-CI method we shall first discuss the recent theoretical studies of CO on Ni (100) of Allison and Goddard./183/ They investigated CO adsorption at a top site having one-fold coordination to one Ni metal atom (see Figure 6). A  $\text{Ni}_{14}$  cluster was used to represent Ni (100). The Ar-like core of the Ni atom was replaced by an effective potential,/184/ while the nine  $3d$  electrons were kept frozen, producing another electrostatic potential as discussed previously. The GVB calculation produced a CO dipole moment of 0.124 Debye in excellent agreement with the experimental value./185/ The best calculated geometry for  $\text{Ni}_{14}\text{CO}$  has bond lengths  $R_{\text{Ni-C}} = 1.94 \text{ \AA}$ ,  $R_{\text{C-O}} = 1.14 \text{ \AA}$  and an Ni-C-O bond angle of  $180^\circ$ , in good agreement with LEED results. Calculations for CO tilted away from the normal showed this geometry to be energetically less favorable. The dissociation energy, after correcting for the zero-point energy and temperature dependence of the enthalpy, is 1.29 eV. The binding energy obtained for the  $\text{Ni}_4\text{OC}$  geometry, i.e. with the O end toward the surface, was much lower, only 0.39 eV. The optimum Ni-O bond length is then 2.03  $\text{\AA}$ . The GVB results agree well with the experimental heat of adsorption of 1.30 eV found by Tracy at low coverages./186/ Good agreement is also found for vibrational frequencies. The calculated frequencies are 49.7 meV for the Ni-CO stretch, 264.0 meV for the C-O stretch, and 40.5 meV for the Ni-CO bend. The calculated frequencies on the  $\text{Ni}_5\text{CO}$  cluster are 54.2, 258.9 and 50.9 meV for these modes.

Another application of the GVB-CI method is the determination of the chemisorption geometries of oxygen and aluminum on the GaAs (110) surface by Barton et al./187/ The 28 core electrons of Ga and As are replaced by model effective core potentials./184/

The bulk GaAs zincblende structure involves tetrahedral bonding about each atom. Each Ga atom is bound to four near As atoms and vice versa. The

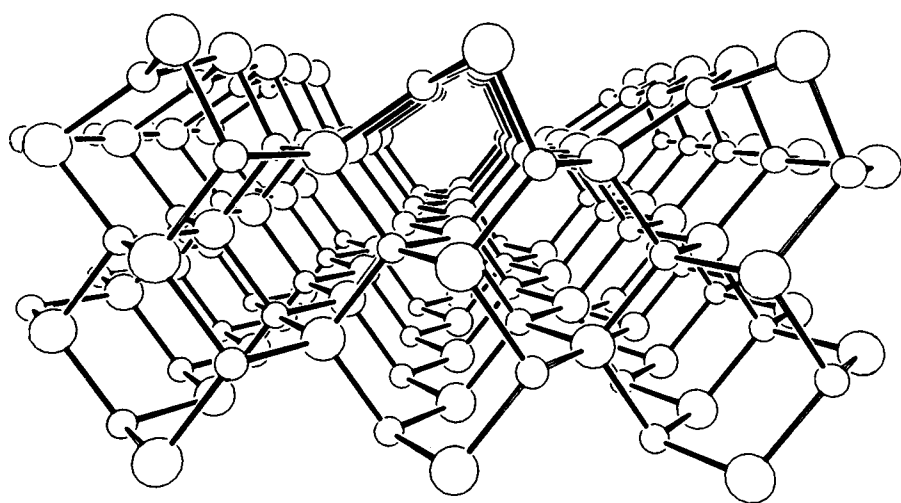


*Fig. 6. The structure of carbon monoxide chemisorbed on Ni(100) in a c(2x2) periodic arrangement. The carbon atoms are bonded to single metal atoms.*



unreconstructed GaAs (110) face consists of parallel zigzag chains of atoms. Each Ga is bound to two As atoms in the surface layer and one As atom in the layer below. GVB calculations by Goddard et al./188/ showed that the two dangling bonds broken at the surface, one on a Ga atom and one on an As atom, coalesce into one lone pair of electrons localized on the As center. The surface As has the character of normal trivalent As with a  $(4s)^2$  lone pair. Such an As atom bonds to the  $p$ -orbitals leading to optimum bond angles near  $90^\circ$ . In contrast, the surface Ga has three bonds with no electron in the fourth valence orbital. The best configuration for Ga is therefore planar, with bond angles near  $120^\circ$  (see Figure 7). Such bond angles around both Ga and As can be accommodated by a reconstruction involving a simple rotation of the surface zig-zag chains by about  $27^\circ$ . With such a reconstruction, the GVB results indeed gave an average bond angle of  $119.4^\circ$  at the Ga site and  $94.9^\circ$  at the As site, consistent with the above local bond analysis. If we define the surface strain as the projection normal to surface of the total displacement of the surface As relative to the surface Ga, the theoretical value of  $0.67 \text{ \AA}$  is found to be in good agreement with the experimental value of  $0.65\text{--}0.70 \text{ \AA}$  obtained by LEED./189,190,191,192/

When oxygen chemisorbs on GaAs (110), the authors predicted that the O atom would be promoted to the singlet state, creating a configuration with an empty  $p$ -orbital prepared to accept an electron pair. They also showed that the energy gained by the O-As bond is larger ( $\sim 2.2\text{eV}$  for  $\text{H}_3\text{AsO}$ ) than the promotion energy of the O atom from the  $^3\text{P}$  to the  $^1\text{D}$  state. Therefore, the only way oxygen can bond to the GaAs (110) surface without disrupting Ga-As bonds is for the O atoms to attach to As atoms. The computed As-O bond length is  $1.63 \text{ \AA}$  (the As-O single bond length is  $1.8 \text{ \AA}$ ), in good agreement with the EXAFS value of  $1.70 \pm 0.05 \text{ \AA}$ ./193/ The calculated dipole moment and vibrational frequency for the As-O bond are 1.00 Debye and  $125 \text{ cm}^{-1}$ . Experimentally,  $(\text{H}_3\text{C})_3\text{AsO}$  has an As-O bond length of  $1.63 \text{ \AA}$ ./194/ a vibrational frequency of  $110 \text{ meV}$ /195/ and a dipole moment of 1.14 Debye./196/ The chemical shifts of the Ga ( $3d$ ) and As ( $3d$ ) levels were also calculated for the As=O case. The authors found that Ga ( $3d$ ) shifts by  $0.8 \text{ eV}$  and the As ( $3d$ ) shifts by  $2.6 \text{ eV}$ , both to higher binding energy. The experimentally reported values for the core-level shifts are  $0.8 \text{ eV}$  for Ga ( $3d$ )/197/ and  $2.9 \text{ eV}$  for the As ( $3d$ )/.198,199,200/

**Ga As (110)**

*Fig. 7. Perspective view of the structure of the GaAs(110) surface which exhibits the 27° rotation of the zig-zag chains of surface atoms.*

The chemisorption of Al on GaAs (110) proceeds in a different way. The Al atom is an electron donor since it is not very electronegative, with a value of 1.5 on the Pauling electronegativity scale./201/ GVB energy minimization for the Al position finds the Al 2.95 Å above a Ga site, tipped very slightly towards an adjacent As (3.90 Å). At the same time it was found that the Ga site reconstructs (the surface strain decreases from 0.67 to 0.55 Å) leading to a 40% increase in net binding energy. When an Al atom binds to a surface Ga atom, the Ga atom will prefer a more tetrahedral geometry to make more use of the electrons available for bonding to its neighbors. This increases the strength of the Al-surface bond.

#### 4.7. Conclusion

The techniques of theoretical surface chemistry have advanced sufficiently to make a significant contribution to the understanding of surface structural chemistry. Even with the simplified "muffin-tin" potential,  $X\alpha$ -scattered wave calculations do a good job of predicting photoemission spectra and their variation with adsorbate geometry. Fully self-consistent *ab-initio* total-energy calculations have been able to successfully determine the structure of adsorbate bonding geometries. These calculations are still very complex, and therefore limited to small clusters. However a number of approaches to increase the efficiency of such calculations are being pursued, and over the next few years such work should make an increasing contribution to the understanding of surface structural chemistry.

### 5. RESULTS OF SURFACE CRYSTALLOGRAPHY

In this section, we present and discuss the accumulated results of surface structure determination. We give a comprehensive tabulation of published results. Only those structures which provide three-dimensional geometrical

information, including bond lengths, bond angles and layer spacings are included. They are brought together in Tables IV to XII, classified by type of surface: clean, unreconstructed metals (Table IV), alloys and reconstructed metals (Table V), chalcogens chemisorbed on metals (Table VI), other atoms adsorbed on metals (Table VII), clean semiconductors (Table VIII), atoms adsorbed on semiconductors (Table IX) and insulators (Table X), carbon monoxide chemisorbed on metals (Table XI) and other molecular adsorption on metals (Table XII). It must be stressed that not all these structures can be termed final or even reliable. However, the vast majority are most likely qualitatively correct.

A variety of surface analytical techniques have contributed to these results, as detailed in the tables. Their names, acronyms and characteristics were summarized in Table I and further described in sections 2 and 3.

We first discuss some basic aspects of surface structures, including ordering principles and notations, followed by an overview of the number and types of surface structures that have been investigated. Finally we shall highlight a few major trends emerging from the structural results.

### 5.1. Ordering Principles

The vast majority of solved surface structures have long-range two-dimensional ordering parallel to the surface. This means that the structures periodically repeat themselves in the two surface dimensions. There are several reasons for this. First, single crystals have been chosen in most cases as substrates, exposing a simple periodic lattice as a template. Second, a surprising number of surface structures can be easily prepared to remain periodic on the experimental time scale, whether in the clean state or after adsorption of atoms or molecules. Third, ordered structures are desirable for investigation because they provide the greatest simplicity in the interpretation of results. And fourth, structural determination itself is simplest in ordered cases.

The driving force for ordering originates, just as with three-dimensional crystal formation, in the mutual atomic interactions. With adsorbates, an important distinction must be made between adatom-adatom and adatom-substrate interactions. In chemisorption the adatom-adatom forces are usually

small compared to the adatom-substrate binding forces, so the adatom locations or sites are determined by the optimum adatom-substrate bonding. But the adatom-adatom interactions still manage to dominate the long-range ordering of the overlayer. These interactions can be studied by examining the phase transitions of the overlayer as a function of temperature or coverage.

The surface coverage of an adsorbate is another important parameter in ordering. We shall use the common definition of coverage where one monolayer corresponds to one adsorbate atom or molecule for each unit cell of the clean, unreconstructed substrate surface. Thus, if an adsorbed undissociated carbon monoxide molecule bonds to alternating top-layer metal atoms exposed at the Ni(100) surface, we have a coverage of a half monolayer.

At very low coverages some adsorbates bunch together in two-dimensional islands. This results from short-range attractive adsorbate-adsorbate interactions combined with easy diffusion along the surface. Other adsorbates repel each other and form disordered overlayers with atoms or molecules adsorbed in sites of a given type (a "lattice gas"). When the coverage is increased so that the mean interadsorbate distance decreases to about 5-10 Å, the mutual interactions often strongly influence the ordering, favoring certain adsorbate configurations over others. As a result, the structure can develop a unit cell that repeats periodically across the surface. This is very evident in the LEED patterns, which depend directly on this unit cell. For example, atomic oxygen on Ni(100) orders very well at a quarter of a monolayer. The oxygen atoms then occupy one quarter of the available "hollow" sites of Ni(100) in a square array labeled (2x2). When the coverage is doubled to a half monolayer, the extra oxygen atoms occupy the empty hollow sites at the center of each of the (2x2) squares, thereby creating a new pattern labeled c(2x2) (c=centered). The corresponding LEED patterns are fundamentally different in that the (2x2) diffraction pattern has twice as many spots as the c(2x2) pattern, which in turn has twice as many spots as the clean-surface pattern, labeled (1x1).

Most nonmetallic adatoms will not compress into a one-monolayer overlayer on the closest-packed metal substrates. There appears to be a short-range repulsion that keeps adatoms apart by approximately a Van der Waals distance. Attempts to compress the overlayer further by increasing the coverage (which is

done by exposing the surface to the corresponding gas) result either in no further adsorption or in diffusion of the adatoms into the substrate, forming compounds.

Some adsorbates do not form strong chemical bonds with substrate atoms. This case is called physisorption. For these adsorbates the adsorbate-adsorbate interactions can dominate the adsorbate-substrate interactions, and the optimum adatom-substrate bonding geometry can be overridden by the lateral adatom-adatom interactions, yielding for example incommensurate structures, in which the overlayer and the substrate have independent lattices. When adsorbates are used which physisorb rather than chemisorb (at suitably low temperatures), one also finds that the Van der Waals distance determines the densest overlayer packing. In this case, Furthermore, with physisorption a larger coverage is possible through multilayer formation.

With metallic adsorbates very close-packed overlayers can be formed, because metal adsorbate atoms attract each other relatively strongly and coalesce with covalent interatomic distances. When the atomic sizes of the overlayer and substrate metals are nearly the same, one observes one-monolayer (1x1) structures, where adsorbate atoms occupy every unit cell of the substrate. With less equal atomic radii, other structures are formed, dominated by the covalent closest packing distance of the adsorbate. Beyond one close-packed overlayer, metal adsorbates frequently form multilayers or also three-dimensional crystallites. Alloy formation by interdiffusion is also observed in a number of cases, even in the submonolayer regime.

## 5.2. Notation for surface lattices

In Tables IV to XII the substrate surface is defined by the Miller indices of the corresponding crystallographic plane.

Unless there is reconstruction, this clean surface is labeled (1x1) to denote that the surface lattice is the same as it would be if the bulk lattice were cut along a crystallographic plane (the "terminated bulk surface"). In the case of reconstruction or with overlayers, the surface symmetry and periodicity is often different from the (1x1) lattice. Often this new periodicity is "commensurate"

with the (1x1) substrate lattice, i.e. there is a simple rational relationship between the two lattices. Such a "superlattice" may be labeled by the notation

$$c(v \times w)R\alpha \text{ or } p(v \times w)R\alpha \quad (47)$$

Here  $v$  and  $w$  are elongation factors of the (1x1) lattice vectors; they indicate that the non-centered ( $v \times w$ ) lattice has unit cell edges that are  $v$  and  $w$  times the lengths of the (1x1) unit cell edges. The quantity  $\alpha$  is an angle through which the (1x1) lattice must be rotated to coincide with the superlattice (the  $R\alpha$  is omitted when  $\alpha=0$ ). The prefix  $p$  denotes a primitive lattice (it is often omitted), while  $c$  denotes a centered lattice. As an example, the above-mentioned (2x2) lattice can also be written  $p(2 \times 2)$ , whereas the  $c(2 \times 2)$  lattice can be given equally by  $(\sqrt{2} \times \sqrt{2})R45^\circ$  or  $p(\sqrt{2} \times \sqrt{2})R45^\circ$ . This so-called Wood notation assumes that the rotation angle  $\alpha$  is the same between the different lattice vectors of the (1x1) lattice and the non-centered ( $v \times w$ ) lattice. For instance, it assumes that the ( $v \times w$ ) lattice is rectangular when the (1x1) lattice is rectangular. (Some authors use the Wood notation even when the rotation angle for the two superlattice vectors is *not* the same. In this case the notation is not well defined. The matrix notation described below should be used for such cases.)

The more general case can be handled with a "matrix notation", which we will not need here. But there is a frequently occurring case which has its own notation. On a substrate with a six-fold rotationally symmetric (1x1) lattice (such as with fcc(111) surfaces), one often finds rectangular superlattices. These can be labeled

$$c(v\sqrt{3} \times w)rect \text{ or } p(v\sqrt{3} \times w)rect \quad (48)$$

indicating that the superlattice unit cell is rectangular and has edges of lengths  $v\sqrt{3}$  and  $w$  times the (1x1) cell edges.

### 5.3. Numbers and types of solved surface structures

If we limit ourselves to observed LEED patterns, we find that over the years about 2000 ordered structures have been reported./202/ Among these, perhaps 180 have been structurally solved by various techniques of surface crystallography. Intensity analyses of low-energy electron diffraction have contributed about 150 of these. The remaining 30 structures were obtained primarily with ion scattering (MEIS, HEIS), SEXAFS or photoelectron diffraction (NPD, ARXPS).

A breakdown of the structural results by type of surface shows results for nearly 50 clean, unreconstructed metal surfaces and about 10 alloys and reconstructed metal surfaces. The structures of about 65 atomic overlayers on metal surfaces have been determined, some 40 of these involving chalcogen atoms. Just over 20 molecular structures have been determined for metal surfaces, half of these being overlayers of undissociated carbon monoxide and the others various hydrocarbons. Turning to semiconductors, some 13 clean, usually reconstructed structures were determined, against nearly 10 atomic overlayer structures. In addition, about 15 insulator surface structures have been investigated.

### 5.4. Major trends among surface structures

#### 5.4.1. Bond length contractions and reconstructions

A general observation at surfaces is a tendency for bond lengths to decrease as the bonding coordination number decreases. This trend fits long-established principles (cf. Pauling), if one relates coordination number to bond order (cf. Mitchell). The clearest manifestation is provided bond length relaxations at clean metal surfaces. As one compares close-packed with less close-packed surfaces of metals, one finds that the latter present a smaller interlayer spacing between the topmost and the second atomic layers, compared with bulk values. The



corresponding bond length contraction is of the order of 2 to 3%. Moreover, the perturbation caused by this surface relaxation propagates a few layers down into the surface. In fact, there is a compensating expansion between the second and third metal layers of the order of 1%, and a small but detectable change in the next layer. Appropriately, such relaxations are largely removed by the chemisorption of atoms on top of these surfaces.

Another manifestation of bond length contractions at surfaces compared to the bulk metal are certain reconstructions, where the topmost metal layer adopts a different structure and symmetry than the underlying layers. For instance, with Ir, Pt and Au(100), the interatomic distance in the topmost layer shrinks by a few percent parallel to the surface. It then becomes more favorable for this layer to collapse into a hexagonally close-packed layer rather than maintaining the square lattice of the underlying layers. Many adsorbates can reverse this reconstruction by removing the driving force towards smaller bond lengths.

In the case of semiconductor surfaces, a more dominant disturbance is the difficulty of surface atoms to compensate for the loss of nearest neighbors. The "dangling bonds" created at the surface cannot easily be satisfied by bonding to neighboring surface atoms, except through more drastic rearrangements of these atoms. Therefore, most semiconductor surfaces reconstruct. Major rebonding between surface atoms occurs in this process. The associated perturbation propagates several layers into the surface until the bulk lattice is recovered. Here again, adsorbates can reverse the reconstruction and induce a return to the bulk structure.

#### 5.4.2. Atomic penetration into substrates

Of major interest for such processes as oxidation and compound formation are results that show several stages of atomic penetration into the substrate lattice. At metal surfaces, this appears to occur mostly with small atoms: oxygen, nitrogen, carbon, as well as hydrogen. In some instances, as the coverage is increased, first an overlayer is formed (e.g. O on Ni(100)). Usually, the adatom occupies high-coordination sites such as "hollow" sites. On less close-packed surfaces, this site puts the adatom nearly coplanar with the surface metal atoms

(e.g. O on Fe(100)). This may be viewed as a single layer of the metal-adatom compound. The next stage of penetration is illustrated by N on Ti(0001), where the adatom occupies interstitial sites between the first and second metal layers, thereby forming a three-layer film of the bulk compound TiN. Deeper penetration is often observed in the form of thicker compound films which show no trace of the structural properties of the parent metal.

In the case of semiconductors, there is more interstitial space available for diffusion of adatoms into the surface. The very few cases whose structure is known show, however, that a substitutional arrangement may be the more stable equilibrium configuration. Thus, with Al on GaAs(110), Al atoms tend to occupy the original Ga and As positions.

#### 5.4.3. Molecular adsorption

In the case of molecules adsorbed at surfaces, it must be first stated that much important information is obtained from high-resolution electron energy loss spectroscopy (HREELS). This technique measures vibration frequencies of surfaces, in a way similar to infra-red absorption spectroscopy in the gas phase. HREELS allows the identification of the molecular species present on the surface, which no surface crystallography method can do.

Carbon monoxide has proved a popular and convenient molecular adsorbate. It provides a rich variety of behavior at surfaces, while being relatively easy to study by various methods. Toward the left of the periodic table, metal surfaces increasingly tend to dissociate CO, then separate C and O atoms bond directly in individual hollow sites, as in atomic adsorption. On other metal substrates, CO remains intact and bonds through its carbon end to the surface, with the C-O axis perpendicular to the surface. However, the adsorption site varies considerably. CO most commonly adsorbs in one-fold coordinated top sites or two-fold coordinated bridge sites. Occasionally, three-fold coordinated hollow sites are also found. The site depends on several factors: the metal, the crystallographic face and the CO coverage. Another factor has recently been found to influence the CO adsorption site -- coadsorbed atomic or molecular species which donate charge to the CO can shift CO to a higher-coordination site.

An example is the coadsorption of benzene, whose  $\pi$ -orbitals give up charge to the metal and to nearby CO species. On Rh(111) this shifts CO adsorption to three-fold sites.

There is a close similarity between the bonding geometry of CO on metal surfaces and in metal-carbonyl complexes. In both cases, for example, the metal-carbon bond length increases markedly with the coordination number of the site where the CO molecule bonds to the metal. The C-O bond length tends to increase slightly at the same time, indicating a C-O bond weakening, which is amplified by nearby electron donors.

If we consider hydrocarbons at surfaces, we must set the saturated ones aside. These do not chemisorb at metal surfaces, but rather physisorb at low temperatures. No detailed structural determination has been performed for saturated hydrocarbons on metals. However, some structures have been obtained using the graphite basal plane as a substrate (cf. Suzanne etc) which give information about Van der Waals packing of such molecules, but no chemical reactions are involved.

With unsaturated hydrocarbons, whether straight-chain or aromatic, two basic types of surface structure can be distinguished. The first consists of the intact molecule bonding to the metal surface by sharing its C-C multiple bonds with the metal. Thus, acetylene and ethylene adsorb with their C-C bonds parallel to the metal surface. The C-C bond is believed to be considerably lengthened, indicating bond-order reduction in favor of the newly formed metal-carbon bonds. The associated rehybridization around the carbon atoms may also result in the hydrogen atoms bending away from the metal. Similarly, benzene bonds with its carbon ring parallel to the surface, and large expansions of the C-C distances have been observed. Such distortions may be a forerunner of molecular dissociation.

The second type of structure for unsaturated hydrocarbons is the result of a hydrogen rearrangement in straight-chain molecules. Alkylidynes are formed, in which a terminal carbon bonds to three metal atoms, while the rest of the chain is saturated and points away from the metal (extra hydrogen atoms may be necessary for these species to form). These alkylidynes, foremost among which is ethylidyne ( $C_2H_3$ ), are quite stable in comparison to other organic species formed at room temperature transition metal substrates. However, at higher

temperatures, a progressive dehydrogenation occurs, leaving CH and CCH fragments, then carbonaceous fragments and ultimately graphite and bulk carbides on the surface.

Both types of hydrocarbon adsorption are also encountered in organometallic complexes. Here again, the geometries found at surfaces often correspond closely to those observed in complexes, including agreement of bond lengths and bond angles. There probably are larger differences between surfaces and complexes in the relative occurrence of the various species.

TABLE IV. Clean Metal Structures (unreconstructed)

(Where multiple layer spacing changes have been investigated these are listed in the table on successive lines.)

Substrate Face	Bulk Spacing(Å)	Surface Spacing(Å)	Expansion(%)	Method
Ag (110) fcc	1.44	1.34	-7	LEED/1/
Ag (110) fcc		1.33±0.04	-7.6	HEIS/2/
		1.50±0.04	4.2	
Ag (111) fcc	2.35	2.35±0.1	0	HEIS/3/
Al (100) fcc	2.02	2.025±0.10	0.0	LEED/4/
Al (100) fcc		2.02	0.0	LEED/5/
Al (100) fcc		2.052	1.0	MEED/6/
Al (110) fcc	1.43	1.30	-9.1	LEED/5/
Al (110) fcc		1.304±0.012	-8.8	LEED/7/
		1.499±0.15	4.8	
		1.404±0.017	-1.8	
		1.429±0.018	0.0	
Al (110) fcc		1.310±0.014	-8.4	LEED/8/
		1.510±0.016	5.6	
		1.463±0.019	2.3	
		1.455±0.022	1.7	
Al (111) fcc	2.33	2.350±0.012	0.9	LEED/9/
Al (111) fcc		2.41±0.05	3.1	LEED/10/
Al (311) fcc	1.23	1.68±0.01	-13.0	LEED/11/
		1.335±0.02	8.8	
Al (331) fcc	0.93	0.82±0.02	-11.7	LEED/12/
		0.89±0.03	-4.1	
		1.05±0.03	10.3	
		0.88±0.03	-4.8	
First-layer registry shifts of $\sim 0.06\pm 0.08$ found				

(continued)

TABLE IV (continued)

Substrate Face	Bulk Spacing(Å)	Surface Spacing(Å)	Expansion(%)	Method
Au (100) fcc (metastable)	2.04	2.04	0.0	LEED/13/
Cd (100) fcc	2.81	2.81	0	LEED/14/
Co (100) fcc	1.77	1.70	-4.0	LEED/15/
Co (111) fcc	2.05	2.05±0.05	0.0	LEED/16/
Co (0001) hcp	2.05	2.05±0.05	0.0	LEED/16/
Co (11 $\bar{2}$ 0) hcp	1.25	1.14±0.04	-8.8	LEED/17/
Cu (100) fcc	1.81	1.785	-1.1	LEED/18/
		1.836	1.7	
		1.832	1.5	
Cu (110) fcc	1.28	1.159	-9.2	LEED/18/
		1.305	2.3	
Cu (110) fcc		1.21±0.02	-5.3	MEIS/19/
		1.32±0.02	3.3	
Cu (110) fcc		1.170±0.008	-8.5	LEED/20/
		1.307±0.010	2.3	
Cu (111) fcc	2.09	2.076±0.02	-0.7	LEED/21/
Cu (311) fcc	1.09	1.035±0.02	-5.0	LEED/22/
Fe (100) bcc	1.43	1.41±0.04	-1.6	LEED/23/
Fe (110) bcc	2.02	2.04±0.04	0.5	LEED/24/
Fe (111) bcc	0.83	0.70±0.03	-15.4	LEED/25/
Fe (111) bcc		0.69±0.025	-16.6	LEED/26/
		0.75±0.025	-9.3	
		0.86±0.03	4.0	
Fe (210)* bcc	0.64	0.81±0.025	-2.1	LEED/27/
		0.50±0.03	-21.9	
		0.57±0.03	-10.9	

(continued)

TABLE IV (continued)

Substrate Face	Bulk Spacing(Å)	Surface Spacing(Å)	Expansion(%)	Method
		0.61±0.03	-4.7	
		0.64±0.03	0.0	
Fe (211)* bcc	1.17	1.05±0.03	-10.3	LEED/28/
		1.23±0.03	5.1	
		1.15±0.04	-1.7	
Fe (310) bcc	0.906	0.76±0.03	-16.1	LEED/29/
		1.02±0.03	12.6	
		0.87±0.04	-4.0	
Ir (100) fcc (metastable)	1.92	1.85±0.01	-3.6	LEED/30/
Ir (110) fcc	1.36	1.26±0.10	-7.4	LEED/31/
Ir (111) fcc	2.22	2.16±0.10	-2.6	LEED/32/
Na (110) bcc	3.03	3.03	0.0	LEED/33/
Na (110) bcc		3.0±0.01	-1.0	LEED/34/
Na (0001) hcp	2.87	2.87	0.0	LEED/35/
Ni (100) fcc	2.22	1.78±0.02	1.1	LEED/36/
Ni (100) fcc		1.604±0.008	-8.9	MEIS/37/
Ni (110) fcc	1.25	1.195±0.01	-4.0	MEIS/38/
Ni (110) fcc		1.18±0.02	-4.8	HEIS/39/
		1.27±0.02	2.4	
Ni (110) fcc		1.14±0.01	-8.1	LEED/40/
		1.28±0.01		
Ni (110) fcc		1.123±0.02	-9.8	LEED/41/
		1.292±0.02	3.8	
Ni (110) fcc		1.138±0.006	-8.6	LEED/42/
		1.284±0.007	3.1	
		1.240±0.009	-0.4	

(continued)

TABLE IV (continued)

Substrate Face	Bulk Spacing(Å)	Surface Spacing(Å)	Expansion(%)	Method
Ni (110) fcc		1.121±0.012	-9.0	MEIS/43/
		1.290±0.019	3.5	
Ni (111) fcc	2.03	2.005±0.025	-1.2	LEED/36/
Ni (111) fcc		2.033±0.020	0.0	HEIS/44/
Ni (311) fcc	1.06	0.894±0.010	-15.9	LEED/45/
		1.106±0.016	4.1	
		1.045±0.017	-1.6	
Mo (100) bcc	1.25	1.424±0.03	-9.5	LEED/46/
Mo (110) bcc	2.22	2.19±0.04	-1.6	LEED/47/
Pd (100) fcc	1.94	1.95±0.05	0.3	LEED/48/
Pd (110) fcc	1.37	1.29±0.03	-5.8	LEED/49/
		1.38±0.03	0.7	
Pd (111) fcc	2.246	2.25±0.05	0.0	HEIS/50/
Pd (111) fcc		2.276	1.3	LEED/51/
		2.216	-1.3	
		2.296	2.2	
		2.296	2.2	
Pt (100) fcc (metastable)	1.96	1.96	0.2	SPLEED/52/
Pt (100) fcc (metastable)		1.963	0.2	HEIS/53/
Pt (100) fcc (metastable)		1.96	0.0	LEED/13/
Pt (111) fcc	2.26	2.29±0.10	1.1	LEED/54/
Pt (111) fcc		2.30±0.02	1.4	HEIS/55/
Pt (111) fcc		2.276±0.02	0.5	SPLEED/56/
Pt (111) fcc		2.265±0.05	0.0	LEED/57/

(continued)



TABLE IV (continued)

Substrate Face	Bulk Spacing(Å)	Surface Spacing(Å)	Expansion(%)	Method
Re (10 $\bar{1}$ 0) hcp		0.67	-16.3	LEED/58/
Rh (100) fcc	1.90	1.91 $\pm$ 0.02	0.5	LEED/59/
Rh (110) fcc	1.34	1.33 $\pm$ 0.02	-0.7	LEED/59/
Rh (111) fcc	2.19	2.16 $\pm$ 0.02	-1.4	LEED/59/
Rh (111) fcc		2.192 $\pm$ 0.10	0.0	LEED/60/
Ru (0001) hcp	2.14	2.10 $\pm$ 0.02	-1.9	LEED/61/
Sc (0001) hcp	2.64	2.59 $\pm$ 0.02	-1.9	LEED/62/
Ta (100) bcc	1.65	1.47 $\pm$ 0.03	-10.9	LEED/63/
		1.67 $\pm$ 0.03	1.2	
Te (10 $\bar{1}$ 0) hcp		0.34 $\pm$ 0.10		LEED/64/
Ti (0001) hcp	2.34	2.29 $\pm$ 0.05	-2.1	LEED/65/
V (100) bcc	1.51	1.41 $\pm$ 0.01	-6.6	LEED/66/
V (110) bcc	2.14	2.13 $\pm$ 0.10	-0.5	LEED/67,68/
W (100) fcc	1.58	1.46 $\pm$ 0.03	-7.6	LEED/69/
W (110) fcc	2.23	2.23 $\pm$ 0.10	0.0	LEED/70/
Zn(0001) hcp	2.44	2.39 $\pm$ 0.05	-2.0	LEED/71/
Zr (0001) hcp	2.57	2.54 $\pm$ 0.05	-1.2	LEED/72/

\* There are relaxations in the layer registries for the stepped iron (211) and (210) surfaces in addition to layer spacing relaxations.

## REFERENCES

1. Zanazzi, E., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Journal of Physics C* **10**, 375.
2. Kuk, Y. and Feldman, L. C. (1984) *Physical Review B* **30**, 5811.
3. Culbertson, R. J., Feldman, L. C., Silverman, P. J., and Boehm, H. (1981) *Physical Review Letters* **47**, 657.
4. Van Hove, M. A., Tong, S. Y., and Stoner, N. (1976) *Surface Science* **54**, 259.
5. (1977) *Surface Science* **62**, 567.
6. Masud, N., Baudoing, R., Aberdam, D., and Gaubert, C. (1983) *Surface Science* **133**, 580.
7. Andersen, J. N., Nielsen, H. B., Petersen, L., and Adams, D. L. (1984) *Journal of Physics C* **17**, 173.
8. Noonan, J. R. and Davis, H. L. (1984) *Physical Review B* **29**, 4349.
9. Nielsen, H. B. and Adams, D. L. (1982) *Journal of Physics C* **15**, 615.
10. Martinez, V., Soria, F., Munoz, M. C., and Sacedon, J. L. (1983) *Surface Science* **128**, 424.
11. Noonan, J. R., Davis, H. L., and Erley, W. (1985) *Surface Science* **152/153**, 142.
12. Adams, D. L. and Sorensen, C. S. (1986) *Surface Science* **166**, 495.
13. Lang, E., Grimm, W., and Heinz, K. (1982) *Surface Science* **177**, 169.
14. Shih, H. D., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Physical Review B* **15**, 5561.
15. Maglietta, M., Zanazzi, E., Jona, F., Jepsen, D. W., and Marcus, P. M. (1978) *Applied Physics* **15**, 409.
16. Lee, B. W., Alsenz, R., Ignatiev, A., and Van Hove, M. A. (1978) *Physical Review B* **17**, 1510.
17. Weltz, M., Moritz, W., and Wolf, D. (1983) *Surface Science* **125**, 473.
18. Davis, H. L. and Noonan, J. R. (1983) *Surface Science* **126**, 245.
19. (1983) *Surface Science* **128**, 281.
20. Adams, D. L., Nielsen, H. B., and Andersen, J. N. (1983) *Surface Science* **128**, 294.

21. Lindgren, S. A., Walldén, L., Rundgren, J., and Westrin, P. (1984) *Physical Review B* **29**, 576.
22. Streater, R. W., Moore, W. T., Watson, P. R., Frost, D. C., and Mitchell, K. A. R. (1978) *Surface Science* **72**, 744.
23. Legg, K. O., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Journal of Physics C* **10**, 937.
24. Shih, H. D., Jona, F., and Marcus, P. M. (1980) *Journal of Physics C* **13**, 3801.
25. Shih, F. E., Jona, F., Jepsen, D. W., and Marcus, P. M. (1981) *Physical Review Letters* **46**, 731.
26. Sokolov, J., Jona, F., and Marcus, P. M. (1986) *Physical Review B* **33**, 1397.
27. Sokolov, J., Jona, F., and Marcus, P. M. (1985) *Physical Review B* **31**, 1929.
28. Sokolov, J., Shih, H. D., Bardi, U., and Jona, F. (1984) *Journal of Physics C* **17**, 371.
29. Sokolov, J., Jona, F., and Marcus, P. M. (1984) *Physical Review B* **29**, 5402.
30. Heinz, K. and Besold, G. (1983) *Surface Science* **125**, 515.
31. Chan, C. M., Cunningham, S. M., Luke, K. L., Weinberg, W. H., and Withrow, S. P. (1978) *Surface Science* **78**, 15.
32. Chan, C. M., Cunningham, S. L., Van Hove, M. A., Weinberg, W. H., and Withrow, S. P. (1977) *Surface Science* **66**, 394.
33. Echenique, P. M. (1976) *Journal of Physics C* **9**, 3193.
34. Andersson, S., Pendry, J. B., and Echenique, P. M. (1977) *Surface Science* **65**, 539.
35. Lindgren, S. A., Paul, J., Walldén, L., and Westrin, P. (1982) *Journal of Physics C* **15**, 6285.
36. Demuth, J. E., Marcus, P. M., and Jepsen, D. W. (1975) *Physical Review B* **11**, 1460.
37. Frenken, J. W. M., van der Veen, J. F., and Allan, G. (1983) *Physical Review Letters* **51**, 1876.
38. van der Veen, J. F., Tromp, R. M., Smeenk, R. G., and Saris, F. W. (1979) *Surface Science* **82**, 468.

39. (1983) *Surface Science* **134**, 329.
40. Gauthier, Y., Baudoin, R., Joly, Y., and Gaubert, C. (1984) *Journal of Physics C* **17**, 4547.
41. Xu, M. L. and Tong, S. Y. (1985) *Physical Review B* **31**, 6332.
42. Adams, D. L., Peterson, L. E., and Sorenson, C. S. (1985) *Journal of Physics C* **18**, 1753.
43. Yalisove, S. M., Graham, W. R., Adams, E. D., Copel, M., and Gustafsson, T. (1986) *Surface Science* **171**, 400.
44. Narusawa, T., Gibson, W. M., and Törnquist, E. (1981) *Physical Review Letters* **47**, 417.
45. Adams, D. L., Moore, W. T., and Mitchell, K. A. R. (1985) *Surface Science* **149**, 407.
46. Clarke, L. J. (1980) *Surface Science* **91**, 131.
47. Morales, L., Garza, D. O., and Clarke, L. J. (1981) *Journal of Physics C* **14**, 5391.
48. Behm, R. J., Christmann, K., Ertl, G., and Van Hove, M. A. (1980) *Journal of Chemical Physics* **73**, 2984.
49. Barnes, C. J., Ding, M. Q., Lindroos, M., Diehl, R. D., and King, D. A. (1985) *Surface Science* **162**, 59.
50. Kuk, Y., Feldman, L. C., and Silverman, P. J. (1983) *Physical Review Letters* **50**, 511.
51. Ohtani, H., Van Hove, M. A., and Somorjai, G. A. (to be published)
52. Feder, R. (1977) *Surface Science* **68**, 229.
53. Davies, J. A., Jackman, T. E., Jackson, D. P., and Norton, P. R. (1981) *Surface Science* **109**, 20.
54. Adams, D. L., Nielsen, H. B., and Van Hove, M. A. (1979) *Physical Review B* **20**, 4789.
55. van der Veen, J. F., Smeenk, R. G., Tromp, R. M., and Saris, F. W. (1979) *Surface Science* **79**, 219.
56. Feder, R., Pleyer, H., Bauer, P., and Müller, N. (1981) *Surface Science* **109**, 419.
57. Hayek, K., Glassl, H., Gutmann, A., Leonhard, H., Prutton, M., Tear, S. P., and Welton-Cook, M. R. (1985) *Surface Science* **152**, 419.

58. Davis, H. L. and Zehner, D. M. (1980) *Journal of Vacuum Science and Technology* **17**, 190.
59. Hengrasmee, S., Mitchell, K. A. R., Watson, P. R., and White, S. J. (1980) *Canadian Journal of Physics* **58**, 200.
60. Van Hove, M. A. and Koestner, R. J. (1984) in "Surface Structure by LEED" (P. M. Marcus and F. Jona, ed.), Plenum, New York.
61. Michalk, G., Moritz, W., Pfnür, H., and Menzel, D. (1983) *Surface Science* **129**, 92.
62. Tougaard, S., Ignatiev, A., and Adams, D. L. (1982) *Surface Science* **115**, 270.
63. Titov, A. and Moritz, W. (1982) *Surface Science* **123**, L709.
64. Meyer, R. J., Salaneck, W.R., Duke, C. B., Paton, A., Griffiths, C. H., Kovnat, L., and Meyer, L. E. (1980) *Physical Review B* **21**, 4542.
65. Shih, H. D., Jona, F., Jepsen, D. W., and Marcus, P. M. (1976) *Journal of Physics C* **9**, 1405.
66. Jensen, V., Andersen, J.N., Nielsen, H. B., and Adams, D. L. (1982) *Surface Science* **116**, 66.
67. Adams, D. L. and Nielsen, H. B. (1981) *Surface Science* **107**, 305.
68. Adams, D. L. and Nielsen, H. B. (1982) *Surface Science* **116**, 598.
69. Debe, M. K. and King, D. A. (1982) *Journal of Physics C* **15**, 2257.
70. Van Hove, M. A. and Tong, S. Y. (1976) *Surface Science* **54**, 91.
71. Unertl, W. N. and Thapliyal, H. V. (1975) *Journal of Vacuum Science and Technology* **12**, 263.
72. Moore, W. T., Watson, P. R., Frost, D. C., and Mitchell, K. A. R. (1979) *Journal of Physics C* **12**, L887.

TABLE V. Alloys and Reconstructed Metals

Substrate and Unit Cell	Structure	Method
Au (110) (fcc) (2x1)	Missing-row reconstruction confirmed by x-ray diffraction. Lateral displacement of second layer rows toward missing row position by $0.12 \pm 0.02$ Å. Anti-phase domains due to monatomic steps observed. Evidence for top-layer expansion or contraction of $\sim 0.25$ Å.	X-Ray Diffraction/1/
Au (110) (fcc) (2x1)	Missing-row reconstruction. Every other top layer ( $1\bar{1}0$ ) row is missing. The second layer lateral displacements are $\sim 0.07$ Å, and the second row is buckled by $-0.24$ Å. The first layer spacing is $-20.1\%$ , and the second and third is $-6.3\%$ relative to the $1.44$ Å bulk layer spacing.	LEED/2/
Au (110) (fcc) (2x1)	Missing-row reconstruction studied by channeling and blocking, top layer spacing contracted by $18\%$ , second layer spacing expanded by $4\%$ from bulk value of $1.44$ Å. Third layer buckled by $\pm 0.10$ Å or $\pm 7\%$ , with atoms under missing row moving up, alternate rows down.	MEIS/3/
Au (110) (fcc) (2x1)	Missing-row reconstruction studied with top layer spacing contracted by $0.15 \pm 0.10$ Å ( $-11\%$ ) from bulk value of $1.44$ Å.	LEIS/4/

(continued)

TABLE V (continued)

Substrate and Unit Cell	Structure	Method
$\alpha$ -CuAl (111) (fcc) ( $\sqrt{3}\times\sqrt{3}$ )R30°	Al substituted in 1/3 of top layer Cu sites, second layer pure Cu, no buckling in top layer, layer spacing is $2.05\pm0.05$ Å, the same as bulk copper. Alloy composition 16% Al atoms in Cu.	LEED/5/
Ir (100) (fcc) (5x1)	The top layer of the surface reconstructs to form a compact hexagonal surface, with 6/5 the density of unreconstructed surface. The layer spacing expands by $14.6\pm5.2\%$ from the bulk value of 1.92 Å. Some top-layer atoms are buckled outward by up to an additional $0.2\pm0.02$ Å so the hexagonal layer can fit the square layer below. This is 1/2 to 2/3 of the buckling required to have top layer atoms in hard-sphere contact with all substrate atoms.	LEED/6/
Ir (100) (fcc) (5x1)	The top layer of the surface reconstructs to form a compact hexagonal surface, with 6/5 the density of unreconstructed surface. The layer spacing expands by $7.3\pm2.6\%$ from the bulk value of 1.92 Å. Some top-layer atoms are buckled outward by up to an additional $0.48\pm0.02$ Å so the hexagonal layer can fit the square layer below. This puts the top layer atoms in hard-sphere contact with all substrate atoms.	LEED/7/

(continued)

TABLE V (continued)

Substrate and Unit Cell	Structure	Method
Ir (110) (fcc) (2x1)	Missing-row reconstruction similar to gold (110). One top layer row is missing, the second layer lateral displacements are $\sim 0.04 \text{ \AA}$ , and the second row is buckled by $-0.23 \text{ \AA}$ . The first layer spacing is $-12.4\%$ , and the second and third $-5.8\%$ relative to the $1.36 \text{ \AA}$ bulk layer spacing.	LEED/8/
NiAl (110) (CsCl) (1x1)	Top layer spacing $1.92 \text{ \AA}$ , contracted $6\%$ from bulk value of $2.04 \text{ \AA}$ . Al atoms buckled out by $0.22 \text{ \AA}$ .	LEED/9/
Ni <sub>3</sub> Al (Cu <sub>3</sub> Au) (100)	Top layer is 50-50 nickel and aluminum, second layer nickel, etc. Top layer spacing is $1.73 \pm 0.03 \text{ \AA}$ with Al atoms buckled outward by $0.02 \pm 0.03 \text{ \AA}$ , second layer spacing is the bulk value of $1.78 \text{ \AA}$ within $\pm 0.03 \text{ \AA}$ .	LEED/10/
Pd (110) (fcc) (2x1)	Missing-row model with third layer assumed bulk-like. Second layer found to be bulk-like, first layer spacing $-0.1\%$ relative to $1.37 \text{ \AA}$ bulk spacing. Saw-tooth model almost as good.	LEED/11/
Pt (100) (fcc)	Top layer reconstruction, quasi-hexagonal surface given by $\begin{pmatrix} 141 \\ -15 \end{pmatrix}$ , also called "1x5" or "4x20" reconstruction. Degree of top-layer buckling and top-layer registry proposed.	LEED/6/

(continued)



TABLE V (continued)

Substrate and Unit Cell	Structure	Method
Pt (110) (fcc) (2x1)	Missing-row model with the top layer expanded by 23% to 1.4 Å from bulk value of 1.387 Å. Lateral shift of second layer rows of 0.05 Å.	LEED/12/
Pt (110) (fcc) (2x1)	Missing-row model is better than buckled-row, paired-row or saw-tooth models. Terminated-bulk positions assumed for remaining atoms.	LEIS/13/
Pt <sub>z</sub> Ni <sub>1-z</sub> (110) (fcc)	LEED study of random substitutional alloy shows Pt enhancement in 1 <sup>st</sup> and 3 <sup>rd</sup> layers, along with a Pt depletion in the 2 <sup>nd</sup> layer.	LEED/14/
W (100) (bcc) (2x1)	Below room temperature the "zig-zag" reconstruction occurs on the W(100) surface. Alternate atoms move along the (011) and (0 $\bar{1}\bar{1}$ ) directions by $\pm 0.22 \pm 0.07$ Å, while the top layer spacing contracts by $0.05 \pm 0.05$ Å from the bulk value of 1.58 Å. There are two domains, since atoms may also move along the (0 $\bar{1}\bar{1}$ ) and (01 $\bar{1}$ ) directions.	LEED/15/
W (100) (bcc) ( $\sqrt{2} \times \sqrt{2}$ )R45°	Below room temperature the "zig-zag" reconstruction occurs on the W(100) surface. Alternate atoms move along the (011) and (0 $\bar{1}\bar{1}$ ) directions by $\pm 0.16$ Å, while the top layer spacing contracts by 0.09 Å (6%) from the bulk value of 1.58 Å.	LEED/16/

## REFERENCES

1. Robinson, I. K., Kuk, Y., and Feldman, L. C. (1984) *Physical Review* **B29**, 4762.
2. Moritz, W. and Wolf, D. (1985) *Surface Science* **163**, L655.
3. Copel, M. and Gustafsson, T. (1986) *Physical Review Letters* **57**, 723.
4. Möller, J., Niehaus, H., and Heiland, W. (1986) *Surface Science* **166**, L111.
5. Baird, R. J., Ogletree, D. F., Van Hove, M. A., and Somorjai, G. A. (1986) *Surface Science* **165**, 345.
6. Van Hove, M. A., Koestner, R. J., Stair, P. C., Bibérian, J. P., Kesmodel, L. L., Bartoš, I., and Somorjai, G. A. (1981) *Surface Science* **103**, 218.
7. Lang, E., Müller, K., Heinz, K., Van Hove, M. A., Koestner, R. J., and Somorjai, G. A. (1983) *Surface Science* **127**, 347.
8. Chan, C.-M. and Van Hove, M. A. (to be published) *Surface Science*
9. Davis, H. L. and Noonan, J. R. (1985) *Physical Review Letters* **54**, 566.
10. Sondericker, D., Jona, F., and Marcus, P. M. (1986) *Physical Review B* **33**, 900.
11. Barnes, C. J., Ding, M. Q., Lindroos, M., Diehl, R. D., and King, D. A. (1985) *Surface Science* **162**, 59.
12. Adams, D. L., Nielsen, H. B., Van Hove, M. A., and Ignatiev, A. (1981) *Surface Science* **104**, 47.
13. Niehaus, H. and Comsa, G. (1984) *Surface Science* **140**, 18.
14. Gauthieer, Y., Baudoing, R., Joly, Y., Rundgren, J., Bertolini, J. C., and Massardier, J. (1985) *Surface Science* **162**, 348.
15. Barker, R. A., Estrup, P. J., Jona, F., and Marcus, P. M. (1978) *Solid State Communications* **25**, 375.
16. Walker, J. A., Debe, M. K., and King, D. A. (1981) *Surface Science* **104**, 405.

TABLE VI. Chalcogen Chemisorption on Metals

(Where relaxation of the metal layer spacing has been investigated this is listed in the table after the adsorbate information.)

Substrate Face	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Bond Length	Method
OXYGEN					
Ag (110)	(2x1)	long bridge	0.20	2.05	SEXAFS/1/
Al (111)	(1x1)	3-fold fcc	$0.70 \pm 0.12$	$1.79 \pm 0.05$	LEED/2/
Al (111)	(1x1)	3-fold fcc	$0.60 \pm 0.10$	$1.75 \pm 0.03$	NEXAFS/3/
Al (111)	(1x1)	sub-surface tetrahedral	$0.60 \pm 0.10$	$1.75 \pm 0.03$	NEXAFS/3/
Al (111)	(1x1)	3-fold fcc	0.70	1.79	LEED/4/
Al (111)	undetermined	3-fold fcc	$0.98 \pm 0.10$	$1.92 \pm 0.05$	EXAFS/5/
Co (100)	c(2x2)	4-fold	0.80	1.94	LEED/6/
Cu (100)	c(2x2)	4-fold	1.40	2.28	LEED/7/
Cu (100)	c(2x2)	4-fold	$0.70 \pm 0.01$	$1.94 \pm 0.01$	SEXAFS/8/
Cu (100)	c(2x2)	4-fold	0.80	1.97	NPD/9/
Cu (110)	(2x1)				ALICISS/10/
confirm surface missing-row model					
Cu (100)	c(2x2)			1.97	ALICISS/11/
Bulk layer spacing 1.28 Å. First layer spacing expanded $25 \pm 1.0\%$ , second and third layers contracted by $10 \pm 5$ .					
Cu (110)	disordered	long bridge	0.35		SEXAFS/12/
Cu (110)	(2x1)	long bridge			HEIS/13/
top layer buckled, [001] rows alternate $+0.27 \pm 0.05$ Å and $-0.02 \pm 0.03$ Å second layer spacing expanded by $0.06 \pm 0.03$ Å					
Cu (410)	(1x1)	quasi 4-fold at step edge	$0.4 \pm 0.2$	$1.85 \pm 0.04$	ARXPD/14/

(continued)

TABLE VI (continued)

Substrate Face	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Bond Length	Method
Cu (410)	(1x1)	quasi 4-fold at step edge and 4-fold	0.4±0.2	1.85±0.04	ARXPD/14/
Fe (100)	(1x1)	4-fold top layer	0.48±0.10 1.54±0.10	2.08±0.02 +7.7%	LEED/15/
Ir (110)	c(2x2)	short bridge top layer	1.37±0.05 1.33±0.07	1.93±0.04 -2.2%	LEED/16/
Ir (111)	(2x2) or (2x1)	3-fold fcc	1.30±0.05	2.04±0.03	LEED/17/
Ni (100)	(2x2)	4-fold	0.86±0.07	1.96±0.03	EXAFS/18/
Ni (100)	c(2x2)	4-fold	0.86±0.07	1.96±0.03	EXAFS/18/
Ni (100)	(2x2)	4-fold	0.90±0.10	1.98±0.05	LEED/19/
Ni (100)	c(2x2)	4-fold	0.92±0.05	1.99±0.02	LEED/20/
Ni (100)	c(2x2)	4-fold	0.85±0.05	1.96±0.02	EXAFS/21/
Ni (100)	c(2x2)	4-fold	0.90±0.04	1.98±0.02	NPD/22/
Ni (100)	c(2x2)	4-fold	0.85±0.04	1.96±0.02	NPD/23/
Ni (100)	c(2x2)	4-fold top layer	0.86±0.10 1.85±0.10	1.96±0.04 +5.1%	HEIS/24/
Ni (100)	c(2x2)	4-fold	0.90±0.10	1.98±0.05	XANES/25/
Ni (100)	c(2x2)	4-fold	0.90	1.98	HREELS/26,27/
Ni (100)	c(2x2)	4-fold	0.922		HREELS/28/
Ni (110)	(2x1)	missing row, long bridge	0.25		ALICISS/29/
Ni (111)	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold hcp	1.20	1.87	HEIS/30/
Ta (100)	(3x1)	sub-surface tetrahedral top layer	-0.43 1.55±0.05	1.95 -6.1%	LEED/31/
W (100)	disordered	4-fold	0.55±0.10	2.30±0.02	LEED/32/

(continued)

TABLE VI (continued)

Substrate Face	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Bond Length	Method
W (110)	(2x1)	3-fold	1.25±0.01	2.09±0.01	LEED/33/
Zr (0001)	(2x2)	sub-surface octahedral	1.37	2.31	LEED/34/
SULFUR					
Co (100)	c(2x2)	4-fold	1.30	2.20	LEED/35/
Cu (100)	(2x2)	4-fold	1.39	2.28	ARXPS/36/
Fe (100)	c(2x2)	4-fold	1.09±0.05	2.30±0.02	LEED/37/
Fe (110)	(2x2)	4-fold	1.43	2.02	LEED/38/
reconstructed					
Ir (111)	( $\sqrt{3} \times \sqrt{3}$ )R30°	3-fold fcc	1.65±0.07	2.28±0.05	LEED/39/
Ni (100)	(2x2)	4-fold	1.30±0.10	2.19±0.06	LEED/19/
Ni (100)	c(2x2)	4-fold	1.28±0.05	2.18±0.03	LEED/20/
Ni (100)	c(2x2)	4-fold	1.30±0.05	2.19±0.03	LEED/40/
Ni (100)	c(2x2)	4-fold	1.30±0.04	2.19±0.02	NPD/22/
Ni (100)	c(2x2)	4-fold	1.39±0.04	2.24±0.02	SEXAFS/41/
Ni (100)	c(2x2)	4-fold	1.37±0.05	2.23±0.03	ARPEFS/42/
Ni (100)	c(2x2)	4-fold	1.35±0.10	2.22±0.06	ARXPS/43/
Ni (100)	c(2x2)	4-fold	1.35	2.22	ARXPS/36/
Ni (100)	disordered	4-fold	1.36±0.03	2.22±0.02	SEXAFS/44/
Ni (110)	c(2x2)	4-fold	0.9	2.34	ARXPS/45/
Ni (110)	c(2x2)	4-fold	0.84±0.03	2.31±0.01	LEED/46/
		top layer	1.372±0.02	+9.6%	
		next layer	1.201±0.02	-4.0%	
Ni (110)	c(2x2)	4-fold	0.87±0.03	2.32±0.01	MEIS/47/
		top layer	1.31±0.04	+4.8%	

(continued)

TABLE VI (continued)

Substrate Face	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Bond Length	Method
Mo (100)	c(2x2)	4-fold top layer	1.04 1.42	-10%	LEED/48/
Pd (100)	c(2x2)	4-fold	1.30±0.05	2.33±0.03	LEED/49/
Pd (111)	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold fcc	1.53±0.05	2.20±0.03	LEED/50/
Pt (111)	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold fcc	1.62±0.05	2.28±0.04	LEED/51/
Rh (100)	(2x2)	4-fold	1.29	2.30	LEED/52/
Rh (110)	c(2x2)	4-fold	0.77	2.45	LEED/53/
Rh (111)	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold fcc	1.53	2.18	LEED/54/
SELENIUM					
Ag (100)	c(2x2)	4-fold	1.91±0.04	2.80±0.03	LEED/55/
Ni (100)	(2x2)	4-fold	1.55±0.10	2.35±0.07	LEED/19/
Ni (100)	c(2x2)	4-fold	1.45±0.10	2.28±0.06	NPD/56/
Ni (100)	c(2x2)	4-fold	1.55	2.35	NPD/57/
Ni (110)	c(2x2)	4-fold	1.10±0.04	2.42±0.02	NPD/58/
Ni (111)	(2x2)	3-fold fcc	1.80±0.04	2.30±0.03	NPD/58/
TELLURIUM					
Cu (100)	(2x2)	4-fold	1.70±0.15	2.48±0.10	LEED/59/
Cu (100)	(2x2)	4-fold	1.90±0.04	2.62±0.03	SEXAFS/60/
Ni (100)	(2x2)	4-fold	1.80±0.10	2.52±0.07	LEED/19/
Ni (100)	c(2x2)	4-fold	1.90±0.10	2.59±0.07	LEED/61/
Ni (100)	c(2x2)	4-fold	1.9		XPD/62/
Ni (100)	c(2x2)	4-fold	1.9		SPLEED/63/

## REFERENCES

1. Puschmann, A. and Haase, J. (1984) *Surface Science* **144**, 559.
2. Martinez, V., Soria, F., Munoz, M. C., and Sacedon, J. L. (1983) *Surface Science* **128**, 424.
3. Norman, D., Brennan, S., Jaeger, R., and Stöhr, J. (1981) *Surface Science* **105**, L297.
4. Neve, J., Rundgren, J., and Westrin, P. (1982) *Journal of Physics C* **15**, 4391.
5. Bachrach, R. Z., Hansson, G. V., and Bauer, R. S. (1981) *Surface Science* **109**, L560.
6. Maglietta, M., Zanazzi, E., Bardi, U., and Jona, F. (1978) *Surface Science* **77**, 101.
7. Onuferko, J. and Woodruff, D. P. (1980) *Surface Science* **95**, 555.
8. Döbler, U., Baberschke, K., Stöhr, J., and Outka, D. A. (1985) *Physical Review B* **31**, 2532.
9. Tobin, J. G., Klebanoff, L. E., Rosenblatt, D. H., Davis, R. F., Umbach, E., Baca, A. G., Shirley, D. A., Huang, Y., Kang, W. M., and Tong, S. Y. (1982) *Physical Review B* **26**, 7076.
10. Niehaus, H. and Comsa, G. (1984) *Surface Science* **140**, 18.
11. Yarmoff, J. A. and Williams, R. S. (1985) *Journal of Vacuum Science and Technology A* **4**, 1274.
12. Döbler, U., Baberschke, K., Haase, J., and Puschmann, A. (1985) *Surface Science* **152/153**, 569.
13. (1983) *Surface Science* **128**, 281.
14. Thompson, K. A. and Fadley, C. S. (1984) *Surface Science* **146**, 281.
15. Legg, K. O., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Physical Review B* **16**, 5271.
16. Chan, C. M., Luke, K. L., Van Hove, M. A., Weinberg, W. H., and Withrow, S. P. (1978) *Surface Science* **78**, 386.
17. Chan, C.-M. and Weinberg, W. H. (1979) *Journal of Chemical Physics* **71**, 2788.

18. Stöhr, J., Jaeger, R., and Kendelewicz, T. (1982) *Physical Review Letters* **49**, 142.
19. Van Hove, M. A. and Tong, S. Y. (1975) *Journal of Vacuum Science and Technology* **12**, 230.
20. Marcus, P. C., Demuth, J. E., and Jepsen, D. W. (1975) *Surface Science* **53**, 501.
21. de Crescenzi, M., Antonangeli, F., Bellini, C., and Rosei, R. (1983) *Physical Review Letters* **50**, 1949.
22. Rosenblatt, D. H., Tobin, J. G., Mason, M. G., Davis, R. F., Kevan, S. D., Shirley, D. A., Li, C. H., and Tong, S. Y. (1981) *Physical Review B* **23**, 3828.
23. Tong, S. Y., Kang, W. M., Rosenblatt, D. H., Tobin, J. G., and Shirley, D. A. (1983) *Physical Review B* **27**, 4632.
24. Frenken, J. W. M., van der Veen, J. F., and Allan, G. (1983) *Physical Review Letters* **51**, 1876.
25. Norman, D., Stöhr, J., Jaeger, R., Durham, P. J., and Pendry, J. B. (1983) *Physical Review Letters* **51**, 2052.
26. Rahman, T. S., Black, J. E., and Mills, D. L. (1981) *Physical Review Letters* **46**, 1469.
27. Rahman, T. S., Mills, D. L., Black, J. E., Szeftel, J. M., Lehwald, S., and Ibach, H. (1984) *Physical Review B* **30**, 589.
28. Strong, R. L. and Erskine, J. L. (1985) *Physical Review Letters* **54**, 346.
29. Niehaus, H. and Comsa, G. (1985) *Surface Science* **151**, L171.
30. Narusawa, T. and Gibson, W. M. (1981) *Surface Science* **114**, 331.
31. Titov, A. V. and Jagodzinski, H. (1985) *Surface Science* **152/153**, 409.
32. Heinz, K., Saldin, D. K., and Pendry, J. B. (1985) *Physical Review Letters* **55**, 2312.
33. Van Hove, M. A. and Tong, S. Y. (1975) *Physical Review Letters* **35**, 1092.
34. Hui, K. C., Milne, R. H., Mitchell, K. A. R., Moore, W. T., and Zhou, M. Y. (1985) *Solid State Communications* **56**, 83.
35. Maglietta, M. (1982) *Solid State Communications* **43**, 395.
36. Bullock, E. L., Fadley, C. S., and Orders, P. J. (1983) *Physical Review B* **28**, 4867.



37. Legg, K. O., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Surface Science* **66**, 25.
38. Shih, F. E., Jona, F., Jepsen, D. W., and Marcus, P. M. (1981) *Physical Review Letters* **46**, 731.
39. Chan, C.-M. and Weinberg, W. H. (1979) *Journal of Chemical Physics* **71**, 3988.
40. Gauthier, Y., Aberdam, D., and Baudoing, R. (1978) *Surface Science* **78**, 339.
41. Stöhr, J., Jaeger, R., and Brennan, S. (1982) *Surface Science* **117**, 503.
42. Barton, J. J., Bahr, C. C., Hussain, Z., Robey, S. W., Klebanoff, L. E., and Shirley, D. A. (1984) *Journal of Vacuum Science and Technology A* **2**, 847.
43. Orders, P. J., Sinkovic, B., Fadley, C. S., Trehan, R., Hussain, Z., and Lecante, J. (1984) *Physical Review B* **30**, 1838.
44. Stöhr, J., Kollin, E. B., Fischer, D. A., Hastings, J. B., Zaera, F., and Sette, F. (1985) *Physical Review Letters* **55**, 1468.
45. Baudoing, R., Blanc, E., Gaubert, C., Gauthier, Y., and Gnuchev, N. (1983) *Surface Science* **128**, 22.
46. Gauthier, Y., Baudoing, R., Joly, Y., Rundgren, J., Bertolini, J. C., and Massardier, J. (1985) *Surface Science* **162**, 348.
47. van der Veen, J. F., Tromp, R. M., Smeenk, R. G., and Saris, F. W. (1979) *Surface Science* **82**, 468.
48. Clarke, L. J. (1981) *Surface Science* **102**, 331.
49. Berndt, W., Hora, R., and Scheffler, M. (1982) *Surface Science* **117**, 188.
50. Maca, F., Scheffler, M., and Berndt, W. (1985) *Surface Science* **160**, 467.
51. Hayek, K., Glassl, H., Gutmann, A., Leonhard, H., Prutton, M., Tear, S. P., and Welton-Cook, M. R. (1985) *Surface Science* **152**, 419.
52. Hengrasmee, S., Watson, P. R., Frost, D. C., and Mitchell, K. A. R. (1979) *Surface Science* **87**, L249.
53. Hengrasmee, S., Watson, P. R., Frost, D. C., and Mitchell, K. A. R. (1980) *Surface Science* **92**, 71.
54. Wong, P. C., Zhou, M. Y., Hui, K. C., and Mitchell, K. A. R. (1985) *Surface Science* **163**, 172.
55. Ignatiev, A., Jona, F., Jepsen, D. W., and Marcus, P. M. (1973) *Surface Science* **40**, 439.

56. Demuth, J. E., Jepsen, D. W., and Marcus, P. M. (1973) *Physical Review Letters* **31**, 540.
57. Rosenblatt, D. H., Kevan, S. D., Tobin, J. G., Davis, R. F., Mason, M. G., Shirley, D. A., Tang, J. C., and Tong, S. Y. (1982) *Physical Review B* **26**, 3181.
58. Rosenblatt, D. H., Kevan, S. D., Tobin, J. G., Davis, R. F., Mason, M. G., Denley, D. R., Shirley, D. A., Huang, Y., and Tong, S. Y. (1982) *Physical Review B* **26**, 1812.
59. Salwén, A. and Rundgren, J. (1975) *Surface Science* **53**, 523.
60. Comin, F., Citrin, P. H., Eisenberger, P., and Rowe, J. E. (1982) *Physical Review B* **26**, 7060.
61. Demuth, J. E., Jepsen, D. W., and Marcus, P. M. (1973) *Journal of Physics C* **6**, L307.
62. Smith, N. V., Farrell, H. H., Traum, M. M., Woodruff, D. P., Norman, D., Woolfson, M. S., and Holland, B. W. (1980) *Physical Review B* **21**, 3119.
63. Lang, J. K., Jamison, K. D., Dunning, F. B., Walters, G. K., Passier, M. A., Ignatiev, A., Tamura, E., and Feder, R. (1982) *Surface Science* **123**, 247.

TABLE VII. Other Atomic Adsorbates on Metal Surfaces

(Where relaxation of the metal layer spacing has been investigated this is listed in the table after the adsorbate information.)

Adsorption system	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Method
Ag (100)-Cl	c(2x2)	4-fold	1.62	LEED/1/
Ag (111)-Au	(1x1)	3-fold fcc	2.35±0.10	HEIS/2/
Ag (111)-I	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold fcc and hcp	2.29±0.06	LEED/3/
	top layer	0.0%	2.36±0.06	
Ag (111)-I	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold	2.34±0.02	SEXAFS/4/
Ag (111)-Xe	hexagonal (1x1)	incommensurate	3.55±0.10	LEED/5/
randomly oriented hexagonal Xe lattice, Xe-Xe spacing 1.51 Å				
Al (100)-Na	c(2x2)	4-fold	2.08±0.12	LEED/6/
Al (100)-Na	c(2x2)	4-fold	2.05±0.10	LEED/7/
Au (111)-Ag	(1x1)	3-fold fcc	2.35±0.10	HEIS/2/
Cu (100)-Cl	c(2x2)	4-fold	1.60±0.03	LEED/8/
	top layer	+3%	1.85±0.03	
Cu (111)-Cs	(2x2)	top	3.01±0.05	LEED/9/
Cu (100)-I	(2x2)	4-fold	1.98±0.02	SEXAFS/10/
Cu (111)-I	( $\sqrt{3}\times\sqrt{3}$ )R30°	3-fold hcp	2.21±0.02	SEXAFS/10/
Cu (111)-Ni	(1x1)	3-fold fcc	2.04±0.02	LEED/11/
Cu (100)-Pb	c(2x2)	4-fold	2.05±0.05	LEED/12/
Cu (100)-3Pb	c(5 $\sqrt{2}\times\sqrt{2}$ )R45°	4-fold	2.4±0.05	LEED/12/
overlayer relaxed with anti-phase domain walls and off-center atoms				
Fe (100)-C+O	c(2x2)	4-fold	0.48	LEED/13/
CO decomposes on Fe(100), C and O occupy random sites in (2x2) lattice				
Fe (110)-H	(2x1)	quasi 3-fold	0.90±0.10	LEED/14/
Fe (110)-2H	(3x1)	quasi 3-fold	1.00±0.10	LEED/14/
Fe (100)-N	c(2x2)	4-fold	0.25±0.05	LEED/15/
	top layer	+8%	1.54±0.05	

(continued)

TABLE VII (continued)

Adsorption system	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Method
Ir (110)-S	(2x2)			LEED/16/
Sulfur is adsorbed in 3-fold sites on the (111) facets left by the clean-surface missing-row reconstruction. The first Ir-Ir spacing is contracted by 3.3%, the second expanded by 1.3%. Ir-S is 2.39 Å for first-layer Ir atoms and 2.26 Å for second-layer Ir atoms.				
Mo (100)-N	c(2x2)	4-fold	1.02	LEED/17/
Mo (100)-Si	(1x1)	4-fold	1.16±0.10	LEED/18/
Ni (100)-Br	c(2x2)	4-fold	1.51±0.03	EXAFS/19/
Ni (100)-C+O	c(2x2)	4-fold	0.93±0.10	LEED/20/
CO decomposes on Ni(100), C and O occupy random sites in (2x2) lattice				
Ni (100)-2C	(2x2)	4-fold	0.10±0.10	LEED/21/
	top layer	+22%	1.96±0.05	
	reconstructed	lateral motions -	0.25±0.35 Å	
Ni (100)-Cu	(1x1)	4-fold	1.80±0.03	LEED/22/
Ni (100)-Na	c(2x2)	4-fold	2.2±0.10	LEED/23/
Ni (100)-Na	c(2x2)	4-fold	2.23±0.10	LEED/24/
Ni (100)-Na	c(2x2)	4-fold	1.9	XPD/25/
Ni (110)-C				SEELFS/26/
Graphitic layer on Ni(110), 3 carbon sites: A - off top, Ni-C = 1.95 Å				
B - 4-fold, Ni-C = 2.49 Å, C - off bridge, Ni-C = 1.95, Ca-Cb = 1.49 Å				
Ni (111)-2C	(1x1)	3-fold hcp and fcc	2.80±0.08	SEELFS/27/
Ni (111)-2H	(2x2)	3-fold hcp and fcc	1.15±0.10	LEED/28/
Pd (111)-Au	(1x1)	3-fold fcc	2.25±0.19	HEIS/29/
Ti (0001)-Cd	(1x1)	3-fold fcc	2.57	LEED/30/
Ti (0001)-Cd	(1x1)	3-fold fcc	2.63	LEED/31/
Two Cd layers, Cd-Cd distance 2.81 Å, Cd-Ti 2.63 Å				

(continued)

TABLE VII (continued)

Adsorption system	Overlayer Unit Cell	Adsorption Site	Adsorbate Spacing	Method
Ti (0001)-N	(1x1)	sub-surface tetrahedral	-1.22±0.05	LEED/32/
W (100)-H	c(2x2)	4-fold	1.32	LEED/33/
W (100)-2H	(1x1)	short bridge		HREELS/34/
W (100)-2H	(1x1)	short bridge	1.74	HREELS/35/
W (100)-2H	(1x1)	2-fold	1.17±0.04	LEED/36/
	top layer	-1.3%	1.56±0.02	
W (100)-N	c(2x2)	4-fold	0.49±0.06	LEED/37/
	top layer	+1.3%	1.60±0.06	

## REFERENCES

1. Jona, F. and Marcus, P. M. (1983) *Physical Review Letters* **50**, 1823.
2. Culbertson, R. J., Feldman, L. C., Silverman, P. J., and Boehm, H. (1981) *Physical Review Letters* **47**, 657.
3. Maglietta, M., Zanazzi, E., Bardi, U., Sondericker, D., Jona, F., and Marcus, P. M. (1982) *Surface Science* **123**, 141.
4. Citrin, P. H., Eisenberger, P., and Hewitt, R. C. (1979) *Surface Science* **89**, 28.
5. Stoner, N., Van Hove, M. A., Tong, S. Y., and Webb, M. B. (1978) *Physical Review Letters* **40**, 243.
6. Van Hove, M. A., Tong, S. Y., and Stoner, N. (1976) *Surface Science* **54**, 259.
7. Hutchins, B. A., Rhodin, T. N., and Demuth, J. E. (1976) *Surface Science* **54**, 419.

8. Jona, F., Westphal, D., Goldman, A., and Marcus, P. M. (1983) *Journal of Physics C* **16**, 3001.
9. Lindgren, S. A., Walldén, L., Rundgren, J., Westrin, P., and Neve, J. (1983) *Physical Review B* **28**, 6707.
10. Citrin, P. H., Eisenberger, P., and Hewitt, R. C. (1980) *Physical Review Letters* **45**, 1948.
11. Tear, S. P. and Roell, K. (1982) *Journal of Physics C* **15**, 5521.
12. Hoesler, W. and Moritz, W. (1982) *Surface Science* **117**, 196.
13. Legg, K. O., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Physical Review B* **16**, 5271.
14. Moritz, W., Imbihl, R., Behm, R. J., Ertl, G., and Matsushima, T. (1985) *Journal of Chemical Physics* **83**, 1959.
15. Imbihl, R., Behm, R. J., Ertl, G., and Moritz, W. (1982) *Surface Science* **123**, 129.
16. Chan, C.-M. and Van Hove, M. A. (1987) *Surface Science* **183**, 303.
17. Ignatiev, A., Jona, F., Jepsen, D. W., and Marcus, P. M. (1975) *Surface Science* **49**, 189.
18. Ignatiev, A., Jona, F., Jepsen, D. W., and Marcus, P. M. (1975) *Physical Review B* **11**, 4780.
19. Lairson, B., Rhodin, T. N., and Ho, W. (1985) *Solid State Communications* **55**, 925.
20. Passler, M. A., Lin, T. H., and Ignatiev, A. (1981) *Journal of Vacuum Science and Technology* **18**, 481.
21. Onuferko, J. H., Woodruff, D. P., and Holland, B. W. (1979) *Surface Science* **87**, 357.
22. Abu-Joudeh, M., Vaishnava, P. P., and Montano, P. A. (1984) *Journal of Physics C* **17**, 6899.
23. Andersson, S. and Pendry, J. B. (1975) *Solid State Communications* **16**, 563.
24. Demuth, J. E., Jepsen, D. W., and Marcus, P. M. (1975) *Journal of Physics C* **8**, L25.
25. Smith, N. V., Farrell, H. H., Traum, M. M., Woodruff, D. P., Norman, D., Woolfson, M. S., and Holland, B. W. (1980) *Physical Review B* **21**, 3119.
26. Papagno, L. and Caputi, L. S. (1984) *Physical Review B* **29**, 1483.

27. Rosei, R., de Crescenzi, M., Sette, F., Quaresima, C., Savoia, A., and Perfetti, P. (1983) *Physical Review B* **28**, 1161.
28. Christmann, K., Behm, R. J., Ertl, G., Van Hove, M. A., and Weinberg, W. H. (1979) *Journal of Chemical Physics* **70**, 4168.
29. Kuk, Y., Feldman, L. C., and Silverman, P. J. (1983) *Physical Review Letters* **50**, 511.
30. Shih, H. D., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Physical Review B* **15**, 5550.
31. Shih, H. D., Jona, F., Jepsen, D. W., and Marcus, P. M. (1977) *Physical Review B* **15**, 5561.
32. Shih, H. D., Jona, F., Jepsen, D. W., and Marcus, P. M. (1976) *Surface Science* **60**, 445.
33. Willis, R. F. (1979) *Surface Science* **89**, 457.
34. Ho, W., Willis, R. F., and Plummer, E. W. (1980) *Physical Review B* **21**, 4202.
35. Woods, J. P. and Erskine, J. L. (1985) *Physical Review Letters* **55**, 2595.
36. Passler, M. A., Lee, B. W., and Ignatiev, A. (1985) *Surface Science* **150**, 263.
37. Griffiths, K., King, D. A., Aers, G. C., and Pendry, J. B. (1982) *Journal of Physics C* **15**, 4921.

TABLE VIII. Semiconductor Surface Structures

Substrate	Unit Cell	Structure	Method
AlP (110) (zincblende)	(1x1)	Surface reconstructed with 25° bond angle rotation preserving bond lengths. Top layer buckled with P +0.06 Å, Al -0.57 Å and layer spacing contracted to 1.33 Å (bulk 1.927 Å). Second layer buckled P -0.035 Å, Al +0.035 Å, layer expanded to 1.92 Å. First layer in-plane displacements of ~ 0.25 Å to preserve bond lengths.	LEED/1/
CdTe (110) (zincblende)	(1x1)	Te is buckled out from the top layer by 0.82±0.05 Å. Top layer Cd atoms contract by ~ 0.5 Å, and there are lateral motions of ~ 0.4 Å to conserve bond lengths.	LEED/2/
GaAs (110) (zincblende)	(1x1)	As is buckled out from the top layer by 0.70 Å. Top layer Ga atoms contract by ~ 0.5 Å, and there are lateral motions of ~ 0.4 Å to conserve bond lengths. The lateral motions may be reduced by ~ 3/4 to give a better agreement with MEIS results without significantly worsening the LEED fit.	LEED/3/
GaAs (110) (zincblende)	(1x1)	As is buckled out from the top layer by 0.69 Å. Top layer Ga atoms contract by ~ 0.5 Å, and there are lateral motions of ~ 0.3 Å to conserve bond lengths.	LEED/4/

(continued)



TABLE VIII (continued)

Substrate	Unit Cell	Structure	Method
GaAs (110) (zincblende)	(1x1)	As is buckled out from the top layer by $0.40 \pm 0.30$ Å. Top layer Ga atoms contract by $\sim 0.20 \pm 0.30$ Å, with no lateral motions.	HEIS/5/
GaAs (111) (zincblende) (Ga terminated)	(2x2)	One quarter of the top layer Ga atoms are missing, and the remaining atoms are almost co-planar with the first As layer, within 0.20 Å. Ga bonding is $sp^2$ rehybridized, instead of the normal $sp^3$ configuration. There are first bi-layer lateral motions of $\sim 0.2$ Å, and some buckling in the third layer to maintain optimum bond-lengths and angles.	LEED/6/
GaP (110) (zincblende)	(1x1)	Surface reconstructed with $27.5^\circ$ bond angle rotation preserving bond lengths. Top layer buckled with Ga +0.09 Å, As -0.54 Å and layer spacing contracted to 1.39 Å. Second layer spacing expanded to 1.93 Å.	LEED/7/

(continued)

TABLE VIII (continued)

Substrate	Unit Cell	Structure	Method
GaP (111) (zincblende) (Ga terminated)	(2x2)	One quarter of the top layer Ga atoms are missing, and the remaining atoms are almost co-planar with the first P layer. Ga bonding is $sp^2$ rehybridized, instead of the normal $sp^3$ configuration. There are lateral motions up to $\sim 2 \text{ \AA}$ , and some buckling in the third layer to maintain optimum bond-lengths and angles.	LEED/6/
GaSb (110) (zincblende)	(1x1)	Sb is buckled out from the top layer by $0.77 \pm 0.05 \text{ \AA}$ . Top layer Ga atoms contract by $\sim 0.5 \text{ \AA}$ , and there are lateral motions of $\sim 0.4 \text{ \AA}$ to conserve bond lengths.	LEED/8,9/
GaSb (110) (zincblende)	(1x1)	Consistent with LEED results (above) for layer displacements and lateral motions.	MEIS/10/
GaSb (110) (zincblende)	(1x1)	Bond-length conserving rotations of $28.5 \pm 2.6^\circ$ bring Sb atoms up, Ga atoms down.	MEIS/11/
Ge (100) (diamond)	(2x1)	The top layer buckles, with one Ge atom moving out by $0.62 \pm 0.04 \text{ \AA}$ and the other moving in by $0.66 \pm 0.04 \text{ \AA}$ . There are lateral displacements of $\sim 0.9 \text{ \AA}$ in the first layer and $\sim 0.1 \text{ \AA}$ in the second layer.	XRD/12/

(continued)

TABLE VIII (continued)

Substrate	Unit Cell	Structure	Method
InAs (110) (zincblende)	(1x1)	The top layer has As buckled outward by $\sim 0.8 \text{ \AA}$ with lateral motions of $\sim 0.6 \text{ \AA}$ in the first three layers to conserve bond lengths.	LEED/13/
InAs (110) (zincblende)	(1x1)	Bond-length conserving rotations of $30.0 \pm 2.4^\circ$ bring As atoms up, In atoms down.	MEIS/11/
InSb (111) (zincblende)	(2x2)	One quarter of the top-layer In atoms are missing in the "vacancy buckling" model. Top layer is not buckled, in plane Sb atoms expand radially by $0.45 \pm 0.04 \text{ \AA}$ from their normal positions and In atoms contract radially by $0.23 \pm 0.05 \text{ \AA}$ from their normal positions. No data on layer spacing changes.	X-Ray Diffraction/14/
InP (110) (zincblende)	(1x1)	Preliminary results show layer buckling with the P atom buckled out, and lateral and vertical shifts in the top layer $\sim 0.4 \text{ \AA}$ . Second and deeper layer shifts were not investigated.	LEED/15/
InP (110) (zincblende)	(1x1)	Top layer buckling with the P atom buckled out by $0.69 \pm 0.10 \text{ \AA}$ , and lateral and vertical shifts in the top layer $\sim 0.4 \text{ \AA}$ . The second layer spacing is contracted by $0.41 \pm 0.10 \text{ \AA}$ with a slight buckling of $0.07 \pm 0.10 \text{ \AA}$ . Only first layer lateral displacements were investigated.	LEED/16/

(continued)

TABLE VIII (continued)

Substrate	Unit Cell	Structure	Method
Si (100) (diamond)	(2x1)	Buckled dimer model is best fit to data. Surface atoms dimerize to take up dangling bonds, surface buckles for best bond angles. There are lateral and vertical ion-core motions for at least 4 layers into the bulk crystal. Authors conclude that (2x1), (2x2) and c(4x2) buckled dimer domains may exist on the surface.	MEIS/17/
Si (100) (diamond)	(2x1)	LEED analysis considering vertical and lateral displacements in the top three atomic layers supports a buckled dimer model.	LEED/18/
Si (111) (diamond)	(1x1) laser annealed	The first two layers are almost coplanar, instead of the normal 0.78 Å separation. The first layer spacing is $0.08 \pm 0.02$ , a contraction of 90%, and the second layer spacing is $2.95 \pm 0.20$ Å, a 25.5% expansion from 2.35 Å.	LEED/19/
Si (111) (diamond)	(2x1)	The top layer is buckled by $0.30 \pm 0.05$ Å, the second layer spacing is $0.70 \pm 0.05$ , a change of +2.9%, and the third layer spacing is contracted by 3.4% to $2.27 \pm 0.02$ from 2.35 Å. There are second-layer lateral shifts of $\sim 0.12$ Å. (More recent results favor the “ $\pi$ -bonded chain” model, see below.)	LEED/20/

(continued)

TABLE VIII (continued)

Substrate	Unit Cell	Structure	Method
Si (111) (diamond)	(2x1)	Analysis of LEED data supports a " $\pi$ -bonded chain" model for the (2x1) reconstruction. Trial geometries based on strain-minimization calculations for the model, involving vertical motions four layers deep with lateral motions along the long side of the unit cell.	LEED/21/
Si (111) (diamond)	(2x1)	Best fit to " $\pi$ -bonded chain" model. This model involves buckling in layers 2 to 6 of up to 0.27 Å and small lateral shifts in the first six layers.	MEIS/22/
GaP (110) (zincblende)	(1x1)	Surface reconstructed with 28.0° bond angle rotation preserving bond lengths. Top layer buckled with Zn +0.08 Å, S -0.51 Å and layer spacing contracted to 1.40 Å. Second layer spacing expanded to 1.91 Å.	LEED/7/
ZnSe (110) (zincblende)	(1x1)	Two different models were consistent with LEED data. First; Se is buckled outward by $\sim 0.70$ Å and the second layer spacing is contracted by $\sim 0.60$ Å, with lateral motions of $\sim 0.7$ Å. Second; Se buckled out by 0.10 Å and the second layer spacing contracted by 0.09 Å, with lateral motions of $\sim 0.07$ Å.	LEED/23/

(continued)

TABLE VIII (continued)

Substrate	Unit Cell	Structure	Method
ZnTe (110) (zincblende)	(1x1)	Te is buckled out from the top layer by $0.71 \pm 0.05$ Å. Top layer Zn atoms contract by $\sim 0.5$ Å, and there are lateral motions of $\sim 0.4$ Å to conserve bond lengths.	LEED/9/

## REFERENCES

1. Duke, C. B., Paton, A., and Bonpace, C. R. (1983) *Physical Review B* **28**, 852.
2. Duke, C. B., Paton, A., Ford, W. K., Kahn, A., and Scott, G. (1982) *Journal of Vacuum Science and Technology* **20**, 778.
3. Duke, C. B. and Paton, A. (1984) *Journal of Vacuum Science and Technology B* **2**, 327.
4. Tong, S. Y., Mei, W. M., and G.Xu, (1984) *Journal of Vacuum Science and Technology B* **2**, 393.
5. Grossman, H. J. and Gibson, W.M. (1984) *Journal of Vacuum Science and Technology B* **2**, 343.
6. Xu, G., Hu, W. Y., Puga, M. W., Tong, S. Y., Yeh, J. L., Wang, S. R., and Lee, B. W. (1985) *Physical Review B* **32**, 8473.
7. Duke, C. B., Paton, A., and Kahn, A. (1984) *Journal of Vacuum Science and Technology A* **2**, 515.
8. Duke, C. B., Paton, A., and Kahn, A. (1983) *Physical Review B* **27**, 3436.
9. Duke, C. B., Paton, A., and Kahn, A. (1983) *Journal of Vacuum Science and Technology A* **1**, 672.

10. Smit, L., Tromp, R. M., and van der Veen, J. F. (1984) *Physical Review B* **29**, 4814.
11. Smit, L. and van der Veen, J. F. (1986) *Surface Science* **166**, 183.
12. Eisenberger, P. and Marra, W. C. (1981) *Physical Review Letters* **46**, 1081.
13. Duke, C. B., Paton, A., Kahn, A., and Bonapace, C. B. (1983) *Physical Review B* **27**, 6189.
14. (1985) *Physical Review Letters* **54**, 1275.
15. Tear, S. P., Welton-Cook, M. R., Prutton, M., and Walker, J. A. (1980) *Surface Science* **99**, 598.
16. Meyer, R. J., Duke, C. B., Paton, A., Tsang, J. C., Yeh, J. L., Kahn, A., and Mark, P. (1980) *Physical Review B* **22**, 6171.
17. Tromp, R. M., Smeenk, R. G., Saris, F. W., and Chadi, D. J. (1983) *Surface Science* **133**, 137.
18. Holland, B. W., Duke, C. D., and Paton, A. (1984) *Surface Science* **140**, L269.
19. Jones, G. J. R. and Holland, B. W. (1985) *Solid State Communications* **53**, 45.
20. Feder, R., Mönch, W., and Auer, P. P. (1979) *Journal of Physics C* **12**, L179.
21. Himpsel, F. J., Marcus, P. M., Tromp, R., Batra, I. P., Cook, M. R., Jona, F., and Liu, H. (1984) *Physical Review B* **30**, 2257.
22. Smit, L., Tromp, R. M., and van der Veen, J. F. (1985) *Surface Science* **163**, 315.
23. Duke, C. B., Paton, A., Kahn, A., and Tu, D. W. (1984) *Journal of Vacuum Science and Technology B* **2**, 366.

TABLE XI. Carbon Monoxide, Di-Nitrogen and Nitric Oxide Chemisorption on Metals

(In all the listed structures the CO or NO molecule is believed to adsorb perpendicular to the surface with the oxygen end away from the surface. For CO structures with multiple non-equivalent adsorption sites these are listed on consecutive lines. If the metal layer spacing has been investigated this is listed after the adsorbate information.)

Substrate Face	Overlayer Unit Cell	Adsorption Site	C-Metal $\perp$ Spacing	C-O Bond Length	Method
Cu (100)	c(2x2)-CO	top	1.90 $\pm$ 0.10	1.13 $\pm$ 0.10	LEED/1/
Cu (100)	c(2x2)-CO	top	1.92 $\pm$ 0.05		NEXAFS/2/
Ni (100)	c(2x2)-CO	top	1.72	1.15	LEED/3/
Ni (100)	c(2x2)-CO	top	1.80 $\pm$ 0.10	1.15 $\pm$ 0.05	ARXPS/4/
Ni (100)	c(2x2)-CO	top	1.80 $\pm$ 0.10	1.15 $\pm$ 0.05	LEED/5/
Ni (100)	c(2x2)-CO	top	1.70 $\pm$ 0.10	1.13 $\pm$ 0.10	LEED/6/
Ni (100)	c(2x2)-CO	top	1.71 $\pm$ 0.10	1.15 $\pm$ 0.10	LEED/7/
Ni (100)	c(2x2)-CO	top	1.80 $\pm$ 0.04	1.13	NPD/8/
Ni (111)	( $\sqrt{3}\times\sqrt{3}$ )R30°-CO	bridge	1.27 $\pm$ 0.05	1.13 $\pm$ 0.05	NPD/8/
Ni (100)	disordered-CO	top			NEXAFS/9/
		molecular axis $\perp$ to surface $\pm 10^\circ$			
Ni (100)	disordered-NO	top			NEXAFS/9/
		molecular axis $\perp$ to surface $\pm 10^\circ$			
Ni (100)	disordered-N <sub>2</sub>	top			NEXAFS/9/
		molecular axis $\perp$ to surface $\pm 10^\circ$			
Pd (100)	(2 $\sqrt{2}\times\sqrt{2}$ )R45°-2CO	bridge	1.36 $\pm$ 0.10	1.15 $\pm$ 0.10	LEED/10/
		top layer	1.945 $\pm$ 0.10	+0.4%	
Pd (111)	( $\sqrt{3}\times\sqrt{3}$ )R30°-2CO	3-fold fcc	1.29 $\pm$ 0.05	1.15 $\pm$ 0.05	LEED/11/
Pd (111)	(3x3)-2CO*	3-fold fcc	1.30 $\pm$ 0.05	1.17 $\pm$ 0.05	LEED/12/
Pt (111)	c(4x2)-2CO	top	1.85 $\pm$ 0.05	1.15 $\pm$ 0.10	LEED/13/
		bridge	1.55 $\pm$ 0.05	1.15 $\pm$ 0.10	

(continued)



TABLE XI (continued)

Substrate Face	Overlayer Unit Cell	Adsorption Site	C-Metal ⊥ Spacing	C-O Bond Length	Method
		top layer	2.26±0.025	0.0%	
Pt (111)	(2√3x4)rect-4CO*	bridge	1.45	1.15	LEED/14/
Rh (111)	(3x3)-2CO*	3-fold hcp	1.30	1.17	LEED/15/
Rh (111)	c(2√3x4)rect-CO*	3-fold fcc	1.50±0.05	1.21±0.10	LEED/16/
Rh (111)	c(4x2)-CO*	3-fold hcp	1.30±0.05	1.17±0.05	LEED/17/
Rh (111)	c(4x2)-NO*	3-fold fcc	1.30±0.05	1.17±0.05	LEED/17/
Rh (111)	(2x2)-3CO	quasi-top	1.87±0.10	1.15±0.10	LEED/18/
		quasi-top	1.87±0.10	1.15±0.10	
		bridge	1.52±0.10	1.15±0.10	
Rh (111)	(√3x√3)R30°-CO	top	1.95±0.10	1.07±0.10	LEED/19/
		top layer	2.19±0.10	0.0%	
Ru (0001)	(√3x√3)R30°-CO	top	2.00±0.10	1.09±0.10	LEED/20/

\* These structures involve carbon monoxide co-adsorbed with other molecules. See table XII for more details.

## REFERENCES

1. Andersson, S. and Pendry, J. B. (1980) *Journal of Physics C* **13**, 2547.
2. McConville, C. F., Woodruff, D. P., Prince, K. C., Paolucci, G., Chab, V., Surman, M., and Bradshaw, A. M. (1986) *Surface Science* **166**, 221.
3. Passler, M., Ignatiev, A., Jona, F., Jepsen, D. W., and Marcus, P. M. (1979) *Physical Review Letters* **43**, 360.
4. Petersson, L. G., Kono, S., Hall, N. F. T., Fadley, C. S., and Pendry, J. B. (1979) *Physical Review Letters* **42**, 1545.
5. Heinz, K., Lang, E., and Müller, K. (1979) *Surface Science* **87**, 595.

6. Tong, S. Y., Maldonado, A., Li, C. H., and Van Hove, M. A. (1980) *Surface Science* **94**, 73.
7. Andersson, S. and Pendry, J. B. (1980) *Journal of Physics C* **13**, 3547.
8. Kevan, S. D., Davis, R. F., Rosenblatt, D. H., Tobin, J. G., Mason, M. G., Shirley, D. A., Li, C. H., and Tong, S. Y. (1981) *Physical Review Letters* **46**, 1629.
9. Stöhr, J. and Jaeger, R. (1982) *Physical Review B* **26**, 4111.
10. Behm, R. J., Christmann, K., Ertl, G., and Van Hove, M. A. (1980) *Journal of Chemical Physics* **73**, 2984.
11. Ohtani, H., Van Hove, M. A., and Somorjai, G. A. (to be published)
12. Ohtani, H., Van Hove, M. A., and Somorjai, G. A. (to be published)
13. Ogletree, D. F., Van Hove, M. A., and Somorjai, G. A. (1986) *Surface Science* **173**, 351.
14. Ogletree, D. F., Van Hove, M. A., and Somorjai, G. A. (1987) *Surface Science* **187**, 1.
15. Van Hove, M. A., Lin, R. F., Blackman, G. A., and Somorjai, G. A. (to be published) *Acta Crystallographica*
16. Van Hove, M. A., Lin, R. F., and Somorjai, G. A. (1986) *Journal of the American Chemical Society* **108**, 2532.
17. Bent, B. E., Mate, C. M., Koestner, R. J., Blackman, G. S., Kao, C. T., Hove, M. A. Van, and Somorjai, G. A. (to be published)
18. Van Hove, M. A., Koestner, R. J., Frost, J. C., and Somorjai, G. A. (1983) *Surface Science* **129**, 482.
19. Koestner, R. J., Van Hove, M. A., and Somorjai, G. A. (1981) *Surface Science* **107**, 439.
20. Michalk, G., Moritz, W., Pfnür, H., and Menzel, D. (1983) *Surface Science* **129**, 92.

TABLE X. Insulator and other Compound Surface Structures

Substrate	Unit Cell	Structure	Method
C (111) (diamond)	(1x1)	Terminated bulk diamond, no relaxation in layer spacing.	LEED/1/
C (111)-H (diamond)	(1x1)	Hydrogen terminated diamond, surface layer spacing relaxed by $-0.05 \pm 0.05$ Å.	MEIS/2/
C (111)-H (diamond)	(1x1)	Hydrogen in (1x1) arrangement. Determined to be in top site, assumed 1.09 Å H-C bond length.	Helium diffraction/3/
C (111) (diamond)	(2x2)	Evidence for " $\pi$ -bonded chain reconstruction.	MEIS/2/
C (0001) (graphite)	(1x1)	Normal graphite layer stacking, first layer contracted 1.4% to 3.30 Å from bulk spacing of 3.35 Å.	LEED/4/
C (0001)-K (graphite)	intercalated	When K is adsorbed on C(0001) it is intercalated between layers, changing the carbon stacking sequence from ABAB... to AAAA... and increasing the layer spacing to 5.35 Å from 3.35 Å.	LEED/5/
C (0001)-Ar (graphite)	incommensurate	Ar forms an incommensurate hexagonal overlayer on graphite $3.2 \pm 0.10$ Å above the graphite surface.	LEED/6/

(continued)

TABLE X (continued)

Substrate	Unit Cell	Structure	Method
C (0001)-Kr (graphite)	$(\sqrt{3} \times \sqrt{3})R30^\circ$	Kr adsorbs in 6-fold hollow sites on the graphite basal plane $3.35 \pm 0.01 \text{ \AA}$ above the graphite surface. No nearest neighbor sites are occupied. At higher coverages next-nearest neighbor sites are occupied.	LEED/7/
CaO (100) (rocksalt)	(1x1)	Top layer contracts by 1.2% to 2.38 from $2.41 \text{ \AA}$ . No top-layer buckling.	LEED/8/
CoO (111) (rocksalt)	(1x1)	Oxygen termination with fcc stacking, top layer contraction of 17% to $1.06 \text{ \AA}$ from bulk $1.27 \text{ \AA}$ .	LEED/9/
CoO (100) (rocksalt)	(1x1)	Terminated bulk structure, top layer spacing is $2.85 \pm 0.08 \text{ \AA}$ .	LEED/10/
MgO (100) (rocksalt)	(1x1)	Top-layer oxygen buckled out by $0.04 \pm 0.05$ , and top layer contracted by $0.02 \pm 0.07 \text{ \AA}$ . Bulk layer spacing $2.10 \text{ \AA}$ .	LEED/11/
MgO (100) (rocksalt)	(1x1)	Terminated bulk structure, no change from bulk layer spacing of $2.10 \text{ \AA}$ .	LEED/12/
MoS <sub>2</sub> (0001) (layer compound)	(1x1)	Normal stacking, S-Mo-S termination, top layer contraction by 5% to $1.51 \text{ \AA}$ from bulk $1.59 \text{ \AA}$ . No second layer contraction.	LEED/13/
Na <sub>2</sub> O (111) (fluorite)	(1x1)	Oxidation of epitaxial Na(110) on Ni(100) substrate. Determine fluorite lattice with Na-O-Na termination.	LEED/14/

(continued)

TABLE X (continued)

Substrate	Unit Cell	Structure	Method
NbSe <sub>2</sub> (0001) (layer compound)	(1x1)	Normal layer stacking with Se-Nb-Se termination, no evidence for relaxation of first two layer spacings.	LEED/13/
NiI <sub>2</sub> (0001)	(1x1)	Bulk NiI <sub>2</sub> is a layer compound with hexagonal metal layers and octohedral coordination in a cubic unit cell. At the surface the bulk Ni-I bond length of 2.78 Å decreases by 0.036±0.010 Å and the I-I separation of 3.89 Å decreases by -0.48±0.010.	SEXAFS/15/
NiO (100) (rocksalt)	(1x1)	Top layer unbuckled, contracted 2% to 2.04 Å from bulk value of 2.08 Å.	LEED/16/
Sb <sub>2</sub> Te <sub>2</sub> Se (0001) (layer com- pound)	(1x1)	This is an elemental layer compound with Te-Sb-Se-Sb-Te stacking. The Te-Te interfaces are cleavage planes. At the surface the bulk Te-Sb layer spacing of 1.733 Å is contracted by ~ 3% to 1.683±0.05 Å.	ARUPS/17/
TiS <sub>2</sub> (0001) (layer com- pound)	(1x1)	This is an elemental layer compound with Ti-S-Ti stacking. The Ti-Ti interfaces are cleavage planes. At the surface the bulk Ti-S layer spacing is contracted by ~ 5% and the first Van der Waal spacing (Ti-S-Ti   Ti-S-Ti) is also contracted by 5%.	LEED/18/

(continued)

TABLE X (continued)

Substrate	Unit Cell	Structure	Method
TiSe <sub>2</sub> (0001) (layer compound)	(1x1)	This is an elemental layer compound with Ti-Se-Ti stacking. The Ti-Ti interfaces are cleavage planes. At the surface the bulk Ti-Se layer spacing is expanded by ~ 5% and the first Van der Waal spacing (Ti-Se-Ti   Ti-Se-Ti) is contracted by 5%.	LEED/18/
ZnO (0001) (wurtzite)	(1x1)	Zn termination, top layer spacing 0.60±0.10 Å, a 25% contraction from the bulk value of 0.80 Å.	LEED/19/
ZnO (10 $\bar{1}$ 0) (wurtzite)	(1x1)	Oxygen buckled outward by ~ 0.8 Å and Zn contracted by 1.2 Å in the first layer. Bulk layer spacing 1.88 Å.	LEED/20/
ZnO (11 $\bar{2}$ 0) (wurtzite)	(1x1)	Terminated bulk structure, no evidence for reconstruction or relaxation in layer spacings.	LEED/20/

## REFERENCES

1. Yang, W. S., Sokolov, J., Jona, F., and Marcus, P. M. (1982) *Solid State Communications* **41**, 191.
2. Derry, T. E., Smit, L., and van der Veen, J. F. (1986) *Surface Science* **167**, 502.

3. Vidali, G., Cole, M. W., Weinberg, W. H., and Steele, W. A. (1983) *Physical Review Letters* **51**, 118.
4. Wu, N. J. and Ignatiev, A. (1982) *Physical Review B* **25**, 2983.
5. Wu, N. J. and Ignatiev, A. (1983) *Physical Review B* **28**, 7288.
6. Shaw, C. G., Fain, S. C. Jr., Chinn, M. D., and Toney, M. F. (1980) *Surface Science* **97**, 128.
7. Bouldin, C. and Stern, E. A. (1982) *Physical Review B* **25**, 3462.
8. Prutton, M., Ramsey, J. A., Walker, J. A., and Welton-Cook, M. R. (1979) *Journal of Physics C* **12**, 5271.
9. Ignatiev, A., Lee, B. W., and Van Hove, M. A. (1977) in "Proceedings of the 7th International Vacuum Congress and 3rd International Conference on Solid Surfaces", Vienna.
10. Felton, R. C., Prutton, M., P.Tear, S., and Welton-Cook, M. R. (1979) *Surface Science* **88**, 474.
11. Welton-Cook, M. R. and Berndt, W. (1982) *Journal of Physics C* **15**, 569.
12. Urano, T., Kanaji, T., and Kaburagi, M. (1983) *Surface Science* **134**, 109.
13. Mrstik, B. J., Kaplan, R., Reinecke, T. L., Van Hove, M. A., and Tong, S. Y. (1977) *Physical Review B* **15**, 897.
14. Andersson, S., Pendry, J. B., and Echenique, P. M. (1977) *Surface Science* **65**, 539.
15. Jones, R. G., Ainsworth, S., Crapper, M. D., Somerton, C., Woodruff, D. P., Brooks, R. S., Campuzano, J. C., King, D. A., and Lambe, G. M. (1985) *Surface Science* **152/153**, 443.
16. Welton-Cook, M. R. and Prutton, M. (1980) *Journal of Physics C* **13**, 3993.
17. Benbow, R. L., Thuler, M. R., Hurych, Z., Lau, K. H., and Tong, S. Y. (1983) *Physical Review B* **28**, 4161.
18. Lau, B., Mrstik, B. J., Tong, S. Y., and Van Hove, M. A. (to be published)
19. Lubinsky, A. R., Duke, C. B., Chang, S. C., Lee, B. W., and Mark, P. (1976) *Journal of Vacuum Science and Technology* **13**, 189.
20. Duke, C. B., Lubinsky, A. R., Lee, B. W., and Mark, P. (1976) *Journal of Vacuum Science and Technology* **13**, 761.

TABLE IX. Atomic Adsorption on Semiconductor Surfaces

(Where substrate relaxations or reconstructions have been investigated this information is listed in the table after the adsorbate information.)

Adsorption System	Overlayer Cell	Adsorbate Site	Layer Spacing	Method
GaAs (110)-Sb	(1x1)	2-fold	note 1	LEED/1/
GaAs (110)-Al	(1x1)	sub surface	note 2	LEED/2/
Ge (111)-Cl	(2x8)	top	$2.07 \pm 0.03$	SEXAFS/3/
Ge (111)-I	(1x1)	top	$2.50 \pm 0.04$	SEXAFS/4/
Ge (111)-Br	undetermined	top	$2.10 \pm 0.04$	X-ray resonance/5/
Ge (111)-Te	(2x2)	hcp hollow		SEXAFS/4/
Si (111)-Au	( $\sqrt{3} \times \sqrt{3}$ )	note 3	0.30	ALICISS/6/
Si (111)-Br	undetermined	top	$2.18 \pm 0.06$	X-ray resonance/7,8/
Si (111)-Ag	(7x7)	top	$0.70 \pm 0.04$	SEXAFS/9/
Si (111)-Ag	( $2\sqrt{3} \times 2\sqrt{3}$ )	subsurface	$-0.68 \pm 0.15$	SEXAFS/9/
	layer 1-2		$1.36 \pm 0.30$	
	layer 2-3		-0.30	
Si (111)-Cl	( $\sqrt{19} \times \sqrt{19}$ )	top	$1.98 \pm 0.04$	SEXAFS/3/
Si (111)-Cl	(7x7)	top	$2.03 \pm 0.03$	SEXAFS/3/
Si (111)-I	(7x7)	top	$2.44 \pm 0.03$	SEXAFS/4/
	top layer	+15%	$0.90 \pm 0.05$	
Si (111)-NiSi <sub>2</sub>	(1x1)	tetrahedral	note 4	LEED/10/
Si (111)-NiSi <sub>2</sub>	(1x1)	tetrahedral	note 5	MEIS/11/
Si (111)-Se	undetermined	4-fold	1.68	X-ray resonance/12/
Si (111)-Te	(7x7)	2-fold	1.51	SEXAFS/4/



Note 1 -- Sb atoms fill As and Ga type sites in a slightly buckled ( $0.10 \text{ \AA}$ ) first layer spaced  $2.3 \text{ \AA}$  above the GaAs surface, with lateral distortions to form a  $sp^3$  bonded chain.

Note 2 -- Aluminum substitutes for sub-surface Ga atoms. For 0.5 monolayers second layer Ga atoms are replaced, for 1.0 monolayers second and third layer Ga atoms are replaced. Above 1.5 monolayers all near surface Ga atoms are replaced by Al, forming an epitaxial AlAs(110) surface. In all cases the first interlayer spacing contracts by  $\sim 0.10 \text{ \AA}$ .

Note 3 -- Modified triplet cluster model - Au triplets substituted for Si atoms with a  $2.9 \text{ \AA}$  Au-Au bond length between the triplet Au atoms.

Note 4 -- Structure of  $\text{NiSi}_2$ , grown on Si(111) substrate. Forms fluorite structure layer compound Si-Ni-Si with nickel in tetrahedral sites. Silicon layer terminates crystal, with a first-layer contraction of  $\sim 25\%$ .

Note 5 -- Ion scattering investigation of  $\text{NiSi}_2$  - bulk Si interface. Determine Si-Ni-Si layer is  $3.06 \pm 0.08 \text{ \AA}$  above the next non-collinear Si atom (the bulk value is  $0.77 + 2.35 = 3.12$ ). Of the two possible terminations this most closely matches the bulk silicon structure.

## REFERENCES

1. Duke, C. B., Paton, A., Ford, W. K., Kahn, A., and Carelli, J. (1982) *Physical Review B* **26**, 803.
2. Kahn, A., Carelli, J., Kanani, D., Duke, C. B., Paton, A., and Brillson, L. (1981) *Journal of Vacuum Science and Technology* **19**, 331.
3. Citrin, P. H., Rowe, J. E., and Eisenberger, P. (1983) *Physical Review B* **28**, 2299.
4. Citrin, P. H., Eisenberger, P., and Rowe, J. E. (1982) *Physical Review Letters* **48**, 802.
5. Bedzyk, M. and Materlik, G. (1985) *Surface Science* **152/153**, 10.
6. Oura, K., Katayama, M., Shoji, F., and Hanawa, T. (1995) *Physical Review Letters* **55**, 1486.

7. Golovchenko, J. A., Patel, J. R., Kaplan, D. R., Cowan, P. L., and Bedzyk, M. J. (1982) *Physical Review Letters* **49**, 1560.
8. Materlik, G., Frohm, A., and Bedzyk, M. J. (1984) *Physical Review Letters* **52**, 441.
9. Stöhr, J., Jaeger, R., Rossi, G., Kendelewicz, T., and Lindau, I. (1983) *Surface Science* **134**, 813.
10. Yang, W. S., Jona, F., and Marcus, P. M. (1983) *Physical Review B* **28**, 7377.
11. van Loenen, E. J., Frenken, J. W. M., van der Veen, J. F., and Valeri, S. (1985) *Physical Review Letters* **54**, 827.
12. Dev, B. N., Thundat, T., and Gibson, W. M. (1985) *Journal of Vacuum Science and Technology A* **3**, 946.

TABLE XII. Other Molecular Adsorption Structures

System	Structure	Method
Cu (100) HCO <sub>2</sub> disordered	The formate radical is in a plane $\perp$ to the surface with the two oxygens closest to the surface, and a formate O-C-O bond angle of 125° is assumed. The O atoms are slightly off-center above two adjacent 4-fold hollow sites. The O atoms are 1.54 Å above the Cu surface and separated by 2.17 Å. The C atom is 2.11 Å above the surface.	NEXAFS/1/
Cu (110) HCO <sub>2</sub> disordered	The formate radical is in a plane $\perp$ to the surface along the [001] direction. The oxygen atoms are closest to the surface, slightly off-center from two adjacent bridge sites, 1.51 Å above the surface and 2.29 Å apart. The O atom is 2.04 Å above the Cu atom.	NEXAFS/2/
Ni (111) C <sub>2</sub> H <sub>2</sub> (2x2)	The acetylene molecules are adsorbed with the C-C bond parallel to the surface and the center of the C-C bond is over a bridge site. The C-C bond is perpendicular to the Ni-Ni bridge. The C-C bond length is 1.50 Å and the carbon atoms are 2.1±0.10 Å above the surface.	LEED/3/
Pd (111) C <sub>6</sub> H <sub>6</sub> +2CO (3x3)	One benzene and two CO's per unit cell. CO is $\perp$ to the surface, adsorbed 1.30 Å above 3-fold fcc sites. Benzene is parallel to the surface, centered 2.25 Å over a 3-fold hcp site.	LEED/4/
Pt (111) C <sub>2</sub> H <sub>2</sub> disordered	Acetylene bonded parallel to surface with C-C bond length 1.45±0.03 Å.	NEXAFS/5/

(continued)

TABLE XII (continued)

System	Structure	Method
Pt (111) C <sub>2</sub> H <sub>3</sub> disordered	Ethylidyne (CCH <sub>3</sub> ) bonded $\perp$ to surface with C-C bond $1.47 \pm 0.03$ Å.	NEXAFS/6/
Pt (111) C <sub>2</sub> H <sub>3</sub> disordered	Ethylidyne (CCH <sub>3</sub> ) bonded $\perp$ to surface with C-C bond $1.49 \pm 0.02$ Å.	NMR/7/
Pt (111) C <sub>2</sub> H <sub>3</sub> (2x2)	Ethylidyne (CCH <sub>3</sub> ) bonded $\perp$ to surface in 3-fold fcc sites, C-C bond $1.50 \pm 0.05$ Å and C-surface $\perp$ distance $1.20 \pm 0.05$ Å.	LEED/8/
Pt (111) C <sub>2</sub> H <sub>4</sub> disordered	Ethylene bonded parallel to surface with C-C bond length $1.49 \pm 0.03$ Å.	NEXAFS/5/
Pt (111) 2C <sub>6</sub> H <sub>6</sub> disordered	Benzene ring parallel to surface, C-C bond length $1.40 \pm 0.02$ Å.	NEXAFS/9/
Pt (111) 2C <sub>6</sub> H <sub>6</sub> +4CO (2 $\sqrt{3}$ x4) rect	Benzene ring parallel to surface over bridge sites, 2.10 Å above surface, with two benzenes and four CO's per unit cell. CO also over bridge sites, 1.45 Å above the metal surface. Benzene ring is expanded with small in-plane distortions consistent with local symmetry.	LEED/10/
Rh (111) C <sub>2</sub> H <sub>3</sub> (2x2)	Ethylidyne (CCH <sub>3</sub> ) is adsorbed with the C-C axis $\perp$ to the surface with a $1.45 \pm 0.10$ Å bond length. The terminal carbon atom is $1.31 \pm 0.10$ Å above a 3-fold hcp hollow site.	LEED/11/

(continued)

TABLE XII (continued)

System	Structure	Method
Rh (111) $C_2H_3+CO$ $c(4 \times 2)$	Ethylidyne ( $CCH_3$ ) is adsorbed in 3-fold fcc sites with a $1.30 \pm 0.05$ Å metal-terminal carbon distance. The C-C axis is $\perp$ to the surface with a $1.45 \pm 0.05$ Å bond length. CO molecules are adsorbed $\perp$ to the surface in 3-fold hcp sites, carbon atom down, with a carbon-metal bond length of $1.30 \pm 0.05$ Å and a C-O bond length of $1.17 \pm 0.05$ Å.	LEED/12/
Rh (111) $C_2H_3+NO$ $c(4 \times 2)$	Ethylidyne ( $CCH_3$ ) is adsorbed in 3-fold hcp sites with a $1.30 \pm 0.05$ Å metal-terminal carbon distance. The C-C axis is $\perp$ to the surface with a $1.45 \pm 0.05$ Å bond length. NO molecules are adsorbed $\perp$ to the surface in 3-fold hcp sites, nitrogen atom down, with a nitrogen-metal bond length of $1.30 \pm 0.05$ Å and a N-O bond length of $1.17 \pm 0.05$ Å.	LEED/12/
Rh (111) $C_6H_6+2CO$ $(3 \times 3)$	One benzene and two CO's per unit cell. CO is $\perp$ to the surface, adsorbed 1.30 Å above 3-fold hcp sites. Benzene is parallel to the surface, centered 2.20 Å over a 3-fold hcp site. Slight in-plane Kekulé distortion of the benzene molecule.	LEED/13/

(continued)

TABLE XII (continued)

System	Structure	Method
Rh (111) C <sub>6</sub> H <sub>6</sub> +CO c(2√3x4) rect	Benzene is coadsorbed with CO, each with one molecule per unit cell, both centered over 3-fold hcp sites, benzene is parallel to and 2.25±0.05 Å above the surface. CO is ⊥ to the surface and the metal-carbon spacing is 1.50±0.05 Å. The benzene molecule has an in-plane Kekulé distortion, with alternating long and short bonds.	LEED/14/

## REFERENCES

1. Outka, D. A., Madix, R. J., and Stöhr, J. (1985) *Surface Science* **164**, 235.
2. Puschman, A., Haase, J., Crapper, M. D., Riley, C. E., and Woodruff, D. P. (1985) *Physical Review Letters* **54**, 2250.
3. Casalone, G., Cattania, M. G., Merati, F., and Simonetta, M. (1982) *Surface Science* **120**, 171.
4. Ohtani, H., Van Hove, M. A., and Somorjai, G. A. (to be published)
5. Stöhr, J., Sette, F., and Johnson, A. L. (1984) *Physical Review Letters* **53**, 1684.
6. Horsley, J. A., Stöhr, J., and Koestner, R. J. (1985) *Journal of Chemical Physics* **83**, 3146.
7. Wang, P.-K., Slichter, C. P., and Sinfelt, J. J. (1985) *Journal of Physical Chemistry* **89**, 3606.
8. Kesmodel, L. L., Dubois, L. H., and Somorjai, G. A. (1979) *Journal of Chemical Physics* **70**, 2180.

9. Horsley, J. A., Stöhr, J., Hitchcock, A. P., Newbury, D. C., Johnson, A. L., and Sette, F. (1985) *Journal of Chemical Physics* **83**, 6099.
10. Ogletree, D. F., Van Hove, M. A., and Somorjai, G. A. (1987) *Surface Science* **187**, 1.
11. Koestner, R. J., Van Hove, M. A., and Somorjai, G. A. (1982) *Surface Science* **121**, 321.
12. Bent, B. E., Mate, C. M., Koestner, R. J., Blackman, G. S., Kao, C. T., Hove, M. A. Van, and Somorjai, G. A. (to be published)
13. Van Hove, M. A., Lin, R. F., Blackman, G. A., and Somorjai, G. A. (to be published) *Acta Crystallographica*
14. Van Hove, M. A., Lin, R. F., and Somorjai, G. A. (1986) *Journal of the American Chemical Society* **108**, 2532.

## 6. FUTURE NEEDS AND DIRECTIONS OF SURFACE STUDIES

It is our hope that this review conveys the rapid developments in surface structural chemistry during the past decade. We have attempted to describe the various techniques that are employed and the types of structural information that has been obtained from experiment.

The theories of surface structure and bonding have been reviewed. It should be clear to the reader that surface structural chemistry is indeed a frontier area for both theorists and experimental researchers. From an experimentalists viewpoint the data base of atomic and molecular surface structures is very small at present. Most investigations have been carried out on flat, low Miller index surfaces of monatomic solids, either clean or with atomic or small molecules as adsorbates.

The structure of clean surfaces of polyatomic solids (alloys, halides, oxides, sulfides, etc.) should be explored along with molecular solid surfaces (organic systems). More open, rough surfaces with high Miller indices should be investigated, including the structure of atoms and molecules bonded to steps and kinks on surfaces. Such sites are known to be key for some important surface chemical processes, but little is known of their structure.

Molecular adsorbates of increasing size should be a fertile area of research. Such research could lead to the study of biological surfaces. The surface structure at solid-gas, solid-liquid, and solid-solid interfaces should be explored and compared with the results for the solid-vacuum interface. Structural properties of thin films, electrodes, and composite materials can be obtained in this way.

Time resolved surface structure analysis can be employed for studies of order-order and order-disorder surface phase transitions, growth, and evaporation. Surface structures and their alterations could then be explored under extreme conditions of high temperature and irradiation by photons or particles.

Once bond distances and angles become available for a large number of systems with diverse chemistry, correlation should lead to a deeper understanding of the nature of the surface chemical bond.



## REFERENCES

1. Ogletree, D. F., Van Hove, M. A., and Somorjai, G. A. (1987) *Surface Science* **187**, 1.
2. Moritz, W. and Wolf, D. (1985) *Surface Science* **163**, L655.
3. Heilmann, P., Lang, E., Heinz, K., and Müller, K. (1984) in "Determination of Surface Structure by LEED" (P. M. Marcus and F. Jona, ed.), Plenum Press, New York.
4. Ogletree, D. F., Somorjai, G. A., and Katz, J. E. (1986) *Review of Scientific Instruments* **57**, 3012.
5. Stair, P. C. (1980) *Review of Scientific Instruments* **51**, 132.
6. McRae, E. G., Malic, R. A., and Kapilow, D. A. (1985) *Review of Scientific Instruments* **56**, 2077.
7. Ogletree, D. F. (1986) *PhD Thesis*, Physics Department, University of California at Berkeley.
8. Weiss, A. H., Rosenberg, I. J., Canter, K. F., Duke, C. B., and Paton, A. (1983) *Physical Review B* **27**, 867.
9. Pierce, D. T. and Coletta, R. J. (1981) in "Advances in Electronic and Electron Physics" **56** (C. Marton, ed.), Academic Press, New York.
10. Masud, N., Kinniburgh, C. G., and Pendry, J. B. (1977) *Journal of Physics C* **10**, 1.
11. Barton, J. J. and Van Hove, M. A. (to be published)
12. Saldin, D. K., Pendry, J. B., Van Hove, M. A., and Somorjai, G. A. (1985) *Physical Review B* **31**, 1216.
13. Rous, P. R., Pendry, J. B., Saldin, D. K., Heinz, K., Müller, K., and Bickel, N. (1986) *Physical Review Letters* **57**, 2951.
14. Orders, P. J., Sinkovic, B., Fadley, C. S., Trehan, R., Hussain, Z., and Lecante, J. (1984) *Physical Review B* **30**, 1838.
15. Barton, J. J., Robey, S. W., and Shirley, D. A. (1986) *Physical Review B* **34**, 778.
16. Norman, D., Stöhr, J., Jaeger, R., Durham, P. J., and Pendry, J. B. (1983) *Physical Review Letters* **51**, 2052.

17. Sette, F., Stöhr, J., Kollin, E. B., Dwyer, D. J., Gland, J. L., Robbins, J. L., and Johnson, A. L. (1985) *Physical Review Letters* **54**, 935.
18. Müller, E. W. (1951) *Zeitschrift Physik* **136**, 131.
19. Ehrlich, G. (1985) in "*The Structure of Surfaces*" (M. A. Van Hove and S. Y. Tong, ed.), Springer Verlag, Berlin.
20. Batra, I. P., Barker, J. A., and Auerbach, D. J. (1983) *Journal of Vacuum Science and Technology A* **2**, 943.
21. Binnig, G., Rohrer, H., Gerber, C., and Weibel, E. (1982) *Applied Physics Letters* **40**, 178.
22. Binnig, G., Rohrer, H., Gerber, C., and Weibel, E. (1983) *Surface Science* **131**, L379.
23. Copel, M. and Gustafsson, T. (1986) *Physical Review Letters* **57**, 723.
24. Ibach, H. and Mills, D. L. (1982) "*Electron Energy Loss Spectroscopy and Surface Vibrations*", Academic Press, New York.
25. Shen, Y. R. (1985) in "*The Structure of Surfaces*" (M. A. Van Hove and S. Y. Tong, ed.), p. 77, Springer Verlag, Berlin.
26. Madey, T. E. (1985) in "*The Structure of Surfaces*" (M. A. Van Hove and S. Y. Tong, ed.), p. 264, Springer Verlag, Berlin.
27. Pendry, J. B. (1974) "*Low Energy Electron Diffraction*", Academic Press, London.
28. Van Hove, M. A. and Tong, S. Y. (1979) "*Surface Crystallography by LEED*", Springer Verlag, Berlin.
29. Clarke, L. J. (1985) "*Surface Crystallography; An Introduction to Low Energy Electron Diffraction*", Wiley-Interscience, Chichester.
30. Li, C. H., Tong, S. Y., and Mills, D. L. (1980) *Physical Review B* **21**, 3057.
31. Einstein, T. L. (1982) *Applications of Surface Science* **11/12**, 42.
32. Rosei, R., de Crescenzi, M., Sette, F., Quaresima, C., Savoia, A., and Perfetti, P. (1983) *Physical Review B* **28**, 1161.
33. Polizzi, S., Antonangeli, F., Chiarello, G., and de Crescenzi, M. (1984) *Surface Science* **136**, 555.
34. Kincaid, B. M., Meixner, A. E., and Platzman, P. M. (1978) *Physical Review Letters* **40**, 1296.

35. Csillag, S., Johnson, D. E., and Stern, E. A. (1981) in *"EXAFS Spectroscopy Techniques and Applications"* (B. K. Teo and D. C. Joy, ed.), Plenum, New York.
36. Rosenblatt, D. H., Tobin, J. G., Mason, M. G., Davis, R. F., Kevan, S. D., Shirley, D. A., Li, C. H., and Tong, S. Y. (1981) *Physical Review B* **23**, 3828.
37. Kevan, S. D., Davis, R. F., Rosenblatt, D. H., Tobin, J. G., Mason, M. G., Shirley, D. A., Li, C. H., and Tong, S. Y. (1981) *Physical Review Letters* **46**, 1629.
38. Barton, J. J., Bahr, C. C., Hussain, Z., Robey, S. W., Tobin, J. G., Klebanoff, L. E., and Shirley, D. A. (1983) *Physical Review Letters* **51**, 272.
39. Teo, B. K. and Joy, D. C. eds. (1981) *"EXAFS Spectroscopy Techniques and Applications"*, Plenum, New York.
40. Bouldin, C. and Stern, E. A. (1982) *Physical Review B* **25**, 3462.
41. Lee, P. A. (1976) *Physical Review B* **13**, S261.
42. Lee, P. A., Citrin, P. H., Eisenberger, P., and Kincaid, B. M. (1981) *Reviews of Modern Physics* **53**, 769.
43. Stöhr, J., Jaeger, R., and Brennan, S. (1982) *Surface Science* **117**, 503.
44. Rothberg, G. M., Choudhary, K. M., den Boer, M. L., Williams, G. P., Hecht, M. H., and Lindau, I. (1984) *Physical Review Letters* **53**, 1183.
45. Stöhr, J. and Jaeger, R. (1982) *Physical Review B* **26**, 4111.
46. Norman, D., Stöhr, J., Jaeger, R., Durham, P. J., and Pendry, J. B. (1983) *Physical Review Letters* **51**, 2052.
47. Bunker, G. and Stern, E. A. (1984) *Physical Review Letters* **52**, 1990.
48. Stöhr, J. (1985) in *"The Structure of Surfaces"* (M. A. Van Hove and S. Y. Tong, ed.), Springer Verlag, Berlin.
49. Pendry, J. B. (1976) *Surface Science* **57**, 679.
50. Plummer, E. W. and Eberhardt, W. (1982) *"Advances in Chemical Physics"*, Wiley, New York.
51. Fadley, C. S. (1983) in *"Progress in Surface Science"* (S. G. Davidson, ed.), Pergamon, New York.
52. Einstein, T. L. *private communication*.

53. Masud, N., Kinniburgh, C. G., and Pendry, J. B. (1977) *Journal of Physics C* **10**, 1.
54. Masud, N. (1980) *Journal of Physics C* **13**, 6359.
55. Maksym, P. A. and Beeby, J. L. (1984) *Surface Science* **140**, 77.
56. Maksym, P. A. and Beeby, J. L. (1981) *Surface Science* **110**, 423.
57. Hirsch, P. B., Howie, A., Nicholson, R. B., Pashley, D. W., and Whelan, M. J. (1965) "*Electron Microscopy of Thin Crystals*", Butterworth, London.
58. Goodman, P. ed. (1981) "*Fifty Years of Electron Diffraction*", D. Reidel, Dordrecht.
59. Duke, C. B. and Laramore, G. E. (1970) *Physical Review B* **2**, 4765.
60. Duke, C. B. and Laramore, G. E. (1970) *Physical Review B* **2**, 4783.
61. Ibach, H. and Mills, D. L. (1982) "*Electron Energy Loss Spectroscopy and Surface Vibrations*", Academic Press, New York.
62. Crowell, J. E., Koestner, R. J., Dubois, L. H., Van Hove, M. A., and Somorjai, G. A. (1982) in "*Recent Advances in Analytical Spectroscopy*" (K. Fuwa, ed.), p. 211, Pergamon Press, New York.
63. Tong, S. Y., Li, C. H., and Mills, D. L. (1981) *Physical Review B* **24**, 806.
64. Ertl, G. and Küppers, J. (1974) "*Low Energy Electrons and Surface Chemistry*", Verlag Chemie, Weinheim.
65. Avouris, P., DiNardo, N. J., and Demuth, J. E. (1984) *Journal of Chemical Physics* **80**, 491.
66. Taft, T. and Zhu, J. (1982) *Ultramicroscopy* **9**, 349.
67. Aberdam, D., Baudoing, R., Blanc, E., and Gaubert, C. (1978) *Surface Science* **71**, 279.
68. Baudoing, R., Blanc, E., Gaubert, C., Gautheir, Y., and Gnuchev, N. (1983) *Surface Science* **128**, 22.
69. Asaad, W. N. (1963) *Nuclear Physics* **44**, 415.
70. McGuire, E. J. (1971) *Sandia Laboratories Report* **SC-RR-710075**
71. Umbach, E. and Hussain, Z. (1984) *Physical Review Letters* **52**, 457.
72. Truhlar, D. G., Abdallah, J. Jr., and Smith, R. L. (1974) *Advances in Chemical Physics* **25**, 211.
73. Beeby, J. L. (1968) *Journal of Physics C* **1**, 1982.
74. Andersson, S. and Pendry, J. B. (1980) *Journal of Physics C* **13**, 2547.
75. Andersson, S. and Pendry, J. B. (1980) *Journal of Physics C* **13**, 3547.

76. Pendry, J. B. (1975) *Journal of Physics C* **8**, 2413.
77. Zimmer, R. S. and Holland, B. W. (1975) *Journal of Physics C* **8**, 2395.
78. Saldin, D. K., Vvedensky, D. D., and Pendry, J. B. (1985) in "*The Structure of Surfaces*" (M. A. Van Hove and S. Y. Tong, ed.), p. 131, Springer Verlag, Berlin.
79. Saldin, D. K., Pendry, J. B., Van Hove, M. A., and Somorjai, G. A. (1985) *Physical Review B* **31**, 1216.
80. Van Hove, M. A., Lin, R. F., and Somorjai, G. A. (1983) *Physical Review Letters* **51**, 778.
81. Levinson, H. J., Plummer, E. W., and Feibelman, P. J. (1979) *Physical Review Letters* **43**, 952.
82. Li, C. H., Lubinsky, A. R., and Tong, S. Y. (1978) *Physical Review B* **17**, 3128.
83. Tong, S. Y. *private communication*.
84. Gadzuk, J. W. (1976) *Surface Science* **60**, 76.
85. Mehl, M. J. and Einstein, T. L. (to be published)
86. Van Hove, M. A. and Somorjai, G. A. (1982) *Surface Science* **114**, 171.
87. Batra, I. P. (1982) in "*Topics in Current Physics*" **29** (W. H. Marlow, ed.), p. 55, Springer Verlag, Heidelberg.
88. Louie, S. G. and Cohen, M. L. (1984) *Annual Review of Physical Chemistry* **35**, 537.
89. Louie, S. G. (1985) in "*Electronic Structure, Dynamics, and Quantum Structural Properties of Condensed Matter*" (J. T. Devreese and P. Van Camp, ed.), p. 335, Plenum, New York and London.
90. Ohnishi, S., Freeman, A. J., and Wimmer, E. (1984) *Physical Review B* **29**, 5267.
91. Smith, J. R. (Ed.) (1980) "*Theory of Chemisorption*", Springer Verlag, Berlin.
92. Fu, C. L., Ohnishi, S., Wimmer, E., and Freeman, A. J. (1984) *Physical Review Letters* **53**, 675.
93. Mattheiss, L. F. and Hamann, D. R. (1984) *Physical Review B* **29**, 5372.
94. Inglesfield, J. E. (to be published)
95. Hoffmann, R. (1963) *Journal of Chemical Physics* **39**, 137.
96. Lohr, L. L. Jr. and Pyykkö, P. (1979) *Chemical Physics Letters* **62**, 333.

97. Messmer, R. P. (1985) *Surface Science* **158**, 40.
98. Wolfsberg, M. and Helmholtz, L. (1952) *Journal of Chemical Physics* **20**, 837.
99. Bennett, A. J., McCarroll, B., and Messmer, R. P. (1971) *Surface Science* **24**, 191.
100. Bennett, A. J., McCarroll, B., and Messmer, R. P. (1971) *Physical Review B* **3**, 1397.
101. Pople, J. A. and Segal, G. A. (1966) *Journal of Chemical Physics* **44**, 3289.
102. Anderson, A. B. and Hoffmann, R. (1974) *Journal of Chemical Physics* **60**, 4271.
103. Anderson, A. B. (1975) *Journal of Chemical Physics* **62**, 1187.
104. Simonetta, M. and Gavezzotti, A. (1980) *Advances in Quantum Chemistry* **12**, 103.
105. Baetzold, R. C. and Hamilton, J. F. (1983) *Progress in Solid State Chemistry* **15**, 1.
106. Ray, N. K. and Anderson, A. B. (1983) *Surface Science* **125**, 803.
107. Froitzheim, H., Hopster, H., Ibach, H., and Lehwald, S. (1977) *Applied Physics* **13**, 147.
108. Ogletree, D. F., Van Hove, M. A., and Somorjai, G. A. (1986) *Surface Science* **173**, 351.
109. Krebs, H. I. and Lüth, H. (1977) *Applied Physics* **14**, 337.
110. Norton, P. R., Goodale, J. W., and Selkirk, E. B. (1979) *Surface Science* **83**, 189.
111. Campbell, C. T., Ertl, G., Kuipers, H., and Segner, J. (1980) *Journal of Chemical Physics* **73**, 5862.
112. Gavezzotti, A., Ortoleva, E., and Simonetta, M. (1982) *Journal of the Chemical Society, Faraday Transactions 1* **78**, 425.
113. Gavezzotti, A. and Simonetta, M. (1980) *Surface Science* **99**, 453.
114. Anderson, A. B. and Hubbard, T. (1980) *Surface Science* **99**, 384.
115. Ibach, H. and Lehwald, S. (1978) *Journal of Vacuum Science and Technology* **15**, 407.
116. Johnson, K. H. (1966) *Journal of Chemical Physics* **45**, 3085.
117. Korringa, J. (1947) *Physica (Utrecht)* **13**, 392.

118. Kohn, W. and Rostoker, N. (1954) *Physical Review* **94**, 1111.
119. Yang, C. Y., Johnson, K. H., Salahub, D. R., Kaspar, J., and Messmer, R. P. (1981) *Physical Review B* **24**, 5673.
120. Yang, C. Y. and Rabii, S. (1975) *Physical Review A* **12**, 362.
121. Kohn, W. and Sham, L. J. (1965) *Physical Review A* **140**, 1133.
122. Slater, J. C. (1951) *Physical Review* **81**, 385.
123. Schwarz, K. (1972) *Physical Review B* **5**, 2466.
124. Batra, I. P. and Brundle, C. R. (1976) *Surface Science* **57**, 12.
125. Danese, J. B. and Connolly, J. W. D. (1974) *Journal of Chemical Physics* **61**, 3063.
126. Danese, J. B. (1974) *Journal of Chemical Physics* **61**, 3071.
127. Johnson, K. H. (1971) *International Journal of Quantum Chemistry* **S4**, 153.
128. Kasowski, R. V. (1947) *Physical Review Letters* **33**, 1147.
129. Slater, J. C. and Johnson, K. H. (1972) *Physical Review B* **5**, 844.
130. Yu, H. L. (1977) *Physical Review B* **15**, 3609.
131. Demuth, J. E. and Eastman, D. E. (1976) *Solid State Communications* **18**, 1497.
132. Messmer, R. P. and Lamson, S. H. (1979) *Chemical Physics Letters* **65**, 465.
133. Norton, P. R., Tapping, R. L., and Goodale, J. W. (1978) *Surface Science* **72**, 33.
134. Messmer, R. P., Lamson, S. H., and Salahub, D. R. (1982) *Physical Review B* **25**, 3576.
135. Demuth, J. E. and Eastman, D. E. (1974) *Physical Review Letters* **32**, 1123.
136. Rösch, N. and Rhodin, T. N. (1974) *Physical Review Letters* **32**, 1189.
137. Howard, I. A. and Dresselhaus, G. (1984) *Surface Science* **136**, 229.
138. Yang, C. Y., Yu, H. L., and Case, D. A. (1981) *Chemical Physics Letters* **81**, 170.
139. Baerends, E. J., Ellis, D. E., and Ros, P. (1973) *Chemical Physics* **2**, 41.
140. Rosén, A., Ellis, D. E., Adachi, H., and Averill, F. W. (1976) *Journal of Chemical Physics* **65**, 3629.
141. Mulliken, R. S. (1955) *Journal of Chemical Physics* **23**, 1833.

142. Mulliken, R. S. (1955) *Journal of Chemical Physics* **23**, 1841.
143. Eastman, D. E. and Cashion, J. K. (1971) *Physical Review Letters* **27**, 1520.
144. Hagstrum, H. D. and Becker, G. E. (1971) *Journal of Chemical Physics* **54**, 1015.
145. Andersson, S. (1977) *Solid State Communications* **21**, 75.
146. Andersson, S. (1977) in "Proceedings of the 7th International Vacuum Congress and 3rd International Conference on Solid Surfaces", p. 815, Vienna.
147. Rosén, A., Baerends, E. J., and Ellis, D. E. (1979) *Surface Science* **82**, 139.
148. Ellis, D. E., Baerends, E. J., Adachi, H., and Averill, F. W. (1977) *Surface Science* **64**, 649.
149. Baerends, E. J. and Ros, P. (1975) *Molecular Physics* **30**, 1735.
150. Allyn, C. L., Gustafsson, T., and Plummer, E. W. (1977) *Chemical Physics Letters* **47**, 127.
151. Davenport, J. (1976) *Physical Review Letters* **36**, 945.
152. Cao, Pei-Lin, Ellis, D. E., and Freeman, A. T. (1982) *Physical Review B* **25**, 2124.
153. Ziegler, A. and Rank, A. (1977) *Theoretica Chimica Acta* **46**, 1.
154. Bagus, P. S. and Hermann, K. (1981) *Journal of Vacuum Science and Technology* **18**, 435.
155. Bagus, P. S. (1981) *Physical Review B* **23**, 2065.
156. Bagus, P. S. (1981) *Physical Review B* **23**, 5464.
157. Bagus, P. S., Brundle, C. R., Hermann, K., and Menzel, D. (1980) *Journal of Electron Spectroscopy and Related Phenomena* **20**, 253.
158. Seeland, M. and Bagus, P. S. (1983) *Physical Review B* **28**, 2023.
159. Bauschlicher, C. W. Jr., Walch, S. P., Bagus, P. S., and Brundle, C. R. (1983) *Physical Review Letters* **50**, 864.
160. Bauschlicher, C. W. Jr., Bagus, P. S., and Yarkony, D. R. (1981) *Proceedings of the Workshop on Recent Developments and Applications of the MC-HF method*, 67. , Lawrence Berkeley Laboratory
161. Brundle, C. R., Bagus, P. S., Menzel, D., and Hermann, K. (1981) *Physical Review B* **24**, 7041.



162. Bagus, P. S. and Roos, B. O. (1981) *Journal of Chemical Physics* **12**, 5961.
163. Bagus, P. S., Nelin, C. J., and Bauschlicher, C. W. Jr. (1983) *Physical Review B* **28**, 5423.
164. Bauschlicher, C. W. Jr. and Bagus, P. S. (1984) *Physical Review Letters* **52**, 200.
165. Melius, C. F., Bisson, C. L., and Wilson, W. D. (1978) *Physical Review B* **18**, 1647.
166. Lehwald, S. and Ibach, H. (1982) in "Vibrations at Surfaces" (R. Caudano, J.-M. Gilles, and A. A. Lucas, ed.), p. 137, Plenum, New York.
167. Stöhr, J., Jaeger, R., and Kendelewicz, T. (1982) *Physical Review Letters* **49**, 142.
168. Van Hove, M. A. and Tong, S. Y. (1975) *Journal of Vacuum Science and Technology* **12**, 230.
169. Marcus, P. C., Demuth, J. E., and Jepsen, D. W. (1975) *Surface Science* **53**, 501.
170. Upton, T. H. and Goddard, W. A. III (1981) *Physical Review Letters* **46**, 1635.
171. Upton, T. H. and Goddard, W. A. III (1981) *Critical Reviews of Solid State and Material Science* **10**, 261.
172. Batra, I. P., Bagus, P. S., and Hermann, K. (1984) *Physical Review Letters* **52**, 384.
173. Goddard, W. A. III, Redondo, A., and McGill, T. C. (1976) *Solid State Communications* **18**, 981.
174. Redondo, A., Goddard, W. A. III, Swarts, C. A., and McGill, T. C. (1981) *Journal of Vacuum Science and Technology* **19**, 498.
175. Ibach, H., Bruchmann, H. D., and Wagner, H. (1982) *Applied Physics A* **29**, 113.
176. Whitten, J. L. (1973) *Journal of Chemical Physics* **58**, 4496.
177. Whitten, J. L. and Pakkanen, T. A. (1980) *Physical Review B* **21**, 4357.
178. Cremaschi, P. and Whitten, J. L. (1981) *Surface Science* **112**, 343.
179. Madhavan, P. and Whitten, J. L. (1982) *Journal of Chemical Physics* **77**, 2673.

180. Hunt, W. J., Hay, P. J., and Goddard, W. A. III (1972) *Journal of Chemical Physics* **57**, 738.
181. Upton, T. H. and Goddard, W. A. III (1982) in "Chemistry and Physics of Solid Surfaces" **3** (R. Vanselow and W. England, ed.), p. 127, CRC Press, Boca Raton.
182. Bobrowicz, F. W. and Goddard, W. A. III (1977) in "Modern Theoretical Chemistry; Methods of Electronic Structure Theory" **3** (H. F. Schaefer, III, ed.), p. 738, Plenum Press, New York.
183. Allison, J. N. and Goddard, W. A. III (1982) *Surface Science* **115**, 553.
184. Melius, C. F., Olafson, B. D., and Goddard, W. A. III (1974) *Chemical Physics Letters* **28**, 457.
185. Meutner, J. S. (1975) *Journal of Molecular Spectroscopy* **55**, 490.
186. Tracy, J. C. (1972) *Journal of Chemical Physics* **56**, 2736.
187. Barton, J. J., Swarts, C. A., Goddard, W. A. III, and McGill, T. C. (1980) *Journal of Vacuum Science and Technology* **17**, 164.
188. Goddard, W. A. III, Barton, J. J., Redondo, A., and McGill, T. C. (1978) *Journal of Vacuum Science and Technology* **15**, 1274.
189. Tong, S. Y., Lubinsky, A. R., Mrstik, B. J., and Van Hove, M. A. (1978) *Physical Review B* **17**, 3303.
190. Kahn, A., So, E., Mark, P., Duke, C. B., and Meyer, R. J. (1978) *Journal of Vacuum Science and Technology* **15**, 1223.
191. Miller, D. J. and Haneman, D. (1978) *Journal of Vacuum Science and Technology* **15**, 1267.
192. Meyer, R. J., Duke, C. B., Patton, A., Kahn, A., So, E., Yeh, J. L., and Mark, P. (1979) *Physical Review B* **19**, 5194.
193. Stöhr, J., Bauer, R. S., McMenamin, J. C., Johansson, L. I., and Brennan, S. (1979) *Journal of Vacuum Science and Technology* **16**, 1195.
194. Wilkins, C. W., Hagen, K., Hedberg, L., Shen, O., and Hedberg, K. (1975) *Journal of the American Chemical Society* **97**, 6352.
195. Rojhtantalab, H. (1976) *Spectrochimica Acta A* **32**, 947.
196. Carlson, R. and Meek, D. W. (1974) *Inorganic Chemistry* **13**, 1741.
197. Brundle, C. R. and Seybold, D. (1979) *Journal of Vacuum Science and Technology* **16**, 1186.

198. Pianetta, P., Lindau, I., Garner, C., and Spicer, W. E. (1975) *Physical Review Letters* **35**, 1356.
199. Spicer, W. E., Pianetta, P., Lindau, I., and Chye, P. W. (1977) *Journal of Vacuum Science and Technology* **14**, 885.
200. Chye, P. W., Pianetta, P., Lindau, I., and Spicer, W. E. (1977) *Journal of Vacuum Science and Technology* **14**, 917.
201. Pauling, L. (1960) "*The Nature of the Chemical Bond*", Cornell University Press, Ithaca, New York.
202. Ohtani, H., Kao, C.-T., Van Hove, M. A., and Somorjai, G. A. (1986) *Progress In Surface Science* **23**, 155.

**On the Change of Spectra Associated with Unbounded  
Similarity Transformations of a Many-Particle  
Hamiltonian  
and the Occurrence of Resonance States  
in the Method of Complex Scaling.**

**Part II. Applications to the Hartree-Fock Scheme  
Based on the Bi-Variational Principle.**

Per-Olov Löwdin  
Quantum Theory Project  
Departments of Chemistry and Physics  
University of Florida  
Gainesville, FL 32611.

and

Piotr Froelich  
Department of Quantum Chemistry  
Uppsala University,  
Box 518, S-75120, Uppsala, Sweden

and

Manoj Mishra  
Department of Chemistry  
Indian Institute of Technology  
Bombay 400076, India.

\*Professor Emeritus at Uppsala University, Uppsala, Sweden;  
NFCR Senior Investigator.

## List of Contents

1. Introduction.
  2. Brief Review of some of the Results in the Previous Studies of the Stability Problem for a Pair of Adjoint Operators.
    - 2.1 General Theory.
    - 2.2 Hartree-Fock Scheme.
  3. Similarity Transformation of a Many-Particle Operator and its Consequences in the Hartree-Fock Scheme.
    - 3.1 General Aspects.
    - 3.2 Similarity Transformations in the Hartree-Fock Scheme; the Tarfala Theorem.
    - 3.3 The Hartree-Fock Scheme for a Complex Symmetric Operator.
    - 3.4 Special Case when  $T^\dagger = T^* = T$ .
    - 3.5 Invariance Property of "Complex Symmetry" under Restricted Similarity Transformations.
  4. Properties of a Many-Particle Hamiltonian under Complex Scaling.
    - 4.1. Method of Complex Scaling; the Dilatation Operator.
    - 4.2. The Coulombic Many-Particle Hamiltonian.
    - 4.3. Application of the Hartree-Fock Method.
  5. Numerical Application of the Complex Scaling Method in the Hartree-Fock Approximation.
  6. Conclusions and Acknowledgements.
- References.

Appendix A. Treatment of One-Particle Subspaces of Order  $N$  which are Stable under Complex Conjugation.

The theory of the change of spectra of a many-particle operator associated with an unbounded similarity transformation - due to the change in the boundary conditions - is briefly reviewed. Special attention is given to operators which remain complex symmetric under certain restricted transformations.

In this paper, the general theory developed in Part I is applied to the Hartree-Fock Scheme for a transformed many-electron Hamiltonian. It is shown that, if the transformation is a product of one-electron transformations, then the Fock-Dirac operator as well as the effective Hamiltonian undergo similarity transformations of the one-electron type. The special properties of the Hartree-Fock scheme for a real self-adjoint Hamiltonian based on the bi-variational principle are discussed in greater detail.

In conclusion, the method of complex scaling as an unbounded similarity transformation of the restricted type is briefly discussed, and some numerical applications containing complex eigenvalues - which may be related to resonance states - are given.

## 1. Introduction.

If an arbitrary self-adjoint operator  $H$ , like the Hamiltonian, undergoes a similarity transformation:

$$H^U = U H U^{-1} \quad (1.1)$$

where  $U$  is an unbounded operator, its spectrum  $\{E\}$  is subject to change: some eigenvalues are persistent, others may be lost, and new eigenvalues may occur also in the complex plane. These changes have been discussed in some detail in a previous paper<sup>1</sup>, which will be referred to as reference A. If the operator  $U$  satisfies the special condition:

$$(U^\dagger)^{-1} = U^*, \quad (1.2)$$

the transformed operator has the important symmetry property:

$$(H^U)^\dagger = (H^U)^* \quad (1.3)$$

i.e. it is complex symmetric, which greatly simplifies the theory, and one speaks of a restricted similarity transformation.

The general stability problem for a pair of adjoint many-particle operators  $T$  and  $T^\dagger$  has been discussed in a previous paper<sup>2</sup>, which will be referred to as reference B. The Hartree-Fock scheme for a pair of such operators has also been discussed<sup>3</sup>, and this paper will be referred to as reference C. Some of the most important results in the references A, B, and C will be briefly reviewed here to make the presentation more self-contained.

This study was started in order to find out whether one could find meaningful complex eigenvalues in the Hartree-Fock scheme for a transformed Hamiltonian in the method of complex scaling. This problem was intensely discussed at the 1981 Tarfala Workshop in the Kebnekaise area of the Swedish mountains. It was found that, if the many-electron Hamiltonian undergoes a similarity transformation  $U$  which is a product of one-electron transformations  $u$  - as in the method of complex scaling - then the Fock-Dirac operator  $p$  as well as the effective Hamiltonian  $H_{\text{eff}}$  undergo one-electron similarity

transformations of the type:

$$\rho^u = u \rho u^{-1}, \quad H_{\text{eff}}^u = u H_{\text{eff}} u^{-1} \quad (1.4)$$

and these results were referred to as the "Tarfala Theorem". If one assumes that the eigenfunctions to these operators all belong to the one-electron Hilbert space  $L^2$ , then it is evident that the spectra of the effective Hamiltonians will remain persistent, and - if one starts from a real self-adjoint Hamiltonian and the ordinary Hartree-Fock scheme - one would hence never get any complex one-electron energies, which is of course a negative result. It was hence somewhat surprising to find that numerical applications of the Hartree-Fock scheme in the complex symmetric case based on the bi-variational principle would sometimes lead to complex eigenvalues even when the original Hamiltonian was real and self-adjoint.

The puzzle depended on the simple fact that most physicists using the method of complex scaling had not realized that the associated operator  $u$  - the so-called dilatation operator - was an unbounded operator, and that the change of spectra - e.g. the occurrence of complex eigenvalues - was due to a change of the boundary conditions. Some of these features have been clarified in reference A, and in this paper we will discuss how these properties will influence the Hartree-Fock scheme. The existence of the numerical examples finally convinced us that the Hartree-Fock scheme in the complex symmetric case would not automatically reduce to the ordinary Hartree-Fock scheme in the case when the many-electron Hamiltonian became real and self-adjoint. Some aspects of this problem have been briefly discussed at the 1987 Sanibel Symposium, and a preliminary report has been given in a paper<sup>4</sup> which will be referred to as reference D.

Quite a few years have now gone since the Tarfala Workshop in 1981, and many new results have appeared in the method of complex scaling. The question of the occurrence of complex eigenvalues corresponding to so-called resonance states in the Hartree-Fock approximation for various types of many-electron Hamiltonians is, however, still not completely solved, and the authors feel that time is now mature to bring up these problems for more intense discussions.



## 2. Brief review of some of the Results in the Previous Studies of the Stability Problem for a Pair of Adjoint Operators.

In this section, we will give a brief review of some of the results obtained in the references A, B, and C, and formulate them in a way which is particularly suitable for our present purpose. We will also introduce some new concepts and some new results.

**2.1. General theory.** - Let us consider a pair of adjoint operators,  $T$  and  $T^\dagger$ , defined on a Hilbert space  $H = \{ x \}$  with the binary product  $\langle x|y \rangle$ , where the second position is linear and the first anti-linear. Following Dirac<sup>5</sup>, the bracket  $\langle x|y \rangle$  is considered as the product of a "bra-vector"  $\langle x|$  and a "ket-vector"  $|y\rangle$ . The ket-bra operator  $F = |b\rangle\langle a|$  is further defined through the relation

$$F = |b\rangle\langle a|, \quad Fx = b \langle a|x\rangle, \quad (2.1)$$

where one may put  $|b\rangle = b$  and consider  $\langle a|$  as a linear functional. The operator  $F$  satisfies the algebraic relation  $F^2 = \langle a|b\rangle F$ , and it has a single non-vanishing eigenvalue  $f = \langle a|b\rangle$ . One has further the relations

$$F^\dagger = |a\rangle\langle b|, \quad \text{Tr } F = \langle a|b\rangle. \quad (2.2)$$

In studying the adjoint pair of operators,  $T$  and  $T^\dagger$ , one starts from the eigenvalue problems:

$$TC = \lambda C, \quad T^\dagger D = \mu D, \quad (2.3)$$

subject to the standard boundary conditions, i.e. that for discrete eigenvalues  $\lambda$  and  $\mu$ , the eigenfunctions  $C$  and  $D$  should be elements of the Hilbert space. Using the elementary relation:

$$\begin{aligned} \mu^* \langle D|C \rangle &= \langle \mu D|C \rangle = \langle T^\dagger D|C \rangle = \langle D|TC \rangle = \langle D|\lambda C \rangle = \\ &= \lambda \langle D|C \rangle \end{aligned} \quad (2.4)$$

i.e.

$$(\mu^* - \lambda) \langle D|C \rangle = 0, \quad (2.5)$$

one obtains

$$\langle D | C \rangle = 0, \quad \text{if} \quad \mu \neq \lambda^*, \quad (2.6)$$

which is the so-called bi-orthogonality theorem. On the other hand, if  $\langle D | C \rangle \neq 0$ , one obtains  $\mu = \lambda^*$ . In such a case, it is convenient to introduce the new eigenfunction  $D' = D \langle C | D \rangle^{-1}$ , and one gets the bi-normalization  $\langle D' | C \rangle = 1$ . The elements  $C$  and  $D'$  are said to be *conjugate* eigenfunctions associated with the eigenvalues  $\lambda$  and  $\lambda^*$ . In the following, we will assume that all the eigenfunctions  $D'$  have been properly "normalized" and omit the prime.

If the operator  $T$  has a spectrum consisting of the discrete eigenvalues  $\lambda_1, \lambda_2, \lambda_3, \dots$ , which are all non-degenerate, and the associated eigenelements are arranged in a row vector  $\mathbb{C} = \{C_1, C_2, C_3, \dots\}$ , etc., one obtains the relations

$$T \mathbb{C} = \mathbb{C} \lambda, \quad T^\dagger \mathbb{D} = \mathbb{D} \lambda^*, \quad (2.7)$$

where  $\lambda$  and  $\lambda^*$  are diagonal matrices. Here we have used the convention that, if  $\mathbf{A}$  and  $\mathbf{B}$  are rectangular matrices - including the special cases of row-vectors, column - vectors, and quadratic matrices - then

$$(\mathbf{AB})_{kl} = \sum_{\alpha} A_{k\alpha} B_{\alpha l} \quad (2.8)$$

i.e. the  $k^{\text{th}}$  row of  $\mathbf{A}$  is multiplied with the  $l^{\text{th}}$  column of  $\mathbf{B}$ . For the eigenelements, one has the bi-orthonormality theorem:

$$\langle D | C \rangle = 1. \quad (2.9)$$

If an eigenvalue  $\lambda$  - as defined by the characteristic equation for  $T$  in any matrix representation - is *degenerate*, the situation is more complicated, and the eigenvalue problems (2.3) have to be replaced by the associated *stability problems*; see ref. B, Sec. 4. In matrix theory, the search for the irreducible stable subspaces of  $T$  is reflected in the block-diagonalization of the matrix  $T$  to the so-called classical canonical form  $\lambda$ , in which one has the eigenvalues  $\lambda$  on the main diagonal and, for degenerate eigenvalues, sequences of 1's in the diagonal immediately above

the main diagonal. This means that one may have blocks of the type:

$$(\lambda), \begin{pmatrix} \lambda & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \text{etc.} \quad (2.10)$$

which are referred to as Jordan blocks; the orders of the Jordan blocks for a multiple eigenvalue  $\lambda$  give the Segré characteristics for the degeneracy. Instead of the first relation (2.7), one gets now the more general relation:

$$T C = C \lambda \quad (2.11)$$

where the matrix  $\lambda$  is on classical canonical form.

The stability problem for the adjoint operator  $T^\dagger$  may now be expressed in a similar form:

$$T^\dagger D = D \mu, \quad (2.12)$$

where we have not yet specified the form of the matrix  $\mu$ , since we are also interested in the metric matrix  $\Delta = \langle D | C \rangle$ . In analogy with (2.4), we obtain

$$\begin{aligned} \mu^\dagger \langle D | C \rangle &= \langle D \mu | C \rangle = \langle T^\dagger D | C \rangle = \\ &= \langle D | T C \rangle = \langle D | C \lambda \rangle = \langle D | C \rangle \lambda, \end{aligned} \quad (2.13)$$

i.e.

$$\mu^\dagger \Delta = \Delta \lambda, \quad (2.14)$$

or

$$\mu = (\Delta^\dagger)^{-1} \lambda^\dagger \Delta^\dagger, \quad (2.15)$$

This means that the matrix  $\mu$  is always a similarity trans-

formation of the adjoint matrix  $\lambda^\dagger$ . Introducing the substitution:

$$\mathbb{D}' = \mathbb{D} (\Delta^\dagger)^{-1} = \mathbb{D} < \mathbb{C} | \mathbb{D} >^{-1}, \quad (2.16)$$

one gets directly:

$$\begin{aligned} T^\dagger \mathbb{D}' &= T^\dagger \mathbb{D} (\Delta^\dagger)^{-1} = \mathbb{D} \mu (\Delta^\dagger)^{-1} = \\ &= \mathbb{D}' \Delta^\dagger \mu (\Delta^\dagger)^{-1} = \mathbb{D}' \lambda^\dagger, \end{aligned} \quad (2.17)$$

as well as

$$< \mathbb{D}' | \mathbb{C} > = \Delta^{-1} < \mathbb{D} | \mathbb{C} > = 1. \quad (2.18)$$

We note that, even if  $\lambda^\dagger$  is not on a classical canonical form, since it has the 1's in the Jordan blocks on the diagonal immediately below the main diagonal, the form (2.17) is most convenient for our purposes, since it is compatible with the relation (2.18). In summary, it is hence possible to write the stability relations in the form:

$$T \mathbb{C} = \mathbb{C} \lambda, \quad T^\dagger \mathbb{D} = \mathbb{D} \lambda^\dagger, \quad (2.19)$$

$$< \mathbb{C} | \mathbb{D} > = 1, \quad (2.20)$$

and we will say that the row-vectors  $\mathbb{C}$  and  $\mathbb{D}$  consist of *conjugate* "eigenelements", which are automaticcally bi-orthonormal. It should be observed that the construction (2.16) is such that, if  $\mathbb{X}$  is an arbitrary basis for the space under consideration, one has

$$\mathbb{D} = \mathbb{X} < \mathbb{C} | \mathbb{X} >^{-1} \quad (2.21)$$

which relation is invariant under linear transformations  $\mathbb{X}' = \mathbb{X} \alpha$ . Putting  $\mathbb{X} = \mathbb{C}$ , one obtains that  $\mathbb{D} = \mathbb{C} < \mathbb{C} | \mathbb{C} >^{-1} = \mathbb{C}_r$ , i.e. that  $\mathbb{D}$  is the reciprocal basis to the basis  $\mathbb{C}$ .

In concluding this section, a few words should be said about the case when the operator  $T$  is *complex symmetric* :

$$T^\dagger = T^*. \quad (2.22)$$

The name comes from the fact that, if one studies such an operator in a dual basis  $(\Phi^*, \Phi)$ , it has the matrix representation

$$T = \langle \Phi^* | T | \Phi \rangle, \quad (2.23)$$

with the property:

$$T^\dagger = \langle T \Phi | \Phi^* \rangle = \langle \Phi | T^\dagger \Phi^* \rangle = \langle \Phi | T^* \Phi^* \rangle = T^*, \quad (2.24)$$

i.e.  $T_{kl} = T_{lk}$ . This means that  $T$  is a *symmetric* matrix with complex elements.

Taking the complex conjugate of the relation  $T C = C \lambda$ , one obtains

$$T^\dagger C^* = C^* \lambda^*, \quad (2.25)$$

where  $\lambda^*$  is still a classical canonical form. If  $P$  is the permutation, which reverses the order of the basic elements  $C$  within each Jordan block, one has further  $\lambda^\dagger = P \lambda^* P^{-1}$ . Putting  $X = C^*$  into relation (2.21), one gets finally the preferred solution:

$$D = C^* \langle C | C^* \rangle^{-1}, \quad (2.26)$$

with the properties

$$T^\dagger D = D \lambda^\dagger, \quad \langle C | D \rangle = 1. \quad (2.27)$$

We note that, due to the bi-orthogonality theorem, the metric matrix  $\Delta = \langle C | C^* \rangle$  is automatically block-diagonalized with different blocks for different eigenvalues, and it is then possible to treat the stable subspaces associated with different eigenvalues separately, which is an essential simplification. In addition, the symmetry property (2.24) indicates another simplification from the computational point of view.

**2.2. Hartree-Fock Scheme.-** For the sake of simplicity, we will consider the eigenvalue relations (2.3) for a non-degenerate eigenvalue  $\lambda$ , for which  $\langle C | D \rangle \neq 0$ . If  $C_a$  and  $D_a$  are two approximate solutions - characterized by the index "a" - the best approximations are obtained by using the bi-variational principle (B.3.6). In the Hartree-Fock scheme for N fermions, the approximate solutions are represented by two Slater determinants:

$$C_a = (N!)^{-1/2} |\psi_k(x_i)|, \quad D_a = (N!)^{-1/2} |\phi_l(x_j)|, \quad (2.28)$$

where  $\Psi = \{\psi_1, \psi_2, \dots, \psi_N\}$  and  $\Phi = \{\phi_1, \phi_2, \dots, \phi_N\}$  are row-vectors of one-particle functions of the combined coordinate  $x_i = (r_i, \zeta_i)$ . If one introduces the overlap matrix:

$$\mathfrak{d} = \langle \Phi | \Psi \rangle \quad (2.29)$$

one has according to (C.2.13) that

$$\langle D_a | C_a \rangle = |\mathfrak{d}|, \quad (2.30)$$

and, since  $\langle D_a | C_a \rangle \neq 0$ , this implies that the matrix  $\mathfrak{d}$  is non-singular with the inverse  $\mathfrak{e} = \mathfrak{d}^{-1}$ . According to (A.2.1) the bi-variational expression may be written in the form  $I = \text{Tr } T \Gamma_a$ , where  $\Gamma_a$  is a skew-projector of the form:

$$\Gamma_a = |C_a\rangle \langle D_a| / \langle D_a | C_a \rangle = |C_a\rangle \langle D_a | C_a \rangle^{-1} \langle D_a|, \quad (2.31)$$

which in turn is completely determined by the one-particle operator  $\rho$ :

$$\rho = |\Psi\rangle \langle \Phi | \Psi \rangle^{-1} \langle \Phi|, \quad (2.32)$$

having the properties

$$\rho^2 = \rho, \quad \text{Tr } \rho = N, \quad \rho \neq \rho^\dagger; \quad (2.33)$$

for details, see eqs. (C.2.16-2.21). It has further the kernel:

$$\rho(x_1, x_2) = \sum_{k,l} \psi_k(x_1) e_{kl} \phi_l^*(x_2), \quad (2.34)$$

where the summation over  $k$  and  $l$  goes from 1 to  $N$ , i.e. over the "occupied" one-particle functions. It is obviously a generalization of the Fock-Dirac density matrix<sup>6</sup> in the conventional Hartree-Fock scheme.

If the  $N$ -particle operator  $T$  is expressed in the form

$$T = T_{(0)} + \sum_i T_i + \sum_{i < j} T_{ij} + \dots \quad (2.35)$$

It is then possible to evaluate the bi-variational expression  $I = \text{Tr } T T_a$  in the form:

$$I = T_{(0)} + \int T_1 \rho(x_1, x_1') dx_1 + \\ + (1/2) \int T'_{12} \rho(x_1, x_1') \rho(x_2, x_2') dx_1 dx_2, \quad (2.36)$$

where

$$T'_{12} = T_{12} (1 - P_{12}), \quad (2.37)$$

and one has used the standard convention that the operators  $T_1$ ,  $T'_{12}$ , ... work only on the unprimed coordinates and that one puts  $x_1' = x_1$ ,  $x_2' = x_2$ , ... before integrations are carried out.

By using the bi-variational principle,  $\delta I = 0$ , one obtains according to (C.2.29 - 2.56) the following conditions:

$$T_{\text{eff}}(1) \psi(1) = \psi(1) h_1, \quad (2.38)$$

$$T_{\text{eff}}^\dagger(1) \phi(1) = \phi(1) h_2, \quad (2.39)$$

where the effective one-electron operator  $T_{\text{eff}}(1)$  has the special form:

$$T_{\text{eff}}(1) = T_1 + \int dx_2 T'_{12} \rho(x_2, x_2'), \quad (2.40)$$

and  $h_1$  and  $h_2$  are quadratic matrices of Lagrangian multipliers. If the similarity transformation  $s_1$  brings  $h_1$  to classical canonical form  $t$ , so that

$$s_1 h_1 s_1^{-1} = t, \quad h_1 = s_1 t s_1^{-1}, \quad (2.41)$$

then the *canonical solutions*  $\Psi' = \Psi \mathbf{s}_1$  satisfy the canonical Hartree-Fock equations:

$$T_{\text{eff}}(1) \Psi'(1) = \Psi'(1) \mathbf{t}. \quad (2.42)$$

Using the technique from the previous section, we will now introduce the conjugate solutions  $\Phi'$  through the relation:

$$\Phi' = \Phi \langle \Psi' | \Phi \rangle^{-1} \quad (2.43)$$

and one gets immediately

$$T_{\text{eff}}^\dagger(1) \Phi'(1) = \Phi'(1) \mathbf{t}^\dagger, \quad \langle \Phi' | \Psi' \rangle = 1, \quad (2.44)$$

and the solutions  $\Phi'$  are  $\Psi'$  automatically bi-orthonormal. The proof for the first relation (2.44) follows from the fact that

$$\begin{aligned} T_{\text{eff}}^\dagger \Phi' &= T_{\text{eff}}^\dagger \Phi \langle \Psi' | \Phi \rangle^{-1} = \Phi \mathbf{h}_2 \langle \Psi' | \Phi \rangle^{-1} = \\ &= \Phi \langle \Psi' | \Phi \rangle \mathbf{h}_2 \langle \Psi' | \Phi \rangle^{-1} = \Phi \langle \Psi' | T_{\text{eff}} \Phi \rangle \langle \Psi' | \Phi \rangle^{-1} = \\ &= \Phi \langle T_{\text{eff}}^\dagger \Psi' | \Phi \rangle \langle \Psi' | \Phi \rangle^{-1} = \Phi \langle \Psi' | \mathbf{t} | \Phi \rangle \langle \Psi' | \Phi \rangle^{-1} = \Phi \mathbf{t}^\dagger \end{aligned} \quad (2.45)$$

in analogy with (2.17). We note that, in this procedure, it is necessary to first evaluate the set  $\Phi = \{\phi_1, \phi_2, \dots, \phi_N\}$ , so that it becomes stable under the operator  $T_{\text{eff}}^\dagger$  according to (2.39). As we will see later, considerable simplifications are rendered on this point in the complex symmetric case.

In the canonical case, we are hence looking for solution to the Hartree-Fock equations of the form :

$$\begin{aligned} T_{\text{eff}}(1) \Psi'(1) &= \Psi'(1) \mathbf{t}, \quad T_{\text{eff}}^\dagger(1) \Phi'(1) = \Phi'(1) \mathbf{t}^\dagger, \\ \langle \Phi' | \Psi' \rangle &= 1, \end{aligned} \quad (2.46)$$



where  $\psi'$  and  $\phi'$  are row-vectors of order  $N$ , and  $t$  is in a classical canonical form with the eigenvalues  $t_1, t_2, \dots, t_N$ . In this case, the fundamental operator  $\rho$  defined by (2.32) takes the simple form

$$\rho = |\psi'\rangle\langle\phi'| \quad (2.47)$$

with the kernel

$$\rho(x_1, x_2) = \sum_k \psi'_k(x_1) \phi'^*_k(x_2), \quad (2.48)$$

where the summation over  $k$  goes from 1 to  $N$ , i.e. over the *occupied* one-particle functions appearing in the Slater determinants.

The purpose of the Hartree-Fock method is, of course, to find approximate eigenfunctions  $C_a$  and  $D_a$  and eigenvalues  $I$  and  $I^*$  to the many-particle operators  $T$  and  $T^\dagger$ . Using (2.36) and (2.40), one may now write the expression for the approximate eigenvalue  $I$  in the form:

$$\begin{aligned} I &= T_{(0)} + \int dx_1 T_{\text{eff}}(1) \rho(x_1, x'_1) - \\ &- (1/2) \int dx_1 dx_2 T'_{12} \rho(x_1, x'_1) \rho(x_2, x'_2) = \\ &= T_{(0)} + \sum_k t_k - (1/2) \int dx_1 dx_2 T'_{12} \rho(x_1, x'_1) \rho(x_2, x'_2). \end{aligned} \quad (2.49)$$

As in the conventional Hartree-Fock method, the approximate eigenvalue  $I$  is hence essentially different from the sum of the  $N$  eigenvalues of the one-particle operator.

### 3. Similarity Transformation of a Many-Particle Operator and its Consequences in the Hartree-Fock Scheme.

**3.1. General Aspects.** - Let us study a N-particle operator  $T$  defined through the expansion (2.35), which undergoes a similarity transformation of the form:

$$T^U = U T U^{-1} \quad (3.1)$$

In reference A, we have characterized the transformed operator  $T$  by a bar, but we will here - for typographical reasons - use the notation  $T^U$  with a superscript. Here the operator  $U$  is a bounded or unbounded linear operator of the product type:

$$U = u(1) u(2) u(3) \dots u(N), \quad (3.2)$$

where  $u(k)$  is a one-particle operator working on the linear space associated with particle  $k$ . For the transformed operator  $T^U$ , one obtains then the expansion

$$T^U = T_{(0)} + \sum_i T^u_i + \sum_{i < j} T^u_{ij} + \dots \quad (3.3)$$

where

$$\begin{aligned} T^u_i &= u(i) T_i u^{-1}(i) \\ T^u_{ij} &= u(i)u(j)T_{ij}u^{-1}(i)u^{-1}(j) \end{aligned} \quad (3.4)$$

If  $u$  is a bounded operator, it may be defined over the entire one-particle Hilbert space, which we will assume to be of type  $L^2$ , so that

$$\langle \phi | \psi \rangle = \int \phi^*(x) \psi(x) dx, \quad (3.5)$$

where the integration over  $x = (\mathbf{r}, \zeta)$  implies integration over the space coordinate  $\mathbf{r}$  and summation over the spin  $\zeta$ . The Hilbert space for the many-particle system is the direct product of the Hilbert spaces for the various particles. In such a case, the many-particle operator  $U$  defined by (3.2) is also a bounded operator, and the transformation (3.1) leaves the spectrum of  $T$  invariant.

In this paper, we are solely interested in unbounded operators  $u$ . In such a case, the operator  $u$  has a specific domain

$D(u)$  in the one-particle  $L^2$  Hilbert space with the complement  $C(u)$ , and - in analogy with (A.2.31) - one gets the following four cases:

$$\begin{aligned}
 f &\in L^2, & uf &\in L^2, & f &\in D(u) \\
 f &\in L^2, & uf &\notin L^2, & f &\in C(u) \\
 f &\notin L^2, & uf &\in L^2, & f &\in A(u^{-1}) \\
 f &\notin L^2, & uf &\notin L^2, & f &\in B(u)
 \end{aligned} \tag{3.6}$$

We note particularly that the set  $A(u^{-1}) = u^{-1}C(u^{-1})$  is situated outside  $L^2$ , but that the image set  $uA(u^{-1}) = C(u^{-1})$  is a subset of  $L^2$ . The corresponding classification holds also for the many-particle operator  $U$  defined by (3.2).

Let us now study what happens to the stability problem of the transformed operators

$$T^U = UTU^{-1}, \quad (T^U)^\dagger = (U^\dagger)^{-1}T^\dagger U^\dagger \tag{3.7}$$

when the operator  $U$  is unbounded. For the sake of simplicity, we will start from the eigenvalue problem (2.3) for a non-degenerate eigenvalue  $\lambda$ . Introducing the transformations

$$C^U = UC, \quad D^U = (U^\dagger)^{-1}D, \tag{3.8}$$

one gets immediately the formal relations

$$T^U C^U = \lambda C^U, \quad (T^U)^\dagger D^U = \lambda^* D^U, \tag{3.9}$$

which are analogous to the original eigenvalue problem but with transformed boundary conditions. This implies that, if the original eigenfunction  $C$  belongs to the domain of  $U$ , then the transformed function  $C^U$  is an element of  $L^2$ , and the eigenvalue  $\lambda$  is persistent. If on the other hand, the eigenfunction  $C$  belongs to the complement of the domain of  $U$ , then the transformed function  $C^U$  is situated outside  $L^2$ , and the corresponding eigenvalue  $\lambda$  is lost. Finally, if the transformed operator  $T^U$  has a discrete non-degenerate eigenvalue  $\lambda'$ , which is not an

eigenvalue to  $T$ , then the eigenfunction  $C^U$  must belong to the complement of the operator  $U^{-1}$ , which implies that the transformed function  $C = U^{-1}C^U$  must be situated outside  $L^2$ , in fact in the set  $A(U^{-1})$ . This means that the eigenvalue  $\lambda'$  becomes a lost eigenvalue under the reverse transformation  $T = U^{-1}T^U U$ . Similar considerations hold for the eigenvalue problem of the transformed operator  $T^\dagger$ . We note that one has the invariance property

$$\langle C^U | D^U \rangle = \langle C | D \rangle, \quad (3.10)$$

provided that all functions involved belong to  $L^2$ , i.e. that they belong to the persistent eigenvalues  $\lambda$  and  $\lambda^*$ , respectively.

Even if the conditions for persistent, lost, and new eigenvalues are completely clear for the exact eigenvalue problems to the operators  $T$  and  $T^\dagger$ , it is considerably more difficult to translate them to the approximate eigenvalue problems associated with the application of the bi-variational principle for the operators  $T$  and  $T^\dagger$  to truncated basis sets. In this connection, the relations (A.1.40-1.49) may turn out to be useful in formulating the problem. Some of the computational aspects, particularly the choice of the dual basis sets, are further discussed in reference A.

**3.2. Similarity transformations in the Hartree-Fock Scheme; the Tarfala Theorem.** - Let us again consider the eigenvalue problem (2.3) for the operators  $T$  and  $T^\dagger$  for a non-degenerate eigenvalue  $\lambda$  with the approximate eigenfunctions  $C_a$  and  $D_a$  given by the Slater determinants (2.28). For the transformed eigenfunctions  $C_a^U = UC_a$ , one obtains according to (3.8):

$$\begin{aligned} C_a^U &= U (N!)^{-1/2} |\psi_k(x_i)| = \\ &= U (N!)^{+1/2} O_{AS} \psi_1(x_1) \psi_2(x_2) \dots \psi_N(x_N) = \\ &= (N!)^{+1/2} O_{AS} \psi_1^u(x_1) \psi_2^u(x_2) \dots \psi_N^u(x_N) = \\ &= U(N!)^{-1/2} |\psi_k(x_i)| \end{aligned} \quad (3.11)$$

where we have used the fact that the product operator (3.2) commutes with the antisymmetric projection operator

$$O_{AS} = (N!)^{-1} \sum_{\underline{p}} (-1)^p P \quad (3.12)$$

and introduced the notation

$$\psi_k^u(x_i) = u(i) \psi_k(x_i), \quad (3.13)$$

for  $k = 1, 2, \dots, N$ . For the transformed eigenfunction  $D_a^U = (U^\dagger)^{-1} D_a$ , one obtains similarly:

$$D_a^u = (U^\dagger)^{-1} (N!)^{-1/2} |\phi_1(x_j)| = (N!)^{-1/2} |\phi_1^u(x_j)|, \quad (3.14)$$

where

$$\phi_1^u(x_j) = \{u^\dagger(j)\}^{-1} \phi_1(x_j). \quad (3.15)$$

For the overlap integral, one obtains further

$$d_{lk}^u = \langle \phi_l^u | \psi_k^u \rangle = \langle \phi_l | \psi_k \rangle = d_{lk}, \quad (3.16)$$

provided that both  $\psi_k$  and  $\phi_l$  belong to the domain of  $u$  and  $(u^\dagger)^{-1}$ , respectively. If all functions  $\Psi = \{\psi_1, \psi_2, \dots, \psi_N\}$  and  $\Phi = \{\phi_1, \phi_2, \dots, \phi_N\}$  belong to the domain of  $u$  and  $(u^\dagger)^{-1}$ , respectively, the transformed determinants  $C_a^u$  and  $D_a^u$  still belong to  $L^2$ , and one has

$$\langle C_a^u | D_a^u \rangle = \langle C_a | D_a \rangle \quad (3.17)$$

In such a case, the special projector  $\Gamma_a^u$  defined by (2.31) or

$$\Gamma_a^u = U \Gamma_a U^{-1}, \quad (3.18)$$

still exists, but we note that this is not the case if a single one of the functions  $\Psi = \{\psi_k\}$  or  $\Phi = \{\phi_l\}$  is outside the domains of  $u$  and  $(u^\dagger)^{-1}$ , respectively. We may now write the relations (3.13) and (3.15) in the condensed form:

$$\Psi^u = u \Psi, \quad \Phi^u = (u^\dagger)^{-1} \Phi. \quad (3.19)$$

The many-particle projector  $\Gamma_a^u$  is in this case fully described by the one-particle projector

$$\rho^u = |\psi^u\rangle \langle \phi^u| \psi^u\rangle^{-1} \langle \phi^u|, \quad (3.20)$$

according to (2.32), and - by using (3.19) - one then obtains the relation

$$\rho^u = u \rho u^{-1}, \quad (3.21)$$

This gives for the associated kernels:

$$\rho^u(x_1, x_2) = u(1) \rho(x_1, x_2) u^{-1}(2). \quad (3.22)$$

If the many-particle projector  $\Gamma_a$  undergoes the similarity transformation (3.18), the one-particle projector  $\rho$  undergoes the corresponding one-particle similarity transformation.

The question is now what happens to the effective one-particle operator  $T_{\text{eff}}^u$  associated with the transformed many-particle operator  $T^u = UTU^{-1}$ . Using (2.40) and (3.4), one obtains

$$T_{\text{eff}}^u(1) = T_1^u + \int dx_2 (T_{12}^u) \rho^u(x_2, x'_2), \quad (3.23)$$

where

$$T_1^u = u(1) T_1 u^{-1}(1),$$

$$(T_{12}^u)' = u(1)u(2) T_{12} u^{-1}(1)u^{-1}(2), \quad (3.24)$$

and  $T_{12}' = T_{12} (1 - P_{12})$ . The treatment of the first term in (3.23) is obvious, and we will concentrate our interest on the second term:

$$t_1^u = \int dx_2 (T_{12}^u)' \rho^u(x_2, x'_2). \quad (3.25)$$

For this purpose, we will start from the corresponding term in the untransformed operator (2.40), and we note that it may be written in the form:

$$t_1 = \int dx_2 T_{12}' \rho(x_2, x'_2) =$$

$$\begin{aligned}
&= \sum_{k,l} \int dx_2 \phi_l^*(x_2) T_{12}' \psi_k(x_2) e_{kl} = \\
&= \sum_{k,l} \langle \phi_l | T_{12}' | \psi_k \rangle_2 e_{kl},
\end{aligned} \tag{3.26}$$

where the summation over  $k$  and  $l$  goes from 1 to  $N$ , and where the index 2 on the bracket  $\langle \cdot | \cdot \rangle$  indicates that we have integrated over the variable  $x_2$ . Similarly, one obtains for the transformed term:

$$\begin{aligned}
t_1^u &= \sum_{k,l} \langle \phi_l^u | (T_{12}')^u | \psi_k^u \rangle_2 e_{kl}^u = \\
&= \sum_{k,l} \langle \phi_l^u | u(1) u(2) T_{12}' u^{-1}(1) u^{-1}(2) | \psi_k^u \rangle_2 e_{kl}^u = \\
&= u(1) \{ \sum_{k,l} \langle u^\dagger \phi_l^u | T_{12}' | u^{-1} \psi_k^u \rangle_2 e_{kl}^u \} u^{-1}(1) = \\
&= u(1) \{ \sum_{k,l} \langle \phi_l | T_{12}' | \psi_k \rangle_2 e_{kl}^u \} u^{-1}(1), \\
&= u(1) t_1 u^{-1}(1)
\end{aligned} \tag{3.27}$$

where we have used the inverse of the relations (3.19) and the fact that  $e_{kl}^u = e_{kl}$ . For the effective one-particle operator, one hence obtains the transformation formula:

$$T_{\text{eff}}^u(1) = u(1) T_{\text{eff}}(1) u^{-1}(1), \tag{3.28}$$

i.e. if the many-particle operator  $T$  undergoes the many-particle similarity transformation (3.1), the effective one-particle operator  $T_{\text{eff}}$  undergoes the corresponding one-particle similarity transformation. This is the content of the so-called *Tarfala Theorem*. It should be observed that the theorem is valid, if and only if all the functions  $\Psi = \{\psi_k\}$  belong to the domain of  $u$  and all the functions  $\Phi = \{\phi_l\}$  to the domain of  $(u^\dagger)^{-1}$ .

For the approximation  $I = \text{Tr } T \Gamma_a$  to the eigenvalue  $\lambda$ , one obtains

$$I^u = \langle D_a^u | T^U | C_a^u \rangle / \langle D_a^u | C_a^u \rangle = \langle D_a | T | C_a \rangle / \langle D_a | C_a \rangle = I \tag{3.29}$$

and the approximation is then invariant under the similarity

transformation and corresponds hence to a *persistent* eigenvalue. At first, this result may look rather disappointing, since we are particularly interested in the occurrence of lost and new eigenvalues, but it should be observed that it depends on the fact that we have required that the Hartree-Fock scheme should be valid for both  $T$  and  $T^U = UTU^{-1}$ .

If anyone of the functions  $\Psi = \{\psi_k\}$  is situated outside the domain of the operator  $u$ , the transformed determinant  $C_a^u = UC_a$  is no longer an element of  $L^2$ , and this may be taken as an indication of a *lost* eigenvalue. However, in such a case the operator  $\rho^u = upu^{-1}$  defined by (3.20) no longer exists, and the entire Hartree-Fock scheme breaks down.

Similarly, if the approximate eigenvalue  $I^U$  would correspond to a *new* eigenvalue of the operator  $T^U$ , one would expect that the associated Slater determinant  $C_a^u$  - as well as  $D_a^u$  - would belong to  $L^2$ , whereas this would not be the case for the transformed determinant  $C_a = U^{-1}C_a^u$ . In such a case, at least one of the functions  $\Psi = \{\psi_k\}$  must be situated in the complement to the domain of the operator  $u^{-1}$ . Introducing the reverse transformation  $T = U^{-1}T^U U$ , one realizes that the Hartree-Fock scheme for the original operator  $T$  breaks down for this particular case; neither the one-particle projector  $\rho$  nor the effective one-particle operator  $T_{\text{eff}}$  do exist.

Unfortunately, the Hartree-Fock scheme is an approximation, and one can never from the very beginning assume that an approximate scheme should reflect the true behaviour of the exact eigenfunctions and eigenvalues. One has hence to be very careful about making final conclusions. We note, however, that - if the Hartree-Fock scheme is valid for both the operators  $T^U$  and  $T = U^{-1}T^U U$ , then one has necessarily  $I = I^U$ , i.e. at least the approximate eigenvalue is persistent.

**3.3. The Hartree-Fock Scheme for a Complex Symmetric Operator.** - Let us now consider the special case when the many-particle operator  $T$  is *complex symmetric*, so that  $T^\dagger = T^*$ . If the eigenvalue  $\lambda$  is non-degenerate, one has according to (2.26) the simple relation  $D = C^* \langle C | C^* \rangle^{-1}$  between the eigen-



functions  $C$  and  $D$  to  $T$  and  $T^\dagger$ , respectively. If the same relation should hold also for the approximate eigenfunctions  $D_a$  and  $C_a$  defined by the Slater determinants (2.28) - which is not necessarily true - then  $D_a$  is proportional to  $C_a^*$ , and one obtains for the associated Fock-Dirac density operators

$$|\phi\rangle\langle\phi|\phi\rangle^{-1}\langle\phi| = \{|\psi\rangle\langle\psi|\psi\rangle^{-1}\langle\psi|\}^*. \quad (3.30)$$

Multiplying this relation to the right by  $|\psi^*\rangle$ , one obtains

$$\phi = \psi^* \alpha, \quad \alpha = \langle\phi|\psi^*\rangle^{-1}\langle\phi|\phi\rangle, \quad (3.31)$$

which is a necessary and sufficient condition that  $D_a$  and  $C_a^*$  are proportional. Substituting the relation (2.31) into the expression for the one-particle projector  $\rho$  defined by (2.32), one obtains

$$\rho = |\psi\rangle\langle\phi|\psi\rangle^{-1}\langle\phi| = |\psi\rangle\langle\psi^*|\psi\rangle^{-1}\langle\psi^*|, \quad (3.32)$$

which is independent of  $\alpha$ . Hence one has

$$\rho^\dagger = \rho^*, \quad (3.33)$$

i.e. also the projector  $\rho$  is complex symmetric.

The many-particle operator  $T$  defined by (2.35) is complex symmetric because the various terms in (2.35) are assumed to be complex symmetric, so that

$$T_i^\dagger = T_i^*, \quad T_{ij}^\dagger = T_{ij}^*, \dots \quad (3.34)$$

and one has further the symmetry property  $T_{ji} = T_{ij}$ . The untransformed effective one-particle operator  $T_{\text{eff}}$  has according to (2.40), (3.26), and (3.32) the form

$$T_{\text{eff}}(1) = T_1 + t_1, \quad (3.35)$$

where

$$\begin{aligned} t_1 &= \int dx_2 T_{12} \rho(x_2, x_2) = \\ &= \sum_{k,1} \int dx_2 \psi_1(x_2) T_{12} \psi_k(x_2) e_k, \end{aligned} \quad (3.36)$$

and  $\mathbf{e} = \Delta^{-1}$  is the inverse of the metric matrix  $\Delta = \langle \Psi^* | \Psi \rangle$ . Since  $\Delta^\dagger = \Delta^*$ , one has also  $\mathbf{e}^\dagger = \mathbf{e}^*$ , i. e.  $e_{kl} = e_{lk}$ .

The operator  $T_{12}$  is complex symmetric, which means that - for any pair of two-particle functions  $\Phi = \Phi(1,2)$  and  $\Psi = \Psi(1,2)$  in  $L^2$ , one has

$$\begin{aligned} \langle \Phi | T_{12} | \Psi \rangle &= \langle T_{12}^\dagger \Phi | \Psi \rangle = \langle \Psi | T_{12}^\dagger \Phi \rangle^* = \\ &= \langle \Phi | T_{12}^* | \Phi \rangle^* = \langle \Psi^* | T_{12} | \Phi^* \rangle \end{aligned} \quad (3.37)$$

Since  $T_{12} = T_{21}$ , the operator  $T'_{12} = T_{12} (1 - P_{12})$  is also complex symmetric. This follows from the fact that

$$\begin{aligned} \langle \Phi | T'_{12} | \Psi \rangle &= \iiint dx_1 dx_2 \Phi^*(1,2) T_{12} (1 - P_{12}) \Psi(1,2) = \\ &= \iiint dx_1 dx_2 \Phi^*(1,2) T_{12} \{ \Psi(1,2) - \Psi(2,1) \} = \\ &= \iiint dx_1 dx_2 \{ \Phi^*(1,2) - \Phi^*(2,1) \} T_{12} \Psi(1,2), \end{aligned} \quad (3.38)$$

where we have simply interchanged the names of the integration coordinates  $x_1$  and  $x_2$  in the second term. Using (3.37), one obtains further

$$\langle \Phi | T'_{12} | \Psi \rangle = \langle (1 - P_{12}) \Phi | T_{12} | \Psi \rangle = \langle \Psi^* | T_{12}' | \Phi^* \rangle \quad (3.39)$$

which proves the statement.

The first term  $T_1$  in the right-hand side of (3.35) is complex symmetric. In order to show that the second term  $t_1$  is also complex symmetric  $t_1^\dagger = t_1^*$ , one has to show that - for any pair of one-particle functions  $f(1)$  and  $g(1)$  in  $L^2$  - one has

$$\langle f | t | g \rangle = \langle t^\dagger f | g \rangle = \langle g | t^\dagger f \rangle^* = \langle g | t^* f \rangle^* = \langle g^* | t | f^* \rangle. \quad (3.40)$$

Using (3.36) and (3.39), one obtains directly

$$\begin{aligned} \langle f | t | g \rangle &= \int dx_1 f^*(1) t_1 g(1) = \\ &= \sum_{k,l} \iint dx_1 dx_2 f^*(1) \psi_l(2) T'_{12} g(1) \psi_k(2) e_{kl} = \end{aligned}$$

$$\begin{aligned}
&= \sum_{k,l} \langle f | \psi_l^* | T_{12}' | g | \psi_k \rangle e_{kl} = \sum_{k,l} \langle g^* | \psi_k^* | T_{12}' | f^* | \psi_l \rangle e_{kl} = \\
&= \sum_{k,l} \iint dx_1 dx_2 g(1) \psi_l(2) T_{12}' f^*(1) \psi_k(2) e_{kl} = \\
&= \int dx_1 g(1) \{ \sum_{k,l} \int dx_2 \psi_l(2) T_{12}' \psi_k(2) e_{kl} \} f^*(1) = \\
&= \langle g^*(1) | t_1 | f^*(1) \rangle = \langle g^* | t | f^* \rangle, \tag{3.41}
\end{aligned}$$

where in the proof we have interchanged the summation indices  $k$  and  $l$  and used the fact that  $e_{kl} = e_{lk}$ . Hence  $t_1$  is complex symmetric, and one has

$$T_{\text{eff}}^\dagger(1) = T_{\text{eff}}^*(1) \tag{3.42}$$

The fact that also the effective one-particle Hamiltonian  $T_{\text{eff}}(1)$  is complex symmetric has deep-going consequences and renders considerable simplifications in the theory.

Let us first evaluate the canonical Hartree-Fock functions  $\psi = \{ \psi_1, \psi_2, \dots, \psi_N \}$  which are solutions to the equations

$$T_{\text{eff}}(1) \psi = \psi t, \tag{3.43}$$

where the matrix  $t$  is on classical canonical for, with the eigenvalues  $t_k$  on the diagonal. In such a case, one can introduce the *conjugate* set  $\phi = \{ \phi_1, \phi_2, \dots, \phi_N \}$  through application of relation (2.26) to the one-particle case, which gives

$$\phi = \psi^* \langle \psi | \psi^* \rangle^{-1}. \tag{3.44}$$

One has further

$$T_{\text{eff}}^\dagger(1) \phi = \phi t^\dagger, \quad \langle \phi | \psi \rangle = 1. \tag{3.45}$$

In such a case, the one-particle projector  $\rho$  defined by (3.32) takes the simplified form

$$\rho = | \psi \rangle \langle \phi |, \tag{3.46}$$

with the kernel

$$\rho(x_1, x_2) = \sum_k \psi_k(x_1) \phi_k^*(x_2) \quad (3.47)$$

which is often very useful in the numerical applications.

In solving the non-linear Hartree-Fock equations (3.43), one is using the iterative SCF-procedure discussed in reference C. If one starts from a one-particle projector  $\rho$  which is complex symmetric, this property is going to stay invariant under the iteration procedure, and one finally reaches a solution with the desired property:  $D_a = C_a^* \langle C_a | C_a^* \rangle^{-1}$ .

**3.4. Special Case when  $T^\dagger = T^* = T$ .** Let us now consider the special case when a complex symmetric operator is *real*, so that  $T^* = T$ . In this case, the operator  $T$  is also self-adjoint,  $T^\dagger = T$ , and one can use the results of the conventional Hartree-Fock method <sup>7</sup>. The eigenvalues are real,  $\lambda = \lambda^*$ , and - if an eigenvalue  $\lambda$  is non-degenerate, the associated eigenfunction  $C$  is necessarily *real* or a real function multiplied by a constant phase factor  $\exp(i\alpha)$ . In both cases, one has  $D = C^* \langle C | C^* \rangle^{-1} = C$ . In the conventional Hartree-Fock theory, the one-particle projector  $\rho$  takes the form

$$\rho = |\Psi\rangle \langle \Psi| \Psi\rangle^{-1} \langle \Psi|, \quad (3.48)$$

and it has hence the special property

$$\rho^\dagger = \rho, \quad (3.49)$$

It is then easily shown that also the effective one-particle operator  $T_{\text{eff}}$  is self-adjoint:

$$T_{\text{eff}}^\dagger(1) = T_{\text{eff}}(1). \quad (3.50)$$

If one starts the iterative SCF-procedure from a self-adjoint operator  $\rho$ , the properties (3.50) and (3.49) are going to be invariant under the iterations and are going to characterize the final solution. Since  $T_{\text{eff}}$  is self-adjoint, the classical canonical matrix  $\mathfrak{t}$  is always on *diagonal* form. We note that, in this case, there is no need to assume that the set  $\Psi = \{\psi_1, \psi_2, \dots, \psi_N\}$  should consist of real functions, and that it is sometimes

convenient to consider complex sets.

It is a peculiarity of the Hartree-Fock scheme that the properties of the approximate eigenfunction  $C_a$  does not always reflect the properties of the exact eigenfunction  $C$ . A well-known example is given by the *symmetry dilemma*<sup>8</sup>, which says that if an eigenfunction  $C$  has a special symmetry property characterized by the projector  $O$ , so that

$$OC = C, \quad TO = OT, \quad (3.51)$$

then the approximate eigenfunction  $C_a$  obtained from the variation principle  $\delta I = 0$  does not automatically have this property. This means that the condition  $OC_a = C_a$  becomes a *constraint* on the variation principle, and one has the choice between looking for an "absolute" minimum  $I$ , in which the Slater determinant  $C_a$  may be a mixture of different symmetry types, and a "local" minimum  $I_{loc} > I$ , in which the constraint is satisfied. It should further be observed that, if one starts from a Slater determinant which satisfies the condition  $OC_a = C_a$ , then the constraint is satisfied during the entire iterative SCF-procedure, and that this approach leads to the so-called restricted Hartree-Fock (RHF) scheme. In this particular case, the matrix  $t$  in the canonical Hartree-Fock equations may contain non-diagonal Lagrangian multipliers  $t_{kl} \neq 0$ , but we note that they are of a different character than those occurring in the Jordan blocks.

Since the operator  $T$  is complex symmetric,  $T^\dagger = T^*$ , one may wonder what happens if one starts the iterative SCF-procedure from a complex set  $\Psi = \{\psi_1, \psi_2, \dots, \psi_N\}$  and a one-particle projector  $\rho$  of the form (3.32), i.e.

$$\rho = |\Psi\rangle\langle\Psi^*| \Psi\rangle^{-1} \langle\Psi^*| \quad (3.52)$$

instead of (3.48). In such a case, the two relations  $\rho^\dagger = \rho^*$  and  $T_{eff}^\dagger = T_{eff}^*$  are going to be invariant under the iterative SCF-procedure and characteristic also for the final solutions. Since the relations (3.44) and (3.45) are now valid, one has also the property  $D_a = C_a^* \langle C_a | C_a^* \rangle^{-1}$ , but we note that the Slater determinant  $C_a$  is not necessarily real, or a real function multiplied by a constant phase factor  $e^{i\alpha}$ . This means also that the approximate

eigenvalue  $\lambda$  defined by the bi-variational principle is not necessarily real.

It is clear that this complex symmetric Hartree-Fock scheme will reduce to the conventional Hartree-Fock scheme, if and only if the relation

$$\Psi^* = \Psi \alpha, \quad (3.53)$$

is satisfied, i.e. if the linear space spanned by the functions  $\Psi = \{\psi_1, \psi_2, \dots, \psi_N\}$  is stable under the operation of *complex conjugation* (\*). In this case, the one-particle operator (3.52) reduces to the conventional form (3.48), and - for the Slater determinant  $C_a = (N!)^{-1/2} |\psi_k(x_i)|$  - one obtains

$$C_a^* = C_a |\alpha|, \quad (3.54)$$

where the determinant  $|\alpha|$  is a constant factor, and the approximate eigenvalue  $I$  becomes real. If one starts the iterative SCF-procedure from a set  $\Psi$ , which satisfies the relation (3.53), the relations  $\rho^\dagger = \rho^*$  and  $T_{\text{eff}}^\dagger = T_{\text{eff}}^*$  are going to stay invariant under the iterations, and one obtains the conventional Hartree-Fock scheme.

On the other hand, if one starts from a set  $\Psi$  which does not satisfy (3.53), the relations  $\rho^\dagger = \rho^*$  and  $T_{\text{eff}}^\dagger = T_{\text{eff}}^*$  are going to stay invariant under the iterative SCF-procedure, and we note that there is nothing in this process which says that the final solution  $\Psi$  must necessarily satisfy relation (3.53). Hence one can expect that the complex symmetric Hartree-Fock scheme and the conventional scheme will sometimes give different results.

**3.5. Invariance of the Property of "Complex Symmetry" under Restricted Similarity Transformations.-** Starting from the general similarity transformation (3.1), one obtains the two relations

$$T^U = UTU^{-1}, \quad (T^U)^\dagger = (U^\dagger)^{-1} T^\dagger U^\dagger. \quad (3.55)$$

One sees immediately that, if the operator  $T$  is originally complex symmetric, so that  $T^\dagger = T^*$ , it will keep this property invariant under the similarity transformation provided that the operator  $U$  satisfies the condition

$$(U^\dagger)^{-1} = U^*, \quad (3.56)$$

in which case one speaks of a *restricted* similarity transformation; see also reference A, eqs. (A.1.54-1.71). Let us further consider the stability problem for the operators  $T$  and  $T^\dagger$  in the form given by the relations (2.19) and (2.20). For the conjugate eigenfunctions  $C^U$  and  $D^U$  to the transformed operators  $T^U$  and  $(T^U)^\dagger$ , one gets in analogy with (3.8) that

$$C^U = U C, \quad D^U = (U^\dagger)^{-1} D, \quad (3.57)$$

$$\langle C^U | D^U \rangle = \langle C | D \rangle = 1. \quad (3.58)$$

In the special case when  $T$  is complex symmetric, one has further according to (2.26) that  $D = C^* \langle C | C^* \rangle^{-1}$ . Since  $\langle C^U | (C^U)^* \rangle = \langle UC | U^* C^* \rangle = \langle C | U^\dagger U^* | C^* \rangle = \langle C | C^* \rangle$ , one obtains

$$D^U = (U^\dagger)^{-1} D = U^* C^* \langle C | C^* \rangle^{-1} = (C^U)^* \langle C^U | (C^U)^* \rangle^{-1}, \quad (3.59)$$

and this important relation is hence invariant under the restricted similarity transformations.

The many-particle operator  $U$  defined by the product (3.2) defines a restricted similarity transformation provided that the one-particle operators  $u$  satisfy the condition:

$$(u^\dagger)^{-1} = u^*. \quad (3.60)$$

Let us now consider the one-particle transformations (3.19) which, in this case, take the form:

$$\psi^u = u\psi, \quad \phi^u = (u^\dagger)^{-1}\phi = u^*\phi. \quad (3.61)$$

If all the functions in the sets  $\psi$  and  $\phi$  belong to the domain of  $u$  and  $u^*$ , respectively, the one-particle operators  $\rho$  and  $\rho^u$  exist,

and one obtains - according to (3.21) and (3.28) - the relations

$$\rho^u = u \rho u^{-1}, \quad T^u = u T_{\text{eff}} u^{-1}. \quad (3.62)$$

Using (3.60), one gets further that

$$(\rho^u)^\dagger = (\rho^u)^*, \quad (T^u_{\text{eff}})^\dagger = (T^u_{\text{eff}})^*, \quad (3.63)$$

i.e. the operators  $\rho$  and  $T_{\text{eff}}$  remain complex symmetric under the restricted similarity transformations. In this particular case, one has

$$T^u_{\text{eff}}(1) \psi^u(1) = \psi^u(1) \mathfrak{t}, \quad \{T^u_{\text{eff}}(1)\}^\dagger \phi^u(1) = \phi^u(1) \mathfrak{t}^\dagger, \quad (3.64)$$

$$\langle \phi^u | \psi^u \rangle = 1 \quad (3.65)$$

and the classical canonical form  $\mathfrak{t}$  stays *invariant* under the similarity transformation as one would expect. According to (3.44), one has further  $\phi = \psi^* \langle \psi | \psi^* \rangle^{-1}$ , and since  $\langle \psi^u | \psi^{u*} \rangle = \langle \psi | \psi^* \rangle$ , one gets also

$$\phi^u = (\psi^u)^* \langle \psi^u | (\psi^u)^* \rangle^{-1}. \quad (3.66)$$

Let us finally assume that one starts from a set  $\psi^u$  which belongs to  $L^2$ , but that this is not necessarily the case for the set  $\psi = u^{-1} \psi^u$ . In such a case, the projector

$$\rho^u = |\psi^u\rangle \langle (\psi^u)^* | \psi^u \rangle^{-1} \langle (\psi^u)^* | \quad (3.67)$$

as well as the operator  $T^u_{\text{eff}}$  exist, and the relations (3.63) are going to be satisfied. This implies that, if one has solved the canonical Hartree-Fock equations (3.43) for this particular case:

$$T^u_{\text{eff}}(1) \psi^u(1) = \psi^u(1) \mathfrak{t}^u, \quad (3.68)$$

where  $\mathfrak{t}^u$  is a classical canonical form, then one can immediately introduce the conjugate solutions  $\phi^u$  through the relation



$$\phi^u = (\psi^u)^* \langle \psi^u | (\psi^u)^* \rangle^{-1}. \quad (3.69)$$

According to (3.45), they satisfy the conditions

$$\{T_{\text{eff}}^u(1)\}^\dagger \phi^u = \phi^u t^\dagger, \quad \langle \phi^u | \psi^u \rangle = 1. \quad (3.70)$$

However, since the projector  $p$  does not necessarily exist, the invariance theorem  $t = t^u$  could very well break down, corresponding to the case of "lost eigenvalues" in the general theory.

Let us now consider the special case when  $T^\dagger = T^* = T$ , i.e. when the original operator  $T$  is self-adjoint and real. In the *conventional* Hartree-Fock scheme - in the following indicated by an index  $c$  (=conventional) - which is based on relation (3.48), one has  $\rho_c = \rho_c^\dagger$  and  $T_{\text{eff},c} = T_{\text{eff},c}^\dagger$ , and it is obvious that the classical canonical form  $t_c$  must be diagonal with only real eigenvalues.

For restricted similarity transformations, the transformed operator  $T^u = UTU^{-1}$  is still complex symmetric, and the same applies - according to (3.33) and (3.42) - to the one-particle operators  $\rho^u$  and  $T_{\text{eff}}^u$ . If one starts from the assumption that

$$\psi^u = u \psi_c, \quad (3.71)$$

one gets according to (3.62) that

$$\rho^u = u \rho_c u^{-1}, \quad T_{\text{eff}}^u = u T_{\text{eff},c}^u u^{-1}, \quad (3.72)$$

and the invariance theorem  $t = t_c$  shows then that, in this case, the classical canonical form  $t$  is diagonal with real eigenvalues. Since we are here interested in *complex* one-particle eigenvalues  $t_k^u$ , this trivial application of the Tarfala theorem leads hence to a negative result.

It is evident that, if one wants to avoid this trivial case, one has to abandon relation (3.71) and assume that the set  $\psi^u$  has more general properties. For this purpose, we will start from the complex symmetric many-particle operator  $T^u$  and consider the associated one-particle operators

$$\rho^u = |\psi^u\rangle\langle(\psi^u)^*| \psi^u\rangle^{-1} \langle(\psi^u)^*|, \quad (3.73)$$

and  $T_{\text{eff}}^u(1)$  defined by (3.23). In this more general case, one is dealing with the *complex symmetric Hartree-Fock scheme*, and it may very well happen that the classical canonical form  $\mathfrak{t}^u$  has *complex* eigenvalues. It is not so easy to demonstrate this fact by a theoretical or numerical example for the simple reason that, even if it is easier to solve the Hartree-Fock equations than the original eigenvalue problem (2.3), no one has so far been able to solve the Hartree-Fock equations *exactly*. In the approximate procedures built on the use of truncated real orthonormal basis sets of order  $m$ , the matrix  $T_{\text{eff}}^u$  is a symmetric matrix of order  $m \times m$  with complex elements, which - as a rule - has complex eigenvalues.

Once one has solved the Hartree-Fock equations associated with the transformed operator  $T^u = UTU^{-1}$ , it is evident that it would be of interest to study also the inverse transformation  $T = U^{-1}T^uU$ , where  $T$  has the properties  $T = T^\dagger = T^*$ . In order to stay within the complex symmetric Hartree-Fock scheme, we will then consider the functions:

$$\psi = u^{-1}\psi^u. \quad (3.74)$$

If at least one of the functions  $\psi^u$  is not in the domain of the operator  $u^{-1}$ , at least one of the functions  $\psi$  would be situated outside  $L^2$ , the complex symmetric Hartree-Fock scheme would break down, and one would have a situation which would be analogous to the case of a *lost* eigenvalue in the theory of the exact eigenfunctions  $C$  and  $D$ .

We note, however, that the Hartree-Fock scheme is an approximation in which the solutions do not necessarily reflect the behaviour of the exact eigenfunctions. In fact, it may very well happen that the canonical Hartree-Fock functions  $\psi^u$  are all situated in the domain of the operator  $u^{-1}$ , and this implies that also the transformed functions  $\psi = u^{-1}\psi^u$  are all situated in the  $L^2$  Hilbert space. In such a case, the one-particle operators  $\rho = |\psi\rangle\langle\psi^*| \psi\rangle^{-1} \langle\psi^*|$  and  $T_{\text{eff}}$  exist, and one has the

transformations:

$$\rho = u^{-1} \rho^u u, \quad T_{\text{eff}}(1) = u^{-1} T_{\text{eff}}^u(1) u, \quad (3.75)$$

which indicate that the operators  $\rho$  and  $T_{\text{eff}}(1)$  are complex symmetric but usually not self-adjoint. It should be observed that the complex symmetric Hartree-Fock scheme is different from the conventional one, and that the former reduces to the latter if and only if the relation (3.53) is satisfied, i.e. if the linear space spanned by the set  $\Psi$  is stable under complex conjugation. In order to obtain complex eigenvalues to the operator  $T_{\text{eff}}^u(1)$ , it is hence necessary to avoid this particular situation, in which  $\Psi^* = \Psi \alpha$ ,  $(u^{-1} \Psi^u)^* = (u^{-1} \Psi^u) \alpha$ , and

$$u u^\dagger (\Psi^u)^* = \Psi^u \alpha, \quad (3.76)$$

i.e. all the functions  $u u^\dagger \Psi_k^u$  for  $k = 1, 2, \dots, N$  should be situated in the linear space spanned by the functions  $\Psi^u = \{\Psi_1^u, \Psi_2^u, \dots, \Psi_N^u\}$ ; for algebraic details, see Appendix A. It is evident that the relation (3.76) is a very special condition for the general set  $\Psi^u = \{\Psi_k^u\}$  which is very seldom satisfied. In practice, one knows for sure that it is not fulfilled as soon as the operator  $T_{\text{eff}}^u(1)$  has at least one complex eigenvalue.

We note finally that, in the numerical studies of the complex symmetric operators  $T^U$  and  $T_{\text{eff}}^u$ , it is usually convenient to use orthonormal basis sets which are *real*, since the associated matrices will then automatically be symmetric with complex elements. In the case of a truncated basis of order  $m$ , most of the eigenvalues will usually turn out to be complex, but we observe that one has to study their behaviour when  $m$  goes to infinity and the set becomes complete, before one can make any definite conclusions as to the existence of true complex eigenvalues to  $T^U$  and  $T_{\text{eff}}^u$ , respectively. The connection between the results of approximate numerical treatments and the exact theory is still a very interesting but mostly unsolved problem.

#### 4. Properties of a Many-Particle Hamiltonian under Complex Scaling.

**4.1. Method of Complex Scaling; the Dilatation Operator.** - The method of complex scaling is based on the use of the dilatation operator  $u = u(\eta)$  - where  $\eta$  is a fixed complex parameter called the "scale factor" which is sometimes expressed in the alternative forms  $\eta = \exp(\theta) = \eta \exp(i\alpha)$  - defined through the property

$$u(\eta) f(x) = \eta^{1/2} f(\eta x). \quad (4.1)$$

Here  $f = f(z)$  is an *analytic* function originally defined on the real axis  $\{x\}$ , which is assumed to be analytic also in the point  $z = \eta x$ . An essential problem in the method of complex scaling in quantum mechanics is hence to study whether a wave function  $\psi = \psi(x)$  defined on the real axis may be continued analytically out in the complex plane to the point  $z = \eta x$ . Since many analytic functions have *natural boundaries*, it is from the very beginning evident that there may be considerable restrictions on the parameter  $\eta$  itself. More generally, one may define the operator  $u$  through the relation

$$u(\eta) f(z) = \eta^{1/2} f(\eta z), \quad (4.2)$$

where it is assumed that the function  $f$  is analytic in both  $z$  and  $\eta z$ . The method of complex scaling has been rigorously developed during the last few decades by a number of mathematicians, and - for a selection of the key references - the reader is referred to our reference A. Here we will consider the method of complex scaling essentially from the physicist's point of view.

There are many ways to show that the dilatation operator  $u$  satisfies the condition (3.60), which is characteristic for the restricted similarity transformations. By putting  $\eta = \exp(\theta)$ , it was proven in eq. (A.3.14) that  $u$  may be written in the form

$$u = \exp(\theta \omega), \quad (4.3)$$

where

$$\omega = (1/2) + x d/dx = (2\pi i/h)(p_x + x p)/2, \quad (4.4)$$

and  $p = (h/2\pi i) d/dx$  is a self-adjoint operator with the special property  $p^* = -p$ . We will here give an alternative derivation for finite values of  $\theta$ .

Let us assume that  $f = f(x)$  is analytic around the point  $x=0$  within a radius  $\rho$ , so that - for  $|x| < \rho$  - one has the Taylor expansion

$$f(x) = \sum_n (x^n/n!) f^{(n)}(0). \quad (4.5)$$

Since  $(x d/dx)x^n = nx^n$ , one gets  $(x d/dx)^p x^n = n^p x^n$ , as well as

$$\begin{aligned} \exp(\theta x d/dx) x^n &= \left\{ \sum_p (\theta^p/p!) (x d/dx)^p \right\} x^n = \\ &= \left\{ \sum_p (\theta^p n^p/p!) \right\} x^n = \exp(\theta n) x^n = (e^{\theta x})^n. \end{aligned} \quad (4.6)$$

Hence one has

$$\exp(\theta d/dx) f(x) = \sum_n (e^{\theta x})^n/n! f^{(n)}(0) = f(e^{\theta x}), \quad (4.7)$$

provided that  $|e^{\theta x}| < \rho$ , so that the series is convergent and the function analytic in the point  $z = e^{\theta x} = \eta x$ . Hence

$$u = e^{\theta \omega} = \exp \{ (\pi i/h) \theta (p_x + x p) \}. \quad (4.8)$$

Since further  $\omega^\dagger = -\omega = \omega^*$ , one gets directly

$$u^\dagger = \exp(\theta \omega)^\dagger = \exp \{ -\theta^* \omega \} = \{ u(\theta^*) \}^{-1} = \{ u^* \}^{-1}, \quad (4.9)$$

which proves our statement. For the specific algebraic rules valid for the dilatation operator  $u$ , etc., the reader is referred to reference A. It should be observed that  $u$  is an *unbounded* operator, as shown by some simple examples.

**4.2. The Coulombic Many-Particle Hamiltonian.**- For a system of atomic nuclei and electrons under the influence of electrostatic forces, the many-particle Hamiltonian takes the form:

$$H = \sum_k (p_k^2/2m_k) + \sum_{k < l} (e_k e_l/r_{kl}) = T + V \quad (4.10)$$

as described in (A.3.7). The method of *complex scaling* may be defined in many different ways, but - in reference A and here - we have defined it through an unbounded similarity transformation of the type (1.1) which gives

$$H^U = U H U^{-1} = \eta^{-2} T + \eta^{-1} V \quad (4.11)$$

In an approximate numerical treatment based on the bi-variational principle and the use of a finite (real or complex) orthonormal basis  $\Phi$ , one has

$$H^U = \eta^{-2} T + \eta^{-1} V, \quad (4.12)$$

If the basis  $\Phi$  is chosen real, one has the additional advantage that the matrices  $T$  and  $V$  are real and symmetric, and - with  $\eta = \rho(\cos \alpha + i \sin \alpha)$  - one can easily separate the real and imaginary parts of the eigenvalue problem. It is evident that, in such a case, the matrix  $H^U$  is automatically *symmetric with complex elements*, and this important matrix property was discovered in the numerical applications long before one started studying the corresponding operator property.

Keeping the basis fixed, one can now study the behaviour of the approximate eigenvalues as a function of the scale factor  $\eta$  and particularly of the "rotation angle"  $\alpha$ . Using the results from the exact theory, it seems reasonable to believe that the eigenvalues which are approximately situated on a ray in the complex plane with the angle  $(-2\alpha)$  belong to the "continuum", whereas those which are situated between this ray and the real axis may correspond to discrete complex eigenvalues associated with physical "resonances". In the exact theory, these complex eigenvalues  $E^U$  are independent of  $\eta$  as soon as they are "uncovered":

$$dE/d\eta = d^2E/d\eta^2 = \dots = d^n E/d\eta^n = \dots = 0, \quad (4.13)$$

whereas in the approximate numerical treatment they show certain stability properties; in practice they are often determined from "stabilization graphs" based on the fact that at

least  $dE/d\alpha = 0$ .

The Coulombic Hamiltonian (4.1) is invariant under translations and rotations, and it is hence convenient to separate the motion of the "center of mass"  $\xi$  and to study the new Hamiltonian:

$$H' = H - p_\xi^2/M_\xi \quad (4.14)$$

in which one may further separate the rotation of the system considered as a whole. One of the fundamental problems in quantum chemistry is that the Hamiltonian  $H'$  does not distinguish between different isomers or configurations, and that one still does not know how to treat the influence of the "nuclear motion" in full detail. In this situation, one often tries to use additional chemical insight obtained from experiments. In the *Born-Oppenheimer approximation*, one assumes that the nuclei  $g$  are very heavy ( $m_g \approx \infty$ ), neglects the kinetic energy of the nuclei, and assumes that they are situated in fixed positions  $\mathbf{r}_g$ . We note that - even if today the calculation of internuclear distances and angles is a highly developed art - the general shapes of the molecules involved are often taken from chemical experience. In this scheme, the Hamiltonian takes the simplified form:

$$H = H_{(0)} + \sum_i H_i + \sum_{i < j} H_{ij}, \quad (4.15)$$

where

$$\begin{aligned} H_{(0)} &= \sum_{g < h} e^2 Z_g Z_h / r_{gh}, & H_i &= p_i^2 / 2m - e^2 \sum_g Z_g / r_{ig}, \\ H_{ij} &= e^2 / r_{ij}, \end{aligned} \quad (4.16)$$

and the indices  $g$  and  $h$  characterize the atomic nuclei and the indices  $i$  and  $j$  the electrons. Scaling also the nuclear coordinates properly (eg. by external scaling), one may now apply the method of complex scaling to the Born-Oppenheimer Hamiltonian (4.15). The results of the general theory of unbounded similarity transformations as outlined in reference A apply also to this case, and - for numerical applications - the reader is referred to the bibliography in this paper.

**4.3. Application of the Hartree-Fock Method.** - Since numerical Hartree-Fock programs dealing with complex numbers are available in many research groups, it seemed natural to apply this scheme also to the scaled Born-Oppenheimer Hamiltonian (4.15). As a consequence, some numerical results were obtained before the theory was developed, and - as we have emphasized in the Introduction - some features seemed rather astonishing.

The Hamiltonian  $H$  defined by (4.15) is self-adjoint and real,  $H = H^\dagger = H^*$ , and this implies that the transformed Hamiltonian  $H^U$  remains *complex symmetric* under complex scaling. Hence the same holds for the one-electron projector  $\rho_u$  defined by (3.52)

$$\rho_u = |\Psi\rangle\langle\Psi^*| \Psi\rangle^{-1}\langle\Psi^*|, \quad (4.17)$$

and for the associated effective Hamiltonian  $H_{\text{eff}}^u(1)$ . In such a case, the *Tarfala theorem* expressed in the relations (3.21) and (3.28) is valid:

$$\rho_u = u \rho u^{-1}, \quad T_{\text{eff}}^u(1) = u T_{\text{eff}}(1) u^{-1}. \quad (4.18)$$

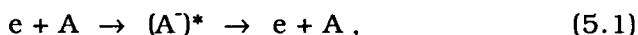
In Sec. 3.3, we have emphasized that the complex symmetric Hartree-Fock scheme usually does not reduce to the conventional one for the special case when  $\eta = 1$ . More generally, in the complex Hartree-Fock scheme, the results obtained from a starting value  $\eta_1$  do not necessarily reduce to the results obtained from another starting value  $\eta_2$  - i.e. each starting value leads to a specific result.

In the numerical applications, it is hence convenient to fix the scale factor  $\eta$ , to use a real orthonormal basis  $\Phi$  of order  $m$ , and to calculate all the  $m$  eigenvalues  $\epsilon_k$  of the effective Hamiltonian  $H_{\text{eff}}(1)$  on the real axis and in the complex plane. One then studies the behaviour of these eigenvalues as the order  $m$  of the basis increases and looks for "stable" complex eigenvalues between the ray  $(-2\alpha)$  and the real axis - often by means of "stabilization graphs". One then repeats the calculation for a series of values of the scale factor  $\eta$ , observing that the results are often discontinuous functions of the rotation angle  $\alpha$ . An example of such a calculation will be given in the next section.



## 5. Numerical Application of the Complex Scaling Method in the Hartree-Fock Approximation.

The calculation of physical resonances associated with the existence of meta-stable states with a finite life-time plays an important role in modern physics. They occur e.g. in scattering problems during the interaction between atoms/molecules with incoming particles or with electromagnetic fields. A simple example may be given by considering a so-called *shape resonance* arising in the course of the scattering of an electron by an atom. This resonance is associated with the temporary capture of the electron by the target atom (A) resulting in the formation of a metastable negative ion  $(A^-)^*$ , which then decays by an electron emission<sup>9</sup>:



This process is most efficient if the incident electron has a *resonance energy*  $K_r$  given by the difference of the average energy  $E_r$  of the composite structure  $(A^-)^*$  and the energy  $E_g$  of the ground state of (A). Conventionally, the energy of  $(A^-)^*$  is expressed as a complex number  $E = E_r - i\Gamma/2$ , where  $E_r$  indicates the middle of the energy level and  $\Gamma$  gives its half-width. In this way, the study of slow electron-atom collisions can be essentially reduced to the study of the resonances of the negative ions.

Resonances may be calculated in many different ways, and - during the last decade - the *complex scaling method* has turned out to be particularly powerful for this purpose, since it has been shown that the complex discrete eigenvalues revealed in this approach are identical with the physical resonances; for a survey of this method, the reader is referred to paper A and references given there. It is evident that, in calculating the eigenvalues of the transformed Hamiltonian  $H^U = UH U^{-1}$ , methods based on the use of *superposition of configurations* (CI) are going to be particularly useful. However, it may also be of some interest to find out how much of the resonances may be found in the Hartree-Fock scheme and how much requires more deep-going correlation calculations. The start of the present work was based on the idea that at least some of the "shape resonances" could be detected already in the Hartree-Fock approximation applied to the scaled Hamiltonian.

It is evident that, in order to solve this problem, the SCF resonances of the composite system  $(A^-)^*$  ought to be studied in full detail. However, in the numerical application described below, we have simplified the problem even further by assuming that one may obtain the resonances of the ionic system  $(A^-)$  at least approximately by studying the "meta-stable" virtual (unoccupied) orbitals in the dilated SCF calculation performed on the target atom (A) alone. The details of such an approach for the study of a resonance formation and decay are given elsewhere<sup>10</sup>.

By using this approach, we have studied the  $^2P$  shape resonance in the e-Be scattering problem. The spectrum of the bi-variationally obtained effective Hamiltonian for the beryllium atom ( $1s^2 2s^2$ ) has been investigated, and this spectrum and its behaviour are presented in Figures 1-4. Even at this crude level of approximation, it seems possible to detect a "trace" of the resonance, i.e. to distinguish between the *resonant* virtual orbital eigenvalue from other virtual eigenvalues describing the dilated scattering solutions<sup>10</sup>. In spite of the crudeness of the model, the energy and width of the resonance show reasonable agreement with those obtained with more accurate methods; see Table 1.

Figure 1.

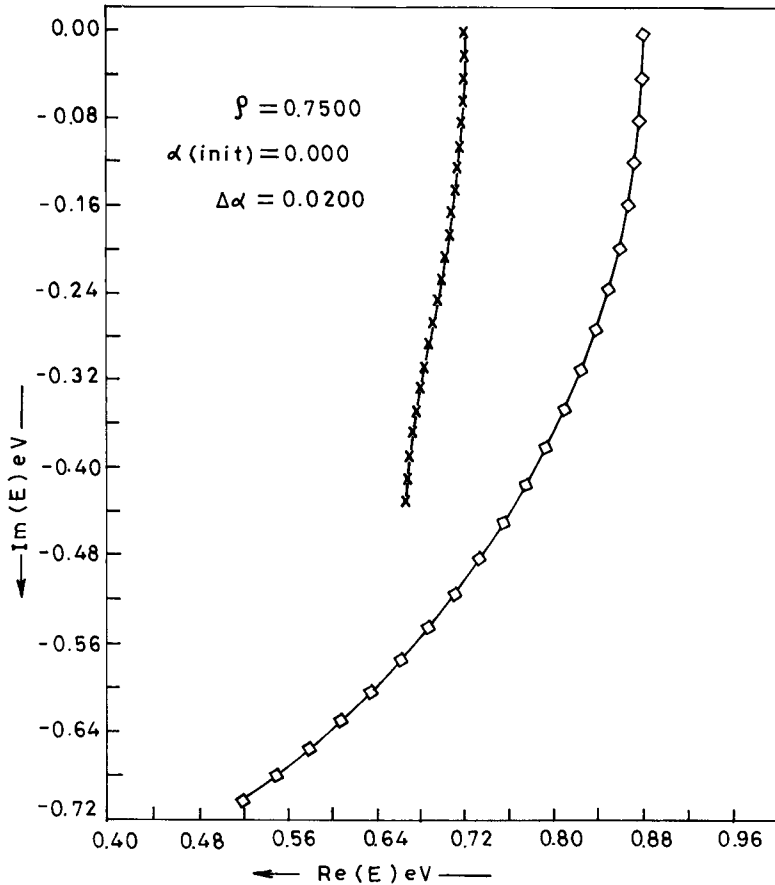


Fig.1. Trajectory for the real and imaginary parts of the resonant orbital energy  $E$  denoted by dots (•) and the nearest scattering root (o) as functions of the rotation angle  $\alpha$  in radians. The starting value is  $\alpha = 0.0$  at the top right ( $\text{Im}[E] = 0.0$ ), and the increment  $\Delta\alpha$  is 0.02. The real factor  $\rho$  in the scale parameter  $\eta = \rho \exp(i\alpha)$  is here fixed at  $\rho = 0.75$ .

Figure 2a.

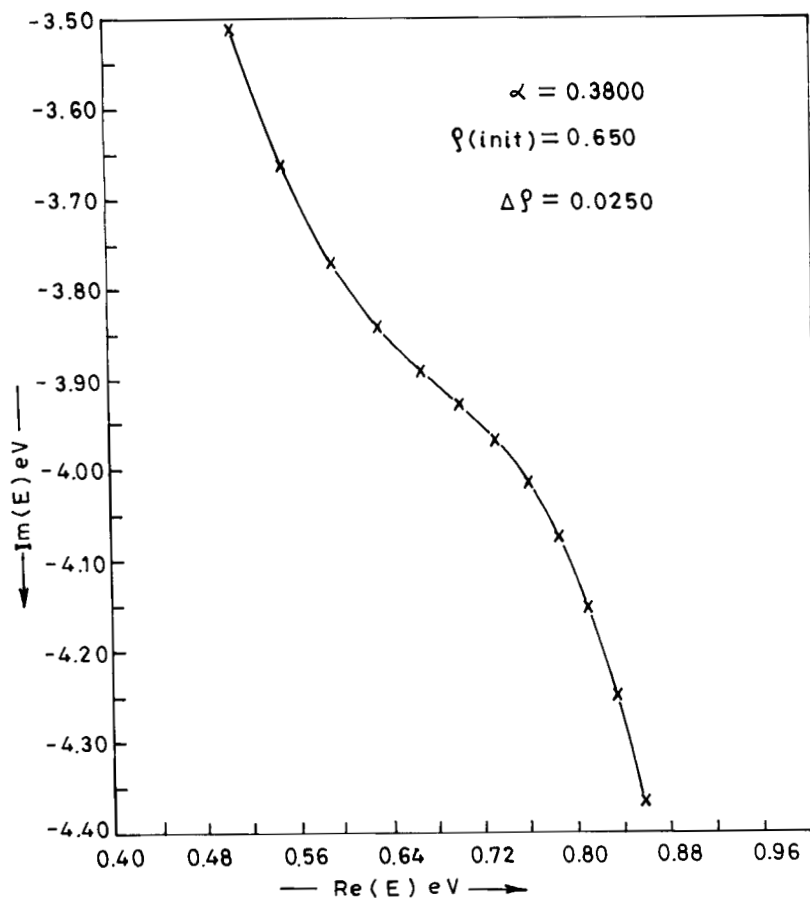


Fig.2a. Trajectory for the real and imaginary parts of the resonant orbital energy  $E$  denoted by dots (•) as a function of the real scale factor  $\rho$ . The starting value is  $\rho = 0.65$  at top left, and the increment  $\Delta \rho$  is 0.025. The rotation angle is fixed at  $\alpha = 0.38$ .

Figure 2b.

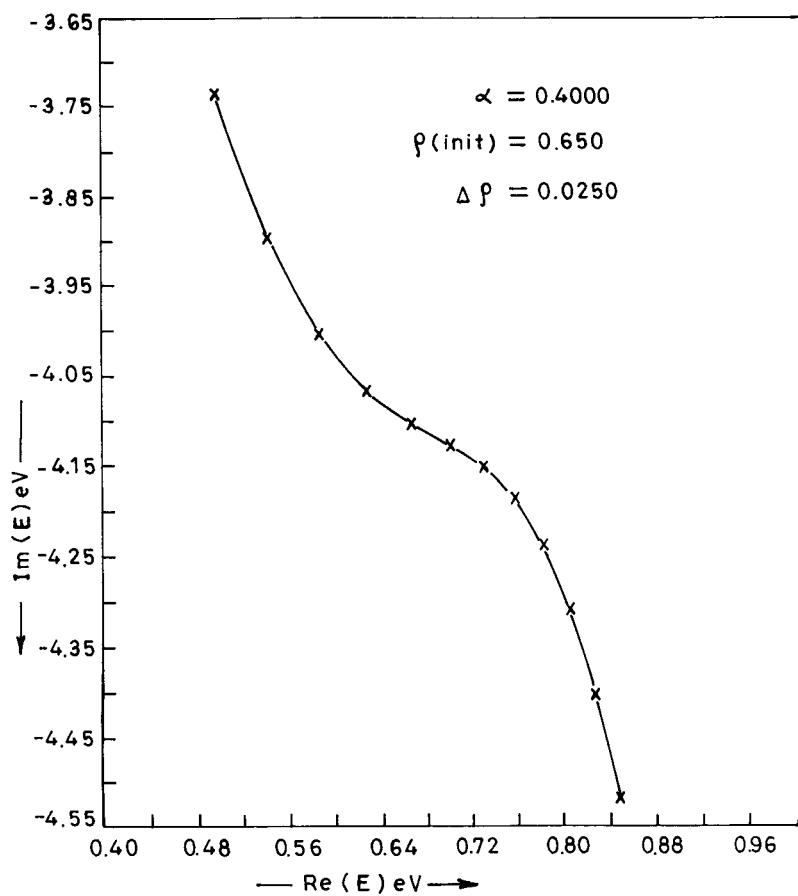


Fig. 2b. The same trajectory as in Fig. 2a, but with the rotation angle fixed at  $\alpha = 0.4$ .

Figure 3a.

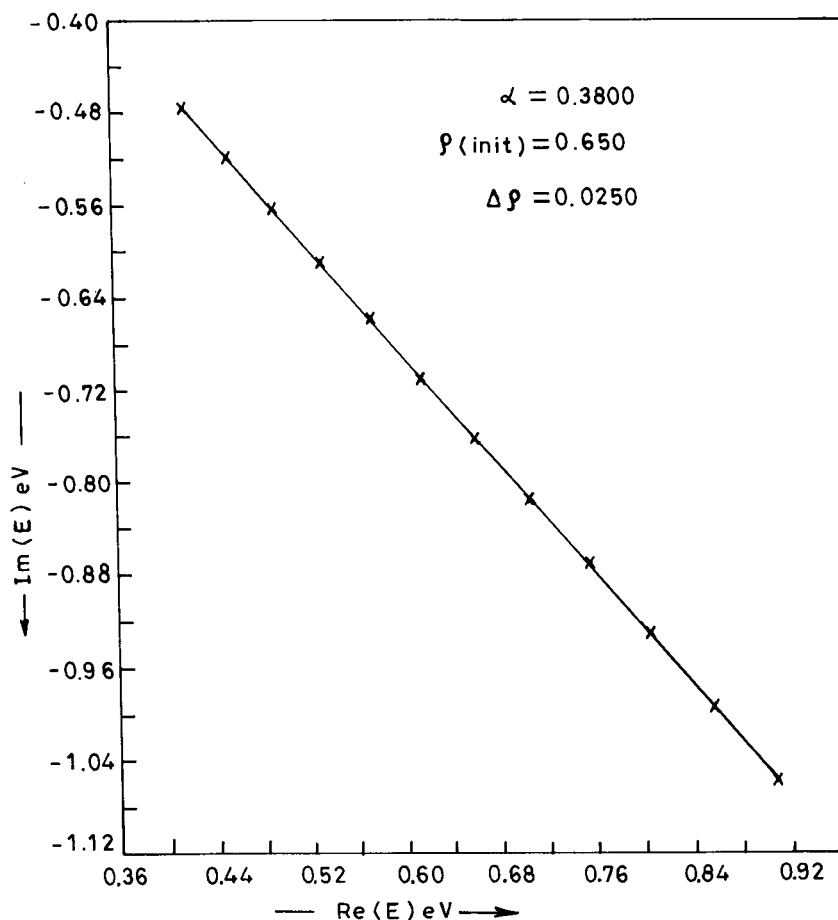


Fig.3a. Trajectory for the nearest scattering root (denoted by o in Fig.1) as a function of the real scale factor  $\rho$ . The starting value is  $\rho = 0.65$  at top left, and the increment  $\Delta \rho$  is 0.025. The rotation angle is fixed at  $\alpha = 0.38$ . Note that the slope of the line corresponds to the value of  $\alpha$  expected from the complex scaling theorem ( $\Delta \text{Im}[E]/\Delta \text{Re}[E] \approx \tan 2\alpha$ ).

Figure 3b.

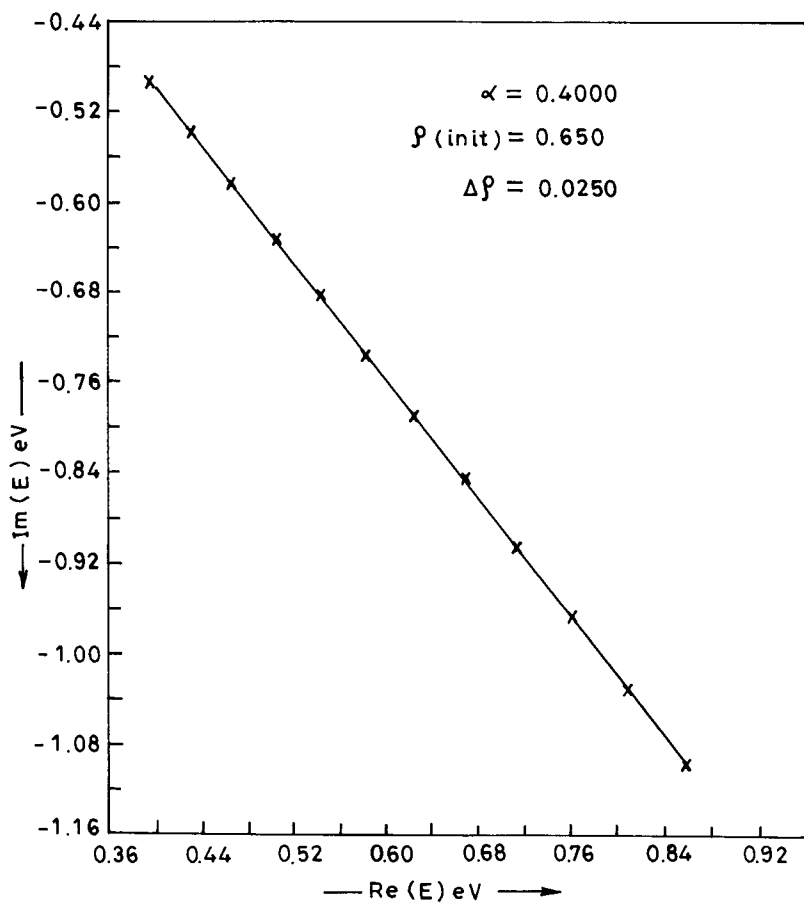


Fig. 3b. The same data as in Fig. 3a, but with the rotation angle fixed at  $\alpha = 0.4$ .

Figure 4.

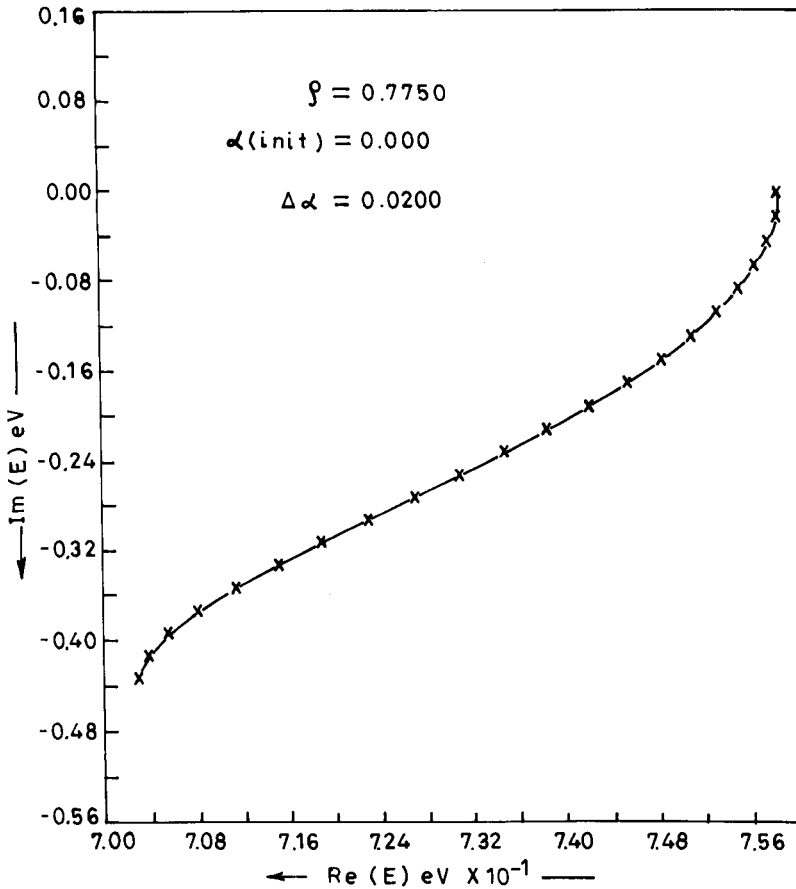


Fig.4. Trajectory for the real and imaginary parts of the resonant orbital energy  $E$  denoted by dots ( $\bullet$ ) as a function of the rotation angle  $\alpha$  with the starting value  $\alpha = 0.0$  at top right and the increment  $\Delta\alpha = 0.02$ . The real factor is here fixed at the optimal value  $\rho = 0.775$  as determined from Fig.2. Note that the distances between consecutive points ( $\bullet$ ) are monotonically increasing in the beginning, but noticeably slowing down at the end which resembles a cusp. This phenomenon is associated with a resonance; for further details see reference 10.



**Table I. Energy and Width of the  $^2\text{P}$  Resonance in the e-Be Scattering from Various Calculations.**

Reference Approach		Energy (eV)	Width (eV)
a	Model Potential	0.60	0.22
b	Static Exchange Phase Shift	0.77	1.61
c	Static Exchange Plus Polarizability Phase Shift	0.20	0.28
d	Static Exchange Phase Shift	0.75	1.64
	Static Exchange Plus Polarizability Phase Shift	0.14	0.13
	Static Exchange Cross Section	1.20	2.6
	Static Exchange Plus Polarizability Cross Section	0.16	0.14
e	Complex SCF	0.70	0.51
f	Coordinate Rotated Static Exchange	0.76	1.11
g	Second-order Dilated Electron Propagator Based on Real SCF	0.57	0.99
h	Complex SCF	0.69	0.51
	Complex CI: Singles, Doubles	0.58	0.38
	Singles, Doubles, and Triples	0.32	0.30
i	Second-order Dilated Electron Propagator Based on Complex SCF	0.62	0.60
	This work Bivariational SCF	0.70	0.84

- a J. Hunt and B.L. Moiseiwitsch, J.Phys. B **3**, 892 (1970).  
b H.A. Kurtz and Y. Öhrn, Phys.Rev. A **19**, 43 (1979).  
c H.A. Kurtz and K.B.Jordan, J.Phys. B **14**, 4361 (1981)  
d C.W. McCurdy, T.N. Rescigno, E.R. Davidson and J.G. Lauderdale, J.Chem.Phys. **73**, 3268 (1980).  
e T.N. Rescigno, C.W. Curdy and A.E. Orel, Phys.Rev. A **17**, 1931 (1978)  
f R.A. Donnelly and J. Simons, J.Chem.Phys. **73**, 2858 (1980).  
g M. Mishra, O. Goscinski, and Y. Öhrn, J.Chem.Phys. **79**, 5495 (1983).  
h J.F. McNutt and C.W. Curdy, Phys.Rev. A **27**, 132 (1983).  
i M. Mishra, O. Goscinski, and Y. Öhrn, J.Chem.Phys. **79**, 5505, (1983).

## 6. Conclusions and Acknowledgements.

The motivation for the start of the research reported in this paper was a desire to find out if at least a "shape resonance" could be detected in the Hartree-Fock approximation applied to a scaled Hamiltonian of the type described by relation (4.12). Using computer programs available for other purposes, programs for solving the dilated SCF equations could be constructed. In starting from a real orthonormal one-electron basis, it was found that all relevant matrices were *complex symmetric*, which greatly facilitated the computations, that many of the one-electron energies were complex, and that some of them showed certain "stability properties" indicating that they may be related to resonances. Even if it was evident that the one-electron spectra  $\{\epsilon_k\}$  were discontinuous functions of the rotation angle  $\alpha$ , it was still of interest to study what happened when this angle approached zero.

If the Hartree-Fock functions  $\psi$  undergo a transformation  $\psi^u = u\psi$ , and all the functions  $\psi$  are situated in the domain of  $u$ , then the one-particle operators  $\rho$  and  $H_{\text{eff}}(1)$  defined by (3.20) and (3.23), respectively, undergo the transformations (3.21) and (3.28) here referred to as the *Tarfala theorem*. It is also assumed that the functions  $\phi$  are situated in the domain of  $(u^\dagger)^{-1}$ , and we note that - in the complex symmetric case - one has  $\phi = \psi^* \langle \psi | \psi^* \rangle^{-1}$ , and that, since  $(u^\dagger)^{-1} = u^*$ , this condition is automatically satisfied. Since the effective Hamiltonian undergoes a similarity transformation, the spectrum  $\{\epsilon_k\}$  for the occupied orbitals ought to be invariant, and - at first sight - the Tarfala theorem seemed hence to be in contradiction to the numerical results already obtained. This puzzle was solved when it was realized that the *complex symmetric Hartree-Fock scheme* does not necessarily reduce to the conventional one when the Hamiltonian becomes real and self-adjoint for  $\alpha = 0$ . This result is also in agreement with some conclusions previously reported<sup>11,12</sup>.

It is interesting to observe that, if one limits oneself to study the original real and self-adjoint Hamiltonian with  $H^\dagger = H = H^*$  and applies the complex symmetric Hartree-Fock method to this particular case without any transformations whatsoever, one

ought still to obtain some complex one-electron energies, which depend on the starting point for the iterative SCF-procedure involved. By relating them to results of the complex scaling method, one could still draw the conclusion that they should be associated with physical "resonances" to the same extent as this method.

Numerical Hartree-Fock programs dealing with complex matrix elements have been available for quite some time in various research groups, and one could hence anticipate that several researchers should have encountered the occurrence of complex one-electron energies. In going over the literature, we have so far found at least one example that this phenomenon has been observed before<sup>13</sup>, but it seems likely that there should be many more. We hope that the analysis given here should help in clarifying the nature of the complex symmetric Hartree-Fock scheme.

In conclusion we note that one of the most urgent problems in this area of research is to find the connection between the exact theory and the numerical approximations based on the use of a finite basis of order  $m$  when the value of  $m$  increases and the basis - in principle - tends to become complete.

### Acknowledgements.

This paper was started during the 1981 Workshop held at the Tarfala Glaciolog station in the Kebnekaise mountain region in Lappland, Sweden. In these international workshops, every day devoted to scientific discussions was usually followed by a day of climbing - a combination leading to a very special form of scientific and personal friendship between the participants. This paper is respectfully dedicated to the memory of the late Professor Walter Schytt, who during many years together with his wife Anna-Nora was host for these workshops and for the International Summer Institutes in Quantum Chemistry and Solid-State Theory arranged by the Uppsala Quantum Chemistry Group in collaboration with the Florida Quantum Theory Project. The warm hospitality of the Schyts is hereby gratefully acknowledged, and one of the theorems in this paper is called the *Tarfala theorem* in their honour. - The authors are also thankful to Mrs. Karin Löwdin for her valuable help in equipping the field part of the Tarfala workshops during many years. They would also like to thank Drs. Jean-Louis Calais and Erkki Brändas for many valuable comments in the discussions.

## Appendix A.

### Treatment of One-Particle Subspaces of Order N, which are Stable under Complex Conjugation.

Let us consider a subspace of the  $L^2$  Hilbert space, which is spanned by the linearly independent set  $\Psi = \{\Psi_1, \Psi_2, \dots, \Psi_N\}$  of order N. Such a subspace is said to be *stable* under complex conjugation (\*), if

$$\Psi^* = \Psi\alpha. \quad (\text{A.1})$$

Introducing the metric matrix  $\Delta = \langle \Psi | \Psi \rangle$  and the reciprocal set  $\Psi_r = \Psi \Delta^{-1}$  having the property  $\langle \Psi_r | \Psi \rangle = \mathbb{1}$ , and multiplying the relation (A.1) to the left by  $\langle \Psi_r |$ , one obtains

$$\alpha = \langle \Psi_r | \Psi^* \rangle = \Delta^{-1} \langle \Psi | \Psi^* \rangle, \quad (\text{A.2})$$

which is our explicit expression for the expansion matrix of order NxN. This leads also to the relations

$$\langle \Psi | \Psi^* \rangle = \Delta\alpha, \quad \langle \Psi^* | \Psi \rangle = \alpha^\dagger \Delta, \quad (\text{A.3})$$

where the second relation is obtained by taking the adjoint of the first. One has now the following theorem:

The necessary and sufficient condition that the subspace spanned by the set  $\Psi$  is *stable* under complex conjugation is that the matrix  $\alpha = \langle \Psi | \Psi \rangle^{-1} \langle \Psi | \Psi^* \rangle$  satisfies the relation

$$\alpha^{-1} = \alpha^*. \quad (\text{A.4})$$

That the condition is *necessary* follows from the fact that, if  $\Psi^* = \Psi\alpha$ , one has  $\Psi = \Psi^*\alpha^*$ , which immediately gives  $\alpha^{-1} = \alpha^*$ .

In order to prove that the condition (A.4) is *sufficient*, we will consider the diagonal elements of the matrix:

$$\Omega = \langle \psi^* - \psi \alpha | \psi^* - \psi \alpha \rangle, \quad (\text{A.5})$$

when  $\alpha$  is defined by the relation (A.2). Using (A.3), one obtains directly:

$$\begin{aligned} \Omega &= \langle \psi^* | \psi^* \rangle - \langle \psi^* | \psi \rangle \alpha - \alpha^\dagger \langle \psi | \psi^* \rangle + \\ &+ \alpha^\dagger \langle \psi | \psi \rangle \alpha = \langle \psi^* | \psi^* \rangle - \alpha^\dagger \Delta \alpha - \alpha^\dagger \Delta \alpha + \\ &+ \alpha^\dagger \Delta \alpha = \langle \psi^* | \psi^* \rangle - \alpha^\dagger \Delta \alpha = \\ &= \langle \psi^* | \psi^* \rangle - \langle \psi^* | \psi \rangle \Delta^{-1} \Delta \Delta^{-1} \langle \psi | \psi \rangle = \\ &= \langle \psi^* | \psi \rangle \{ \langle \psi^* | \psi \rangle^{-1} \langle \psi^* | \psi^* \rangle - \langle \psi | \psi \rangle^{-1} \langle \psi | \psi^* \rangle \} = \\ &= \langle \psi^* | \psi \rangle \{ (\alpha^{-1})^* - \alpha \} = 0, \end{aligned} \quad (\text{A.6})$$

i.e. all the matrix elements of  $\Omega$  - diagonal as well as non-diagonal - are identically vanishing. Hence, one has necessarily  $\psi^* - \psi \alpha = 0$ , which completes the proof.

In relation (3.76), we are considering a stability relation of the type

$$v \psi_1^* = \psi_1 \beta, \quad (\text{A.7})$$

where - for the sake of generality - we have replaced  $\psi^u$  by  $\psi_1$  and  $\alpha$  by  $\beta$ . Here the operator  $v = u u^\dagger$  has the special property  $v^{-1} = (u^\dagger)^{-1} u^{-1} = u^*(u^\dagger)^* = (u u^\dagger)^* = v^*$ , which gives

$$v = v^\dagger = (v^{-1})^* \quad (\text{A.8})$$

and we note that also  $v$  is connected with a restricted transformation. Multiplying relation (A.7) to the left by  $\langle \psi_1 | v^{-1}$ , one obtains the following special form for the coefficient matrix:

$$\beta = \langle \psi_1 | v^{-1} | \psi_1 \rangle^{-1} \langle \psi_1 | \psi_1^* \rangle. \quad (\text{A.9})$$

One has now the theorem:

The necessary and sufficient condition that the subspace spanned by the set  $\psi_1$  is stable under complex conjugation combined with the operator  $v = uu^\dagger$  is that

$$\beta^{-1} = \beta^*. \quad (\text{A.10})$$

A simple proof is obtained by introducing the transformed functions  $\psi = u^{-1}\psi_1$  and reducing (A.10) to (A.4). However, even the direct proof is simple.

For this purpose, we will proceed as follows. In order to prove that the condition (A.10) is *sufficient*, we observe that, if  $\beta^{-1} = \beta^*$ , one has an alternative form for the matrix  $\beta$  defined by (A.9):

$$\beta = (\beta^*)^{-1} = \langle \psi_1 | \psi_1^* \rangle^{-1} \langle \psi_1^* | v | \psi_1^* \rangle \quad (\text{A.11})$$

Let us now consider the diagonal elements of the matrix

$$\begin{aligned} \Omega_1 &= \langle u^\dagger \psi_1^* - u^{-1} \psi_1 \beta | u^\dagger \psi_1^* - u^{-1} \psi_1 \beta \rangle = \\ &= \langle u^\dagger \psi_1^* | u^\dagger \psi_1^* \rangle - \langle u^\dagger \psi_1^* | u^{-1} \psi_1 \rangle \beta - \\ &- \beta^\dagger \langle u^{-1} \psi_1 | u^\dagger \psi_1^* \rangle + \beta^\dagger \langle u^{-1} \psi_1 | u^{-1} \psi_1 \rangle \beta = \\ &= \langle \psi_1^* | v | \psi_1^* \rangle - \langle \psi_1^* | \psi_1 \rangle \beta - \beta^\dagger \langle \psi_1 | \psi_1^* \rangle + \\ &+ \beta^\dagger \langle \psi_1 v^{-1} | \psi_1 \rangle \beta. \end{aligned} \quad (\text{A.12})$$

Using the form (A.9) for  $\beta$  in the last term, one finds that the last two terms cancel. Using the alternative form (A.11) for in the second term, one obtains

$$\Omega_1 = \langle \psi_1^* | v | \psi_1^* \rangle - \langle \psi_1^* | v | \psi_1^* \rangle = 0, \quad (\text{A.13})$$

Hence all the matrix elements of  $\Omega_1$  - diagonal as well as non-

diagonal - are identically vanishing, and one can draw the conclusion that  $u^\dagger \psi_1^* - u^{-1} w_1 \beta = 0$ , i.e. that  $v \psi_1^* = \psi_1 \beta$ . The proof is then completed. It is then fairly easy to test whether the stability conditions (3.76) are valid or not.

It should be observed that there are also other forms than (A.9) for the expansion matrix  $\beta$ . Multiplying the relation (A.7) to the left by  $\langle \chi |$ , where  $\chi = \{\chi_1, \chi_2, \dots, \chi_N\}$  is an arbitrary set of  $N$  linearly independent functions, one obtains

$$\beta = \langle \chi | \psi_1 \rangle^{-1} \langle \chi | v | \psi_1^* \rangle \quad (\text{A.14})$$

provided that the matrix  $\Delta_1 = \langle \chi | \psi_1 \rangle$  is non-singular. For the special choice  $\chi = v^* \psi_1$ , one obtains e.g. that

$$\begin{aligned} \Delta_1 &= \langle v^* \psi_1 | \psi_1 \rangle = \langle \psi_1 | v^{-1} | \psi_1 \rangle = \langle \psi_1 | (u^\dagger)^{-1} u^{-1} | \psi_1 \rangle = \\ &= \langle u^{-1} \psi_1 | u^{-1} \psi_1 \rangle, \end{aligned} \quad (\text{A.15})$$

which is a positive definite matrix. As before, one proves that the relation  $\beta^{-1} = \beta^*$  must necessarily be valid, and one has hence also the alternative form

$$(\beta^*)^{-1} = \langle \chi^* | v^{-1} | \psi_1 \rangle^{-1} \langle \chi^* | \psi_1^* \rangle, \quad (\text{A.16})$$

which corresponds to the introduction of the set  $v^* \chi^*$  instead of  $\chi$  in the original formula (A.14). We note, however, that - in the case of a general set - one can no longer expect that the condition  $\beta = (\beta^*)^{-1}$  should be sufficient for the required stability property, and that our proof is limited to the specific choice  $\chi = v^* \psi_1$ . This completes the appendix.

## References.

1. P.-O. Löwdin, *Advances in Quantum Chemistry* (Academic Press), vol. **19** (1987). This is the first part of this paper and is here referred to as reference A.
2. P.-O. Löwdin, *J.Math.Phys.* **24**, 70 (1983). This paper is here referred to as reference B.
3. P. Froelich and P.-O. Löwdin, *J.Math. Phys.* **24**, 88 (1983). This paper is here referred to as reference C.
4. P. Froelich and P.-O. Löwdin, submitted to *Int. J.Quantum Chem.* This paper is here referred to as reference D.
5. P.A.M. Dirac, "The Principles of Quantum Mechanics", (4th ed. Clarendon Press, Oxford 1958).
6. V. Fock, *Z.Physik* **61**, 1266 (1930); P.A.M. Dirac, *Proc. Cambridge Phil.Soc.* **26**, 376 (1930), **27**, 240 (1931).
7. D.R. Hartree, *Proc. Cambridge Phil.Soc.* **24**, 89 (1928); V. Fock, *Z.Physik* **61**, 126 (1930); J.C. Slater, *Phys.Rev.* **35**, 210 (1930).
8. P.-O. Löwdin, *Revs. Mod. Phys.* **35**, 496 (1963).
9. G.J. Schultz, *Revs.Mod.Phys.* **45**, 378 (1973); H.S. Taylor, *Adv. Chem.Phys.* **18**, 91 (1970).
10. M. Mishra, Y. Öhrn, and P. Froelich, *Phys.Lett.* **84**, 4 (1981); M. Mishra, P. Froelich, and Y. Öhrn, *Chem.Phys.Lett.* **81**, 339(1981); M. Mishra, O. Goscinski, and Y. Öhrn, *J.Chem.Phys.* **79**, 5494 (1983); **79**, 5505 (1983); M. Mishra, H.A. Kurtz, O. Goscinski, and Y. Öhrn, *J.Chem.Phys.* **79**, 1896 (1983).
11. E. Brändas, P. Froelich, C.H. Obcemea, N. Elander, and M. Rittby, *Phys.Rev. A* **26**, 3656 (1982); P. Froelich and E. Brändas, *Int.J.Quantum Chem.* **23**, 91 (1983); E. Brändas, *Int.J.Quantum Chem. QC Symp.* **20**, 119(1986); P. Krylstedt et.al., *J.Phys. B* **20**, 1295 (1987).
12. P. Froelich, *J.Math. Phys.* **24**, 2762 (1983).
13. C. W. McCurdy, J.L. Lauderdale, and R.C. Mowrey, *J.Chem.Phys.* **75**, 1835 (1981).



This Page Intentionally Left Blank

NEWTON BASED OPTIMIZATION METHODS  
FOR OBTAINING MOLECULAR CONFORMATION

John D. Head\*  
Michael C. Zerner

Quantum Theory Project  
University of Florida  
Department of Chemistry  
Gainesville, Florida 32611

and

Guelph Waterloo Centre for Graduate Work in Chemistry  
University of Guelph  
Department of Chemistry  
Guelph, Ontario, N1G 2W1 CANADA

---

\*Present address: University of Hawaii, Department of Chemistry,  
2545 The Mall, Honolulu, Hawaii 96822 USA.

ADVANCES IN QUANTUM CHEMISTRY, VOLUME 20

Copyright © 1989 by Academic Press, Inc.  
All rights of reproduction in any form reserved.

## TABLE OF CONTENTS

1.	Introduction
2.	Characterization of Minima
3.	Optimization Algorithms
4.	Line Search Strategy
5.	Search Directions
5.1	Steepest Descent
5.2	Newton's Method
5.3	Quasi-Newton Methods
5.4	Approximate Analytic Hessians
5.5	Restricted Step Method
5.6	The Augment Hessian Method
5.7	Conjugate Gradients
6.	Analysis of Search Direction Methods
7.	Computational Procedures
8.	Results
8.1	General Conclusions
8.2	Newton's Method
8.3	Quasi-Newton Methods with Unit Hessian
8.4	Quasi-Newton Methods with Non-Unit Starting Hessians
8.5	Approximate Hessians
8.6	Augmented Hessian
9.	Conclusions
	Appendix
	Acknowledgements
	References

## 1. INTRODUCTION

One of the most important advances in molecular quantum mechanics of the last fifteen years has been the increased ease and accuracy of predicting molecular conformation through purely theoretical models [1-4]. The stable geometries of molecules as yet unmade can be determined and molecular properties of these systems studied. The conformations and properties of electronically excited molecules can be examined theoretically, whereas experimentally these geometries are often very elusive. Transition states, the species that separates reactants from products, and that have but fleeting existence -- if at all, experimentally -- are also obtainable using molecular quantum mechanics.

Much of this progress has been centered around the discovery that energy derivatives can often be explicitly evaluated [3-7], freeing potential energy searches from inefficient minimization algorithms such as uniaxial methods, simplex methods and even numerical derivative methods [8].

We might classify methods that search potential energy surfaces as [9]:

1. Methods without derivatives,
2. Methods with numerical gradients (first derivatives),
3. Methods with analytic gradients, and
4. Methods with analytic gradients, second derivatives (Hessian).

with obvious extension and divisions. Methods of the first type such as the simplex methods [10] are simply not competitive with the more modern methods which obtain gradients (except for diatomic molecules). Since first derivatives for nearly any wavefunction can be obtained analytically, at least with the help of the coupled perturbed Hartree-Fock Theory (CPHF) [11], methods that use only numerical gradients will not be competitive except, again, for very small systems. Analytic second derivatives are generally time consuming to obtain, even for Hartree-Fock

(molecular orbital) type wavefunction. The evaluation of second derivatives inevitably involves at least the CPHF and, as such, is a factor of  $N$  more difficult than obtaining only the gradients, where  $N$  is the size of the basis set. Because of this, methods of the third type, those that use analytic first derivatives and numerical second, are important and are likely to remain important.

In this paper we examine various optimization procedures that build up information on the second derivatives as they minimize the energy. We focus in this work only on energy minima, although the connection with techniques for finding saddle point, such as transition states, is reasonably straight forward, at least in concept [12,13]. The descriptions of the various potentially useful methods is reasonably complete, although we have chosen to give examples only for the methods we have found most promising. For comparisons we do calculate the exact Hessian matrix and perform true Newton steps, although, except in especially difficult cases, we do not recommend this time consuming procedure. On the other hand, if vibrational spectra about a minima are to be examined, or if a transition state is sought, a "good" Hessian is required. The update procedures that we examine are not reliably "good" enough for these purposes.

In the next two sections we briefly detail the problem. Sections 4 and 5 describe two aspects of optimization algorithms, the line search and the search direction. These sections are somewhat lengthy as befits "art," for there is still a good deal of artistry in optimizing functions that do not explicitly show dependence on their independent variables. In section 4 we give an overview of the various suggested procedures in the context of a normal mode analysis where the "physics" of the situation is somewhat more transparent.

In Sections 7 and 8 we present results for the more successful approaches. The calculations presented are of the Hartree-Fock self consistent field type utilizing the INDO/1 [14,15] model Hamiltonian. Although the force constants from

such calculations are generally twice as large as those obtained from ab-initio results [16], this nearly constant scaling can be shown not to affect the generality of the conclusions presented in the last section of this paper.

## 2. CHARACTERIZATION OF MINIMA

We assume that the energy hypersurface  $E(\underline{x})$  is a unique single-valued function of the nuclear coordinates  $\underline{x}$ . We further assume  $E(\underline{x})$  and its derivatives with respect to  $\underline{x}$  are continuous. Under these conditions, the second order Taylor series for the energy hypersurface  $E(\underline{x})$  is well defined.

$$E(\mathbf{x}^k + 1) = E(\mathbf{x}^k) + \mathbf{g}^k \mathbf{s}^k + \frac{1}{2} \mathbf{s}^{k+} \mathbf{G} \mathbf{s}^k \quad (1a)$$

$$\mathbf{g}_i^k = \frac{\partial E}{\partial x_i} \Big|_{\mathbf{x} = \mathbf{x}^k} \quad (1b)$$

$$\mathbf{G}_{ij} = \frac{\partial^2 E}{\partial x_i \partial x_j} \Big|_{\mathbf{x} = \mathbf{x}^k} \quad (1c)$$

where  $\mathbf{g}^k = \mathbf{g}(\mathbf{x}^k)$  is a column vector of the energy gradients, and  $\mathbf{G}^k$  is the matrix of second derivatives, called the Hessian Matrix.  $\mathbf{s}^k$  is the "step" vector, given by

$$\mathbf{s}^k = \mathbf{x}^{k+1} - \mathbf{x}^k \quad (1d)$$

A strong local minimum with

$$E(\mathbf{x}^*) < E(\mathbf{x}) \text{ for all } \mathbf{x} \neq \mathbf{x}^* \quad (2a)$$

occurs at  $\mathbf{x}^*$  when

$$\mathbf{g}^* = 0 \text{ or } \mathbf{g}^{*+} \mathbf{s} = 0 \quad (2b)$$

$$\mathbf{G}^* > 0 \text{ or } \mathbf{s}^+ \mathbf{G}^* \mathbf{s} > 0 \text{ for all } \mathbf{s} \quad (2c)$$

Condition (2c) requires the Hessian matrix to be positive definite; that is, the eigenvalues of  $\mathbf{G}^*$  are all greater than zero.

The minimum usually found in a geometry optimization using cartesian coordinates is a "weak" minimum, distinguished by the equality of

$$E(\mathbf{x}^*) \leq E(\mathbf{x}), \mathbf{x} \neq \mathbf{x}^* \quad (3a)$$

The equality of Eq. (3a) arises because of the translational and rotational invariance of the energy. At the weak minimum we also have

$$\mathbf{g}^* = 0 \quad (3b)$$

$$\mathbf{G}^* \geq 0 \quad (3c)$$

where the Hessian matrix  $\mathbf{G}^*$  has 6 (or 5) zero eigenvalues corresponding to the rotational and translational modes; the remaining  $3N - 6(5)$  eigenvalues should all be positive [17]. The 6(5) coordinates corresponding to translations and rotations can be removed by a suitable choice of internal coordinates. However, the energy expression is invariant to this removal, and to any choice of linearly independent coordinates. This invariance property means that there is no particular advantage in an optimization procedure for any given choice of coordinates providing the initial steps (see below) are the same. Of course, it may be easier to constrain chemically meaningful coordinates, such as a bond length or angle, in an internal set of coordinates.

A potential energy surface in general will have many extreme or stationary points, all characterized by  $\mathbf{g}^* = 0$ . If  $\mathbf{s}^+ \mathbf{G}^* \mathbf{s} > 0$ , Eq. (2c), the point  $\mathbf{x}^*$  represents a minimum. Such a minimum may be local or the global minimum; the latter by definition will have the lowest energy in the hypersurface. All the methods presented in this paper are only capable of finding local minima. In contrast, a transition state with positions given by  $\mathbf{x}^\ddagger$  is characterized as an extreme point having all eigenvalues of the Hessian matrix positive except one and only one which is negative [18]. On such a "saddle point" the energy increases in all directions except along the "reaction path." Along this

direction  $E(\mathbf{x}^*)$  is a maximum separating two local minima. Other extreme points with two or more negative eigenvalues of the Hessian are of less immediate interest in chemistry.

### 3. OPTIMIZATION ALGORITHMS

Optimization methods are generally based on the typical algorithm as follows [9]: Choose an initial set of coordinates (variables)  $\mathbf{x}^0$  and calculate the energy  $E(\mathbf{x}^0)$  and gradients  $\mathbf{g}(\mathbf{x}^0)$ , then

1. Choose a search direction  $\mathbf{s}^k$ , then
2. Perform a line search to determine how far along this direction to move to reduce the value of  $E(\mathbf{x}^k)$ . In the simplest case one chooses an  $\alpha$  value in  $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha \mathbf{s}^k$ , then
3. Test for convergence. If the system has not converged, increase  $k$  to  $k + 1$ , and go to Step 4. If convergence is obtained, the optimization is complete.
4. Calculate or estimate  $\mathbf{G}^{k+1}$  a new Hessian at  $\mathbf{x}^{k+1}$  and repeat Steps 1 through 4.

The information gained about  $\mathbf{G}^{k+1}$  in Step 4 is used to choose the search direction  $\mathbf{s}^k$  of Step 1. Generally the first search direction is chosen along the direction of steepest descent although other choices, as discussed below, are possible. An efficient algorithm balances the work of finding the best search directions ( $\mathbf{s}^k$ ) against that of finding the optimal distance to move along that search direction ( $\alpha$ ).

### 4. LINE SEARCH STRATEGY

The aim of the line search is to determine the appropriate displacement along a search direction. This is, if at the point  $\mathbf{x}^k$  we have found a search direction  $\mathbf{s}^k$ , the  $k + 1$  point is that



given by

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{s}^k \quad (4)$$

where  $\alpha^k$  is obtained from the line search and should be such that  $E(\mathbf{x}^{k+1})$  is "reasonably" less than the previous energy  $E(\mathbf{x}^k)$ . The decrease in energy should be reasonable enough to ensure the optimization algorithm converges to the local minimum. The weak requirement that  $E(\mathbf{x}^{k+1}) < E(\mathbf{x}^k)$  can, of course, by itself result in very slow convergence.

Figure 1 illustrates a typical variation of energy with respect to  $\alpha$  along the search direction  $\mathbf{s}^k$ . In an exact line search, the minimizing value  $\alpha_m$ , which gives a minimum in  $E(\mathbf{x}^{k+1})$ , is located. This often requires several energy and gradient evaluations along  $\mathbf{s}^k$ . Such calculations may be more useful in selecting alternate search directions. The more efficient approach is to perform an inexact or "partial" line search where an  $\alpha$  value is selected to give a reasonable decrease in the function. For example, in Figure 1, an appropriate  $\alpha$

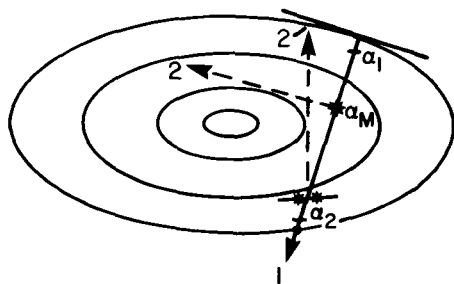


Figure 1

might lie in the interval  $(\alpha_1, \alpha_2)$ . The problem now is to determine from  $E(\mathbf{x}^{k+1})$  and  $\mathbf{g}^{k+1}$  appropriate value of  $\alpha_1$  and  $\alpha_2$ .

The energy along the search direction  $\mathbf{s}^k$  up to quadratic terms is given by

$$E(\mathbf{x}^k + \alpha \mathbf{s}^k) = E(\mathbf{x}^k) + \alpha \mathbf{g}^+(\mathbf{x}^k) \cdot \mathbf{s}^k + \frac{1}{2} \alpha^2 \mathbf{s}^{k+} \mathbf{G} \mathbf{s}^k \quad (5)$$

The energy at  $\mathbf{x}^k + \alpha \mathbf{s}^k$  will be lowered providing

$$\mathbf{g}^+(\mathbf{x}^k) \cdot \mathbf{s}^k < 0 \quad (6)$$

Inequality (6), called "the descent condition," selects the sign required for the search direction  $\mathbf{s}^k$  to ensure a reduction in the

energy. The minimum  $E(\mathbf{x}^k + \alpha_M \mathbf{s}^k)$  value is found by differentiating (5) with respect to  $\alpha$

$$\frac{\partial E}{\partial \alpha} = \mathbf{g}^+(\mathbf{x}^k) \cdot \mathbf{s}^k + \alpha \mathbf{s}^{k+} \mathbf{G} \mathbf{s}^k = \mathbf{g}^+(\mathbf{x}^k + \alpha \mathbf{s}^k) \cdot \mathbf{s}^k \quad (7)$$

Thus the exact line search requires

$$\mathbf{g}^+(\mathbf{x}^k + \alpha_M \mathbf{s}^k) \cdot \mathbf{s}^k = 0 \quad (8)$$

that is, the gradient  $\mathbf{g}$  at  $\mathbf{x}^k + \alpha_M \mathbf{s}^k$  is perpendicular to the search direction  $\mathbf{s}^k$  (see, for example, Figure 2). An estimate for the reduction of the energy can be obtained from Eqs. (5 - 8).

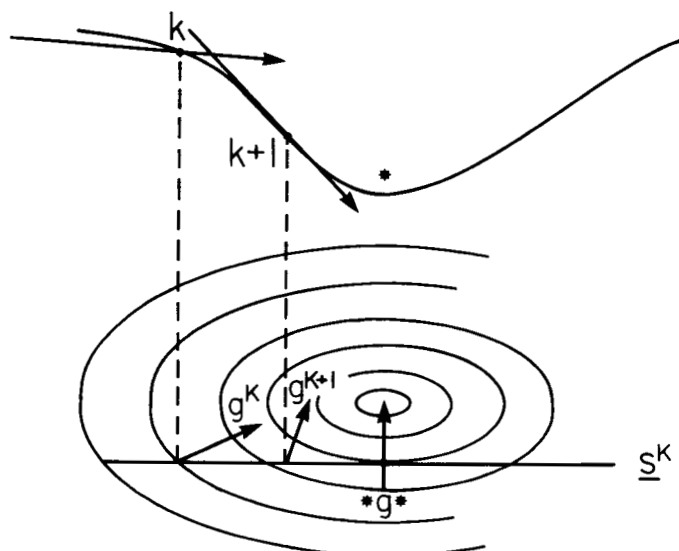


Figure 2

$$E^k - E^{k+1} = - \left( \frac{\alpha_M - \frac{1}{2}\alpha}{\alpha_M} \right) \alpha \mathbf{s}^{k+} \cdot \mathbf{g}^k \quad (9)$$

with the greatest reduction  $-1/2\alpha_M \mathbf{s}^{k+} \cdot \mathbf{g}^k$  when  $\alpha$  equals  $\alpha_M$ . When  $\alpha > \alpha_M$ , there is a reduction in the energy if

$$0 < \left( \frac{\alpha_M - \frac{1}{2}\alpha}{\alpha_M} \right) \leq \frac{1}{2} \quad (10)$$

which implies that

$$0 \leq -\rho \alpha s^{k+} \cdot g^k \leq E^k - E^{k+1} \leq \frac{1}{2} \alpha_M s^{k+} \cdot g^k \quad (11a)$$

when

$$0 \leq \rho \leq \frac{1}{2} \quad (11b)$$

Equations (11), originally suggested by Goldstein [19] motivate the test for whether  $\alpha$  is less than  $\alpha_2$ . For  $\alpha$  to be inside the right extreme of the line search interval  $(\alpha_1, \alpha_2)$ ,

$$E^k - E^{k+1} \geq -\rho \alpha s^{k+} \cdot g^k \quad (12)$$

when  $\rho = 1/2$  the line search becomes exact.

For the left extreme a test similar to Eq. (12) could be used. However, a more successful test in our experience is based on the gradients [20]. Since the Hessian  $G$  is positive definite, Eq. (7) gives

$$g^{k+1+} \cdot s^k \geq g^{k+} \cdot s^k \quad (13)$$

indicating for the new point  $x^{k+1}$  the gradient vector component along  $s^k$  should decrease. A successful test for the left extreme  $\alpha_1$  is

$$|g^{k+1+} \cdot s^k| \leq -\sigma g^{k+} \cdot s^k ; \quad \sigma \in (0,1) \quad (14a)$$

An exact line search is obtained when  $\sigma = 0$ .

Occasionally, and especially when starting with a poor initial geometry the gradient norm  $|g^{k+1}| > |g^k|$ , even when  $E^{k+1} < E^k$ . Such a situation is depicted in Figure 2. When this occurs we use the "cosine test" of Eq. (14b)

$$\frac{g^{k+1} \cdot s^k}{|g^{k+1}| |s^k|} < \sigma \cdot \frac{g^{k+} \cdot s^k}{|g^k| |s^k|} \equiv \cos \theta_k \quad (14b)$$

rather than the test of Eq. (14a); that is, the cosine of the angle between the step direction and the gradient along that

direction. This is because the new gradient becomes perpendicular to the direction of the step as the line search becomes more exact regardless of the magnitude of the gradient.

With left and right extremes of  $\alpha$  as chosen from Eqs. (12) and (14) the optimization procedure can be shown to converge, providing  $\mathbf{g}^{k+} \cdot \mathbf{s}^k \neq 0$ . Numerically, however, the condition that  $\mathbf{g}^{k+} \cdot \mathbf{s}^k \sim 0$  results more often than appreciated. In the Newton-like methods described below, this pathology is associated with ill-conditioned Hessian matrices. We discuss further this point in Section 6.

One additional constraint is needed in the discussion of line searches. The condition,

$$|\alpha_2 \mathbf{s}^k| \leq \Delta \quad (15)$$

safeguards against large coordinate changes which might accidentally lead to points outside the region of interest (in the neighborhood in which we have initiated the search), or beyond the neighborhood that we might reasonably expect quadratic behavior.

In our update procedure, after considerable numerical experimentation, we have adopted the following line search algorithm which is a variant of Fletcher [9a]. Initially an interval  $(\alpha_1^T, \alpha_2^T)$  is chosen, where the superscript T emphasizes trial values.  $\Delta$  of Eq. (15) is taken as 0.4 a.u. when  $\mathbf{x}$  is chosen as cartesian coordinates.  $\alpha_2^T$  is initially taken as 0.

We then calculate  $\alpha_2^T$  from Eq. (15).  $\alpha$  is initially set to 1 as suggested by a quadratic potential, unless  $\alpha_2^T < 1$  in which case we set  $\alpha = \alpha_2^T$ . We then evaluate  $E(\mathbf{x}^k + \alpha \mathbf{s}^k)$  and perform the right extreme test given by Eq. (12). If this test fails  $\alpha > \alpha_2$  (as opposed to  $\alpha_2^T$ ) and is reduced in value using the interpolation formula

$$\alpha_{\text{new}} = \alpha_1^T + \frac{1}{2} (\alpha - \alpha_1^T) \left[ 1 + \frac{E(\mathbf{x}^k) - E(\mathbf{x}^k + \alpha \mathbf{s}^k)}{(\alpha - \alpha_1^T) E'(\alpha_1^T)} \right]^{-1} \quad (16a)$$

$$E'(\alpha_1^T) = \mathbf{g}^+(\mathbf{x}^k + \alpha_1^T \mathbf{s}^k) \cdot \mathbf{s}^k = \left( \frac{\partial E}{\partial \alpha} \right)_{\alpha = \alpha_1^T} \quad (16b)$$

The energy at  $\mathbf{x}^k + \alpha_{\text{new}} \mathbf{s}^k$  is now computed and the right line search test is repeated with  $\alpha_2 = \alpha$  and  $\alpha = \alpha_{\text{new}}$  until Eq. (12) is satisfied.

The gradient  $\mathbf{g}(\mathbf{x}^k + \alpha \mathbf{s}^k)$  is now evaluated and used in the left extreme test, Eq. (14).

If the left extreme test fails then  $\alpha$  is too small. A trial  $\alpha_{\text{new}}$  is computed from the extrapolation formula

$$\alpha_{\text{new}} = \alpha + \frac{(\alpha - \alpha_1^T) E'(\alpha)}{E'(\alpha_1^T) - E'(\alpha)} \quad (16c)$$

so that  $\alpha < \alpha_{\text{new}}$  providing  $E'(\alpha_1) > E'(\alpha)$ . The energy is re-evaluated at  $\mathbf{x}^k + \alpha_{\text{new}} \mathbf{s}^k$  and both the right and left line searches are repeated with  $\alpha_1^T = \alpha$  and  $\alpha = \alpha_{\text{new}}$ . When both tests have been successfully completed, the partial line search is finished and a new search direction is sought.

The line search algorithm outlined is parameter dependent. After much numerical experimentation, we choose  $\sigma = 0.9$  and  $\rho = 0.03$ , parameters which constitute a fairly weak line search. In our experience it is more effective to change search directions than to determine a more exact minimum of energy along a given search direction. In practice  $\alpha$  is usually equal to unity. When a line search fails, however, the subsequent treatment of  $\alpha$  can have a marked affect on the efficiency of the optimization.

For steepest descent search directions, after the first step in any update procedure where the Hessian matrix is as yet unknown,  $\alpha^T \approx 0.4$  is most useful.

## 5. SEARCH DIRECTIONS

### 5.1 Steepest Descent.

The search direction in steepest descent is given by the negative of the gradient

$$\mathbf{s}^k = -\mathbf{g}^k \quad (17)$$

While the steepest descent search direction  $\mathbf{s}^k$  can be shown to converge to the minimum with a proper line search, in practice the method has slow and often oscillatory behavior. Most Quasi-Newton procedures, however, make this choice for the initial step, for there is no information yet on  $\mathbf{G}$ .

When working with different coordinates,  $\mathbf{q}$ , the Taylor series expansion suggests

$$\delta^+ \cdot \mathbf{g} = \delta_q^+ \cdot \mathbf{g}_q \quad (18)$$

where

$$\delta_q = \mathbf{B} \delta \quad (19)$$

defines the new set of coordinate steps  $\delta_q$ . From Eq. (18)

$$\mathbf{g} = \mathbf{B}^+ \mathbf{g}_q \quad (20)$$

leading to

$$\mathbf{s}_q^k = \mathbf{B} \mathbf{s}^k = -\mathbf{B}\mathbf{B}^+ \mathbf{g}_q \quad (21)$$

as the proper replacement for Eq. (17). Failure to use Eq. (21) correctly for the steepest descent can lead to symmetry breaking [21]. It has also on occasion led to incorrect conclusions on the importance of linear coordinate transformations.

### 5.2 Newton's Method.

Newton's method includes the quadratic nature of the energy hypersurface in the search direction computation. Consider the Taylor's series for the change in gradient given by

$$\mathbf{g}(\mathbf{x}^{k+1}) = \mathbf{g}(\mathbf{x}^k) + \mathbf{G}^k \cdot \mathbf{s}^k \quad (22)$$

at an extremum  $\mathbf{g}(\mathbf{x}^*)$  is zero. Hence a step in the direction of minimum is predicted by

$$\mathbf{s}^k = (\mathbf{x}^* - \mathbf{x}^k) = -\mathbf{G}^{-1} \mathbf{g}(\mathbf{x}^k) \quad (23)$$

Equation (23) is used to generate search direction  $\mathbf{s}^k$  for all the Newton-like algorithms. The actual step taken  $\delta^k = \alpha \mathbf{s}^k$  is determined by the line search and allows for the non-quadratic part of the function.

In molecular quantum mechanics, the analytical calculation of  $\mathbf{G}$  is very time consuming. Furthermore, as discussed later, the Hessian should be positive definite to ensure a step in the direction of the local minimum. One solution to this later problem is to precondition the Hessian matrix and this is discussed for the restricted step methods. The Quasi-Newton methods, presented next, provides alternative solution to both of these problems.

### 5.3 Quasi-Newton Methods.

The Quasi-Newton, or variable metric, methods of optimization are essentially based on the algorithm

1. Set search direction  $\mathbf{s}^k = -\mathbf{H}^k \mathbf{g}^k$ .
2. Perform line search to give  $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{s}^k$ .
3. Update  $\mathbf{H}^k$  to  $\mathbf{H}^{k+1}$

where the search direction is essentially given as a Newton step, Eq. (23), but with  $\mathbf{H}^k$  approximating the inverse Hessian.  $\mathbf{H}^k$  avoids the explicit evaluation of  $\mathbf{G}^{-1}$  and is updated on each algorithm cycle.

The inverse Hessian  $\mathbf{H}$  is updated by relating the differences in the gradients

$$\mathbf{y}^{k+1} = \mathbf{g}^{k+1} - \mathbf{g}^k \quad (24)$$

to the differences in the coordinates

$$\delta^k = \alpha^k s^k = x^{k+1} - x^k \quad (25)$$

For a quadratic function, the coordinates and gradients are related by

$$H^{k+1} \gamma^k = \delta^k \quad (26)$$

Equation (26) is the "Quasi-Newton" condition; it is fundamental to all the updating formula. There have been many updates proposed, and we briefly review some of the more important ones.

The simplest are based on

$$H^{k+1} = H^k + E^k = H^k + a |u\rangle\langle u| \quad (27)$$

where  $|u\rangle$  is a vector;  $|u\rangle\langle u|$  forms a dyadic, and thus the correction matrix is of rank one. Substituting  $H^{k+1}$  into the Quasi-Newton condition gives

$$H^k \gamma^k + a |u\rangle\langle u| \gamma^k = \delta^k \quad (28)$$

and hence  $|u\rangle$  is proportioned to  $|g^k\rangle - H^k |\gamma^k\rangle$ . Taking, for simplicity, the scale  $a = \langle u|\gamma^k\rangle^{-1}$  we obtain from Eqs. (27) and (28)

$$H^{k+1} = H^k + \frac{(|\delta\rangle - H|\gamma\rangle)(\langle\delta| - \langle\gamma|H)}{\langle\gamma|\delta - H\gamma\rangle} \quad (29)$$

where the  $k$  superscript has been dropped for convenience. Rank one formula of this type have been proposed by Broyden [22], Davidon [23], Fiacco and McCormick [24], and Murtagh and Sargent [25], and examined for their utility in molecular quantum mechanics by many workers [2,26]. A common approach used is to take  $H^1$  as a unit matrix, and thus the first step  $s^1$  is in the direction of steepest descent. An alternative procedure, potentially useful when the initial search is near the minimum of energy, is to start the Hessian as a diagonal matrix of known force constants for bonded atoms that have been tabulated for the model at hand (27,28). In such a case, internal coordinates should be used, Eqs. (20) and (21). In later cycles more of the quadratic nature of the energy is included in  $H^k$ . Unfortunately,



this simple rank one update procedure has two serious disadvantages. The positive definiteness of the inverse Hessian may be lost on updating. Secondly, the denominator in Eq. (29) may become close to zero, with resulting loss in numerical precision. Murtagh and Sargent [29] have developed successful algorithms for detecting these problems but not correcting them. Often the only solution is to reset the inverse Hessian back to the unit matrix. This results in the loss of all the previous curvature information. Nevertheless, the Murtagh Sargent variant of the rank one formula has in the past been the most successful in molecular geometry searches.

There is a whole class of rank two update formula, the two most important members are

$$\mathbf{H}_{\text{DFP}}^{k+1} = \mathbf{H} + \frac{|\delta\rangle\langle\delta|}{\langle\gamma|\delta\rangle} - \frac{\mathbf{H}|\gamma\rangle\langle\gamma|\mathbf{H}}{\langle\gamma|\mathbf{H}|\gamma\rangle} \quad (30)$$

where the subscript DFP indicates Davidson, Fletcher and Powell [30] who developed this update; and

$$\begin{aligned} \mathbf{H}_{\text{BFGS}}^{k+1} = & \mathbf{H} + \left( 1 + \frac{\langle\gamma|\mathbf{H}|\gamma\rangle}{\langle\delta|\gamma\rangle} \right) \frac{|\delta\rangle\langle\delta|}{\langle\delta|\gamma\rangle} \\ & - \left( \frac{|\delta\rangle\langle\gamma|\mathbf{H} + \mathbf{H}|\gamma\rangle\langle\delta|}{\langle\delta|\gamma\rangle} \right) \end{aligned} \quad (31)$$

suggested by Broyden, Fletcher, Goldfarb and Shanno (BFGS) [31]. The complete family of rank two update is given by a linear combination of the DFP and BFGS formula [32]

$$\mathbf{H}_{\phi}^{k+1} = \phi \mathbf{H}_{\text{DFP}}^{k+1} + (1 - \phi) \mathbf{H}_{\text{BFGS}}^{k+1} \quad (32)$$

and is known as the Broyden rank two "family." These rank two updates have several desirable features, but perhaps the most important property is that they retain the positive definiteness of the Hessian if the appropriate line search is used. The advantage of a particular rank two formula over another in molecular problems is not clear, and for exact line searches all

the formula give the same search directions. As is indicated in the result section, the BFGS formula in our experience tends to be the most successful with partial line searches.

An alternative type of update formula has been suggested by Greenstadt [33]. The updating formula has the usual form

$$\mathbf{H}_G^{k+1} = \mathbf{H}_G^k + \frac{|\epsilon\rangle\langle\gamma| + |\gamma\rangle\langle\epsilon|}{\langle\gamma|\gamma\rangle} - \frac{\langle\epsilon|\gamma\rangle|\gamma\rangle\langle\gamma|}{\langle\gamma|\gamma\rangle^2} \quad (33)$$

where  $|\epsilon\rangle = |\delta\rangle - \mathbf{H}|\gamma\rangle$ . In the Greenstadt formula, the positive definiteness of the inverse Hessian is not retained, nor is it invariant to coordinate changes. In the above we have given update formula for the inverse Hessian for direct use in such equations as Eq. (23). In the Appendix we summarize the update formulas for the Hessian itself. Such formulas may be of greater utility in methods such as the augmented Hessian and restricted step methods, discussed below.

#### 5.4 Approximate Analytic Hessians

Exact second derivatives methods require the solution of the coupled perturbed Hartree-Fock equations, CPHF [11,34,35]. At the Hartree-Fock level this requires several steps in addition to the usual SCF procedure and the evaluation of the first derivatives.

1. An ao to mo transformation. Only those integrals with two occupied and two virtual indices are required,  $\langle ij|ab\rangle$  and  $\langle ij|ba\rangle$ . This is a  $N^5$  step. This transformation step can be avoided by working directly in the ao basis [36] but at the expense of increased I/O.
2. Evaluation of the integral second derivatives, a step proportional to the number of nuclear coordinates times the number integrals, about  $N^5$ .
3. The solution of the CPHF equations, a  $N^4$  step for each nuclear coordinate, or again about  $N^5$ .

Obtaining the second derivatives for a SCF wave function is thus a factor of N times more difficult than it is for evaluating the SCF energy and the first derivatives.

The resulting CPHF equations to be solved are

$$(1 - A) U^{(1)} - B = 0 \quad (34a)$$

$$\sum_{b,j} (\delta_{ai,bj} - A_{ai,bj}) U_{bj}^{(1)} - B_{ai} = \quad (34b)$$

$$= \sum_{bj} \frac{\Delta E_{ai,bj} - \tilde{A}_{ai,bj}}{\epsilon_i(0) - \epsilon_a(0)} U_{bj}^{(1)} - B_{ai} = 0 \quad (34c)$$

$$\Delta E_{ai,bj} \equiv [\epsilon_i(0) - \epsilon_a(0) - \langle ai || aj \rangle] \quad (34c)$$

$$\tilde{A}_{ai,bj} \equiv \langle ab || ij \rangle + \langle aj || ib \rangle (1 - \delta_{ai,bj}) \quad (34d)$$

where  $U^{(1)}$  is the matrix of first order changes in the molecular coefficients

$$C(X) = C(X_0)U(X_0 + \delta X) = CU(X_0 + \delta X) \quad (35a)$$

$$U(X_0 + \delta X) = U(X_0) + \delta XU^{(1)} + \dots \quad (35b)$$

and

$$B_{ai} = \frac{\tilde{Q}_{ai}^{(1)}}{\epsilon_i - \epsilon_a} \quad (35c)$$

with  $\tilde{Q}^{(1)}$ , the matrix of first derivatives given by

$$\begin{aligned} \tilde{Q}_{ai}^{(1)} &= \tilde{H}_{ai}^{(1)} - \tilde{S}_{ai}^{(1)} \epsilon_i - \sum_k \tilde{S}_k^{(1)} \langle a || ik \rangle \\ &+ \sum_{\mu\nu\sigma\lambda} C_{\mu\sigma}^* C_{\nu i} P_{\lambda\sigma} \left[ \frac{\partial}{\partial y} (\mu\lambda || \nu\sigma) \right]_y = 0 \end{aligned} \quad (36a)$$

$$\tilde{H}_{ai}^{(1)} = \sum_{\mu, \nu} C_{\mu\sigma}^* \left[ \frac{\partial H_{\mu\nu}}{\partial y} \right]_{y=0} C_{\nu i}, \text{ etc.} \quad (36b)$$

From Eq. (35c), regrouping terms,

$$\sum_{bj} (\Delta E_{ai,bj} - \tilde{A}_{ai,bj}) U_{bj}^{(1)} - \tilde{Q}_{ai}^{(1)} = 0 \quad (37a)$$

or

$$U^{(1)} = (\Delta E - \tilde{A})^{-1} \tilde{Q}^{(1)} \quad (37b)$$

Expanding  $U^{(1)}$  itself as a second perturbation sequence

$$\begin{aligned} U^{(1)} &= \Delta E^{-1} \tilde{Q}^{(1)} + \Delta E^{-1} \tilde{A} \Delta E^{-1} \tilde{Q}^{(1)} + \dots \\ &= V^{(1)} + V^{(2)} + \dots \end{aligned} \quad (38a)$$

$$V_{ai}^{(1)} = [\epsilon_i - \epsilon_a + \langle ai || ai \rangle]^{-1} \tilde{Q}_{ai}^{(1)} \quad (39)$$

Equations such as (38) are immediately suggested by the original CPHF Eq. (34),

$$U = (1 - A)^{-1} B = \sum_{n=0} A^n B \quad (40)$$

the only difference being the denominator shift by  $\langle ai || ai \rangle$ .

Whereas Eq. (40) is slow to converge, Eq. (38a) is not.

Replacement of  $U^{(1)}$  by  $V^{(1)}$  in

$$\begin{aligned} G_{xy} &= \sum_{\mu, \nu} P_{\mu\nu} \left( \frac{\partial^2 H_{\mu\nu}}{\partial x \partial y} + \frac{1}{2} \sum_{\sigma, \lambda} P_{\lambda\sigma} \frac{\partial^2}{\partial x \partial y} \langle \mu\lambda || \nu\sigma \rangle \right) \\ &+ \frac{\partial^2 V_{NUC}}{\partial x \partial y} - \sum_{\mu, \nu} W_{\mu\nu} \frac{\partial^2 S_{\mu\nu}}{\partial x \partial y} + \sum_{\mu, \nu} \frac{\partial P_{\mu\nu}}{\partial y} \frac{\partial H_{\mu\nu}}{\partial x} + \end{aligned} \quad (40a)$$

$$+ \sum_{\mu\nu} \sum_{\sigma\lambda} \frac{\partial P_{\mu\nu}}{\partial y} P_{\lambda\sigma} \frac{\partial}{\partial x} \langle \mu\lambda | | \nu\sigma \rangle \sum_{\mu\nu} \frac{\partial W_{\mu\nu}}{\partial y} \frac{\partial S_{\mu\nu}}{\partial x}$$

with

$$W_{\mu\nu} = \sum_i \epsilon_i C_{\mu i}^* C_{\nu i} \quad (40b)$$

$$\begin{aligned} \frac{\partial P_{\mu\nu}}{\partial y} &= \left( \frac{\partial P_{\mu\nu}}{\partial y} \right)_{y=0} = \sum_i^{\text{occ}} \sum_p C_{\mu p} C_{\nu i} U_{pi}^{(1)*} \\ &\quad + C_{\mu i}^* C_{\nu p} U_{pi}^{(1)} \end{aligned} \quad (40c)$$

generally give the second derivatives to within  $\pm 0.01$  au/bohr, far better than any of the update procedures. For the first iterates,  $V^{(1)}$ , only the  $\langle ai|ai \rangle$  and  $\langle ai|ia \rangle$  integrals are required, "a" virtual and "i" occupied. All the integral second derivatives are required. For semi-empirical methods the reduction of two of the three additional steps after the SCF and first derivative evaluation is enormous, a full factor of  $N$  as seen in the results section. For ab-initio theories, reduction of the integral transforms is approximately a factor of 4. Realized efficiencies then depends on the relative time/rate required for the integral transform step and the derivative integral steps, and this is often a function of the computer used.

### 5.5 Restricted Step Methods.

The Quasi-Newton methods are very powerful techniques for locating local minima. The success of these procedures can largely be attributed to the retention of the positive definiteness of the Hessian. In this section we present an alternative approach which has the added capability of handling Hessians which may contain negative eigenvalues. Such procedures will prove especially useful for starting geometries which are well removed from the desired minima. For example, a starting

geometry may correspond to a transition state structure with one negative eigenvalue in the Hessian. The restricted step algorithm should be capable of optimizing towards both the reactant and the product geometries.

The restricted step method of the type discussed below were originally proposed by Levenberg and Marquardt [37,38] and extended to minimization algorithms by Goldfeld, Quandt and Trotter [39]. Recently Simons [13] discussed the restricted step method with respect to molecular energy hypersurfaces. The basic idea again is that the energy hypersurface  $E(\mathbf{x})$  can reasonably be approximated, at least locally, by the quadratic function

$$E(\mathbf{x} + \delta) = E(\mathbf{x}) + \mathbf{g}^+(\mathbf{x}) \cdot \delta + \frac{1}{2} \delta^+ G(\mathbf{x}) \delta \quad (41)$$

where the step  $\delta$  is restricted in some way. In our case we restricted the step norm  $|\delta|$  to be smaller or equal to some trust region size  $\Delta$  which is a measure of the extent of the quadratic approximation.

$$|\delta| \leq \Delta \quad (42)$$

The minimum on such a quadratics surface can readily be identified. it is either a minimum within the region, with the gradient vector  $\mathbf{g}$  null and the Hessian matrix  $G$  positive definite or the minimum is on the boundary of the region.

The step  $\delta$  towards the minimum is found by minimizing  $E(\mathbf{x} + \delta)$  with respect to  $\delta$  imposing the constraint from Eq. (42). Introducing a Lagrange's undetermined multiplier  $\lambda$  we obtain the functional

$$F = E(\mathbf{x} + \delta) - \frac{\lambda}{2} (\delta^+ \delta - \Delta^2) \quad (43a)$$

which on differentiation becomes

$$\nabla F = \mathbf{g} + G\delta - \lambda\delta = 0 \quad (43b)$$

and

$$\delta = -(G - \lambda I)^{-1} \mathbf{g} \quad (43c)$$

$\lambda$  is a parameter which ensures that the matrix  $(G - \lambda I)$  is a positive definite and that  $\delta$  is essentially given by a Newton-like step. Analysis of Eqs. (43) reveals that a downhill search requires

$$\lambda < \lambda_{\min} \quad (44)$$

where  $\lambda_{\min}$  is the lowest eigenvalue of  $G$ . When  $\lambda$  is zero we obtain the usual Newton step. If  $\lambda$  equals zero satisfies the inequality (44) and the step size  $|\delta|$  is smaller than  $\Delta$ , then we have the situation where the minimum is within the quadratic region.

The restricted step method requires the two parameters  $\Delta$  and  $\lambda$  to be defined. Various numerical tests have been suggested for varying the trust region size  $\Delta$  depending upon the quality of the quadratic surfaces [9,37,38].

Until now we have defined this trust region  $\Delta$  as fixed, as explained with the discussion on Eq. (15). It may be the nature of the energy surfaces we have examined, but very little seems gained in searching for energy minima by examining the trust region in detail. In addition, solving for  $\lambda$ , which determines  $\Delta$  is a non-trivial problem [38]. The augmented Hessian procedure described next circumvents this particular problem by introducing a natural choice for  $\lambda$ .

In our experience the Newton and Quasi-Newton methods with the line search described is faster to reach a minimum than the restricted step method with confidence region. This latter method is a more "conservative" method as discussed in the results section. The restricted step methods, however, are very effective in searching for transition states [40,41].

## 5.6 The Augmented Hessian Method

The starting point in the augmented Hessian approach is the eigenvalue equation

$$\begin{pmatrix} G & \alpha g \\ \alpha g^+ & 0 \end{pmatrix} \begin{pmatrix} v \\ \beta \end{pmatrix} = \lambda \begin{pmatrix} v \\ \beta \end{pmatrix} \quad (45a)$$

where the gradient  $g$  is scaled by  $\alpha$ . The step  $s$  is found by diagonalizing Eq. (45a) for the lowest eigenvalue  $\lambda$  and lowest normalized eigenvector  $v^2 + \beta^2 = 1$ . By expanding Eq. (45a) two equations are obtained.

$$Gv + \alpha\beta g = \lambda v \quad (45b)$$

$$\alpha \langle g | v \rangle = \lambda \beta \quad (45c)$$

Rearranging (45b) gives the step

$$s = \frac{v}{\alpha\beta} = - (G - \lambda I)^{-1} g \quad (45d)$$

to be compared with the restricted step equation, Eq. (43c), and Newton's equation, Eq. (22b). The eigenvalues of the augmented Hessian ( $\lambda_i^{m+1}$ ) will bracket those of the Hessian itself, ( $\lambda_i^m$ ). In particular, the lowest eigenvalue  $\lambda_1^{m+1}$  must be zero or negative.

In the diagonalization one of the following three situation are obtained.

1.  $\lambda < 0$  and  $\beta = 0$ . In this case  $v = v(\text{low})$  is the eigenvector of  $G$  corresponding to the lowest eigenvalue.  $\langle v(\text{low}) | g \rangle$  is zero and  $\lambda = \lambda(\text{low})$ . A displacement along  $v(\text{low})$  will reduce the energy. Alternatively, if a displacement along  $v(\text{low})$  is not desired, then  $v(\text{low})$  is shifted to a higher eigenvalue by  $G' = G + \lambda(\text{shift}) |v(\text{low})\rangle \langle v(\text{low})|$  and the Augmented Hessian reformed.
2.  $\lambda < 0$  and  $\beta = 1$ .  $G$  is now positive definite. The step will vary from Newton like for  $\lambda$  small to steepest descent for  $\lambda$  large. The step length will usually be inside the restricted region.



3. Intermediate between the two conditions is  $\lambda < 0$  and  $\beta \neq 0$ .  $G$  has a negative eigenvalue but  $s$  given by Eq. (45d) will give a reduction in energy along the negative eigenvector mode.

The step size,  $|s|$ , depends in a non-linear way on  $\alpha$ . Since  $|s|$  is a monotonically decreasing function of  $\alpha$ , interpolation can be easily used to find a step of the appropriate length.

The augmented Hessian method requires an exact Hessian, or an update method on the Hessian itself. The update formula for the Hessian analysis to the inverse Hessian appear in the Appendix. At each step in the update procedure, the procedure is defined such that the Hessian times its inverse is unity.

### 5.7 Conjugate Gradients

Of some historic value, and for completeness, we mention that an alternative method for selecting search directions is to use the conjugate gradient approach [9]. The initial direction is taken as that given by steepest descent

$$s^1 = -g^1 \quad (46a)$$

then for  $k > 1$ ,  $s^{k+1}$  is given by

$$s^{k+1} = -g^{k+1} + \sum_{j=1}^k \beta_j s^j \quad (46b)$$

where  $\beta_j$  is determined by requiring all the search directions be "conjugate" to each other; that is

$$s^i + G s^j = 0 \quad i \neq j \quad (46c)$$

In the Fletcher-Reeves algorithm [42], Eq. (46b) is simplified to

$$s^{k+1} = -g^{k+1} + \beta^k s^k \quad (46d)$$

with

$$\beta^k = \frac{\langle g^{k+1} | g^{k+1} \rangle}{\langle g^k | g^k \rangle} \quad (46e)$$

where Eq. (46d) is exact for a quadratic function. Other algorithms present different ways of forming  $\beta^k$  which are equivalent for a quadratic function. The principle drawback of conjugate-gradient methods are that they are designed for use with exact line searches. They do have the interesting feature of not explicitly requiring the Hessian in the equations. It is now fairly well established, however, that when there is gradient information available, the conjugate gradient approaches are not competitive with the quasi-Newton methods [9].

## 6. ANALYSIS OF SEARCH DIRECTION METHODS

The Hessian matrix  $G$  is real and symmetric. It can then be resolved for analysis in terms of its eigenvalues  $\lambda_i$  and its eigenvectors  $|\omega_i\rangle$  as

$$G = \sum_i |\omega_i\rangle \lambda_i \langle \omega_i| \quad (47)$$

Since the set  $|\omega_i\rangle$  is complete, the gradient  $|g\rangle$  and the step  $|s\rangle$  can be decomposed as

$$|g\rangle = \sum_i \epsilon_i |\omega_i\rangle, \quad \epsilon_j = \langle \omega_j | g \rangle \quad (48a)$$

$$|s\rangle = \sum_i \eta_i |\omega_i\rangle, \quad \eta_i = \langle \omega_i | s \rangle \quad (48b)$$

From Eq. (47) and Eq. (44)

$$|s\rangle = - \sum_i \frac{\epsilon_i}{\lambda_i} |\omega_i\rangle \quad (48c)$$

$$\eta_i = -\epsilon_i/\lambda_i \quad (48d)$$

The descent condition  $\langle s|g \rangle < 0$  becomes

$$\langle s|g \rangle = \sum_i \epsilon_i \eta_i = \sum_i -\epsilon_i^2/\lambda_i < 0 \quad (49)$$

and is certainly obeyed if each  $\epsilon_i \eta_i < 0$  or if  $\lambda_i > 0$  for all  $i$ . This is, of course, the case, if  $G$  or its approximations are positive definite. When  $G$  is indefinite, the restricted step method modifies the  $\lambda_i$ , Eq. (43), so they are all positive. The most common techniques when they run into trouble, however, reset all  $\lambda_i$  to 1 as in the Murtagh Sargent procedure [29]. Some techniques diagonalize  $G$ , and, if  $\lambda_i$  is negative, simply change its sign. Such approaches do not systematically modify the Hessian and reduce the overall efficiency of the optimization algorithm.

The angle,  $\theta$ , between the step and the gradient is given by

$$\cos \theta = \frac{-\langle s|g \rangle}{(\langle s|s \rangle \langle g|g \rangle)^{1/2}} \geq \frac{\lambda_{\min}}{\lambda_{\max}} \quad (50)$$

where  $\lambda_{\min}$  and  $\lambda_{\max}$  are the minimum and maximum eigenvalues of  $G$ . When  $\cos \theta$  approaches zero, the step and the gradient become perpendicular, leading to slow convergence. In practice this may occur when the a gradient vector has major components associated with the large eigenvalues of  $G$ , while the step has major components elsewhere. Again the restricted step method, or the augmented Hessian method, which shift all  $\lambda$  equally, give a more favorable  $\lambda_{\min}/\lambda_{\max}$  ratio for descent.

In terms of the normal mode analysis the Taylor series expansion for the energy can be rewritten as

$$E(\mathbf{x}^1) = E(\mathbf{x}^0) - \alpha(1 - 1/2\alpha) \sum_i \eta_i^2 \lambda_i \quad (51a)$$

$$E(\mathbf{x}^1) = E(\mathbf{x}^0) - \alpha(1 - 1/2\alpha) \sum_i \epsilon_i^2 \lambda_i^{-1} \quad (51b)$$

Eqs. (51) demonstrate the need for  $G$  to be positive definite. We see explicitly from Eq. (51a) that for similar step displacements  $\eta_1$  the major energy contribution comes from those modes with large eigenvalues. Conversely, Eq. (51b) indicate the major energy contribution comes from the small eigenvalue modes when the gradient components  $\varepsilon_i$  are comparable. In an optimization, the large eigenvalue modes are minimized quickly at first, until all the gradient components  $\varepsilon_1$  become similar; then larger step displacements along softer modes occur. This is illustrated in the Results Section.

If  $\varepsilon_1 = 0$ , then there will be no displacements in these directions whether  $\lambda_1$  is zero or not. Since the energy is invariant to translations and rotations, the center of coordinates and moments about these coordinates will be conserved. In the former case it can be shown that  $\lambda_1 = 0$ ; in the case of rotation the corresponding  $\lambda_1$  may not equal zero except at the extreme point, but the corresponding  $\varepsilon_1 = 0$ . Since the energy is also invariant to symmetry operations that leave the Hamiltonian invariant, the gradient vector will have non-vanishing components only along totally symmetric modes, "A" (see [43] for a more detailed proof).

$$|g\rangle = \sum_{i \in A} \varepsilon_i |\omega_i\rangle \quad (52)$$

Since motion along totally symmetric modes cannot change the point group symmetry, this symmetry is preserved throughout the optimization. The exception is at the extreme points themselves, where a higher symmetry might be obtained.

Symmetry greatly reduces the number of natural parameters in the optimization problem. The number of parameters is simply the number of totally symmetric vectors in  $G$ . At the same time, if the local minimum being sought is at lower symmetry than the starting geometry it will not be found. For example, a planar molecule will remain planar throughout the optimization even if a non-planar structure is of lower energy.

## 7. COMPUTATIONAL PROCEDURES

All calculations reported are performed within the INDO/1 semi-empirical framework [14,15]. These methods have also been programmed for the Quantum Theory Project ab-initio programs, and similar conclusions are drawn.

The energy derivatives with respect to the nuclear coordinates are obtained from analytical formula. The "exact" second derivative matrix is obtained from the CPHF equations. For comparison, these results are compared with numerical second derivatives in which each coordinate  $x_i$  is symmetrically displaced in turn by  $\delta x_i$  to give the central difference second derivative

$$G_{ij} = \frac{\left[ \frac{\partial E}{\partial x_j} \right]_{x + \delta x_i} - \left[ \frac{\partial E}{\partial x_j} \right]_{x - \delta x_i}}{2\delta x_i} \quad (53)$$

Hence in a system of N-nuclei the formation of the numerical Hessian matrix requires 6N full SCF calculations, one for each coordinate displacement. Second derivatives are obtained much more rapidly through analytic expressions such as those of Eq. (40), than through the use of numerical expressions such as that of Eq. (53). A projection technique is used to eliminate from G the small non-zero eigenvalues associated with the rotation and translation normal modes [44]. The exact inverse Hessian matrix H is obtained by diagonalizing G and discarding those eigenvectors with zero eigenvalues when back transforming to H.

Selection of a convergence criterion in an optimization procedure is an important consideration. In theory, optimization is complete when  $\mathbf{g} = 0$ . The more practical alternative is to ask for a maximal gradient component, or gradient norm  $\langle \mathbf{g} | \mathbf{g} \rangle^{1/2}$  or step size  $\langle \delta | \delta \rangle^{1/2}$ , or for a minimum change in the energy, or for all of these. An assumption, often not true, is made that the norms of the gradient and step size monotonically decrease through the optimization.

In practice, when all cartesian gradients are less than  $10^{-3}$  atomic units, bond lengths are usually within  $\pm 0.01$  Å and bond angles within  $\pm 2^\circ$  of their final optimized values. Torsional dihedral angles require a much finer optimization. In some cases, large changes in coordinates can be driven along very soft modes by very small gradients, Eq. (48). The physical meaning of optimizing along such soft modes, however, is not clear. Its justification lies only in the aesthetics of the mathematics.

## 8. RESULTS

### 8.1 General Conclusions.

Table I summarizes results for an arbitrary starting and the final geometry for hydrogen peroxide. The starting geometry is not greatly different from the final geometry, the largest cartesian coordinate difference is approximately 0.02 Å along the z direction. The relatively small starting gradients, also given in Table I, reflect the closeness of the initial geometry to the optimized geometry.

A summary of the Newton iterations for hydrogen peroxide is given in Table II. We report the variation of the 6 non-zero eigenvalues of the exact Hessian matrix  $G$ , four of which transform as  $a_1$  irreducible representations and two as  $a_2$ . We see that the eigenvalues stay roughly constant throughout the optimization except for the smallest one which is initially negative indicating that we are more or less in the quadratic region of the energy surface. The eigenvalues also have a large range in values, the very small eigenvalue 0.0178 corresponds to a torsional mode of the hydrogens around the oxygen axis.

Table II also gives the scalar products  $\epsilon_i$  of the Hessian vibrational eigenvectors with the gradient  $|g\rangle$  and the step component  $\eta_i$  calculated by Eq. (52). For the Newton optimization

TABLE I:  $\text{H}_2\text{O}_2$  COORDINATES AND GRADIENTS.

---

a) Trial Coordinates (Å)			
	x	y	z
O	1.1527552	-0.1322834	0.1511810
O	-1.1527552	0.1322834	0.1511810
H	1.8897626	1.0771647	-1.1338576
H	-1.8897626	-1.0771647	-1.1338576
Initial Gradients (a.u.)			
	$g_x$	$g_y$	$g_z$
O	-0.0079031	0.0238236	-0.0226673
O	0.0079031	-0.0238236	-0.0226673
H	-0.0124657	-0.0210846	0.0226673
H	0.0124657	0.0210846	0.0226673
b) Final Geometry (Å)			
	x	y	z
O	1.1563486	-0.1356317	0.1695288
O	-1.1563486	0.1356377	0.1695288
H	1.8934178	1.0810208	-1.1522053
H	-1.8934178	-1.0810208	-1.1522053

---

TABLE II:  $\text{H}_2\text{O}_2$  - ALL QUANTITIES ARE IN ATOMIC UNITS.  
 $\lambda$ 'S ARE THE HESSIAN EIGENVALUES.

a) $\lambda$ Values						
	$a_1$	$a_1$	$a_1$	$a_1$	$a_2$	$a_2$
cycle	1	2	3	4	5	6
1	4.5053	2.1933	.1630	-.0178	2.3411	.2592
2	4.3564	2.0450	.1807	.0159	2.1600	.2818
3	4.3572	2.0334	.1788	.0181	2.1518	.2873
4	4.3586	2.0346	.1784	.0178	2.1538	.2874
5	4.3586	2.0346	.1784	.0178	2.1538	.2874

b) $\varepsilon_1 = \langle g   \omega_1 \rangle^a$ , Eq. (48a)				
	1	2	3	4
1	-4.804 D-3	6.702 D-2	-8.171 D-4	-3.186 D-4
2	9.044 D-4	2.131 D-3	-5.109 D-4	5.550 D-4
3	9.220 D-6	-6.721 D-4	4.507 D-5	5.067 D-5
4	9.406 D-7	-5.526 D-6	5.309 D-6	9.319 D-6
5	5.250 D-8	-3.705 D-7	8.705 D-7	1.522 D-6

c) $\eta_1 = \langle s   \omega_1 \rangle^a$ , Eq. (48b)				
	1	2	3	4
1	1.066 D-3	-3.055 D-2	5.014 D-3	-1.786 D-2
2	-2.076 D-4	-1.042 D-3	2.827 D-3	-3.493 D-2
3	-2.116 D-6	3.305 D-4	-2.521 D-4	-2.799 D-3
4	-2.158 D-7	2.176 D-6	-3.021 D-5	-5.227 D-4
5	-1.204 D-8	1.821 D-7	-4.896 D-6	-8.530 D-5

d) Summary of the Optimization			
	Energy (a.u.)	$ g $ (a.u./bohr)	Max $ g_i $
1	-34.828473	0.067195	0.023824
2	-34.829517	0.002435	0.000994
3	-34.829530	0.000676	0.000276
4	-34.829530	0.000012	0.000007

<sup>a</sup>All Computed  $|\varepsilon_i|$  and  $|\eta_i|$  along the  $a_2$  modes ( $i = 5$  and  $6$ ) are less than  $10^{-14}$  a.u.



all searches were successful with  $\alpha = 1$ . The scalar products  $\eta_1$  and  $\epsilon_1$  are essentially zero for the  $a_2$  vibrational modes as anticipated from symmetry conservation. Initially the largest gradient and step scalar products occur for the  $2a_1$  vibrational mode which corresponds to a symmetric OH stretching mode. Although the  $2a_1$  gradient component  $\epsilon_1$  remains the largest for the next two optimization cycles, the largest step component  $\eta_1$  occurs along the soft  $4a_1$  mode for the reasons discussed in the previous section.

Table III presents an analysis of the energy change between the second and third optimization cycles.  $\epsilon_i \eta_i = \langle s | \omega_i \rangle \langle \omega_i | g \rangle$  is the descent condition and should be negative for each "i." Here only the symmetric modes contribute (see Section 6). For this cycle the major decrease in energy is along the soft fourth mode. This is so even though the largest gradient component (Table II) remains along the second mode. There is from this specific analysis and the more general one of the previous section an apparent danger in using only the magnitude of the gradient as a check for convergence of the optimization procedure.

TABLE III: ENERGY DESCENT FOR  $H_2O_2$   
FROM CYCLE 2 TO CYCLE 3 ( $10^{-6}H$ ).

	1	2	3	4
$\langle s   \omega_i \rangle \langle \omega_i   G_2 \rangle^a$	-0.18775	-2.22050	-1.44431	-19.38615
$1/2 \lambda_i \langle s   \omega_i \rangle^2^b$	0.09388	1.11019	0.72207	9.69983
$\Delta E_i$	-0.09387	-1.10931	-0.72224	- 9.68632
Total:	-11.61174			

<sup>a</sup>This is the descent condition and should be negative.

<sup>b</sup>The contribution of the first two terms in the Taylor series expansion of the energy.

TABLE IV: GEOMETRY OPTIMIZATION OF  $\text{H}_2\text{O}_2$  USING THE BFGS ALGORITHM.

	$\lambda_1^a$	$\lambda_2$	$\lambda_3$	$\lambda_4$
1	1.0	1.0	1.0	1.0
2	2.1776	1.0	1.0	0.9849
4	3.9361	1.4179	1.0	0.9400
6	4.4429	2.0445	0.9996	0.8710
8	4.8876	2.3818	1.0324	0.1471
10	6.2775	2.6888	0.3846	0.0552
15	4.4232	2.1109	0.1948	0.0178
Exact	4.3586	2.0346	0.1784	0.0178

<sup>a</sup>Atomic units or Hartrees.

TABLE V: STEP SIZES (a.u.) IN THE BFGS OPTIMIZATION OF THE GEOMETRY OF  $\text{H}_2\text{O}_2$ .

Cycle/Mode	1	2	3	4
2	3.924 D-3	-----	-----	-6.861 D-3
4	1.707 D-3	4.171 D-3	-----	3.028 D-3
6	-1.194 D-5	1.200 D-4	-4.148 D-5	9.083 D-4
8	7.146 D-5	1.015 D-4	-2.094 D-4	-6.410 D-4
10	-5.193 D-5	-4.750 D-5	-5.344 D-4	-2.389 D-3
15	-1.012 D-5	-1.380 D-5	-3.654 D-5	2.907 D-4

A similar analysis for a BFGS quasi-Newton optimization of hydrogen peroxide is given in Tables IV and V starting with the same initial geometry *a* in Table I. Only the symmetric vibrational modes are included in Tables IV and V since only these are updated. A more detailed discussion of the Hessian eigenvalue evolution for the BFGS scheme is given below, but in Table IV we see that it takes over eight Hessian updates before the low frequency normal mode is realized. However the eighth cycle energy is only 2 $\mu\text{H}$  above the converged energy.

## 8.2 Newton's Method.

In this section, we present a summary of some Newton optimizations. The number of energy and gradient calculations required to reach the minimum forms a frame of reference for considering the efficiency of the different optimization

procedures. The molecules studied are  $\text{H}_2\text{CO}$ ,  $\text{NH}_3$  and  $\text{C}_2\text{H}_6$ . The initial geometry for each molecule has been chosen to have low symmetry ( $C_s$ ) by making a 0.08 a.u. coordinate displacement along each vibrational mode from the minimum. A second initial geometry for  $\text{NH}_3$  is considered with 0.05 a.u. displacements along each mode from the minimum.

The number of line searches, energy and gradient evaluations are given in Table VI for both the Newton and Quasi-Newton methods. Table VI clearly indicates that the use of an exact inverse Hessian requires less points to arrive at the optimum geometry. However in Table VI we have not included the relative computer times required to form the second derivative matrix. If this is taken into account, then the Newton's method with its requirement for an exact Hessian matrix is considerably slower than the quasi-Newton procedures.

TABLE VI: NUMBER OF OPTIMIZATION CYCLES REQUIRED FOR DIFFERENT UPDATE FORMULAE<sup>a</sup>.

	Exact	Rank 1	BFGS	DFP	Greenstadt
$\text{H}_2\text{CO}$	4(5,5)	9(11,10)	11(12,12)	12(13,13)	16(18,17) <sup>b</sup>
$\text{NH}_3^d$	6(8,7)	7(11, 8)	7( 8, 8)	11(12,12)	12(13,13) <sup>b</sup>
$\text{NH}_3^e$	3(4,4)	7( 9, 8)	7( 8, 8)	10(11,11)	11(13,12) <sup>b</sup>
$\text{C}_2\text{H}_6$	6(7,7)	17(19,18)	19(20,20)	29(30,30) <sup>c</sup>	19(27,20) <sup>b</sup>

<sup>a</sup>The notation x(y,z) refers to x line searches, y energy evolutions and z gradient evaluations.

<sup>b</sup>Energy changing less than 5.D-07 07 Ha but above true minimum energy value.

<sup>c</sup>Optimization not quite converged.

<sup>d</sup>Displaced 0.08 a.u. along each normal mode.

<sup>e</sup>Displaced 0.05 a.u. along each normal mode.

TABLE VII: EXACT HESSIAN OPTIMIZATION OF  $\text{H}_2\text{CO}$  - HESSIAN EIGENVALUES (a.u.).

1	5.281D+00	1.592D+00	7.879D-01	1.996D-01	1.922D-01	1.104D-01
2	4.209D+00	2.146D+00	1.060D+00	1.996D-01	1.906D-01	1.417D-01
3	4.249D+00	2.046D+00	1.023D+00	2.217D-01	1.847D-01	1.584D-01
4	4.279D+00	2.033D+00	1.023D+00	2.396D-01	1.830D-01	1.589D-01
5	4.279D+00	2.033D+00	1.023D+00	2.396D-01	1.830D-01	1.589D-01

Eval. No.		Energy	$ g $	Max $ g_i $	$\alpha$	Opt Type
E	G					
1	1	-24.316168	0.369646	0.264229	0.00000D+00	1
2	2	-24.336801	0.084631	0.066710	1.00000D+00	1
3	3	-24.339573	0.009807	0.007186	1.00000D+00	1
4	4	-24.339593	0.000050	0.000034	1.00000D+00	1
5	5	-24.339593	0.000000	0.000000	1.00000D+00	1

TABLE VIII: THE OPTIMIZATION  $\text{NH}_3$  WITH THE EXACT HESSIAN - HESSIAN EIGENVALUES (a.u.).

1	3.405D+00	2.209D+00	1.341D+00	1.684D-01	3.866D-02	-3.840D-03
3	3.179D+00	2.072D+00	1.431D+00	1.689D-01	8.842D-02	5.403D-02
4	2.440D+00	2.355D+00	1.354D+00	1.553D-01	1.492D-01	1.122D-01
5	2.216D+00	2.213D+00	1.468D+00	1.750D-01	1.567D-01	1.564D-01
6	2.270D+00	2.270D+00	1.454D+00	1.609D-01	1.543D-01	1.543D-01
7	2.274D+00	2.274D+00	1.451D+00	1.599D-01	1.543D-01	1.543D-01
8	2.274D+00	2.274D+00	1.451D+00	1.599D-01	1.543D-01	1.543D-01

Eval. No.		Energy	$ g $	Max $ g_i $	$\alpha$
E	G				
1	1	-12.499372	0.358164	0.198326	0.00000D+00
2		-12.492400			1.83049D-01
3	2	-12.502913	0.321531	0.176294	5.95025D-02
4	3	-12.521926	0.032392	0.021575	1.00000D+00
5	4	-12.523048	0.016203	0.012235	1.00000D+00
6	5	-12.523215	0.000832	0.000694	1.00000D+00
7	6	-12.523216	0.000012	0.000010	1.00000D+00
8	7	-12.523216	0.000000	0.000000	1.00000D+00

Tables VII and VIII examine the nature of the Hessian matrix for  $\text{H}_2\text{CO}$  and  $\text{NH}_3$  as the geometry varies during the optimization. In both cases the starting geometry is far from the quadratic region about the minima. After the first geometry update in  $\text{H}_2\text{CO}$  (cycle 2) the Hessian eigenvalues are within 10% of their final

value. The next step (cycle 3) shows gradients reduced by a factor of 10, and Hessian eigenvalues within 1% of their final values. A step with these values leads to a reduction of the gradient by a factor of 200. Similar conclusions can be reached for  $\text{NH}_3$  from the information in Table VIII. Highly accurate Hessian matrices at the early stages of an optimization are probably unnecessary. In the final stages of an optimization, when the quadratic region is being approached, an accurate Hessian has a big impact. This is also demonstrated in Table VI when comparing the optimization of  $\text{NH}_3$  with two different starting geometries. When the starting geometry is far enough removed from the neighborhood of the first geometry each successive step takes the molecule through regions of the energy hypersurface with quite different quadratic behavior, see below.

In the case of  $\text{NH}_3$ , Table VIII, the final Hessian shows two degenerate modes, a requirement of  $C_{3v}$  symmetry. The initial distorted geometry has quite different eigenvalues for the "parents" of these modes, and it is quite interesting to follow the evolution of the eigenvalues of these modes through the optimization procedure.

The starting geometry is poor enough to produce very large gradients and a negative eigenvalue in the Hessian. The step, with  $\alpha = 1$ , produced from this gradient is sufficiently large as to produce coordinate displacements larger than allowed by Eq. (15). The geometry and energy,  $E_2$ , obtained from the step scaled by the value 0.183 [from Eq. (15)] fail the right hand line search test. The step is then scaled by 0.0595 obtained from the quadratic interpolation formula (16a) leading to an energy  $E_3$  lower than the starting energy  $E_1$ . After this, optimization is rapid.

### 8.3 Quasi-Newton Methods with Unit Hessian.

Here we consider the relative efficiencies of the different inverse update formula. The calculations take the initial inverse Hessian as a unit matrix and the initial  $\alpha$  in the line

search as 0.4. Table VI provides a comparison of the number of line searches, energy and gradient evaluation for the different formula. The results of Table VI as well as studies on other systems, indicates that the rank one and BFGS methods are the most successful formula and are similar in their optimization efficiency for small molecules. Our experience and those of others [9] suggest that popular DFP formula may be quite successful with very exact line searches, although still not as good as the rank one and BFGS methods. The Greenstadt method was found not to converge for many of the systems studied.

A study of the eigenvalues of the updated Hessian are revealing. Tables IX, X, XI and XII report these values for  $\text{NH}_3$  with the initial displacement 0.08 a.u. along each normal mode as described previously. As expected, the rank one update summarized in Table IX updates one eigenvalue each cycle until all six modes are realized. The negative eigenvalue shown in the first mode on the fourth cycle is an example of a serious flaw. A step taken with this approximate inverse Hessian led to an increase in energy and a succession of right line search failures. The eighth evaluation is finally taken with a very small step,  $\alpha = 0.007$ . Most optimization, of course, do not work with the normal modes and do not diagonalize the Hessian as we have done here to examine the problem. If this were done, then this particular negative eigenvalue could have been set back to its previous value, or even back to unity. Working in normal modes might be desirable in ab-initio calculations, however, where this diagonalization, of order  $N^3$ , is a small step compared with the formation of the integrals and Fock matrices, both  $N^4$  steps, where  $N$  is the size of the basis.

The rank two BFGS and DFP formula form two non-unit eigenvalues on the first update, again reflecting the rank two nature of the update. Both formula, as expected, preserve the positive definite nature of the Hessian.

TABLE IX: RANK 1 UPDATE OPTIMIZATION OF  $\text{NH}_3$  - HESSIAN  
EIGENVALUES (a.u.).

1	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00
2	2.724D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00
3	3.227D+00	2.714D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00
4	-2.354D-02	2.717D+00	1.802D+00	1.000D+00	1.000D+00	1.689D-01
8	2.719D+00	2.331D+00	1.458D+00	1.000D+00	1.000D+00	1.689D-01
9	2.587D+00	2.303D+00	1.456D+00	1.000D+00	9.994D-01	1.581D-01
10	2.370D+00	2.267D+00	1.438D+00	9.999D-01	9.862D-01	1.552D-01
11	2.359D+00	2.262D+00	1.435D+00	9.997D-01	5.302D-01	1.506D-01

Exact

2.274D+00 2.274D+00 1.452D+00 1.600D-01 1.543D-01 1.543D-01

Eval. No.

E	G	Energy	$ g $	Max $ g_i $	$\alpha$
1	1	-12.499372	0.358164	0.198326	0.00000D+00
2	2	-12.521284	0.050917	0.031968	4.00000D-01
3	3	-12.522080	0.028010	0.017822	1.00000D+00
4	4	-12.522281	0.019552	0.012914	1.00000D+00
5		-12.476193			8.67976D-01
6		-12.520752			8.01833D-02
7		-12.522073			1.55110D-02
8	5	-12.522235	0.020010	0.012987	3.67105D-03
9	6	-12.523207	0.004431	0.002909	1.00000D+00
10	7	-12.523216	0.000612	0.000343	1.00000D+00
11	8	-12.523216	0.000086	0.000049	1.00000D+00

TABLE X: BFGS OPTIMIZATION OF  $\text{NH}_3$  - HESSIAN EIGENVALUES (a.u.).

1	1.00D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00
2	2.723D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	9.784D-01
3	3.719D+00	1.755D+00	1.000D+00	1.000D+00	1.000D+00	6.725D-01
4	2.836D+00	2.378D+00	1.032D+00	1.000D+00	1.000D+00	2.641D-01
5	2.785D+00	2.377D+00	1.527D+00	1.000D+00	1.000D+00	1.618D-01
6	2.577D+00	2.364D+00	1.514D+00	1.000D+00	9.994D-01	1.541D-01
7	2.423D+00	2.265D+00	1.476D+00	1.000D+00	9.952D-01	1.627D-01
9	2.434D+00	2.347D+00	1.472D+00	1.000D+00	9.254D-01	1.586D-01

Exact

2.274D+00 2.274D+00 1.452D+00 1.600D-01 1.543D-01 1.543D-01

Eval. No.

E	G	Energy	$ g $	Max $ g_i $	$\alpha$
1	1	-12.499372	0.358164	0.198326	0.00000D+00
2	2	-12.521284	0.050917	0.031968	4.00000D-01
3	3	-12.522072	0.028891	0.018206	1.00000D+00
4	4	-12.522491	0.019688	0.010447	1.00000D+00
5	5	-12.523027	0.016294	0.010447	1.00000D+00
6	6	-12.523199	0.007948	0.005281	1.00000D+00
7	7	-12.523215	0.000941	0.000621	1.00000D+00
8	8	-12.523216	0.000101	0.000059	1.00000D+00



TABLE XI: DFP UPDATE OPTIMIZATION OF  $\text{NH}_3$   
- HESSIAN EIGENVALUES (a.u.).

1	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00
2	2.732D+00	1.000D+00	1.000D+00	1.00D+00	1.000D+00	9.918D-01
3	2.721D+00	1.926D+00	1.000D+00	1.000D+00	1.000D+00	7.735D-01
4	4.004D+00	2.584D+00	1.112D+00	1.000D+00	1.000D+00	4.347D-01
5	5.743D+00	2.602D+00	1.285D+00	1.000D+00	1.000D+00	2.399D-01
6	6.785D+00	2.601D+00	1.294D+00	1.000D+00	9.999D-01	1.578D-01
7	5.846D+00	2.600D+00	1.310D+00	1.001D+00	9.999D-01	1.555D-01
8	3.321D+00	2.589D+00	1.328D+00	1.000D+00	9.999D-01	2.168D-01
9	2.620D+00	2.175D+00	1.277D+00	1.000D+00	9.953D-01	2.842D-01
10	2.629D+00	2.475D+00	1.215D+00	1.001D+00	9.884D-01	2.120D-01
11	2.814D+00	2.584D+00	1.242D+00	1.001D+00	9.838D-01	1.611D-01
12	2.638D+00	2.431D+00	1.258D+00	1.000D+00	9.828D-01	1.571D-01

Exact

2.274D+00 2.274D+00 1.452+00 1.600D-01 1.543D-01 1.543D-01

Eval. No.

E	G	Energy	$ g $	Max $ g_i $	$\alpha$
1	1	-12.499372	0.358164	0.198326	0.00000D+00
2	2	-12.521284	0.050917	0.031968	4.00000D-01
3	3	-12.522077	0.028333	0.017964	1.00000D+00
4	4	-12.522434	0.018602	0.013570	1.00000D+00
5	5	-12.522807	0.019927	0.009532	1.00000D+00
6	6	-12.523011	0.021627	0.012255	1.00000D+00
7	7	-12.523144	0.016718	0.009988	1.00000D+00
8	8	-12.523190	0.009809	0.005927	1.00000D+00
9	9	-12.523210	0.003410	0.001953	1.00000D+00
10	10	-12.523215	0.000453	0.000453	1.00000D+00
11	11	-12.523216	0.000173	0.000173	1.00000D+00
12	12	-12.523216	0.000096	0.000096	1.00000D+00

TABLE XII: GREENSTADT UPDATE IN THE OPTIMIZATION OF  $\text{NH}_3$  -  
HESSIAN EIGENVALUES (a.u.).

1	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00
2	2.724D+00	1.000D+00	1.000D+00	1.000D+00	1.000D+00	9.969D-01
3	2.772D+00	1.758D+00	1.000D+00	1.000D+00	1.000D+00	8.639D-01
4	1.025D+01	2.304D+00	1.034D+00	1.000D+00	1.000D+00	6.917D-01
5	-1.302D+01	2.313D+00	1.142D+00	1.000D+00	9.999D-01	5.936D-01
6	2.892D+00	2.269D+00	1.111D+00	1.000D+00	9.998D-01	7.870D-01
7	-1.729D+02	2.339D+00	1.000D+00	1.000D+00	9.998D-01	3.647D-01
8	1.060D+01	2.337D+00	1.117D+00	1.000D+00	9.998D+00	4.593D-01
9	3.653D+00	2.342D+00	1.131D+00	1.000D+00	9.998D-01	4.937D-01
10	7.378D+00	2.309D+00	1.101D+00	1.001D+00	9.994D-01	4.248D-01
11	-3.148D+02	2.320D+00	1.096D+00	1.001D+00	9.998D-01	3.999D-01
12	-4.799D+01	2.346D+00	1.110D+00	1.000D+00	9.898D-01	3.974D-01
13	-4.799D-02	6.525D+01	2.183D+00	1.001D+00	4.0588-01	2.509D-01

Exact

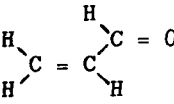

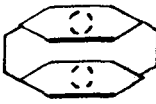
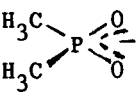
2.274D+00 2.2274D+00 1.452D+00 1.600D-01 1.543D-01 1.543D-01

Eval. No.		Energy	$ g_i $	Max $ g_i $	$\alpha$
E	G				
1	1	-12.499372	0.358164	0.198326	0.00000D+00
2	2	-12.521284	0.050917	0.031968	4.00000D-01
3	3	-12.522079	0.028129	0.017874	1.00000D+00
4	4	-12.522402	0.019060	0.013480	1.00000D+00
5	5	-12.522565	0.023526	0.011952	1.00000D+00
6	6	-12.522591	0.019286	0.010941	1.00000D+00
7	7	-12.522774	0.014169	0.007835	1.00000D+00
8	8	-12.522889	0.029717	0.021467	1.00000D+00
9	9	-12.522968	0.018153	0.011784	1.00000D+00
10	10	-12.523059	0.011297	0.006420	1.00000D+00
11	11	-12.523154	0.009138	0.005761	1.00000D+00
12	12	-12.523155	0.010399	0.006741	1.00000D+00

We notice that even at optimization, none of these procedures yield particularly accurate Hessian matrices. For most of the systems we have examined, however, the BFGS method is most accurate, followed by the rank one method. Least accurate is the Greenstadt method. The modes least accurately represented are the softer modes. If this system had been given an initial geometry that had  $C_{3v}$  symmetry, then only the  $2a_1$  modes 3 and 4 of the tables, would be updated.

Summarizing, the BFGS and rank one formula appear to be the most successful for updates, at least for small molecules. The rank one formula, however, has a serious disadvantage of not preserving the positive definiteness of the Hessian.

TABLE XIII: A COMPARISON OF THE MURTAGH-SARGENT (MS) PROCEDURE WITH A BROYDE-FANNO-GOLDFARB-SHANNO (BFGS) PROCEDURE. THE NOTATION IS THE SAME AS THAT OF TABLE VI.

System	Max. Comp. Grad. (a.u./bohr)	MS	BFGS
	0.20	>50	24(27,26)
	0.39	11(14,13)	10(11,11)
	0.16	16(20,18)	10(11,11)
	0.15	>50	43(44,44)

Larger molecules have very many soft modes and it is often difficult to optimize their structures. For these systems, however, the BFGS formula has a strong advantage over the rank methods that we have examined. Some examples are presented in Table XIII. Pyrimidine and 2,2' paracyclophane are more "typical" examples. The number of SCF cycles in these two cases are 14 versus 11 and 20 versus 11 for the Murtagh-Sargent (MS) versus BFGS methods, respectively. The other two molecules in this study were not optimized in 50 searches by the rank one MS

method, partially a consequence of our initial distortions that destroy all symmetry. In the  $P(O)_2(CH_3)_2$  case, for example, the rotations of the two  $CH_3$  groups are soft and very strongly coupled.

#### 8.4 Quasi-Newton Methods with Non-Unit Starting Hessians.

We consider starting the optimization procedure with a more accurate Hessian than the unit matrix described in the previous section. Three approximations are suggested. The first, not examined here, is to have a table of force constants between bonded atoms for "typical" situations, and to form the starting Hessian from this table [27,28]. The second is to start the optimization with an exact Hessian. Although a time-consuming process, the numbers of subsequent geometry searches, as summarized in Table XIV, is nearly halved in the case of  $H_2CO$ , and reduced from 8 to 6 in the case of  $NH_3$ . An approximate Hessian can also be formed in an SCF procedure by displacing the coordinates as described in Section VII, but freezing the density at the initial geometry. This procedure avoids 6N SCF calculations although new integrals involving the displaced atom must be calculated at each displacement. Starting with this approximate Hessian does not show a great advantage over starting with the unit matrix, as typified in Table XIV. Since it is more time-consuming than the latter, it is not worthwhile.

Starting with an exact Hessian matrix (or one formed from a Table of force constants) at an arbitrary setting geometry is of questionable value. If the starting geometry is sufficiently near the optimized geometry, and "quadratically related" to it, an exact Hessian is advantageous ( $H_2CO$  of Table XIV) and may allow optimization along the soft modes not easily obtained otherwise. Far from the sought structure, this time-consuming procedure is not worthwhile, for the Hessian itself is changing too rapidly through the optimization. What might be desirable is a better (more rapidly obtained) approximate Hessian used until the optimization had led to a geometry in the neighborhood of the

minimum and then a stable update procedure such as BFGS employed. Tables XV and XVI follow the course of an optimization of  $\text{NH}_3$  utilizing an exact inverse Hessian and an approximate one, respectively, and should be compared with Tables VIII - XII. For ab-initio calculations such an approximate Hessian might be effectively obtained using analytic expressions from the much faster semi-empirical methods. In addition, large basis set ab-initio calculations might have the Hessian matrix estimated at the minimum basis set level.

TABLE XIV: NUMBER OF BFGS OPTIMIZATION CYCLES USING DIFFERENT INITIAL HESSIAN MATRICES.

	Unit H	Approx. H	Exact H	Newton
$\text{H}_2\text{CO}$	11(12,12)	9(10,10)	6(7,7)	4(5,5)
$\text{NH}_3^a$	7( 8, 8)	6( 7, 7)	5(6,6)	3(4,4)

<sup>a</sup>Displaced 0.05 a.u. along each normal mode.

TABLE XV: BFGS OPTIMIZATION OF  $\text{MH}_3$  WITH INITIAL EXACT INVERSE HESSIAN - HESSIAN EIGENVALUES (a.u.).

1	2.907D+00	2.238D+00	1.409D+00	1.634D-01	9.583D-02	7.978D-02
2	2.522D+00	2.223D+00	1.422D+00	1.633D-01	9.628D-02	8.576D-02
3	2.401D+00	2.256D+00	1.454D+00	1.633D-01	1.495D-01	9.585D-02
4	2.431D+00	2.256D+00	1.454D+00	1.633D-01	1.495D-01	9.585D-02
5	2.377D+00	2.228D+00	1.457D+00	1.638D-01	1.595D-01	9.687D-02
6	2.331D+00	2.257D+00	1.447D+00	1.633D-01	1.561D-01	1.199D-01

Exact

2.274D+00 2.274D+00 1.451D+00 1.599D-01 1.543D-01 1.543D-01

Eval. No.

E	G	Energy	$ g $	Max $ g_i $	$\alpha$
1	1	-12.514386	0.204926	0.112413	0.00000D+00
2	2	-12.522959	0.024481	0.014681	1.00000D+00
3	3	-12.523022	0.013808	0.009355	1.00000D+00
1	4	-12.523215	0.001333	0.000805	1.00000D+00
5	5	-12.523216	0.000101	0.000066	1.00000D+00
6	6	-12.523216	0.000018	0.000012	1.00000D+00

TABLE XVI: BFGS OPTIMIZATION OF  $\text{NH}_3$  WITH INITIAL APPROXIMATE INVERSE HESSIAN - HESSIAN EIGENVALUES (a.u.)

1	3.173D+00	2.493D+00	1.668D+00	3.570D-01	2/244D-01	1.760D-01
2	2.804D+00	2.477D+00	1.598D+00	3.570D-01	2.260D-01	1.765D-01
3	2.905D+00	2.498D+00	1.462D+00	2.621D-01	2.277D-01	1.764D-01
4	3.240D+00	2.505D+00	1.564D+00	2.285D-01	1.801D-01	1.761D-01
5	2.732D+00	2.469D+00	1.540D+00	2.284D-01	1.765D-01	1.587D-01
6	2.559D+00	2.359D+00	1.498D+00	2.282D-01	1.765D-01	1.672D-01
7	2.512D+00	2.335D+00	1.550D+00	2.034D-01	1.766D-01	1.659D-01
Exact						
	2.274D+00	2.274D+00	1.451D+00	1.599D-01	1.543D-01	1.543D-01
Eval. No.						
E	G	Energy	$ g $	Max $ g_i $	$\alpha$	
1	1	-12.514386	0.204926	0.112413	0.00000D+00	
2	2	-12.522727	0.033809	0.016685	1.00000D+00	
3	3	-12.523121	0.007852	0.006127	1.00000D+00	
1	4	-12.523202	0.003959	0.002449	1.00000D+00	
5	5	-12.523216	0.002398	0.001415	1.00000D+00	
6	6	-12.523216	0.000363	0.000239	1.00000D+00	
7	7	-12.523216	0.000035	0.000027	1.00000D+00	

### 8.5 Approximate Hessians

As seen in Table XV, the use of the V(1) approximate Hessians, Eq. (38), is remarkable. For all practical purposes the number of geometry searches required for optimization using V(1) is about the same number as that required using the exact Hessian. For semi-empirical methods there is an impressive savings of computer time over both the BFGS procedure, and the exact evaluation of the Hessian. For large systems the calculation of V(1) is still practical, although the exact Hessian is not. [2,2'] paracyclophane optimizes in but 3 complete SCF calculations using the V(1) approximation, to be compared with the 11 BFGS SCF cycles (and 10 searches) reported in Table XIII.

For ab-initio methods the use of V(1) might be expected to lead to a savings of a factor of two - four in computer time for large systems for the reasons previously discussed. Although the

use of the V(1) method is of advantage over the use of the exact Hessian for geometry searches, the BFGS method is likely still to be much faster overall except in pathological cases.

#### 8.6 Augmented Hessian.

We [44], as well as others [40,41], have found the augmented Hessian or related restricted step methods of great use for determining transition states in molecules. For minima, this method is most useful when an accurate estimate of the second derivative matrix is available and negative eigenvalues occur in it. Otherwise, use of this method slows down the optimization as shown in Table XVII. The reason for this is the restricted nature of the step caused by the denominator shift in Eq. (45d). The augmented Hessian method is very successful with the V(1) estimate of the Hessian, as indicated in this Table. The augmented Hessian is not, in general, reliable with any of the update procedures we have examined.

TABLE XVII: A SUMMARY OF SOME OPTIMIZATIONS.

For each type of optimization the number of cycles is given. The number in paranthesis is time (min:sec) on a VAX11/780, and because of system load, are in possible error of 20%. The test calculations are INDO SCF.

System	NH <sub>3</sub>	CH <sub>3</sub> OH	Cis 1,2-Br,F Cyclopropane
Initial <sup>a</sup>			
Grad Norm	0.169	0.160	0.143
BFGS, Eq. 31	8	25(4:45)	2(10:24)
Newton Method <sup>b</sup>			
Approx. Hessian,	6	4(1:42)	9(5:46)
Aug. Approx.			
Hessian, Eq. 45	5	9(2:28)	17(10:30)
Newton Method,			
Exact Hessian,	3	7(4:30)	8(37:32)
Aug. Hessian,			
Exact Hessian Eq. 45	3	9(5:28)	17(84:42)

<sup>a</sup>Initially all symmetry is broken. For convergence all gradients have been reduced to below  $5 \times 10^{-4}$  a.u./bohr.

<sup>b</sup>The approximate Hessian is formed from the V(1) method of Eq. 38.

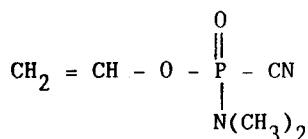
A more interesting use of the augmented Hessian is illustrated for HCN. A transition state for the HCN to HNC has been found. The initial geometry for HCN is then formed by displacing the coordinates 0.05 a.u. along the negative eigenvalue normal mode. The BFGS method with unit starting Hessian is slow to find the minimum with this starting conformation, requiring 12 (17, 14) or 12 line search, 17 energy evaluations and 14 gradients evaluations. The augmented Hessian approach required 8 (12,11), representing a considerable reduction of effort. Both methods insure the positive definiteness of the Hessian throughout the cycling - as evidenced that both obtained the minimum - but clearly the extra curvature introduced by the augmented Hessian, Eqs. (45), has helped.

## 9. CONCLUSIONS

In this paper we have reviewed and studied several Hessian update procedures for optimizing the geometries of polyatomic molecules. Of the several methods we have studied, many of which we have reported on here in detail, the most reliable and efficient is the BFGS. The weak line search outlined in this work is sufficient to ensure successful optimization even when starting geometries are generated through molecular mechanics [45] or ball and stick models [46].

Although the number of iterations in the optimizations procedure can be reduced by starting update procedures with an exact Hessian or Newton step, the time involved in forming this Hessian is seldom worth the effort unless the gradients are sufficiently small as to suggest that the current geometry is close to the optimized geometry. At this stage in the optimization, a single Newton step can save considerable effort. These findings are dramatically summarized for





in Table XVIII, a case with many soft modes. It should be noted, however, for reference that either the rank one or BFGS methods after 10 cycles have found all bond lengths to within  $\pm 0.02$  a.u. of their final value, and all bond angles to within  $\pm 3^\circ$ . The difficult modes involve optimization of dihedral angles that are coupled with pseudo-rotations.

TABLE XVIII: OPTIMIZATION OF  $\text{CH}_2 = \text{CH} - \text{O} - \text{P}(\text{O})(\text{N}(\text{CH}_3)_2) - \text{CN}$ , a.u.

Murtagh-Sargent Rank 1			BFGS		BFGS + 10th Cycle Exact Hessian	
	$\Delta E$ (a.u.)	$ g $		$\Delta E$ (a.u.)	$ g $	
1	0.0163	0.0050		0.0163	0.0050	
5	0.0149	0.0026		0.0134	0.0030	
10	0.0141	0.0024		0.0093	0.0046	
11	---	---		---	---	
15	0.0109	0.0024		0.0032	0.0030	
25	0.0101	0.0020		0.0002	0.0026	
					0.0000	$6 \times 10^{-6}$

In addition to examining Newton steps, we have also examined methods of the restricted step or augmented Hessian type. Although of enormous utility in searching for transition states, the very nature of their construction makes these methods not effective for finding minima unless the starting geometries have negative Hessian eigenvalues. In such a case augmented Hessian steps could be determined until the second (non-zero) eigenvalues of the augmented Hessian is positive, and subsequent steps determined using the Newton methods outlined.

We have also examined here the use of approximate solutions of the coupled perturbed Hartree-Fock equations for estimating the Hessian matrix. This Hessian appears to be more accurate than any updated Hessian we have been able to generate during the normal course of an optimization (usually the structure has optimized to within the specified tolerance before the Hessian is very accurate). For semi-empirical methods the use of this approximation in a Newton-like algorithm for minima appears optimal as demonstrated in Table 17. In ab-initio methods searching for minima, the BFGS procedure we describe is the best compromise.

# APPENDIX

Update formula for the direct Hessian G.

Rank 1

$$G^{k+1} = G^k + \frac{|\gamma - G\delta\rangle\langle\gamma - G\delta|}{\langle\gamma - G\delta|\delta\rangle} \quad (\text{A.1})$$

corresponding to Eq. (29).

BFGS, corresponding to Eq. (31)

$$G_{\text{BFGS}}^{k+1} = G^k + \frac{|\gamma\rangle\langle\gamma|}{\langle\gamma|\delta\rangle} - \frac{G|\delta\rangle\langle\delta|G}{\langle\delta|G|\delta\rangle} \quad (\text{A.2})$$

DFP, corresponding to Eq. (30)

$$G_{\text{DFP}}^{k+1} = G^k + \left(1 + \frac{\langle\delta|G|\delta\rangle}{\langle\gamma|\delta\rangle}\right) \frac{|\gamma\rangle\langle\gamma|}{\langle\gamma|\delta\rangle} - \left(\frac{|\gamma\rangle\langle\delta|G + G|\delta\rangle\langle\gamma|}{\langle\gamma|\delta\rangle}\right) \quad (\text{A.3})$$

Note that at each cycle of the update,  $G^{k+1} H^{k+1} = H^{k+1}$   
 $G^{k+1} = 1$  providing the same update is used for  $G^k$  and  $H^k$ .

## ACKNOWLEDGEMENTS

This work is supported in part through a grant from the Division of Sponsored Research at the University of Florida, and through a grant from the US Air Force (Tyndall) F08635-83-C0136.

## REFERENCES

1. P. Pulay, in Modern Theoretical Chemistry, Vol. 4, ed. H.F. Schaefer III (Plenum Press, New York, 1977), p. 153.
2. H.B. Schlegel, Adv. in Chemical Physics, 67, 249 (1987).
3. Geometric Derivatives of Energy Surfaces and Molecular Properties, ed. P. Jorgensen and J. Simons, NATO ASI series C, Vol. 166 Reidel, Boston (1985).
4. P. Pulay, Mol. Phys. 17, 197 (1969); 18, 473 (1970).
5. I.W. McIver, Jr. and A. Komornicki, Chem. Phys. Lett. 10, 303 (1971).
6. H.B. Schlegel, S.L. Wolfe, J. Chem. Phys. 63, 3632 (1975); 67, 4181 (1977); 67, 4194 (1977).
7. H.B. Schlegel, J. Chem. Phys. 77, 3676 (1982).
8. M.C. Zerner, in Computational Methods for Molecular Structure Determination: Theory and Technique, NRCC Proceedings No. 8.
9. a) R. Fletcher, in Practical Method of Optimization, Vol. 1, (Wiley, New York, 1980);  
b) P.E. Gill, W. Murray and M.H. Wright, in Practical Optimization, (Academic Press, New York, 1981).
10. a) W. Spendley, G.R. Hext and F.R. Himsforth, Technometrics 4, 441 (1962);  
b) J.A. Nelder and R. Mead, Comp. Jour. 7, 308 (1965).
11. J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, Int. J. Quantum Chem. S13, 225 (1979).
12. a) K. Muller, Angew. Chem. 19, 1 (1980);  
b) S. Bell and J.S. Crighton, J. Chem. Phys. 80, 2464 (1984).

13. a) J. Simons, P. Jørgensen, H. Taylor and J. Ozment, J. Phys. Chem. 87, 2745 (1983).  
b) A. Banerjee, N. Adams, J. Simons and R. Shepard, J. Phys. Chem. 89, 52 (1985).
14. J.A. Pople, D.L. Beveridge and P.A. Dobosh, J. Chem. Phys. 47, 2026 (1967).
15. A.D. Bacon and M.C. Zerner, Theoret. Chim. Acta, 53, 21 (1979).
16. G. Fogarsí and P. Pulay, J. Mol. Struct. 39, 275 (1977).
17. E.B. Wilson, Jr., J.C. Decius and P.C. Cross, in Molecular Vibrations, (McGraw-Hill, New York, 1955).
18. J.N. Murrell and K.J. Laidler, Trans. Faraday Soc. 64, 371 (1968).
19. A.A. Goldstein, SIAM J. Control 3, 147 (1965).
20. M.J.D. Powell, in Optimization in Action, ed. L.C.W. Dixon (Academic Press, New York, 1976), p. 117.
21. P. Scharfenberg, Theor. Chim. Acta 53, 279 (1979).
22. C.G. Broyden, Maths. Comput. 21, 368 (1967).
23. W.C. Davidon, Computer J. 10, 406 (1968).
24. A.V. Fiacco and G.P. McCormick, in Non-linear Programming, (Wiley, New York, 1968).
25. B.A. Murtagh and R.W.H. Sargent, in Optimization, ed. R. Fletcher, (Academic Press, London, 1969).
26. For example, the ab-initio packages GAUSSIAN 80, HONDO, TEXAS and GAMES.
27. H.B. Schlegel, Theoret. Chim. Acta 66, 333 (1984).
28. R.M. Badger, J. Chem. Phys. 2, 128 (1934); 3, 227 (1935).
29. B.A. Murtagh and R.W.H. Sargent, Comput. J. 13, 185 (1970).
30. a) W.C. Davidon, AEC Res. and Dev. Report ANL-5990 (1959);  
b) R. Fletcher and M.J.D. Powell, Comput. J. 6, 163 (1963).
31. a) C.G. Broyden, J. Inst. Maths Applns. 6, 76, 222 (1970);  
b) R. Fletcher, Coput. J. 13, 317 (1970);  
c) D. Godfarb, Maths. Comput. 24, 23 (1970);  
d) D.F. Shanno, Maths. Comput. 24, 647 (1970).

32. C.G. Broyden, Maths. Comput. 21, 368 (1967).
33. J.C. Greenstadt, Maths. Comput. 24, 1 (1970).
34. a) R.M. Steens, R. Pitzer and W.N. Libscomb, J. Chem. Phys. 38, 550 (1963).  
b) J. Gerratt and I.M. Mills, J. Chem. Phys. 49, 1719 (1968); 49, 1730 (1968).  
c) T.C. Coves and M. Karplus, J. Chem. Phys. 50, 3649 (1969).
35. J.D. Head and M.C. Zerner, Chem. Phys. Letters 131, 359 (1986).
36. Y. Osamura, Y. Yamaguchi, P. Saxe, M.A. Vincent, J.F. Gaw and H.F. Schaeffer III, Chem. Phys. 72, 131 (1982).
37. K. Levenberg, Quant. Appl. Math. 2, 164 (1944).
38. D.W. Marquardt, SIAM J. 11, 431 (1963).
39. S.M. Goldfeld, R.E. Quandt and H.F. Trotter, Econometrica 34, 541 (1966).
40. C.J. Cerjan and W.M. Miller, J. Chem. Phys., 75, 2800 (1981).
41. J. Simons, P. Jorgensen, H. Taylor and J. Ozment, J. Phys. Chem. 87 2745 (1983); D. O'Neal, H. Taylor and J. Simons, J. Phys. Chem., 88, 1510 (1984); H. Taylor and J. Simons, J. Phys. Chem., 89, 52 (1985); D.T. Nguyen and D.A. Case, J. Phys. Chem., 89, 4020 (1985).
42. R. Fletcher and C.M. Reeves, Comput. J. 1, 149 (1964).
43. P. Pechukas, J. Chem. Phys. 64, 1516 (1976).
44. J.D. Head and M.C. Zerner, to be published.
45. N.L. Allinger, J. Amer. Chem. Soc. 99 8127 (1977).
46. D.G. Purvis, QUIPU Program System, Quantum Theory Project, University of Florida.

# USE OF CLUSTER EXPANSION METHODS IN THE OPEN-SHELL CORRELATION PROBLEM

Debashis Mukherjee

Theory Group, Department of Physical Chemistry  
Indian Association for the Cultivation of Science  
Calcutta 700 032, India

Sourav Pal

Physical Chemistry Division  
National Chemical Laboratory  
Pune 411 008, India

1. INTRODUCTION
2. 'THEORETICAL MODEL CHEMISTRY' AND ITS RELEVANCE  
TO THE OPEN-SHELL FORMALISMS
3. CLUSTER EXPANSION AND SIZE-EXTENSIVITY
4. CATEGORIZATION OF THE OPEN-SHELL MANY-BODY  
FORMALISMS
  - 4.1 The Earlier Open-Shell Perturbative  
Developments
  - 4.2 The Propagator Methods
  - 4.3 The Open-Shell Cluster Expansion  
Approach
  - 4.4 Certain General Comments
5. SEMI-CLUSTER EXPANSION THEORIES FOR THE OPEN-SHELL  
STATES
  - 5.1 General Considerations
  - 5.2 Variational Methods
  - 5.3 Non-variational Methods
6. FULL CLUSTER EXPANSION THEORIES IN HILBERT SPACE
  - 6.1 General Preliminaries
  - 6.2 Single Root Hilbert Space Formalisms
  - 6.3 Multiroot Hilbert Space Formalisms
7. FULL CLUSTER EXPANSION THEORIES IN FOCK SPACE
  - 7.1 Preliminaries for a Fock Space Approach

- 7.2 Similarity Transformation-Based Fock Space Theories
- 7.3 Bloch Equation-Based Fock Space Theories
- 7.4 Variational CC Theory in Fock Space
- 8. SIZE-EXTENSIVE FORMULATIONS WITH INCOMPLETE MODEL SPACE (IMS)
  - 8.1 Preliminaries
  - 8.2 General Formulation
  - 8.3 The Case of IMS having valence holes and particles
  - 8.4 Applications of IMS-based CC Theories
  - 8.5 Size-extensivity of energies with IMS
- 9. CONCLUDING REMARKS
- ACKNOWLEDGEMENTS
- REFERENCES

## 1. INTRODUCTION

Our aim in this review is to highlight mainly the formal developments of the open-shell many-body cluster expansion theories that have taken shape over the last one and a half decades. We have now come to believe that a proper understanding and interpretation of most of the spectroscopic and dynamical phenomena involving excited or ionized molecules and molecular fragments require models for open-shell states transcending the simple orbital description. The language of occupation number representation - coupled with the hole-particle description of the operators, the use of Wick's theorem and the attendant diagrammatic depiction of the various terms/1/- makes the many-body approach a very attractive alternative to the traditional CI-based theories. The open-shell many-body perturbation theory(MBPT)/2-9/ and the propagator or the Green's function methods/10-16/ have now developed into vast computational technology, generating softwares which are cost-effective and competitive with the CI-based codes. The cluster expansion based open-shell many-body formalisms possess even greater potentiality since they not only share and embed all the desirable features and technical advantages of the open-shell MBPT or propagator theories, but also are inherently non-perturbative in nature. They provide more compact descriptions of the electron correlation and are more flexible. Moreover, the cluster expansion

of the open-shell functions provides us with valuable insight regarding the structure of general many-electron wave-functions. In fact, over the last few years there have been very fruitful developments /91-96,138-144/ on the size-extensivity of approximate functions (which generate energies scaling properly with the size of the molecules ) from a careful analysis of the structure of the cluster expansion for an open-shell function of arbitrary complexity, which have as yet no counterparts in the other many-body branches.

The developments of the cluster expansion theories appear to have reached a stage where a clear perspective is beginning to emerge, although no comprehensive review of the various facets of the approach and a critical evaluation of the seemingly disparate formalisms put forward is available in the literature. There are, however, several reviews on closed-shell coupled cluster theories where the open-shell cluster expansion theories are also touched upon/18,19,21,22/. A few reviews on the open-shell MBPT describe in broad terms the cluster expansion techniques in so far as they relate to MBPT /20,23/. A concise survey of what we shall call 'full cluster expansion theories' appears in a recent article by Lindgren and Mukherjee/94(a)/.

We shall try to present the various theories from a general and unified point of view. In our exposition, we shall not necessarily follow the chronological order of the developments. Rather, we shall follow a logical course which will bring out the evolution of the cluster expansion strategy, gradually encompassing finer aspects of the electron correlation that are peculiar to the open-shell states. This attitude leads to a natural subdivision of the currently available open-shell cluster-expansion formalisms into various categories. The historical order will be maintained as far as is possible in the delineation of each category.

## 2. 'THEORETICAL MODEL CHEMISTRY' AND ITS RELEVANCE TO THE OPEN-SHELL FORMALISMS

Before describing the unifying theme, which will classify the various theoretical developments in the cluster expansion formalisms, it is pertinent to summarize first certain essential criteria that an



acceptable theoretical model, describing large systems like molecules of interest in chemistry, should satisfy. In the modern many-body parlance the necessity of maintaining certain qualities of the wave-function satisfying these criteria has been emphasized by Bartlett *et al*/17,18/, Pople *et al*/24/ and Kutzelnigg/46/, although an awareness about these aspects can be traced back to such early practitioners as Bruckner/25/, Goldstone/26/, Hubbard and Hugenholtz/27/, Sinanoglu/28/, Primas/29/, and Coester/30/.

A 'Theoretical Model Chemistry'/18,24/ has an underlying approximation scheme which is of potentially uniform precision for all the phenomena it wants to describe, so that the viability of the model can be tested by an appeal to the experimental results. The criteria that a theoretical model should satisfy are the following :

- (a) Size-extensivity, i.e., the calculated energies should scale properly with the size of the molecule (or, equivalently, with the number of electrons)
- (b) Generality, i.e., the model should not depend for its validity on some special configurations or their symmetry.
- (c) The model should respect the proper space and spin symmetry required by the symmetry of the system.
- (d) The model should guarantee separability of the molecule into proper fragments if the situation so demands.
- (e) The model should be versatile enough, so that it applies equally well to ground, excited, ionized or open-shell states.

We may recall that the desirability of ensuring size-extensivity for a closed-shell state was one of the principal motivations behind the formulation of the MBPT for the closed-shells. The 'linked cluster theorem' of Bruckner/25/, Goldstone/26/ and Hubbard/27/, proving that each term in the perturbation series for energy can be represented by a 'linked' (connected) diagram directly reflects the size-extensivity of the theory. Hubbard/27/ and Coester/30/ even pointed out immediately after the inception of MBPT/25,26/, that the size-extensivity is intimately related to a cluster expansion structure of the associated wave-operator that is not just confined only to perturbative theory. The corresponding non-perturbative scheme for the closed-shells was first described by Coester and Kummel/30,31/ in nuclear physics and this was transcribed to quantum chemistry

by Cizek/32/, and then has been further developed by Paldus, Bartlett/18,21(a)/ and others/33-37/. This formalism, called the coupled-cluster (CC) method for closed shells, is now widely accepted as the most versatile and powerful electron correlation theory for closed shell systems such as the molecular closed shell ground state/18,21/.

As Bartlett/18/ has pointed out, for a closed shell single determinant Hartree-Fock (HF) function as the starting point for MBPT or CC theory, separability into neutral fragments is not automatically guaranteed, since sometimes the UHF rather than the RHF dissociates properly. When the criteria (a) and (d) above are both satisfied, the model may be called size-consistent/24/.

Although the closed shell CC theory has been sometimes used to describe some special open-shell situations/39,40/ (like single configuration triplets/39/) or dissociating species/40-42/, from the point of view of generality, and on physical grounds it is natural to look for open-shell generalizations of the cluster expansion formalisms as developed for the closed shells.

The criteria (a) to (e) noted above pose special problems for the open-shell states, having no corresponding analogues for the closed-shell situation. In any theoretical model for the open-shell situation, one first identifies a conceptually minimum set of determinants which are essential to describe the basic physico-chemical properties which the model sets out to describe (such as space-spin symmetry, treatment of quasidegeneracy, separability into proper fragments). The determinants of this starting or 'reference' function may be said to span a 'model' space/3/. Correlation is brought in by superposing other configurations belonging to the 'virtual' space. It is customary to designate the orbitals which are completely filled in all the model space functions as 'core' orbitals, those which are partially occupied (i.e., occupied in some model space functions and unoccupied in others) as the 'valence' orbitals. The rest of the orbitals are called 'particles'. The above nomenclature owes its origin to the early developments of the open-shell MBPT (see /2-9,20/ for extensive reviews), and forms a convenient starting point for all the later modern open-shell formalisms that have followed. The total number of electrons in the model space functions is more than those in the core, according to this scheme. Sometimes it is more convenient to take a core having more electrons than

in the model space functions, and vacant core orbitals are interpreted as 'valence holes' in the core. We shall, at the moment, side-step this complication. Most of the discussions to follow remain valid for this latter situation as well.

The unifying theme for describing electron correlation in the open-shell states may now be laid down. We wish to look upon the overall effect of electron correlation as stemming from additive contributions from the correlation among the core-electrons, the correlation among the valence electrons and the interactions of core and valence electrons leading to core-polarization and core-relaxation effects. The various open-shell developments to be described emphasize these various effects in various degrees. More importantly, we can envisage several different, although inter-related, aspects of the size-extensivity requirements which are special to the open-shell situations. They define different models for the open-shell many-body formalisms:

(a1) Core-extensive models : The model is explicitly size-extensive with respect to the core electrons  $N_c$  only. This is quite useful for a good description of electron correlation for a series of open-shell systems with a fixed number  $N_v$  of the valence electrons, when  $N_c \gg N_v$ . The bulk of the size-extensivity then comes from the size-extensivity of the core-correlation. The valence correlation and the core-valence interaction need not be so rigorously size-extensive, since  $N_v$  does not change in the series. Computation of the energies of the states in an isovalent series like Be, Mg etc. with a few valence electrons, treatment of electron-detached or attached states of closed shell molecules, low-lying excited states of closed-shell molecules with only a few valence electrons ( $Li_2$ ,  $Na_2$ ,  $K_2$ , ...) can be conveniently performed with such a model without much size-extensivity error.

(a2) All-electron extensive models: These models are explicitly size-extensive with respect to the total number of electrons  $N (=N_c + N_v)$  only. Here, the dissection of the total correlation energy into core correlation, valence correlation and core-valence interaction plays a subsidiary role. Addition and deletion of valence electrons to the system demands that one starts from scratch once again.

(a3) Core-valence extensive models : These models are explicitly size-extensive with respect to  $N_c$  and  $N_v$  separately. This is obviously the most general

size-extensive model, since it not only automatically ensures the feature (a2) above, but also the feature (a1), when  $N_V$  is fixed. In addition, when electrons are deleted from or added to the valence orbitals, thereby changing  $N_V$  for a fixed  $N_C$ , it looks upon the consequent change in energy as a change of core-valence interaction accompanying the change in  $N_V$  and an additional valence correlation effect. One needs to compute only these additional energies, which are differential correlation energies, and can thus by-pass a more elaborate calculation de novo. This aspect is of paramount interest in formulating a size-extensive theory of such energy differences of spectroscopic interest as ionization potential (IP), electron affinity (EA) or excitation energy (EE) of the closed shell systems, since the common core-correlation energy cancels out on taking energy differences and thus need not be computed at all. Another important advantage of this model is that, when the correlation effect is short range (as is usually the case), successive addition of valence electrons will gradually lead to a saturation in the change in electron correlation, and the overall valence correlation involving many valence electrons may be computed from a knowledge of valence correlation involving fewer electrons. The core-valence extensive formalisms would thus become progressively more accurate with increasing number of valence electrons.

In view of the finer classifications (a1) to (a3) of the size-extensivity, the aspects of the size-consistency for the open-shell theories are also analogously modified, depending on which of the features (a1) to (a3) is chosen for size-extensivity. Whatever is the choice, however, the size-consistency is most conveniently ensured if the model space is selected in such a way that the starting functions constructed from the model space functions dissociate into the desired fragments. When the core electrons are much more rigidly bound compared to the valence electrons, and the core separates into closed-shell fragment cores, the various fragments that one needs to describe in the different dissociation limits can all be generated from the model space functions if the model space contains all the determinants that can be formed by allocating  $N_V$  valence electrons in the valence orbitals in all possible ways. This is the space, for example, which is often used in the CI methods, and is called the 'complete active space' (CAS) /43/ or a 'full-valence' model space. In the

current many-body terminology, one calls this a 'complete model space' /44/, although this nomenclature is unfortunately not in conformity with the definition of 'completeness' in a configurational expansion as given by Löwdin/108/. Most of the many-body formalisms used to date start with a complete model space. Size-consistency then follows as a consequence of the size-extensivity of the model underlying the theory. If the model space is incomplete/44/, great care is needed to choose proper model functions ensuring proper dissociation into fragments, and all possible fragments cannot certainly be described in a size-consistent manner with a set of fixed model space functions spanning an imcomplete model space /45/. It is, however, possible to ensure separability into specific fragments/96/ by a suitable choice of the model space. For a general but comprehensive discussion of the aspects of size-extensivity, see e.g., a recent paper by Chaudhuri et al/147/.

### 3. CLUSTER EXPANSION AND SIZE-EXTENSIVITY

There is an intimate connection between the cluster-expansion of a wave-function and the property of size-extensivity. To describe this aspect in the simplest manner, it is pertinent to recall first the closed-shell ground state. The ways to encompass the open-shell states can then be indicated as appropriate extensions and generalizations of the closed-shell cluster expansion strategy.

For an N-electron closed-shell state with a dominant reference function  $\Phi_0$ , the exact wave-function  $\Psi$  can be written as a superposition of various n-fold excited determinants  $\Phi_n^i$  on  $\Phi_0$ .  $\Psi$  may thought to be generated from  $\Phi_0$  by a wave operator  $\Omega$  :

$$\Psi_0 = \Phi_0 + \sum_{n=1}^N \sum_i C_n^i \Phi_n^i = \Omega \Phi_0 \quad (3.1)$$

In the many-body description,  $\Phi_0$  is usually taken as the 'vacuum', and 'holes' and 'particles' are appropriately defined as those occupied and vacant in

$\Phi$  respectively.  $\Phi_n^i$ 's may then be viewed as nh-np excited determinants obtained from  $\Phi$ .

The amplitudes  $C_n^i$  in the expansion, eq.(3.1), are a measure of the importance of n-electron correlation in  $\Psi$ . As the number of amplitudes  $C_n^i$  increases astronomically with the increase in  $N$ , any practicable computational scheme has to truncate the expansion, eq. (3.1). A systematic rationale for truncation is provided by a model for electron correlation. Thus, for example, Sinanoglu/28/ advocated that the pair-correlation is the dominant electron correlation for a closed-shell state with the HF function taken as  $\Phi$ . Unfortunately, however, a restriction to the approximation  $\Omega \approx 1 + \Omega_2$ , leads to an energy which becomes increasingly poorer with growing  $N$ . The approximation is hence not size-extensive. Sinanoglu/28/ demonstrated that, even if the pair-correlation is dominant, the amplitudes  $C_4, C_6, C_8 \dots$  etc are not negligible, and they are determined by simultaneous correlation of two, three, .. pairs of electrons correlating simultaneously/46/. Inclusion of these leads to the size-extensivity of energy. This finding ties up neatly with the earlier insight offered by Hubbard/27/ and Coester/30/ gleaned from MBPT, which leads to the cluster expansion for  $\Psi$  /29-30/ from  $\Phi$  :

$$\Psi_0 = \exp(T) \Phi_0 = \Omega \Phi_0 \quad (3.2)$$

where  $T$  is a combination of various nh-np excitation operators  $T_n$  :

$$T = \sum_{n=1}^N T_n \quad (3.3)$$

$T$  introduces true n-particle correlation, and products like  $T^k T^m$  etc., arising out of the expansion eq.(3.2), generate simultaneous presence of k-particle and m-particle correlations in a (k+m)-fold excited determinants etc. The truncation of  $T \approx T_2$  then corresponds to the pair-correlation model<sup>2</sup> of Sinanoglu, while incorporating higher excited states with several disjoint pair excitations induced through the powers  $T^k$ . The amplitudes for  $T$  may be called 'linked' or 'connected' clusters for  $N$  electrons. The difficulty of a linear variation method such as CI lies in its inability to realize the cluster expansion structure eq.(3.2), in a simple and practicable manner.

Following an analysis due to Primas/29/, we can indicate now one hallmark of a size-extensive representation of a wave-function as revealed in the expansion, eq.(3.2), above. Once we identify the most important true  $n$ -electron correlations, as reflected in the nature of the operators  $T_n$ , the total energy for a size-extensive  $\Psi$  must become additively separable in the asymptotic limit of no interactions between the clusters. Let us note carefully that we are talking here of the separability into various groups of interactions (clusters) and not of separability into fragments. Unless this feature is explicitly built-in, no approximate wave-function can have this property automatically. It is straightforward to verify that, once we identify the linked clusters with groups A,B, etc., then in the limit of no interaction between the groups, the representation, eq. (3.2) leads to an additive separation of the energy with respect to these groups. From eq.(3.2), we find that

$$E_0 = \langle \Phi_0 | \exp(-T) H \exp(T) | \Phi_0 \rangle \quad (3.4)$$

Let us assume that

$$H \rightarrow H_A + H_B + \dots \quad (3.5)$$

when the groups do not interact, and write  $T$  as

$$T \rightarrow T_A + T_B + \dots \quad (3.6)$$

The operators with different labels commute in the limit of no interactions between the groups, and we have

$$E_0 \rightarrow E_A + E_B + \dots \quad (3.7)$$

with

$$E_A = \langle \Phi_0 | \exp(-T_A) H_A \exp(T_A) | \Phi_0 \rangle \text{ etc.} \quad (3.8)$$

As we have taken the groupings A,B etc., to refer to true linked clusters  $T_A, T_B$  etc., the operators  $T_A, T_B$  must appear as physically connected entities in the occupation number representations.  $E_A, E_B$  etc., will also then appear as connected entities - as a consequence of the multi-commutator expansion generated by eqs. (3.8). Since the groupings are

arbitrary,  $E_0$  itself is a connected entity.

The linked-cluster theorem for energy, from the above analysis, is a consequence of the connectivity of  $T$ , and the exponential structure for  $\Omega$ . Size-extensivity is thus seen as a consequence of cluster expansion of the wave function. Specific realizations of the situation are provided by the Bruckner-Goldstone MBPT/25,26/, as indicated by Hubbard/27/, or in the non-perturbative CC theory as indicated by Coester/30,31/, Kummel/31/, Cizek/32/, Paldus/33/, Bartlett/21(a)/ and others/30-38/. There are also the earlier approximate many-electron theories like CEPA/47/, Sinanoglu's Many Electron Theory/28/ or the CI methods with 'cluster correction'/46/.

For the open-shell cluster expansion theories, we have the three different choices (a1) to (a3), as described in Sec.2, for ensuring the three different types of size-extensivity. Unlike the closed-shell situation, we have here the core electrons as well as valence electrons and we have the option of explicitly maintaining the size-extensivity of the core electrons only, or of  $N_c + N_v$  or  $N_c$  and  $N_v$  separately.

For the core-extensive theories (i.e., with the feature (a1)), there must be an explicit cluster expansion structure with respect to the core electrons, and no such cluster expansion maintained for the valence electrons. The wave-operator  $\Omega$  should have then either of the following forms

$$\Omega \sim W \exp(T) \quad (3.9)$$

$$\Omega \sim \exp(T)W \quad (3.10)$$

where  $\exp(T)$  has the np-nh cluster expansion structure with respect to the core, as in eq. (3.2), and  $W$  is a linear combination of several operators, as in a CI, introducing valence- and core-valence correlations. There may be several ways to realize the operators  $\exp(T)$  and  $W$  /50-63/.

Silverstone and Sinanoglu/48/, already in 1965, indicated how a cluster-expansion satisfying size-extensivity with respect to all the electrons (feature (a2)) can be effected. There is also a related work by Roby/49/. Transcribed into occupation number representation, this amounts to a cluster expansion of the closed-shell type with respect to each determinant



in the model space

$$\Psi_k = \sum_i C_{ik} \exp(T^i) \Phi_i \quad (3.11)$$

where the set  $\{\Phi_i\}$  spans the model space and  $C_{ik}$ 's are the combining coefficients for the  $k$ -th root. This is the general cluster expansion structure of a size-extensive theory with feature (a2). Each  $T^i$  excites various  $n$ -fold excitations from the corresponding  $\Phi_i$ 's and there is no subdivision into core correlation, valence correlation and core-valence interaction. The various formalisms/64-66/ in this category will differ in their realization of the wave-function, eq.(3.11).

The size-extensive theories with the feature (a3) can be developed by converting the representation of the operator  $W$  from a linear structure (as in eq. (3.9) or (3.10)) to an exponential-like cluster structure. There are several ways of achieving this/67-78/, and they represent the various approaches to the solution of the problem.

We may mention here that the closed-shell CC theory/32/ results from the corresponding MBPT/25,26/ only when the latter is a Rayleigh-Schrodinger (RS) series. In a Brillouin-Wigner (BW) perturbation expansion, there is a parametric dependence of the wave-operator on the total energy, and no cluster decomposition is possible for non-interacting groups A,B, etc., owing to the dependence of each of the cluster operators/operators on the total energy. Consequently only the RS type of development in open-shell MBPT theories can be fully size-extensive. Any part of the expansion of the wave-operator lacking RS structure will not satisfy size-extensivity/109/.

#### 4. CATEGORIZATION OF THE OPEN-SHELL MANY BODY FORMALISMS

##### 4.1. The Earlier Open-shell Perturbative Developments

The first core-extensive open-shell MBPT /2/ (i.e., with the feature (a1)) was developed by Bloch

and Horowitz. In this formulation, there is an energy-dependent effective hamiltonian  $H_{\text{eff}}(\Delta E_k)$  which acts on an appropriate combination  $\Psi_{\text{eff}}^{(0)}$  of the model space functions to generate  $\Delta E_k$  - the energy shift with respect to the correlated  $k$  core. Diagrammatically speaking, no term of  $H_{\text{eff}}$  has a disconnected closed diagram representing the core-energy, although the diagrams may be disconnected, since there is no explicit size-extensivity with respect to valence electrons. The model space is chosen to be complete. Brandow/3,20/ has discussed the Bloch-Horowitz theory in detail, and based his RS MBPT formulation on this as the starting point. Brandow's theory/3/ is the first fully size-extensive many-body formalism with the feature (a3). The diagrams representing the terms of this RS effective hamiltonian  $H_{\text{eff}}$  (having now no dependence on the unknown  $\Delta E_k$ 's) are all fully connected, indicating that the valence correlation and the core-valence interaction are calculated in a size-extensive manner. The series for  $H_{\text{eff}}$  is known as the linked valence expansion/3/.

The essential requirement for arriving at the proof of the above 'linked valence' expansion/3/ was the assumption that the model space is complete. With a complete model space, the model space functions separate into proper fragments as well, and Brandow's theory is thus additionally size-consistent. Brandow chose the intermediate normalization convention for the perturbed functions. It can then be shown that, as a consequence,  $H_{\text{eff}}$  is nonhermitian, whose non-orthogonal eigenvectors are just the model space projections of the exact functions/3/. There is a transformation, due to des Cloizeaux/79/, by which an equivalent hermitian effective hamiltonian can be derived, if desired. There have been several other formulations of RS MBPT/4-9/ with the feature (a3). Brandow/3/ also attempted to generalize his formalism when the model space contains valence particles as well as valence holes. He implicitly assumed that a model space containing determinants, corresponding to all possible occupancy of fixed number of valence holes and valence particles among the valence hole orbitals and valence particle orbitals, is also complete and consequently claimed that the linked valence expansion will be valid. This model space, however, is incomplete /20,44,71,94/, since with respect to the real electrons there are still other determinants in which the electrons can be allocated among the valence orbitals in many other ways. Thus, in this case the linked valence expansion does not

generally hold good, as Brandow himself later pointed out/20/.

Hose and Kaldor/44/ had been the first to have analyzed in detail the structure of the RS MBPT in intermediate normalization when the model space is generally incomplete. They showed that there are disconnected diagrams in this case. In their formulation, they abandoned the classification of the orbitals into core, valence and particles, and instead chose each ket  $\Phi_i$  on which  $H_{\text{eff}}$  acts as the vacuum. Until very recently /91-96,147/, it was not clear whether the disconnected diagrams appearing in the formalism spell a breakdown of the size-extensivity. Also, since in this formulation, the core energy is not explicitly subtracted, the size-extensivity- if at all respected- will be of the type (a2). Clearly, if the model space is made complete the formalism will go over to a size-extensive theory of the type (a3), without disconnected diagrams. This may be looked upon as a perturbative realization of the Silverstone-Sinanoglu expansion/48/. Chaudhuri *et al* have recently discussed/147/ and demonstrated/146/ that the energies obtained by diagonalizing the  $H_{\text{eff}}$  of the Hose-Kaldor formalism entail disconnected terms indicating a breakdown of size-extensivity. We shall discuss again the aspect of size-extensive theories with incomplete model space in Secs. 4.3 and 8.4, where the emphasis is on the cluster expansion approach.

#### 4.2.The Propagator Methods

The propagator, or equation of motion methods, and the related Green's function techniques/10-16/ were primarily designed to generate EE,IP or EA of a closed shell ground state. The size-extensivity that they ensure is of the type (a1), since the self-energy operator-playing a role rather analogous to that of  $H_{\text{eff}}$  - is dependent on the unknown shift  $\Delta E_k$ , and in this sense they are structurally analogous to the BW series of the Bloch-Horowitz type/9/.The kinship with the Bloch-Horowitz series is apparent, since there is again no closed core diagram in the series for the self-energy. There are major structural differences also- notably in their choice of model space (for a discussion on this point, see, e.g., Brandow/81/, Robb *et al*/82/and Mukherjee and Kutzelnigg/84/), but the feature of core-extensivity is shared by both. The formulation of an EOM-type of theory satisfying

size-extensivity is rather recent, and we refer to the works of Kuo *et al*/83/ and to a forthcoming paper by Mukherjee and Kutzelnigg/84/ for a detailed discussion of the subject.

#### 4.3 The Open-Shell Cluster Expansion Approach

The development of the open-shell many-body cluster expansion methods has followed a rather circuitous and tangled route. The early developments by Mukherjee and co-workers/68,69/ directly emphasized the size-extensivity feature (a3). Theirs was essentially a complete model space formulation, and it was tailored for computing both total state energies and energy differences/69/. The formalism necessitates an explicit consideration of the differential correlation energies as more and more electrons are added to the valence orbitals. This added flexibility requires that we need consider not only the desired complete model space for a fixed  $N_v$ , but also all the lower valence (henceforth called 'subduced') model spaces with fewer valence electrons  $m$ , ( $0 \leq m \leq N_v$ ) as well. The formalism is thus not just confined to a specific Hilbert space of a fixed number of electrons /68,69/, but is defined in a Fock-space which is a direct sum of several Hilbert spaces of different electron numbers. Mukherjee *et al* have discussed several Fock-space strategies/68,69/ to arrive at size-extensive formulations. These methods may be viewed as open-shell generalizations of the closed-shell CC method.

In nuclear physics, Offermann *et al* and Ey/70/ proposed a related but different formalism which also shares the size-extensivity feature (a3), and have the flexibility of describing differential correlation energy. This is also a Fock space oriented approach. Somewhat later, Lindgren/71/ generalized his formulation of the Open-shell MBPT/4/ to derive an open-shell coupled cluster method which may be viewed as a highly summed up version of the MBPT. Lindgren started from a complete model space with fixed  $N_v$ , and looked for a set of equations which will guarantee the connected nature of the cluster amplitudes. Although Lindgren's cluster expansion contained the flexibility of describing the differential correlation energy similar to that in the theory of Mukherjee *et al* /68,69/, he confined his development to Hilbert space

of fixed  $N_v$ . This generally leads to an underdetermined system of equations for the cluster amplitudes since their number far exceeds the number of equations/66,69/. Lindgren, however, augmented his formalism by a set of sufficiency conditions for ensuring the connectivity of cluster amplitudes which gave as many equations as are unknowns. Haque and Mukherjee/69/ commented that these conditions essentially imply adopting a Fock-space strategy, and in fact Haque demonstrated/69/ that Lindgren's equations can be derived by assuming that his wave-operator for the  $N_v$ -valence problems behaves as a wave-operator for the lower valence problems. This point has been further clarified in the recent analysis of Lindgren and Mukherjee/94(a)/ and Mukherjee/94(b)/.

The next set of open-shell cluster expansion theories to appear on the scene emphasized the size-extensivity feature (a1), and all of them were designed to compute energy differences with a fixed number of valence electrons. Several related theories may be described here - (i) the level-shift function approach in a time-dependent CC framework by Monkhorst/56/ and later generalizations by Dalgaard and Monkhorst/57/, also by Takahasi and Paldus/105/, (ii) the CC-based linear response theory by Mukherjee and Mukherjee/58/, and generalized later by Ghosh et al/59,60,107/,(iii) the closely related formulations by Nakatsuji/50,52/ and Emrich/62/ and (iv) variational theories by Paldus et al/54/ and Saute et al/55/ and by Nakatsuji/50/.

The next two formulations advocated maintaining size-extensivity with respect to the total electron number  $N$ , i.e., they subscribed to the size-extensivity feature (a2). The method of Banerjee and Simons/65/ started with a CAS-SCF function and used a cluster expansion inducing correlation of valence electrons only. The core remains frozen in this theory, leading to some computational simplifications, but at the expense of generality. Since only one combination of determinants was chosen as the reference function - as in an MR-CI-only one root could be calculated. There was thus no  $H_{eff}$  in this theory. Extensions of this approach to include core-excitations were initiated by Laidig et al /115(a)/ and Hoffman and Simons/115(b)/. The theory of Jeziorski and Monkhorst/66/ was general, and this sought to compute all the roots of an  $H_{eff}$  defined in a complete model space. The cluster expansion used was the Silverstone-Sinanoglu expansion (eq. (3.11))/48/,

and each ket on which  $\exp(T^i)$  acts was taken as the vacuum in the same spirit as that of Hose and Kaldor/44/.  $H_{\text{eff}}$  was proved to be connected by a rather involved algebraic route. All these theories work in Hilbert space of a fixed number of electrons and are thus conceptually distinct from the Fock-space based theories of the type (a3).

Kutzelnigg/76/ and Kutzelnigg and Koch/77/ revived the cluster expansion approach of the type (a3) in a series of papers where the name 'Fock Space Quantum Chemistry' appeared first. They emphasized that the structure of  $H$  in a occupation number representation is much simpler than in a Hilbert space since this is independent of the number of electrons and, once the valence orbitals are identified, a similarity transformation of  $H$  by  $\Omega$  can bring  $L = \Omega^{-1} H \Omega$  to a form where valence orbitals couple only with the valence orbitals through  $L$ . A diagonalization of this part of  $L$  having valence orbitals only in a model space of a specified  $N_v$  will then furnish selected eigenvalues of  $H$  for an  $N_v$ -valence problem. In effect, this means that one defines  $\Omega$  in Fock space, transforms  $H$  to  $L$  in Fock space, and projects onto a model  $N_v$ -valence Hilbert space to finally construct  $H_{\text{eff}}$  on the Hilbert space. Operationally speaking, this method advocates the same strategy as that expounded by Mukherjee *et al*/67-69/ or Lindgren/71/ as amplified by Haque and Mukherjee/69/. But there is a subtle difference. While the earlier works/67-69/ used explicit determinantal projectors for the equations determining the cluster amplitudes and  $H_{\text{eff}}$ , Kutzelnigg and Koch emphasized the operator aspects of the same equations without determinantal projectors. The Fock-space strategy is thus more explicit in the latter. Kutzelnigg and Koch/77/ considered complete model space theories only, with valence orbitals as of the particle type. A related approach containing holes and particles was initiated by Stotarczyk and Monkhorst/85/, which, however, -strictly speaking- is not an effective hamiltonian theory (see, e.g., Lindgren and Mukherjee/94(a)/ for a critique on this point). See also Sinha *et al*/119/.

It should be emphasized here that the convenience of working in a complete model space for arriving at the connectedness of  $H_{\text{eff}}$  was first shown by Brandow/3/ and this view-point has since influenced most of the later developments of the open-shell MBPT and CC theory. When the valence orbitals are rather well-spaced in energy, this may create a serious numerical instability in actual applications. Whenever

there are some virtual space determinants close in energy to some model space functions, they will mix strongly resulting in the divergence of the perturbation series for  $H_{\text{eff}}$  or an ill-conditioning of the open-shell CC equations. The offending virtual determinants for which such strong mixing occurs are called 'intruder states'/86/. Their presence has been a perennial problem in both nuclear physics/86/ and quantum chemistry/44,87/. As we have already discussed, Hose and Kaldor/44/ formulated an incomplete model space version of MBPT. Their motivation was to include only those functions in the model space which do not mix too strongly with the virtual space functions. Haque and Mukherjee/88/ formulated a hermitian version of MBPT for incomplete model space. Jeziorski and Monkhorst/66/ indicated how their CC formulation for the complete model space may be generalized to the incomplete model space situation leading to the CC analogue of the Hose-Kaldor MBPT/44/. In all these developments, disconnected diagrams appear in  $H_{\text{eff}}$ , and they are not size-extensive/146,147/.

Recently, Lindgren/88/ introduced the concept of a quasi-complete model space, which is a very special type of incomplete model space, for which he claimed that  $H_{\text{eff}}$  is connected. In a quasi-complete model space, the valence orbitals are bunched in different groups, the occupancy of orbitals in each group is held fixed and the various possible occupancies within each group are exhausted. Looked at from this classification, a model space consisting of  $nh$ -mp determinants- as discussed by Brandow/3,20/ in his generalized MBPT for valence holes and particles- is quasi-complete, with valence holes and valence particles in different groups. It is known that  $H_{\text{eff}}$  is disconnected in this case. Unfortunately for a general quasi-complete case also,  $H_{\text{eff}}$  is not connected/90/. Size-extensivity of the energies is thus not respected in these formalisms/146,147/.

Very recently, Mukherjee/91/ made a re-investigation of the problem of the appearance of disconnected terms in  $H_{\text{eff}}$  for  $nh$ - $np$  determinants as constituting a model space, and traced the origin of the disconnected to the tacit assumption of maintaining the intermediate normalization of the eigenfunctions. By abandoning the intermediate normalization, and by adopting a different normalization, Mukherjee proved that there are no disconnected terms in  $H_{\text{eff}}$ . Using this same strategy, Mukherjee/92/ proved for the first time that for a

general incomplete model space a connected  $H_{\text{eff}}$  results if the intermediate normalization is abandoned. We may term those normalization conventions for which  $H_{\text{eff}}$  is connected as size-extensive normalizations /92,93/. To prove the connectedness of the cluster amplitudes and also of  $H_{\text{eff}}$ , it was found essential to adopt a Fock-space strategy and to postulate that the same wave-operator  $\Omega$  applies to all the various valence sectors of the Fock space. The wave-operator is thus 'valence-universal'. Connected MBPT and CC theory for a general model space from a unified viewpoint has also been presented by Mukherjee/93/. Some further generalizations and the necessary and sufficient conditions for arriving at a connected  $H_{\text{eff}}$  have been discussed by Lindgren and Mukherjee/94/. The generalizations of /76/ and /77/ to tackle incomplete model space were done by Kutzelnigg *et al*/95/ and Mukherjee *et al*/96/, where a special type of incomplete model space - called the 'isolated incomplete model space'/95/ - has been introduced which possesses the interesting property of supporting a kind of intermediate normalization as well as generating a connected  $H_{\text{eff}}$ . There has been considerable activity in this field over the last two years/138-147/.

#### 4.4 Certain General Comments

For the analysis of the various formalisms, manipulation of the equations, generating normal product of terms via Wick's theorem, and particularly for indicating how the proofs of the several different linked cluster theorems are achieved, we shall make frequent use of diagrams. For the sake of uniformity, we shall mostly adhere to the Hugenholtz convention/1/. All the constituents of the diagrams will be operators in normal order with respect to suitable closed-shell determinant taken as the vacuum. We shall refer to the creation/annihilation operators with respect to this vacuum after the h-p transformation. The hamiltonian  $H$  will also be taken to be in normal order with respect to  $\Phi_0$ :

$$\begin{aligned}
 H = & \langle \Phi_0 | H | \Phi_0 \rangle + \sum_{A,B} \langle A | h | B \rangle \{ a_A^\dagger a_B \} \\
 & + (1/2!) \sum_{A,B}^2 \langle AB | v | CD \rangle \{ a_A^\dagger a_B^\dagger a_D a_C \} \quad (4.4.1) \\
 & C,D
 \end{aligned}$$



Here  $\{...\}$  denotes the normal ordering, and  $\langle A|h|B\rangle$  and  $\langle AB|v|CD\rangle$  are the one- and antisymmetrized two-electron<sup>a</sup> integrals. Their skeletons are depicted in Fig.1. A,B, etc., stand for general spin orbitals. We shall denote the orbitals by the Greek letters  $\alpha, \beta$ , etc., particle orbitals by p,q,etc., hole valence by  $\bar{\alpha}, \bar{\beta}$ , etc., and particle valence by p,q, etc. Fig. 2 shows a typical classification scheme containing valence holes and valence particles. In a parallel terminology, the valence orbitals (valence holes and/or valence particles) are called 'active' and the non-valence holes and particles are called 'inactive'/8/. Fig. 2 displays this terminology as well. We shall sometimes use these two terminologies interchangeably.

The inactive holes and particles are indicated by single arrows on the lines in the usual manner. The active holes and particles are depicted as lines with double arrows. These are illustrated in Fig. 3. We shall also choose the convention of drawing the diagrams horizontally whereby the sequence of operators appearing from right to left will be depicted as a sequence of appropriate vertices sitting in the same order starting from the right.

## 5. SEMI-CLUSTER EXPANSION THEORIES FOR THE OPEN-SHELL STATES

### 5.1 General Physical Considerations

When the number of valence electrons are far less in comparison to the number of core electrons, the bulk of the size-extensivity of the total energy

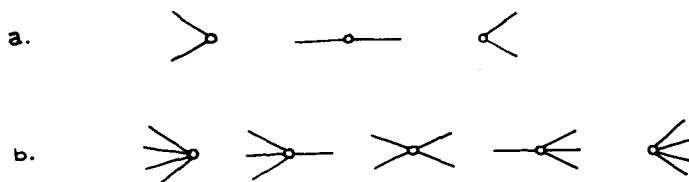


Fig.1 (a) The skeletons for the operator  $h$ .  
(b) Corresponding skeletons for  $v$

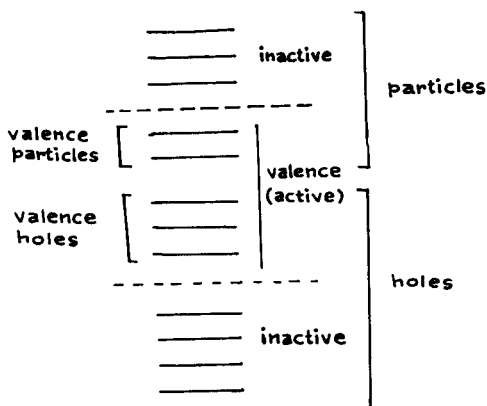


Fig.2 Classification of orbitals into holes and particles. Also displayed are further divisions into inactive and active orbitals.



Fig.3 The arrow conventions for inactive and active holes and particles.

of an open-shell state stems from the size-extensivity of the core-electron energy. This is also the situation where the major portion of the electron correlation remains relatively unchanged on going over from the core system to the open-shell state. For example, for the low-lying valence excited or ionized states derived from a closed shell ground state, it is a good approximation to hold that the bulk of electron correlation remains unaffected. The inclusion of the correlation between the valence electrons and the core-valence interaction can be taken care of by a valence wave-operator  $W^V$ . Thus for a fixed number of valence electrons, if we introduce a model space spanned by a set of  $M$  determinants  $\{\Phi_i\}$ , then a semi-cluster representation of the associated exact functions  $\Psi_k$ , incorporating a size-extensive description of the core-correlation, can be given in the following manner:

$$\begin{aligned}\Psi_k &= \sum_{i=1}^M C_{ik} \exp(T) W^V \Phi_i \\ &= \sum_{i=1}^M C_{ik} W^C W^V \Phi_i\end{aligned}\quad (5.1.1)$$

where  $W^C = \exp(T)$  is the wave-operator of the core  $\Phi_0$  and each  $\Phi_i$  is obtained by the action of  $N_i$  valence creation operators (valence holes and/or particles)  $Y_i^\dagger$  on the core  $\Phi_0$ , taken as the vacuum :

$$\Phi_i = Y_i^\dagger \Phi_0 \quad (5.1.2)$$

$W^V$  is a combination of various nh-np excitations from the set  $\{\Phi_i\}$ . If we denote the set of operators  $Y_i^\dagger$  as of the type  $W_{1+n}$ , then the operators inducing nh-np excitations can be said to be of the types  $W_{1+n}$ . writing eq.(5.1.1) in long hand, we find

$$\Psi_k = W^C \left[ \sum_{i=1}^M C_{ik} \Phi_i + \sum_{m \in i} C_{mk} \Phi_m \right] \quad (5.1.3)$$

where the various nh-np excited determinants  $\{\Phi\}$  with respect to  $\Phi_0$ 's are generated by the operator manifold  $\{Y_i^\dagger\}$  contained in the set  $\{W_{1+n}\}$ . As the sets  $\{Y_i^\dagger\}$  and  $\{W_{1+n}\}$  contain creation operators only, they commute with  $W^C$ , and we can rewrite eq.(5.1.3) as

$$\Psi_k = \Omega_k \Psi_0 \quad (5.1.4)$$

where  $\Psi_0 = W^C \Phi_0$  is the correlated core, and  $\Omega_k$  is an 'excitation operator' defined by the relation

$$\Omega_k = \sum_{i=1}^M C_{ik} Y_i^\dagger + \sum_{m \in i} C_{mk} Y_m^\dagger \quad (5.1.5)$$

Eq.(5.1.5) indicates that  $\Psi_k$  receives contribution from the correlating virtual determinants  $\{\Phi\}$  in two ways : (a) contribution originating from the set  $\{W_{1+n}\}$  generating nh-np excitations out of  $\{\Phi_i\}$ , and (b) that from the product excitations of the form  $W_{1+n} T_k$  generating additional mh-mp excitations. If the core-reorganization and core-correlation changes attendant with the addition of valence holes and/or particles are not very severe, then a

truncation of the set  $\{W_{1+n}\}$  to a few small  $n$  will be a good approximation, and most of the additional correlation will be well-simulated by product excitations of the type  $W_{1+n}^T$ .

For generating the working equations for the various semi-cluster expansion formalisms, it will be convenient to denote by the set  $\{\Phi\}$  the union of  $\{\Phi_i\}$  and  $\{\Phi\}$ , and the associated operator manifold as  $\{Y^\dagger\}^m$ . The unknown quantities are the sets  $\{C_{ik}\}$  and  $\{C_{ak}\}^{mk}$  denoted henceforth collectively as  $C \equiv \{C_{ik}, C_{ak}\}^{mk}$ . We may envisage several approaches for their determination.

## 5.2 Variational Methods

As the name indicates, these methods determine the coefficients  $C$  by an appeal to the variational principle for the state energies  $E_k$ . We shall discuss two principal formalisms which make use of this strategy.

In one variant, due to Paldus *et al*/54/ and Saute *et al*/55/, a linked cluster theorem is proved whereby  $E_k$  comes as a sum of  $E_0$ , the correlated core-energy, and a ratio of the form  $N_k/D_k$ , where  $N_k$  and  $D_k$  are connected entities (see eq.(5.2.4)) :

$$E_k = \langle \Psi_k | H | \Psi_k \rangle / \langle \Psi_k | \Psi_k \rangle = E_0 + N_k/D_k \quad (5.2.1)$$

Eq.(5.2.1) indicates that each  $E_k$  has a constant energy shift  $E_0$ . Since  $\Psi_0$  is assumed to be known, the variational principles for  $E_k$  and  $\Delta E_k$  (the energy difference) are synonymous.  $\Delta E_k$  is core-extensive since there are no disconnected terms containing  $E_0$ .

With  $\Phi_0$  as the Hartree-Fock function for the ground state, eq.(5.2.1) can generate a connected expression for such energy differences as IP, EA or EE directly, depending on the manifold in  $\Omega_k$ . Thus, for IP/EA, the natural choice for  $W_1$  are operators  $\{a^\dagger/a\}$ , while for EE the corresponding  $W_1$  manifold is the set  $\{a^\dagger a\}$ . Paldus *et al*/5 / also showed that the same expressions result even when the orthogonality of  $\Psi_k$  and  $\Psi_0$  is explicitly imposed.

Introducing the hermitian matrices  $\Sigma$  and  $\Delta$  through their elements :

$$\Sigma_{ab} = \langle \Phi_0 | W^C Y_a H Y_b^\dagger W^C | \Phi_0 \rangle_c \quad (5.2.2)$$

$$\Delta_{ab} = \langle \Phi_0 | W^C Y_a Y_b^\dagger W^C | \Phi_0 \rangle_c \quad (5.2.3)$$

eq.(5.2.1) can be alternatively written as

$$\Delta E_k = C_k^\dagger \Sigma C_k / C_k^\dagger \Delta C_k \quad (5.2.4)$$

which is ratio of connected bilinear hermitian forms. Variation of  $\Delta E_k$  with respect to the coefficients  $C_k$  leads to an eigenvalue equation of the form :

$$\Sigma C_k = \Delta E_k \Delta C_k \quad (5.2.5)$$

Using a perturbative analysis to estimate the importance of various operators, Paldus et al /54/ advocated the following truncation scheme:

$$\begin{aligned} \langle \Psi_k | O | \Psi_k \rangle \\ \simeq \langle \Phi_0 | [W_1^\dagger + W_2^\dagger + T_2^\dagger W_1^\dagger] O [W_1 + W_2 + W_1 T_2] | \Phi_0 \rangle_c \end{aligned} \quad (5.2.6)$$

where  $O = H$  and the unit operator  $I$  for  $\Sigma$  and  $\Delta$  respectively. Saute et al/55/ applied this truncation scheme for IP,EA and EE calculations of model  $\pi$ -systems in Pariser-Parr-Pople framework, and showed the viability of the formalism.

Nakatsuji/50/ developed an alternative variant, where essentially the same ansatz as eq. (5.1.4) is invoked though no attempt was made to derive a linked cluster theorem in an explicit manner. Assuming that the cluster amplitudes of the operator  $T$  are known, the coefficients  $C_k$  are determined directly from a variation of the expression

$$\langle \Psi_k | H - E_k | \Psi_k \rangle = 0 \quad (5.2.7)$$

leading to an eigenvalue equation of the form

$$\sum_b \langle \Psi_0 | Y_a H Y_b^\dagger | \Psi_0 \rangle C_{bk} = \Delta E_k \langle \Psi_0 | Y_a Y_b^\dagger | \Psi_0 \rangle \quad (5.2.8)$$

Nakatsuji also showed, somewhat in the line of Paldus et al/54/, that no generality is lost in assuming that

$\Psi_k$  and  $\Psi_0$  are orthogonal.

In actual numerical implementations, Nakatsuji and his co-workers/50/ used an approximation scheme, in which matrix-elements of the type  $\langle \Phi_0 | T_2^\dagger Y H Y T_2 | \Phi_0 \rangle$  are neglected, since they are both higher order in effect and difficult to compute. IP and EE computations on several prototypical systems produced encouraging results. It, however, appears that Nakatsuji in his later applications/50-52/ tended to favour a non-variational, projection method—presumably in view of its relative ease of applications. This has been covered in Sec.5.3.

### 5.3.Non-Variational Methods

There are several non-variational semi-cluster expansion formalisms of seemingly disparate structures/56-63/. All but one of them are closely related/56-62/, and they all assume the ansatz, eq.(5.1.4), either implicitly or explicitly. These theories use the projection method for obtaining  $C_k$ , which generates a nonhermitian eigenvalue problem, with  $\Delta E_k$ 's as the eigenvalues. A different ansatz for  $\Psi_k$ , leading to a hermitian eigenproblem has also been suggested/63/.

Mukherjee and Mukherjee proposed what they termed as a CC-based linear response theory (CC-LRT)/58/. The method invokes the apparatus of the linear response formalism for calculating energy differences directly. The essential idea is to compute the dynamic linear response of a closed-shell ground state subjected to a coupling with a photon field of frequency  $\omega$ , and obtain the energy differences as poles of the linear response function as a function of  $\omega$ . The interaction hamiltonian  $H_{int}$  is taken to be of the form

$$H_{int} = iV^{el} [C-C^\dagger] \quad (5.3.1)$$

where  $V^{el}$  involves electronic operators only, and  $C/C^\dagger$  are annihilation/creation operators for a photon field of frequency  $\omega$ . In the absence of a coupling, the composite system of the molecular ground state and photon satisfies the equations

$$[H + H_{ph}] \Psi_0 X_{ph} = [E_0 + n\omega] \Psi_0 X_{ph}, \quad (5.3.2)$$

$$H_{ph} = \omega C^\dagger C \quad (5.3.3)$$

$X_{ph}$  is an eigenstate of  $H_{ph}$  with eigenvalue  $n\omega$ .  $\Psi_0$  is assumed to be of the  $ph$  CC-form. In the presence of interaction, the eigenstate of the composite system will not have a definite number of photons, and the extent of correlation as a result of coupling will also be modified. These changes can be induced by the action of a second cluster operator of the exponential form :  $\exp(S)$ . The operator  $S$  destroys/creates zero, one, two, ..., photons and simultaneously induce various  $nh$ - $mp$  excitations out of  $\Phi_0$ . The nature of the electronic part of the cluster operator in  $S$  is dictated by the nature of the energy difference we are interested in. For IP/EA calculations,  $V_{el}$  in eq.(5.3.1) will destroy/create an electron from  $\Psi_0$ , and consequently  $S$  should involve  $nh$ -( $n\pm 1$ ) $p$  excitations. Similarly, for EE,  $V_{el}$  will conserve the number of electrons, and  $S$  should involve  $nh$ - $np$  excitations. For computing the linear response, it suffices to retain only the terms linear in  $C/C^\dagger$  :

$$S = i[S^-C - S^+C^\dagger] \quad (5.3.4)$$

where  $S^-$  and  $S^+$  are the electronic parts of  $S$ , as discussed above.

To arrive at the working equations for CC-LRT, we proceed as follows/58,60/. Starting from the Schrodinger equation in the presence of the perturbation :

$$\begin{aligned} [H + H_{ph} + H_{int}] \exp(T+S) \Phi_0 X_{ph} \\ = \tilde{E} \exp(T+S) \Phi_0 X_{ph} \end{aligned} \quad (5.3.5)$$

we pre-multiply by  $\exp(-T)$ , and find

$$\begin{aligned} [\bar{H} + H_{ph} + \bar{H}_{int}] \exp(S) \Phi_0 X_{ph} \\ = \tilde{E} \exp(S) \Phi_0 X_{ph} \end{aligned} \quad (5.3.6)$$

where the single bar operators  $\bar{O}$  are defined by

$$\bar{O} = \exp(-T) O \exp(T). \quad (5.3.7)$$

If we now premultiply eq. (5.3.6) further by  $\exp(-S)$ , and truncate the resultant operators after the terms linear in photon destruction/creation operators (hence the term 'linear' response theory),

then we obtain

$$[\bar{H} + [\bar{H}, S] + H_{ph} + [H_{ph}, S] + \bar{H}_{int}] \Phi_0 X_{ph} \simeq E^{(1)} \Phi_0 X_{ph} \quad (5.3.8)$$

where  $E^{(1)}$  is the energy of the composite in the linear response approximation.  $S^\pm$  in  $S$  of eq.(5.3.4) can be written in terms of operators  $\{Y_a^\dagger\}$ :

$$S^\pm = \sum_a X_a^\pm Y_a^\dagger \quad (5.3.9).$$

To determine the vectors  $X_a^\pm$ , we project eq.(5.3.8) onto the various functions  $\langle \Phi_0 X_{\pm ph} |$ , where  $X_{\pm ph}$  are obtained from  $X_{ph}$  by creating/destroying a photon:

$$\langle \Phi_a | [H, S^\pm] + \bar{V}^{el} \pm \omega S^\pm | \Phi_0 \rangle = 0 \quad (5.3.10)$$

Utilizing eq.(5.3.9), and introducing the nonhermitian matrix  $A$  and the vector  $V$  through the relations

$$A_{ab} = \langle \Phi_a | [\bar{H}, Y_b^\dagger] | \Phi_0 \rangle \quad (5.3.11)$$

$$V_a = \langle \Phi_a | \bar{V}^{el} | \Phi_0 \rangle \quad (5.3.12)$$

we have

$$[A \pm \omega I] X^\pm + V = 0 \quad (5.3.13)$$

The coefficients  $X^\pm$  are thus given by

$$X^\pm = -[A \pm \omega I]^{-1} V \quad (5.3.14)$$

Determining  $X^\pm$  is equivalent to determining the operator  $S$ , and hence of the first order perturbed function  $S \Phi_0 X_{ph}$ . According to the general theory of linear response, the amplitudes of  $S^\pm$ , i.e., the vectors  $X^\pm$ , will become singular when  $\omega$  matches an elementary excitation such as IP, EA or EE. This happens whenever the matrices  $[A \pm \omega I]$  become singular.

If  $\omega_k$ 's are eigenvalues of  $A$  obtained from



$$A Z_k = \omega_k Z_k \quad (5.3.15)$$

then the vectors  $X^\pm$  will become singular for  $\omega = \mp \omega_k$ . Since the positive poles of the linear response function correspond to an absorption of photon, the solution of eq.(5.3.15) furnishes us directly with  $\Delta E_k$ 's. Eq.(5.3.15) is the working equation for CC-LRT. Explicit expression for  $A$  was derived for EE in /58/ and for IP in /59/. Let us note that the actual form for  $\bar{V}_{el}$  has not been needed to get  $\omega_k$ 's.

It may not be immediately apparent that CC-LRT does indeed utilize the ansatz, eq.(5.1.4). To see this explicitly, we follow the analysis of Ghosh et al/60/. Using the ansatz, eq.(5.1.4), we first write an equation for  $\Delta E_k$  as in EOM/13,15,16/

$$[H, \Omega_k] \Psi_0 = \Delta E_k \Omega_k \Psi_0 \quad (5.3.16)$$

Since  $W^C$  and  $\Omega_k$  commute, we find from eq.(5.3.16), using eq.(5.3.7), the relation

$$[\bar{H}, \Omega_k] \Phi_0 = \Delta E_k \Phi_0 \quad (5.3.17)$$

Projecting onto the states  $\langle \Phi_i |$  and using eqs.(5.1.5) and (5.3.11) we have

$$A C_k = \Delta E_k C_k \quad (5.3.18)$$

which establishes that the energy obtained from CC-LRT is indeed the same as that obtained from an EOM strategy using the method of projection and the ansatz, eq.(5.1.4). CC-LRT was derived in this alternative manner by Adnan et al/97/ and Ghosh et al/60/, which shows that CC-LRT may be viewed as a renormalized nh-mp TDA; the renormalization enters via the dressed molecular hamiltonian incorporating the ground state correlation.

It should be mentioned that Emrich/62/, in the context of many-body nuclear structure theory, also arrived at eq.(5.3.18) for EE, but unfortunately he missed the earlier works. There is also a related earlier paper by Coester/128/.

Pilot numerical applications of CC-LRT were performed by Adnan et al/98/ for singlet EE, and by Ghosh et al/59/ for IP calculations on model  $\pi$ -electron systems. Ghosh et al/58/ also spin-adapted the formalism for EE to encompass both singlet and

triplet excitations, and pilot calculations with this scheme were performed by Adnan *et al*/97/. In the ab-initio framework, the method has since been applied by Mukhopadhyay *et al*/99/ and Roy *et al*/100/ for IP calculations covering both the outer and inner valence region. The theory was able to predict the satellite spectra in the inner valence region of molecules like Nitrogen and Water satisfactorily.

In these applications, the earlier strategy of Adnan *et al*/97,98/ and Ghosh *et al*/59/ was utilized.  $T \approx T_2$  approximation was used and was computed from the closed shell CC theory. The matrix elements of  $H$  are computed next. The construction of  $H$  is facilitated by the diagrammatic handling of the various terms. As typical examples, the one-body hole-hole term of  $H$  and the two-body p-h - p-h term of  $H$  are shown in Figs. 4(a) and 4(b). In general, any term of  $H$  of a given shape (dictated by the nature of the lines joined to it) is obtained as a collection of connected diagrams of the same shape obtained by joining  $H$  and  $T$  vertices in every possible manner. Mukherjee and Mukherjee/58/ listed the expression for all the one- and two-body terms of  $H$ , and denoted them respectively by  $\bar{F}$  and  $\bar{V}$  /101/. Once  $T$  is known,  $\bar{F}$  and  $\bar{V}$  can be computed once for all and stored. Three and higher body terms of  $H$  should, in principle, be included but they have been neglected as yet. The diagrams entering the matrix  $A$  for IP and EE, with  $\Omega_k$  confined to  $W_1$  and  $W_2$  manifolds, are shown in Figs.5.

Closely related to CC-LRT is the non-variational formulation of Nakatsuji/50-52/ and Hirao/53/. Starting with the ansatz, eq.(5.1.4), they projected the Schrodinger equation for  $\Psi_k$  onto the functions  $\langle \Phi_a |$ , and obtained the eigenvalue equation

$$\sum_b \langle \Phi_a | H Y_b^{\dagger} W^C | \Phi_b \rangle C_{bk} = E_k \sum_b \langle \Phi_a | Y_b^{\dagger} W^C | \Phi_b \rangle C_{bk} \quad (5.3.19)$$

Extensive numerical investigations of this formalism were undertaken by Nakatsuji/52/ and Hirao/53/ for IP and EE computations. For IP calculations, the operator manifold taken by them were  $W_1$  and  $W_2$ , and product excitations of the form  $W_1 T_2$  were also included. A similar approximation scheme for EE was also used, although all the spin-adapted  $W_2$  operators for the triplet EE calculations were not included. This should be contrasted with the scheme of

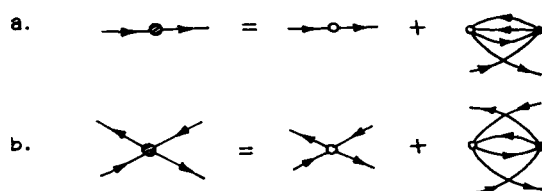


Fig.4 (a) Typical terms of the h-h matrix elements of  $\bar{F}$ .  
 (b) ph-ph matrix elements of  $\bar{V}$

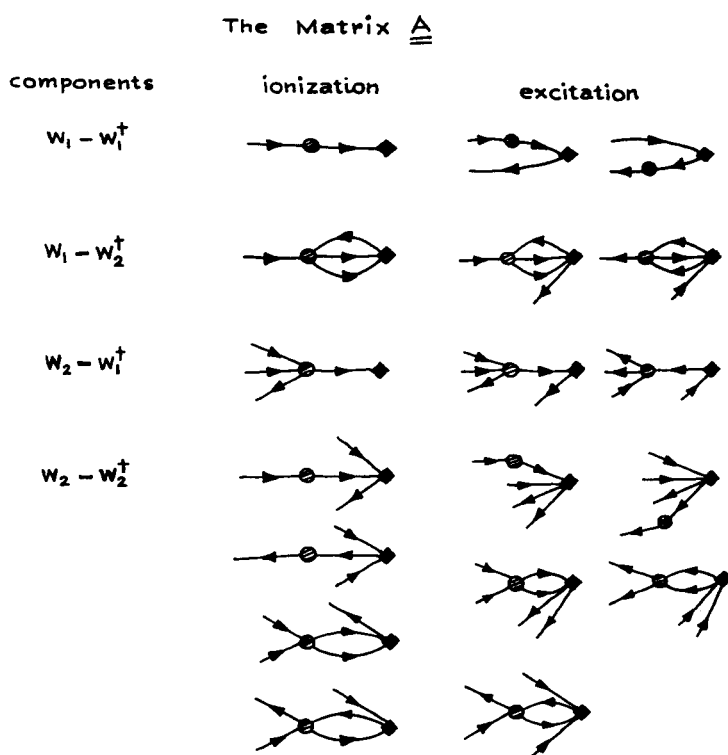


Fig.5 Typical matrix elements of A coupling various blocks for IP and EE computations.

Ghosh *et al*/58/, where the full space of  $W_2$  was utilized. Since only a few eigenvalues and the corresponding eigenvectors of eq.(5.3.19) are of interest, Hirao and Nakatsuji/102/ developed a generalization of Davidson's algorithm/103/ for nonsymmetric matrices. Hirao/53/ also devised a direct cluster expansion method for constructing the vector  $AC^T$ , where  $C^T$  is a trial eigenvector, from the molecular integrals directly as an intermediate step for the iterative solution of the eigenvalue equation. (The feasibility of such a procedure in CC-LRT was emphasized even earlier by Adnan *et al*/98/.) Inner and outer valence IP and valence and Rydberg excited states of several prototypical systems were computed /50-53/.

The non-variational formalisms described above generate non-hermitian eigenproblems. A hermitian version of CC-LRT was suggested in a propagator theory framework by Prasad *et al*/63/. The formalism, however, is more complex than CC-LRT and involves more computational work for the same degree of accuracy.

The origin of the non-hermiticity of the matrix  $A$  in CC-LRT can be traced to the non-hermiticity of  $H$  which is obtained from  $H$  by a similarity, rather than unitary, transformation via  $\exp(T)$ . Prasad *et al*/63/ suggested instead a unitary cluster operator  $\exp(\sigma)$  for correlating  $\Phi_0$  :

$$\Psi_0 = \exp(\sigma)\Phi_0 \quad \text{with } \sigma = T - T^\dagger \quad (5.3.20)$$

We shall review next the time-dependent version of the CC theory of Monkhorst/56/ which also generates a linear response function closely corresponding to the CC-LRT response function. This formalism also thus falls in the category of non-variational method. To underline the similarity of this method with CC-LRT, we shall denote analogous entities by the same symbols and no confusion should arise if the context is remembered.

If the molecule in its ground state is subjected to a coupling with a time-dependent field with the interaction hamiltonian :

$$H_{int}(t) = V^{el} [\exp(i\omega t) + \exp(-i\omega t)] \exp(\alpha t) \quad (5.3.21)$$

with  $\alpha \rightarrow 0^+$ , corresponding to adiabatic switching from infinite past, the time-dependent Schrodinger equation in presence of the coupling is given by

$$[H + H_{\text{int}}(t)] \Psi(t) = i \partial \Psi(t) / \partial t \quad (5.3.22)$$

Eq.(5.3.21) indicates that

$$\Psi(t) = \lim_{\alpha \rightarrow 0} U(t, -\infty) \Psi_0 \quad (5.3.23)$$

Following Langhoff *et al*/104/, Monkhorst/56/ separated the secular and normalization terms in  $\Psi(t)$  by factorizing  $U(t, -\infty)$  into a regular part  $U_r(t, -\infty)$  and a vacuum term  $\langle \Psi_0 | U | \Psi_0 \rangle$  :

$$\Psi(t) = U_r(t, -\infty) |\Psi_0\rangle \langle \Psi_0 | U(t, -\infty) | \Psi_0 \rangle \quad (5.3.24)$$

The vacuum term generates the complex level shift involving the secular and normalization components :

$$\langle \Psi_0 | U(t, -\infty) | \Psi_0 \rangle = \exp[-i(E_0 t + e(t))] \quad (5.3.25)$$

where  $e(t)$  is given by

$$e(t) = \lim_{\alpha \rightarrow 0} \int_{-\infty}^t \langle \Psi_0 | H_{\text{int}} U(t', -\infty) | \Phi_0 \rangle dt' \quad (5.3.26)$$

Using eq.(5.3.26), eq.(5.3.22) may be written as

$$[H + H_{\text{int}}(t) - (E_0 + e(t))] U_r \Psi_0 = i \partial U_r(t, -\infty) / \partial t \Psi_0 \quad (5.3.27)$$

Monkhorst/56/ (see also Dalgaard and Monkhorst/57/) postulated the following ansatz for  $U(t, -\infty)$  :

$$U_r(t, -\infty) = \exp(S(t)) \quad (5.3.28)$$

with  $S$  as various nh-np excitations. This, coupled with the CC ansatz for  $\Psi_0$ , leads to the time-dependent CC equation :

$$[H + H_{\text{int}}(t) - (E_0 + e(t))] \exp(T + S(t)) \Phi_0 = i \partial S(t) / \partial t \exp(T + S(t)) \Phi_0 \quad (5.3.29)$$

To compute the linear response function, it suffices

to expand  $S(t)$  up to the harmonic terms involving  $\omega$  :

$$S(t) \approx S^+ \exp(i\omega t) + S^- \exp(-i\omega t) \quad (5.3.30)$$

Premultiplying eq.(5.3.39) by  $\exp(-T-S)$ , retaining terms up to those linear in  $S$ , and projecting onto the states  $\langle \Phi_a |$ , we have

$$\langle \Phi_a | [\bar{H}, S^{\pm}] + \bar{V}^{el} \pm \omega S^{\pm} | \Phi_0 \rangle = 0 \quad (5.3.31)$$

which shows that the linear response function obtained from  $S(t)\Psi_0$  will have the same pole-structure as that obtained from CC-LRT, since eq.(5.3.31) is entirely equivalent to eq.(5.3.10) of CC-LRT. Monkhorst/56/ discussed only the case of EE calculation and did not elaborate on the scheme further. The computation of the transition moments was also touched upon, and later elaborated by Dalgaard and Monkhorst/57/. It should be mentioned here that the computation of the transition moments in the context of residues of dynamical polarizability was also indicated in the CC-LRT framework by Mukherjee and Mukherjee/58/. The time-dependent linear response version may be viewed as the fourier transformed counter part of the CC-LRT. The proper treatment of the secular terms in the time dependent treatment is a rather tricky business. This is entirely bypassed in a time-independent approach.

Takahashi and Paldus /105/ recently formulated the spin-adapted version of the time-dependent formalism for computing singlet as well as triplet EE. In a pilot  $\pi$ -electron calculation, they computed the full  $\bar{H}$  with the two-body term of  $T$ . They did not precompute  $\bar{H}$ , however, which is itself a scalar in both point group and spin. They instead spin-adapted each diagram containing  $H$ ,  $T_2$  and  $W/W^\dagger$ . They noted the close similarity of their approach with that of Ghosh *et al* /58/ and endorsed Ghosh *et al* /58-60/ on the desirability of a prior computation of  $\bar{H}$ . Calculations on the singlet and triplet EE with  $S \approx S_1$  have earlier been performed by Sekino and Bartlett/61/.

Recently Geerstsens and Oddershede /106/ utilized the CC ground state function for calculating EE using polarization propagator. This approach may also be viewed as a semi-cluster expansion strategy. However a clear connection with the wave function approach is difficult to establish in a propagator theory (unless it is consistent /63,84/) and we shall not elaborate further on this theory.

## 6. FULL CLUSTER EXPANSION THEORIES IN HILBERT SPACE

## 6.1. General Preliminaries

As explained in Sec.2, the full cluster-expansion theories in Hilbert space are designed to compute wave-functions for the open-shell states that are explicitly size-extensive with respect to the total number of electrons  $N$ . The underlying cluster structure of all these developments is what was envisaged by Silverstone and Sinanoglu/48/:

$$\Psi_k = \sum_{i=1}^M C_{ik} \exp(T^i) \Phi_i \quad (6.1.1)$$

As discussed in Sec 2., the most commonly used strategy is to take a set of model space functions  $\{\Phi_i\}$  that is 'complete' with respect to valence occupancies. The cluster operators  $T^i$  in that case generate from the model space determinants  $\Phi_i$ , the virtual determinants  $X_1$  by exciting core orbitals to valence/particle orbitals and, in addition, some valence orbitals to particle orbitals. Since it is possible to produce a particular determinant  $X_1$  from several  $\Phi_i$ 's by the action of the appropriate components of the  $T^i$  operators, the overall amplitude of a particular  $X_1$  comes out as a combination of several  $T^i$  amplitudes. If we explicitly indicate by an additional label  $l$  the determinant  $X_1$  that is generated by a  $T^i$ , then we may write

$$T^i = \sum_l T^{i,l} \quad (6.1.2)$$

where

$$T^{i,l} \Phi_i = t^{i,l} X_1 \quad (6.1.3)$$

Using eqs. (6.1.2) and (6.1.3), eq.(6.1.1) can be written as

$$\begin{aligned} \Psi_k &= \sum_i C_{ik} \Phi_i + \sum_l \sum_i [C_{ik} t^{i,l}] X_1 + \dots \\ &= \sum_i C_{ik} \Phi_i + \sum_l d_{lk} X_l + \dots \end{aligned} \quad (6.1.4)$$

Let us now note carefully the chief structural

difference between the cluster expansion for the closed shells and the Silverstone-Sinanoglu cluster expansion for the open shells. If we confine our attention to only one state  $\Psi_k$ , configurational coefficients  $d_{ik}$  accompanying the determinants  $X_i$  are the only unknown variables appearing in the Schrodinger equation for  $\Psi_k$ . Eq.(6.1.4) shows that these coefficients are some particular linear combinations of the  $t_{i,l}^{i,l}$  amplitudes. The set of amplitudes  $\{t_{i,l}^{i,l}\}$  for all  $i$  and  $l$  are thus linearly dependent in  $\Psi_k$  and there is no dynamical way to determine them from the Schrödinger equation for  $\Psi_k$  only. In their applications/48,110/, Silverstone and Sinanoglu used an 'anonymous parentage approximation' scheme where they assumed that all the  $t_{i,l}^{i,l}$  amplitudes for a given  $l$  are independent of  $i$ , which reduced the number of variables in  $T$  to the number of independent coefficients of  $X_i$  of  $\Psi_k$ . Introducing projectors  $P_i = |\Phi_i\rangle\langle\Phi_i|$ , eq.(6.1.1) be written as

$$\Psi_k = \sum_{i=1}^M C_{ik} \exp(T^i) \Phi_i = \Omega \Psi_k^0 \quad (6.1.5)$$

where

$$\Psi_k^0 = \sum_{i=1}^M C_{ik} \Phi_i, \quad (6.1.6)$$

$$\Omega = \sum_{i=1}^M [\exp(T^i)] P_i \quad (6.1.7)$$

In the anonymous parentage approximation,  $T^i \simeq T$  for all  $i$ , and eq.(6.1.7) reduces to

$$\Omega \simeq \exp(T) \quad (6.1.8)$$

which has the form of the cluster expansion for the closed shells. Silverstone and Sinanoglu advocated the choice of  $\Psi_k^0$  as the CAS-SCF function (called by them as the generalized RHF function /110/). Since  $T$ 's can generate only the virtual set  $\{X_i\}$ , the coefficients  $\{C_{ik}\}$  accompanying  $\{\Phi_i\}$  in the exact  $\Psi_k$  remain frozen their CAS-SCF values.<sup>1</sup> This introduces an additional approximation.

Another way to circumvent the linear dependency of the  $T$  amplitudes is suggested by the more recent Multireference CI (MR-CI) developments/111,112/. In an



MR-CI approach, one superposes virtual determinants  $X_i$  that are obtained by single, double, ... excitations out of the set  $\{\Phi_i\}$  spanning the reference space. Denoting the various operators inducing excitations as  $Y_i^\dagger$ , it is found that the set  $\{Y_i^\dagger \Phi_i, \forall i \text{ and } m\}$  is linearly dependent, and one must choose only a linearly independent subset thereof for performing the CI. An analogous cluster expansion can thus be envisaged, where  $\Omega$  is taken to be of the form eq.(6.1.8), but  $T$  is now confined to an expression of the form

$$T = \sum_m' t^m Y_m^\dagger \quad (6.1.9)$$

where the sum  $\sum'$  runs over the linearly independent excitations. Since this choice can be made in many ways, a unique prescription demands that we specify explicitly which of the operators  $\{Y_i^\dagger\}$  are chosen as forming the linearly independent set. The simplest way to achieve this is to consider each  $X_i$  in turn, and select from among the possible  $Y_i^\dagger$  operators of the form  $Y_{i+1}^\dagger$  only one, which fixes the  $i$ . The genesis of  $X_i$  is thus biased towards a fixed  $i$ , and we may call this scheme as a 'preferred parentage approximation'.

The way out of this difficulty was shown in the context of open-shell MBPT by Brandow/3/, whose strategy is not, however, confined to perturbative theories only. If we consider all the  $M$  functions  $\Psi_k$  that can be generated from  $M$  linearly independent starting functions  $\Psi_k^0$ , then the totality of the Schrödinger equations for  $M$  such  $\Psi_k$ 's have precisely the same number of unknown combining coefficients  $\{d_{ik}\}$  as the amplitudes  $\{t^{i,1}\}$ , since the number of functions  $\{\Phi_i\}$  and the number of roots  $E_i$  are both equal to  $M$ . In this case, then, the amplitudes  $\{t^{i,1}\}$  can, in principle, be rigorously determined from the  $M$  Schrödinger equations. Furthermore, if we determine the coefficients  $\{C_{ik}\}$  from the Schrödinger equation for  $\Psi_k$ 's, rather than from a prior CAS-SCF calculation, then the sets  $\{C_{ik}\}$  and  $\{t^{i,1}\}$  are both exact in principle. The effective hamiltonian approach envisaged by Brandow/3/ achieved precisely this in the MBPT framework. It should be noted that Brandow's approach used a size-extensive description of the type (a3), rather than (a2). But this approach to generate the functions  $\Psi_k$  for  $k=1, M$  is equally relevant to the theories of the type (a2). Since  $T^1$ 's can only excite, each  $\Psi_k$  can be written as

$$\Psi_k = \Psi_k^0 + \theta_k \quad (6.1.10)$$

where  $\theta_k$  contains only the set  $\{X_1\}$ , and is given by

$$\theta_k = \sum_i C_{ik} \sum_{n=1} (T^i)^n / n! \Phi_i \quad (6.1.11)$$

The functions  $\Psi_k$  satisfy the intermediate normalization with respect to the unperturbed functions  $\Psi_k^0$ :

$$\langle \Psi_k | \Psi_k^0 \rangle = 1 \quad \forall k=1, M \quad (6.1.12)$$

If we now consider the generalized Silverstone-Sinanoglu strategy discussed above, then the most natural way to proceed is by way of an effective hamiltonian formalism. We introduce a single (state-universal) wave-operator  $\Omega$ , whose action on  $\Psi_k^0$ 's produce the functions  $\Psi_k$ , defined by eq.(6.1.1) or (6.1.7), and write Schrödinger equations for  $\Psi_k$ 's as

$$H\Omega\Psi_k^0 = E_k\Omega\Psi_k^0 \quad \forall k=1, M \quad (6.1.13)$$

Introducing two idempotent and mutually exclusive projectors  $P$  and  $Q$  for the spaces spanned by  $\{\Phi_i\}$  and  $\{X_1\}$ , eq.(6.1.13) can be converted into an equivalent operator equation. Eq.(6.1.12) implies

$$P\Psi_k = \Psi_k^0 \quad (6.1.14)$$

so that  $\Psi_k^0$ 's are the model space projections of the eigenfunctions. If we assume that the functions  $\Psi_k^0$  are linearly independent, then the matrix of coefficients  $\{C_{ik}\}$  is invertible, and eq.(6.1.13) can be converted to

$$H\Omega P = \Omega H_{\text{eff}} P \quad (6.1.15)$$

where  $H_{\text{eff}}$  is given by

$$H_{\text{eff}} = C E C^{-1} \quad (6.1.16)$$

with  $E$  as the diagonal matrix of dimension  $M$  with entries  $E_k$  in the diagonals. Since  $\Omega$  converts  $\Psi_k^0$  to  $\Psi_k$ , eq.(6.1.14) also implies that

$$P\Omega P = P \quad (6.1.17)$$

which is the operator equivalent of the intermediate normalization for  $\Psi_k$ . Using eq.(6.1.17), it follows that

$$H_{\text{eff}} = PH\Omega P, \quad (6.1.18)$$

so that the eq.(6.1.13) is equivalent to

$$H\Omega P = \Omega PH\Omega P \quad (6.1.19)$$

Equation of this form for an open-shell situation was apparently first considered by Bloch/116/, and is now known as the Bloch equation/4/. Eq.(6.1.16) indicates that the diagonalization of  $H_{\text{eff}}$ , which is defined in the model space only, will furnish us with the eigenvalues  $E_k$  and the associated coefficients  $\{C_{ik}\}$ .  $H_{\text{eff}}$  can hence be called an effective hamiltonian. The unknown component of  $\Omega$  in eq.(6.1.19) is  $Q\Omega P$ , since  $P\Omega P$  is already fixed by eq.(6.1.19). This can be determined from the  $Q$ -projection of eq.(6.1.19)

$$QH\Omega P = Q\Omega PH\Omega P \quad (6.1.20)$$

Eqs. (6.1.18) and (6.1.19) form the starting point of several open-shell many-body formalisms. Thus Lindgren/4/ developed his version of the open-shell MBPT by expanding these equations order by order in perturbation theory, and Offermann et al, Ey/70/ and Jeziorski and Monkhorst/66/ based their corresponding coupled cluster formalisms on them. Mukherjee et al/67-69/ and Kutzelnigg/76/ and Kutzelnigg and Koch/77/ also made use of these equations - implicitly - they first premultiplied eq.(6.1.14) by  $\Omega^{-1}$  and then projected the resultant equations by  $P$  and  $Q$ .

For the cluster expansion of the type (a2), we may thus conceive of two approaches. One is to invoke a single root strategy, and use either anonymous parentage or preferred parentage approximation. This approach by-passes the need for  $H_{\text{eff}}$ , but by its very nature cannot generate a potentially exact formalism. The other is to use the multi-root strategy through the Bloch equation, and thereby produce a formally exact theory. We shall review these two types of schemes in Secs. 6.2 and 6.3 respectively.

## 6.2 Single Root Hilbert Space Formalisms

Banerjee and Simons/64/ developed a single root

formalism that utilizes a wave-operator structurally very similar to that in the closed-shell CC theory. They started from a CAS-SCF wave-function  $\Psi_k^0$ , and wanted to incorporate only the valence correlations by exciting the valence orbitals to the particle orbitals. The core thus remained frozen. For the cluster wave operator  $\Omega$ , they chose the ansatz, eq.(6.1.8), and used a selection scheme for the various operators in  $T$  which is essentially the same as the preferred parentage approximation.

They chose  $T$  to be of the form

$$T = \sum_n T_n \quad (6.2.1)$$

where  $T_n$  is an  $n$ -body excitation operator with the structure

$$\begin{aligned} T_n &= 1 / (n!)^2 \sum_{\substack{p,q \\ \bar{p},\bar{q}}} \langle pq \dots | t_n | \bar{p}\bar{q} \dots \rangle_a \{ a_p^\dagger a_q^\dagger \dots a_{\bar{q}}^- a_{\bar{p}}^- \} \\ &= \sum_{m,n} t_{m,n} Y_{m,n}^\dagger \end{aligned} \quad (6.2.2)$$

where the sets  $m = (p, q, \dots, \bar{p}, \bar{q})$  are chosen in such a way that  $\{T_n \Phi_1\}$  form a linearly independent set. For determining the cluster amplitudes of  $T$ , they used a projection method :

$$\langle \Psi_k^0 | Y_{m,n} \exp(-T) H \exp(T) | \Psi_k^0 \rangle = 0, \quad \forall \text{ linearly independent } m \quad (6.2.3)$$

Several simplifications occur in eq.(6.2.3) as a consequence of the approximations involved. Firstly, one-valence excitation amplitudes of  $T_1$  are expected to be small, since upto the first order the single excitations  $\langle \Psi_k^0 | Y_{m,n}^\dagger$  do not mix with the CAS-SCF  $\Psi_k^0$  (the Generalized Brillouin Theorem)/64/. They can thus be neglected. This is entirely analogous to the situation for the closed-shells, where the  $T_1$  operator is neglected for the HF function  $\Phi_0$  as the starting point. Secondly, the components of the  $T$  operator commute among themselves, again just as in the closed-shell case, and as a result the Hausdorff expansion in powers of  $T$  in eq (6.2.3) terminates at the quartic term.

The energy  $E_k$  can be found out from

$$E_k = \langle \Psi_k^0 | \exp(-T) H \exp(T) | \Psi_k^0 \rangle \quad (6.2.4)$$

In their applications/64,65/, Banerjee and Simons used unitary-group generators/113,114/ in a spinfree form to adapt the resulting CC equations to proper spins, and applied this formalism for studying various small prototypical systems - starting with  $H_2$  molecule at large internuclear separation, to the singlet-triplet splitting of  $CH_2$  and obtaining the potential curve for the symmetrical  $H_2$  abstraction reaction from  $BeH_2$ .

The neglect of the core excitations may not be computationally justified in many cases - particularly the semi-internal ones/64/, as has been found in many MR-CI results, but their inclusion will make the resultant formalism much more involved. Such a generalization has nevertheless been attempted recently by Laidig *et al*/115(a)/, who computed the potential curves for  $N_2$  including the semi-internal excitations as well at the linearized level (i.e.,  $\exp(-T) H \exp(T)$  truncated at the first commutator). Although Laidig *et al* couch their formulation in terms of the equations of the multi-root approach of Jeziorski and Monkhorst/66/, the approximations used by them are precisely the same as those implied by the preferred parentage approximation, and they generate a single root theory. Laidig *et al*/115(a)/ include in  $T$  the semi-internal excitations, which makes the virtual space  $\{X_1\}$  somewhat bigger. The equations they invoke for solving the  $T$  amplitudes are of the following form

$$\langle X_1 | H + [H, T] | \Psi_k^0 \rangle = 0 \quad (6.2.5)$$

$T$  is constructed from only those operators  $Y_i^\dagger$  which produce a linearly independent set  $\{Y_i^\dagger \Phi_i\}$ . For obtaining the coefficients  $\{C_{ik}\}$ , Laidig *et al* advocated a projection onto the functions  $\{\Phi_i\}$ , leading to

$$\begin{aligned} \sum_j \tilde{H}_{ij} C_{jk} &= \sum_j \langle \Phi_i | (H + [H, T]) | \Phi_j \rangle C_{jk} \\ &= E_k C_{ik} \end{aligned} \quad (6.2.7)$$

which shows that a particular root of the matrix  $\tilde{H}$  is the desired eigenvalue  $E_k$ . The other roots are spurious. This contrasts with the multi-root strategy,

where all  $T_i$ 's are used to generate a set of  $\Psi_k^0$ 's. Very recently Hoffman and Simons/115(b)/ used a similar strategy, where they used a unitary cluster operator  $\exp(T-T^\dagger)$  instead of  $\exp(T)$  to produce a hermitian matrix  $L$ .

### 6.3 The Multi-root Hilbert Space Formalisms

Jeziorski and Monkhorst/66/ developed a rigorous cluster expansion formalism for the manifold of states  $\{\Psi_k\}$ , starting from a complete model space, using the extended Silverstone-Sinanoglu formalism outlined in eq.(6.1.7) and (6.1.10) to (6.1.20).  $\Omega$  of eq.(6.1.7) has a hybrid structure: there are creation/annihilation operators in  $T_i$ 's, and there are in addition explicit  $N$ -electron determinantal projectors  $P^1$ . Following Hose and Kaldor/44/, Jeziorski and Monkhorst/66/ chose each  $\Phi_i$  on which the operator  $\exp(T_i)$  acts as the vacuum for calculating the matrix-elements. With this choice, it is sufficient to indicate in each  $T_i$  only the orbitals vacated in  $\Phi_i$  and the orbitals replaced to reach a virtual determinant  $X_i$ . Thus, in this scheme it is more useful to classify the orbitals into holes and particles with respect to each  $\Phi_i$  for any operator acting on  $\Phi_i$ . This, somewhat unorthodox, procedure has both advantages and limitations, as we shall observe at an appropriate place later.

An  $n$ -body  $T_i$  operator can be written as

$$T_n^i = 1/(n!)^2 \sum_{\substack{c_1, c_2 \dots c_n \\ a_1, a_2 \dots a_n}} \langle c_1 c_2 \dots c_n | t_n^i | a_1 a_2 \dots a_n \rangle a_{c_1}^\dagger a_{c_2}^\dagger \dots a_{c_n}^\dagger a_{a_1} a_{a_2} \dots a_{a_n} \quad (6.3.1)$$

An excited state function  $X_i$  reached by the product  $a_{c_1}^\dagger a_{c_2}^\dagger \dots a_{c_n}^\dagger a_{a_1} a_{a_2} \dots a_{a_n}$  acting on  $\Phi_i$  may be denoted by  $|c_1 c_2 \dots c_n(i)\rangle$ . For the determination of cluster amplitudes of  $T_i$ 's, Jeziorski and Monkhorst started from the Bloch equation/116/, eq.(6.1.19), projected it on the ket  $|\Phi_i\rangle$  and premultiplied it with  $\exp(-T_i)$ , to finally obtain

$$\begin{aligned} \langle c_1 c_2 \dots c_n(i) | \exp(-T_i) H \exp(T_i) | \Phi_i \rangle \\ = \sum_{j \neq i} \langle c_1 c_2 \dots c_n(i) | \exp(-T_i) \exp(T_j) | \Phi_j \rangle \\ \langle \Phi_j | \exp(-T_i) H \exp(T_i) | \Phi_i \rangle \quad (6.3.2) \end{aligned}$$

where use is made of the fact that  $\langle \Phi_i | \exp(-T^i) = \langle \Phi_j |$  for  $i \neq j$  for a complete model space. The last term on the right of eq. (6.3.2) may be symbolically denoted as  $\langle \Phi_j | H_{\text{eff}} | \Phi_i \rangle$ . The left side of eq. (6.3.2) is connected if  $T^i$  is connected, but the connectedness of the right side is not immediately manifest. In fact, a special proof is needed to prove that  $T^i$ 's are connected. The difficulty lies in proving that  $(\exp(-T^i) \exp(T^j))$  is connected. Since  $T^i/T^j$  have no reference to orbitals unaffected by them, the book-keeping of the terms becomes recondite in nature. The use of multiple vacua makes clear diagrammatic rules hard to formulate. This is unfortunate, since a proof of the connectivity of an operator becomes quite easy in the diagrammatic representation. Moreover, since the vacua  $\Phi_i$  may not be spin singlets, the operators  $T^i$  are not spin-scalars, and spin-adaptations of the CC equations (6.3.2) are quite complex. A positive advantage of the formalism is that, with respect to  $\Phi_i$  as the vacuum, the ranks of the operators in  $T^i$  are controlled by the orbitals vacated and replaced in  $\Phi_i$ , so that the operators in  $T^i$  thus carry no redundant or spectator labels, and this has advantages in book-keeping. The perturbative expansion of eq. (6.3.2) generates the complete model space version of the Hose-Kaldor formalism/44/.

Laidig and Bartlett/118(a)/ have implemented the theory in its linearized form. They applied it to the symmetric dissociating potential curve for the ground state of water and the symmetric  $H_2$  abstraction from  $BeH_2$ . Each  $T^i$  is truncated at the two-body level. The results were compared to the corresponding MR-CISD values. To our knowledge, a quadratic extension of the formalism has not been applied computationally. A spin-adapted version of the linearized model has also been recently formulated/118(b)/.

## 7. FULL CLUSTER EXPANSION THEORIES IN FOCK SPACE

### 7.1. Preliminaries For a Fock Space Approach

In this section, we shall motivate towards the need for a Fock-space approach to generate core-valence extensive cluster expansion theories (i.e., of type (a3)), introduced in Sec.2. Since we have to maintain size-extensivity of the energy

explicitly with respect to both  $N_c$  and  $N_v$ , there should be a cluster expansion involving cluster operators capable of correlating the core and also of introducing core-valence interaction and valence correlation effects involving  $1, 2, \dots, N_v$  valence electrons. The degree of flexibility in the wave-operator is thus much more than what is needed to just generate the functions  $\Psi_k^{(N_v)}$  from an  $N_v$ -valence model space spanned by  $\{\Phi_i\}$ . Consequently, if we confine ourselves to a particular  $N_v$ -valence Hilbert space, then the number of cluster amplitudes in the wave-operator become linearly dependent. This is somewhat analogous to what was encountered in Sec. 6.2 and 6.3 describing the single root and multiroot cluster-expansions of the type (a2): the Silverstone-Sinanoglu expansion involves linearly dependent cluster amplitudes for one function  $\Psi_k$ ; this redundancy is eliminated if we consider all the  $M$  functions  $\{\Psi_k\}$  that can be generated from the  $M$ -valence model space. In a similar manner, the redundancy in the cluster amplitudes of the cluster expansion of the type (a3) can be eliminated if we consider not only the Schrödinger equation for the  $N_v$ -valence problem but also all the lower  $n$ -valence subduced problems, with  $0 \leq n \leq N_v$ . The simplest procedure is to postulate the existence of a single wave-operator which generates the various exact functions  $\Psi_k^{(n)}$  for all the  $n$ -valence problems/68/, with  $0 \leq n \leq N_v$ .

$$H\Omega\Psi_k^{(n)} = E_k^{(n)}\Omega\Psi_k^{(n)} \quad ; \quad \forall k=1, M^{(n)}; \quad 0 \leq n \leq N_v \quad (7.1.1)$$

$\Omega$  can thus be regarded as a valence universal wave-operator/69/. Denoting by  $S^{(m)}$ , the cluster operator of valence rank  $m$  (i.e., having  $m$  valence-destruction operators),  $\Omega$  should be taken of the form

$$\Omega \sim \exp\left(\sum_{m=0}^{N_v} S^{(m)}\right) \quad (7.1.2)$$

for treating up to the  $N_v$  valence problem. The amplitudes of  $S^{(0)}$  correspond to the core-correlation cluster amplitudes, and  $S^{(1)}$ ,  $S^{(2)}$ , etc., successively introduce valence correlation and core-valence interaction effects containing one, two, ...,  $N_v$  valence electrons. The model space for the  $n$ -valence problem is spanned by the set  $\{\Phi_i^{(n)}\}$ .

There are two possible ways to generate the CC-



equations from eq.(7.1.1). One way is to compute for the cluster amplitudes of S by solving the projected equations of the form

$$\langle X_1^{(n)} | \Omega^{-1} H \Omega | \Phi_1^{(n)} \rangle = 0 \quad \forall 1, i \text{ and } n \quad (7.1.3)$$

The energies are finally obtained from the equations

$$\sum_j \langle \Phi_1^{(n)} | \Omega^{-1} H \Omega | \Phi_j^{(n)} \rangle C_{jk}^{(n)} = E_k^{(n)} C_{ik}^{(n)} ; \quad \forall n, k \quad (7.1.4)$$

Another way is to start from the precursor of the Bloch equation, eq.(6.1.15), generalized to Fock space/69,76,93/ and use the Q and P projections for the n-valence sector. Writing the Fock space Bloch equation as

$$H \Omega P^{(n)} = \Omega P^{(n)} H_{\text{eff}} P^{(n)} \quad \forall n \quad (7.1.5)$$

with  $H_{\text{eff}}$  a valence-universal Fock space effective hamiltonian, the equations determining the cluster amplitudes can be obtained from

$$Q^{(n)} H \Omega P^{(n)} = Q^{(n)} \Omega P^{(n)} H_{\text{eff}} P^{(n)} \quad \forall n \quad (7.1.6)$$

and  $H_{\text{eff}}$  is obtained from

$$P^{(n)} H \Omega P^{(n)} = P^{(n)} H_{\text{eff}} P^{(n)} \quad \forall n \quad (7.1.7)$$

when  $\Omega$  satisfies 'intermediate normalization':

$$P^{(n)} \Omega P^{(n)} = P^{(n)}, \quad \forall n \quad (7.1.8)$$

Both methods have been tried, and each has its own operational advantages. We shall classify these methods into two categories:

(a) those utilizing eqs.(7.1.3) and (7.1.4) will be called 'similarity transformation-based Fock space theories' (See, Sec.7.2) and (b) those using eqs. (7.1.6) and (7.1.7) will be called Bloch equation-based Fock space theories (Sec. 7.3). It should, however, be observed that the starting equations for the two types of theories are eqs. (7.1.5), since a pre-multiplication by  $\Omega^{-1}$  followed by  $Q^{(n)}$  and  $P^{(n)}$  projection furnishes eqs.(7.1.3) and (7.1.4) respectively. For truncated calculations, the two approaches will, however, differ and their convergence

behaviour will also be different. A variational strategy is also possible where the cluster amplitudes in  $\Omega$  are determined by varying the energies with respect to the cluster amplitudes. This is covered in Sec.7.4.

The various Fock space methods that we shall describe principally differ in their choice of the actual form of the cluster operator. Following the earlier analysis of Primas/29/, the form of  $\Omega$  will be of exponential type, symbolically represented by eq.(7.1.1), but it may not strictly be exponential. There have been several choices for  $\Omega$ , each having its own advantages.

The Fock space approach has the potential advantage of exploiting the fact that operators written in the occupation number representation are independent of electron number, so that all the manipulations involving  $H$  and  $\Omega$  can be performed at the operator level first, which is somewhat simpler and more transparent than working with the matrix-elements involving functions  $\{\Phi_i^{(n)}\}$  and  $\{X_i^{(n)}\}$ . Furthermore, since the same operators appear in all the  $n$ -valence sectors of the Fock-space, the projections onto the wave-functions can be performed at the very end, and it is possible to treat systems with varying  $n$  (i.e., different degrees of ionization) on the same footing, and in an explicitly size-extensive manner with respect to the valence electrons. As it has turned out, only the Fock-space approach has the potentiality to furnish explicitly connected size-extensive theories for a general incomplete model space/91-96/, which shows its flexibility and generality.

## 7.2 Similarity Transformation-Based Fock-Space Theories

Mukherjee et al/68-69/ formulated an explicitly connected CC theory for complete model space by invoking a valence universal cluster operator  $\Omega$  of the form

$$\Omega = \exp(S) \quad (7.2.1)$$

with  $S$  given as a sum over various  $S^{(m)}$  operators of valence rank  $m$  :

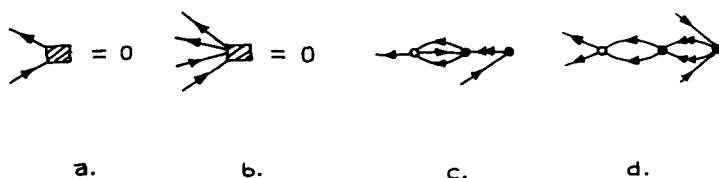


Fig.6 (a),(b) L blocks for the zero-valence problem  
 (c) A typical diagram for (a). (d) A typical  
 diagram for (b). Note that a two-valence S can  
 contribute to a zero-valence block, as in 6(d).

$$S = \sum_{m=0}^{N_v} S^{(m)} \quad (7.2.2)$$

In their first exposition, Mukherjee *et al*/67/ did not include in  $S^{(m)}$  operators with passive valence lines, which was rectified in the subsequent papers/68,69/. Actually it is the presence of passive valence lines with different valence labels attached to them that distinguishes the different scattering amplitudes that involve the same change in occupancy but connect different pairs of functions  $X_1^{(m)}$  and  $\Phi_1^i$ . To prove the connectivity of the cluster operator  $S^i$ , Mukherjee *et al*/68-69/ chose a hierarchical strategy of proceeding from the zero-valence core problem and going upwards by successively considering one, two,... valence problems, stopping finally at the desired  $N_v$  valence level. Since  $S$  contains both valence creation and destruction operators, the components of  $S$  do not commute among themselves and, as a result, in the expression of  $L = \Omega^{-1} H \Omega$  in eqs.(7.1.13) and (7.1.14) using eq.(7.2.1), we have many  $S$ - $S$  contractions. The multi-commutator Hausdorff formula for the transformed hamiltonian  $L$  connected terms only, and it has operators of varying valence rank. Diagrammatically, each component of  $L$  can be depicted as a composite block of a given shape with valence destruction operators entering from the right, and hole, valence and particle operators emanating from the left. The block consists of all the connected diagrams that can be obtained by joining  $H$  with various  $S$  vertices from left and right in a specific manner : From the multi-commutator expansion, it follows that we should start out from  $H$  and begin connecting the  $H$  vertex

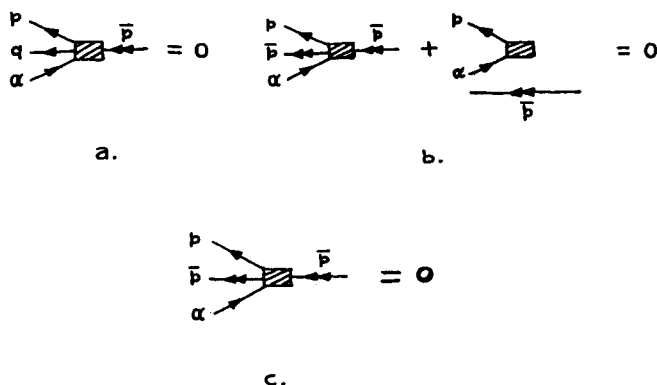


Fig.7 (a) One-valence two body block with no common valence orbitals before and after scattering. (b) One-valence block with a common valence label. (c) Reduction of (b) due to a lower valence equation (Fig.6(a)).

with  $S$  vertices on its immediate left or right and move out systematically connecting the next  $S$  neighbours and proceed without leaving any intermediate  $S$  uncontracted during the process.

For the zero-valence core problem, eq.(7.1.3) can be diagrammatically depicted as in Figs.6(a) and (b) for illustration. Although they look like the closed shell CC equations in shape, there is a major difference. Owing to the  $S$ - $S$  contractions, many  $S^{(n)}$  operators of valence rank greater than zero will also contribute to eq.(7.1.3) for  $n=0$ . A few typical diagrams are shown in Figs.6(c) and (d). This shows that the cluster operators for the various  $m$ -valence Hilbert space sectors are all coupled together, and eqs.(7.1.3) for all  $n$  have to be simultaneously considered to get the cluster amplitudes. Proceeding to the one valence problem, eq.(7.1.3) generates diagrams which may be of two types. In one type, the valence orbitals in  $\Phi_i$  are vacated, to be replaced by another valence or particle orbital. The  $L$  diagrams in this case are just blocks of valence rank 1, with no valence label on the entering lines from the right of the block ever appearing on the lines to the left. A typical such diagram will look as in Fig.7(a). For the others type,  $X_1^{(n)}$  and  $\Phi_i^{(n)}$  share some common

valence orbital. In this case, eq.(7.1.3) generates two types of blocks, whose sum vanishes. In one type, some valence lines actually enter the block, but they leave on the left carrying the same labels ('passive scattering'), as shown in the first block of Fig.7(b), and there is another type of block in which the valence lines do not enter the diagram at all ('spectator scattering'), as shown for the second block of Fig.7(b) for the one-valence situation. Since the spectator line does not contribute to the matrix-elements, the last block independently vanishes from Fig. 6(a), corresponding to the equations for the core-problem. Thus, eq.(7.1.3) in this case also reduces to a block of valence rank 1 equated to zero, as shown in Fig.7(c). Thus irrespective of whether we have valence orbitals common to  $X_1^{(1)}$  and  $\Phi_1^{(1)}$ , the resultant eqs.(7.1.3) are all one-valence  $i_L$  blocks equated to zero. Proceeding upwards, it easily follows that the eqs.(7.1.3) for an n-valence problem will look like a block of valence rank n, equated to zero. Since in eq.(7.1.3), the diagrams entering the block are all connected, and are of same shape, the  $S^{(m)}$  operators are all connected. This is the first linked cluster theorem for an open shell CC theory of the type (a3) using similarity transformation.

Since the operator  $S$  and its powers always excite to the virtual space, the functions  $\Psi_k^{(n)}$  are generated in the intermediate normalization. The valence universal nonhermitian  $H_{eff}$  can be diagrammatically depicted as blocks with incoming and outgoing lines labelled by valence only. Typical one-valence and two valence blocks are shown in Figs. 8(a) and (b), and some representative diagrams contributing to these are shown in Figs. 8(c) and (d). We note here that  $S^{(2)}$  has appeared in the one-valence block, Fig. 8(c), and this indicates how the various m-valence blocks are coupled together. If we drop the constant closed zero-valence part of  $H_{eff}$ , whose value is equal to the correlated core energy, from eqs.(7.1.4), then we shall get energy differences with respect to the core  $\Delta E_k^{(n)}$  directly.

Mukherjee et al/68/ utilized this strategy to calculate ground state energy, its IP and excitation energy for a model  $\pi$ -electron problem by starting from the doubly charged molecular ground state as the core. Mukherjee et al/67/ also advocated a strategy whereby one may generate a hermitian  $H_{eff}$ . In this formulation, one needs a valence universal unitary cluster operator  $\exp(\sigma)$ , with  $\sigma$  given by

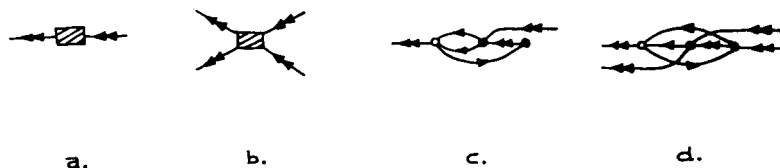


Fig.8 (a),(b) One- and two-valence parts of  $H_{\text{eff}}$ .  
 (c),(d) Typical terms contributing to (a),(b).

$$\sigma = S - S^\dagger \quad (7.2.3)$$

where  $S$  has the same form as in eqs.(7.2.1) and (7.2.2). Since  $\sigma$  is more complex than  $S$ , the corresponding equations will contain more terms. This approach was adopted in /120/ to generate a formal theory of a hermitian valence-shell hamiltonian, where the connectivity of the effective hamiltonian was also explicitly established. There is also a related work by Westhaus *et al*/121/, where the valence universality of  $\Omega$  is not mentioned and the logical basis of the connectivity remains somewhat obscure. As we have already discussed, in the formulation using the ansatz, eq.(7.2.1), the calculation of the energy differences  $\Delta E_k^{(n)}$  remains coupled with the calculation of the correlated core energy  $E_0$ . This contrasts with the situation of RS MBPT for open-shells/3-9/, where the core energy  $E_0$  can be determined without reference to the open-shell problems, which enter in the next stage. This decoupling was termed 'the core-valence separation' by Brandow/3/. A generalization of the cluster ansatz for which allows such a core-valence separation was formulated later by Mukherjee *et al*/121/. In this version,  $\Omega$  is written in the form

$$\Omega = \exp(T_C)\exp(T_V) \quad (7.2.4)$$

where  $T_C$  has the form of the closed-shell cluster operator, and  $T_V$  contains all the operators  $S^{(n)}$  with  $1 \leq n \leq N_V$ . The closed-shell cluster amplitudes are first solved from the closed-shell CC equations for the core  $\Phi_0$ :

$$\langle X_1^{(0)} | \exp(-T_C) H \exp(T_C) | \Phi_0 \rangle = 0 \quad (7.2.5)$$

The various  $m$ -valence problems are handled simultaneously, with the core cluster-amplitudes

frozen at their closed-shell values. Introducing the dressed hamiltonian  $\bar{H}$  through the relation :

$$\bar{H} = \exp(-T_c) H \exp(T_c) \quad (7.2.6)$$

and separating its vacuum part  $E_0$  explicitly,

$$\bar{H} = E_0 + H_c \quad (7.2.7)$$

we find that the equations for the various  $m$ -valence sectors can be written as

$$\langle X_i^{(m)} | H_c | \Phi_i^{(m)} \rangle = 0, \quad \forall i, l, \quad 1 \leq m \leq N_V \quad (7.2.8)$$

The energy difference  $\Delta E_k^{(m)}$  are obtained from the eigenvalue equations

$$\sum_j \langle \Phi_i^{(m)} | H_c | \Phi_j^{(m)} \rangle C_{jk}^{(m)} = \Delta E_k^{(m)} C_{ik}^{(m)} \quad (7.2.9)$$

The calculation of the core cluster amplitudes, eq.(7.2.5), are thus completely decoupled from that of the valence cluster amplitudes, eq.(7.2.8).

It is interesting to note that  $\Omega$  in eq.(7.2.4) can be regarded as a valence universal wave-operator, i.e.,  $\Omega$  is also the wave-operator for the core-problem/94(a)/. This assertion follows from the simple observation that all  $T_V$  operators have destruction operators and hence they annihilate  $\Phi_0$ .

Mukhopadhyay et al/122/ generalized the ansatz (7.2.4) still further. They suggested the use of a recursive procedure whereby the wave-operator  $\Omega^{(n)}$  at the  $n$ -valence level is recursively generated from the lower valence  $\Omega$ 's. Thus, they define

$$\Omega^{(n)} = \Omega^{(n-1)} \exp(T^{(n)}) \quad (7.2.10a)$$

$$\Omega^{(0)} = \exp(T_c) \quad (7.2.10b)$$

Clearly, with this choice, not only are the equations for the core cluster amplitudes decoupled from the rest, but also the equations for the  $T^{(n)}$  cluster-amplitudes have only  $T^{(m)}$  amplitudes with  $m < n$  frozen at their  $m$ -valence values. It should again be noted that  $\Omega^{(n)}$  acting on a  $P^{(m)}$  for  $m < n$  behaves as  $\Omega^{(m)}$  so that, formally speaking, one may still claim that this approach uses implicitly a valence-universal wave-operator.

A somewhat inconvenient feature of the ordinary

exponential ansatz is that the components of the cluster operator ( $S$  or  $T_V$ ) do not commute among themselves. There are thus contractions of the type  $(\overline{S} S)$  or  $(\overline{T_V} T_V)$ , and this couples the various  $m$ -valence sectors together. Clearly, if  $\overline{S} S$  or  $\overline{T_V} T_V$  contractions could be avoided, no  $S$  or  $T_V$  operator of valence rank  $> m$  can appear for an  $m$ -valence problem, and a decoupling of the various valence sectors would be possible.

Mukherjee/69/ and Haque and Mukherjee/69/ advocated the use of normal ordered exponential ansatz, in the manner of Lindgren/71/, to avoid the  $S$ - $S$  contractions. They showed that  $\Omega$  is of the form

$$\Omega = \{\exp(S)\} \quad (7.2.11)$$

with  $\{\dots\}$  indicating the normal ordering with respect to  $\Phi_0$ .  $\Omega$  has the interesting property that

$$\Omega P^{(m)} = \Omega^{(m)} P^{(m)} \quad \forall m \quad (7.2.12)$$

where  $\Omega^{(m)}$  is a normal ordered exponential containing  $S^{(n)}$  operators up to valence rank  $m$  only. Moreover, eq.(7.2.11) implies that

$$\Omega = \exp(S^{(0)}) \exp(\bar{S}) \quad (7.2.13)$$

where  $\bar{S}$  contains all the  $S^{(m)}$  operators except  $S^{(0)}$ . Eq.(7.2.12) shows that  $\Omega^{(0)} = \exp(S^{(0)})$  is the cluster expansion operator for the closed shell problem, and the  $S^{(0)}$  amplitudes can be found out in a manner completely decoupled from the rest of the  $\bar{S}$  operators. Introducing the operator  $H_C$  through eq.(7.2.7), the equations determining the  $S^{(0)}$  cluster amplitudes for  $1 \leq m \leq N_V$  can be written as

$$\langle X_1^{(m)} | \Omega^{(m)-1} H_C \Omega^{(m)} | \Phi_1^{(m)} \rangle = 0, \quad \forall 1, i, \quad 1 \leq m \leq N_V \quad (7.2.14)$$

It has been shown in /69/ (see also /123/) that the operator  $\Omega^{(m)-1} H_C \Omega^{(m)}$  is explicitly connected, so that, invoking the same arguments as those used in proving the connected nature of the  $S$  operators in an ordinary exponential, the connectedness of  $S$  operator can be proved. The blocks representing the matrix-elements of eq.(7.2.14) will now have diagrams in which  $S$ - $S$  contractions do not appear, and as a result a complete decoupling of the various  $m$ -valence



sectors is possible. Haque and Mukherjee/69/ showed that the diagrams in the blocks of eq.(7.2.14) are all generated by contracting H with the S vertices on its right, avoiding S-S contractions, and joining the resultant composite with the S vertices on its left in the same way but allowing contractions between the S operators this time.

Kutzelnigg/76/ and Kutzelnigg and Koch/77/ advocated that the operator transformation  $L = \Omega^{-1} H \Omega$  be performed in Fock space entirely, independent of the particular  $N_v$ -valence model space about which one is interested. By defining the 'closed' operators as the ones with 'active' (valence) lines only, and 'non-diagonal' operators as the ones in which there are active lines on one side and there is at least one inactive line on the other, the Fock space decoupling conditions, equivalent to eq.(7.1.3), can be stated on an operator level. Kutzelnigg/76/ and Kutzelnigg and Koch/77/ classified the operators depending on the types of lines attached to them as follows:

- (i)  $X_C$  : a diagonal operator with active lines only
- (ii)  $X_A$  : a non-diagonal operator in which the outgoing lines are all active.
- (iii)  $X_B$  : an off-diagonal operator in which the incoming lines are all active
- (iv)  $X_O$  : an operator with active and inactive lines on both sides.

Kutzelnigg/76/ and Kutzelnigg and Koch/77/ emphasized that the classification of the operators into the categories C, A, B and O is essentially a Fock space concept, so that the condition that  $L_B$  vanishes will automatically generate operator equivalent of the eq.(7.1.3).  $L_C$  is then the Fock-space effective hamiltonian  $H_{eff}$ , and appropriate  $n$ -valence energies  $E_k^{(n)}$  can be obtained by a projection of  $L_C$  onto the  $P^{(n)}$  space and its subsequent diagonalization. Clearly, to ensure that the  $L_B$  operators of all valence ranks ( $n$ ) are zero, we need in the Fock-space wave-operator  $\Omega$  cluster operators S, which are of the type  $S_B$ , of all valence ranks

$$\Omega = \exp(S_B) \quad (7.2.15)$$

$$S_B = \sum_{n=0} S_B^{(n)} \quad (7.2.16)$$

The equations determining the cluster amplitudes are

then given by

$$L_B^{(n)} = 0 ; \quad \forall n \quad (7.2.17)$$

It should be noted that Mukherjee *et al*/67,68/ in their formulation did use a Fock-space wave-operator, although they expressed their decoupling conditions (7.1.3) in terms of explicit determinantal projections. As a result, when valence lines are present, eq.(7.1.3) leads to a sum of various blocks equated to zero, as in Fig.7(b), for example. The lower valence blocks with spectator lines, however, are individually zero as follows from the decoupling conditions for the lower valence blocks, and as a result each individual block appearing in the decoupling condition, eq.(7.1.3), is zero. These blocks are indeed the  $L_B$  operators of various ranks. Thus, operationally speaking, the Fock-space approach of Kutzelnigg/76/ and Kutzelnigg and Koch/77/ is entirely equivalent to the Fock-space formalism of Mukherjee *et al*/67,68/. But conceptually speaking, there is a subtle difference : Kutzelnigg and Koch emphasize the decoupling condition, on the operator level, and as a result directly comes to eq.(7.2.17) without the intermediary of going through an expression of the type of Fig. 7(b) which is bound to arise when explicit projectors are used.

Kutzelnigg and Koch/77/ also introduced a unitary Fock space cluster operator  $\Omega = \exp(\sigma)$  with an antihermitian Fock space  $\sigma$ , as in eq.(7.2.3). In this case, the transformed Fock space hamiltonian  $L$  should be brought into a form such that

$$L_B^{(n)} = 0 ; \quad L_A^{(n)} = 0 \quad \forall n \quad (7.2.18)$$

so that there would be decoupling from both sides. Clearly  $\sigma$  should be of the form

$$\sigma = S_B - S_B^+ \quad (7.2.19)$$

and  $S_B^+$  is an operator of the type A.

Since Kutzelnigg and Koch utilized a wave-operator of the exponential structure, the various valence sectors will be generally coupled.

Numerical applications/77,124/ on the low-lying energy levels of small molecules such as  $H_2$  produced encouraging results for the unitary choice of  $\Omega$ .

It should be mentioned that an earlier work of

Reitz and Kutzelnigg/78/ was a forerunner of the Kutzelnigg-Koch strategy where IP of closed-shell molecules were sought by using a single unitary cluster transformation for both the ground state and the ionized states. The method has since been implemented to compute principal IP's of a series of small atoms and molecules/125/.

Haque and Mukherjee/126/ developed a Fock-space method for generating hermitian  $H_{\text{eff}}$  in which the cluster wave-operator  $\Omega$  is not unitary, following a suggestion of Jorgensen/127/, who chose  $\Omega$  to preserve the norm of the functions in the model space. Thus  $\Omega$  satisfies

$$P\Omega^+\Omega P = P \quad (7.2.20)$$

and it follows that

$$H_{\text{eff}} = P\Omega^{-1}H\Omega P = P\Omega^+H\Omega P \quad (7.2.21)$$

i.e.,  $H_{\text{eff}}$  is hermitian. The condition (7.2.20) may be called an 'isometry' condition. Haque and Mukherjee developed a Fock-space version of this approach, in which  $\Omega$  is written in the form

$$\Omega = \{\exp(S+X)\} \quad (7.2.22)$$

where  $S$  has the same form as in eq.(7.2.11), and  $X$  are 'closed' operators inducing valence to valence scatterings. Using this  $\Omega$ , they derived a set of equations like eq.(7.2.14). The closed operator  $X$  is chosen in such a way that  $H_{\text{eff}}$  becomes hermitian. In a straightforward extension of the Jorgensen approach/127/, Haque and Mukherjee/126/, used a Fock-space equivalent of eq.(7.2.20), which reads

$$P^{(n)}\Omega^+\Omega P^{(n)} = P^{(n)} \quad (7.2.23)$$

Since the first term in the expansion of  $\Omega$  is the unit operator, eq. (7.2.20) indicates that all the  $n$ -valence connected diagrams obtained by contracting powers of  $S^+$  and  $X^+$  operators with those of  $S$  and  $X$  will vanish. It can be shown that there is no loss of generality if  $X$  is chosen as hermitian. Haque and Mukherjee /126/ explicitly used the condition that  $H_{\text{eff}}^+ = H_{\text{eff}}$ , for defining the operators  $X^{(n)}$ . There was also a related earlier work by Kvasnicka /72/, which hermitized the CC theory of Lindgren/71/. For more discussions on hermitian  $H_{\text{eff}}$ , see Sec.7.4.

## 7.3 Bloch Equation Based Fock Space Theory:

Offermann *et al* /70/ formulated an open-shell CC formalism in Nuclear Physics in which a Bloch wave-operator  $\Omega^{(n)}$  for an n-valence problem is obtained recursively from those of the lower valence problems:

$$\Omega^{(0)} = \exp(S^{(0)}) P^{(0)} \quad (7.3.1)$$

$$\Omega^{(1)} = \exp(S^{(0)}) (1 + S^{(1)}) P^{(1)} \quad (7.3.2)$$

$$\Omega^{(2)} = \exp(S^{(0)}) (1 + S^{(1)} + 1/2 S^{(1)} \otimes S^{(1)} + S^{(2)}) P^{(2)} \quad (7.3.3)$$

and so on, where  $\otimes$  denotes an outer product between two operators ( i.e., no contractions between the operators are allowed). Defined this way,  $\Omega^{(n)}$  in their work embeds in it information about all the lower Valence sectors and defines a Fock space strategy. A similar idea was also pursued by Ey/70/. The cluster amplitudes are obtained recursively, by solving the zero valence problem first, starting with the Bloch equation for the zero valence problem, solving the one valence problem next, with  $S^{(0)}$  amplitudes kept frozen, and so on- finally arriving at the desired  $N_v$  valence situation. It should be noted that the ansatz for  $\Omega^{(n)}$ , obtained recursively following the chain (7.3.1), (7.3.2) .., is similar to the one used in eq.(7.2.10) with the difference that the outer product convention in eq.(7.3.3) eliminates contractions between S operators. Since  $\Omega^{(n)}$  is recursively generated from  $\Omega^{(m)}$  of lower valence ranks, it might appear that there is no underlying assumption of valence universality in this formulation. However, interestingly enough,  $\Omega^{(n)}$  defined in this scheme has the property that

$$\Omega^{(n)} P^{(m)} = \Omega^{(m)} P^{(m)} \quad \forall m < n \quad (7.3.4)$$

so that one may again formally claim that the method has a valence universal wave-operator.

Lindgren/71/ developed an open-shell CC theory by postulating an wave-operator  $\Omega$  in normal order :

$$\Omega = \{\exp(S)\} \quad (7.3.5)$$

where the structure of the cluster operator  $\Omega$  was inferred by an appeal to the open-shell MBPT/4/,

starting from Bloch equation/116/. He called operators causing transitions within the model space as 'closed' and those causing transition outside the model-space as 'open'. Lindgren's starting equation was

$$[\Omega, H_0] P^{(N_v)} = V \Omega P^{(N_v)} - \Omega P^{(N_v)} V \Omega P^{(N_v)} \quad (7.3.6)$$

which originates from eq. (6.1.15), and the assumption of intermediate normalization. Inserting eq. (7.3.5) to eq. (7.3.6) and expanding  $\Omega$  (or equivalently  $S$ ) in order of perturbation theory, Lindgren/71/ concluded that  $S$  must be the sum over various  $S^{(m)}$  operators of rank  $0 \leq m \leq N_v$ . In the original formulation Lindgren/71/ confined himself to the particular  $N_v$ -valence model space, and consequently the cluster amplitudes of various valence ranks are not necessarily linearly independent. This situation is entirely analogous to the one discussed in Sec. 7.1. Using Wick's theorem, eq. (7.3.6) with eq. (7.3.5), leads to

$$[S, H_0] \Omega P^{(N_v)} = \{V \Omega\} P^{(N_v)} - \{\Omega V_{\text{eff}}\} P^{(N_v)} \quad (7.3.7)$$

where the contractions  $V \Omega$  etc mean the connected expressions of the form

$$\{V \Omega\} = \sum_n \frac{1}{n!} \{V \begin{array}{c} \text{---} \\ \text{---} \end{array} S \begin{array}{c} \text{---} \\ \text{---} \end{array} S \dots\} \quad (7.3.8)$$

Since the  $S^{(N_v)}$  operators alone are sufficient to determine the eigenfunctions  $\Psi_k^{(N_v)}$ , (7.3.7) are an under-determined set of coupled equations for  $S^{(n)}$  with  $0 \leq n \leq N_v$ . Lindgren thus needed some auxiliary conditions to determine the lower valence cluster amplitudes. For this he appealed to MBPT, and demanded that the order by order expansion of  $S$  in powers of  $V$  should produce the wave-operator of the open shell MBPT. Since  $S$  has to be connected then, Lindgren postulated a set of sufficiency conditions which would guarantee the connectivity. Thus he demanded that the following relations must hold:

$$[S, H_0] P^{(N_v)} = \{V \Omega\} P^{(N_v)} - \{\Omega V_{\text{eff}}\} P^{(N_v)} \quad (7.3.9)$$

It should be noted that the validity of eq.

(7.3.9) implies the validity of eq. (7.3.7), but not vice-versa. Since  $V_{\text{eff}}$  is closed, taking the closed part of eq. (7.3.9) leads to

$$V_{\text{eff}} = P \prod_v^{(N_v)} \{V\Omega\} P \prod_v^{(N_v)} \quad (7.3.10)$$

The method has since been used to compute energies of quasi-degenerate Be iso-electronic sequence ground state/131/.

Superficially it might appear that Lindgren's formulation is a Hilbert space approach involving redundant S operators, and the redundancy is eliminated by a set of augmented equations which at the same time ensures the connectedness of  $H_{\text{eff}}$ . However, as noted by Mukherjee, and Haque and Mukherjee/69/, use of the sufficiency conditions (7.3.9) amounts in effect to assuming that  $\Omega$  is a valence-universal wave-operator. In fact Haque has explicitly demonstrated/123/ that the use of a valence-universal  $\Omega$  in the Fock-space Bloch equation leads automatically to eqn (7.3.9) with the ad-hoc sufficiency requirement. We give the sketch of a general proof here, since it shows that the extra information content of a Fock-space  $\Omega$ , as opposed to a Hilbert space, can be used to advantage for ensuring the connectivity of the cluster amplitudes of S/93/. For a valence-universal  $\Omega$ , the Fock-space Bloch equation (6.1.15) leads to

$$\{H\Omega\}P^{(n)} = \{\Omega H_{\text{eff}}\}P^{(n)} \quad \forall n \quad (7.3.11)$$

Calling the composites  $H\Omega$  and  $\Omega H_{\text{eff}}$  as Z and  $\theta$  respectively, eq. (7.3.11) implies that

$$\{Z^{(0)}\Omega^{(0)}\}P^{(0)} = \{\Omega^{(0)}\theta^{(0)}\}P^{(0)} \quad (7.3.12)$$

where the superscript (0) indicates operators of valence rank 0. Since  $\Omega^{(0)}$  contains various nh-np excitations in S which are linearly independent, it follows that

$$\{Z^{(0)}\} = \{\theta^{(0)}\} \quad (7.3.13)$$

For the one-valence problem, similarly

$$\{Z^{(1)}\Omega^{(0)} + Z^{(0)}\Omega^{(1)}\}P^{(1)} = \{\Omega^{(0)}\theta^{(1)} + \Omega^{(1)}\theta^{(0)}\}P^{(1)} \quad (7.3.14)$$

and because of eq. (7.3.13), we have

$$\{Z^{(1)}\Omega^{(0)}\}P^{(1)} = \{\Omega^{(0)}\theta^{(1)}\}P^{(1)} \quad (7.3.15)$$

and it again follows that

$$\{Z^{(1)}\} = \{\theta^{(1)}\} \quad (7.3.16)$$

It thus follows generally, proceeding upwards in valence rank, that

$$\{Z^{(k)}\} = \{\theta^{(k)}\} \quad \forall k \quad (7.3.17)$$

from which eq. (7.3.9) follows easily on summing eq. (7.3.17) for the various ranks  $k$ , and using the partition of  $H$  into  $H_0 + V$ .

The above analysis indicates that Lindgren's eq. (7.3.9) is restrictive, since eq. (7.3.17) shows that we may write a more general Fock-space equation

$$[S, H_0]P^{(n)} = \{\bar{V}\Omega\}P^{(n)} - \{\bar{V}\Omega_{\text{eff}}\}P^{(n)} \quad \forall n \quad (7.3.18)$$

Thus, with eq. (7.3.9) it will not be possible to compute energy differences  $\Delta E_k$ , such as IP etc. In contrast, the Haque-Mukherjee formulation/69/, however, uses eq. (7.3.18), and as a result IP, EA, EE etc., can be calculated. Lindgren, however, later generalized his equation suitably to compute IP etc./129,130/, which means that eqn (7.3.18) was implicitly used.

Haque-Mukherjee formulation/69/ also indicates that there is an automatic decoupling of the equations for the various  $n$ -valence sectors. There are no  $S^{(m)}$  amplitudes with  $m > n$  for the  $n$ -valence problem. This decoupling, which is rigorous, has since been termed as the 'subsystem embedding condition' (SEC)/69/. This leads to a computationally advantageous scheme. Haque and Kaldor/132/, and Kaldor/133/ made several applications of the formalism on small atoms and recently on molecules/134/. In these applications, low-lying excited states of diatomic species, valence IP's of closed shell ground states and low-lying excitation energies of simple atoms and molecules are computed. It appears from these calculations that

three body cluster operators at the one and two valence levels may be important for a quantitative accuracy. Kaldor et al/148/ have very recently also applied the formalism for studying the potential surface of  $\text{NH}_3 + \text{H}^+$  reaction.

Sinha et al/135/ also applied the Haque-Mukherjee formalism for IP calculations, but they envisaged a more general computational strategy where not only the principal IP values at the outer and inner valence region but also the satellite peaks with dominant shake-up effects were aimed at. For this, they included all the hole orbitals upto the inner valence region as active. Since the dimension of the model space for the one-hole valence problem is just equal to the number of hole orbitals, it may at first appear surprising that satellite peaks can at all be obtained in this scheme in addition to the main peaks. We should, however, note that the CC equations are inherently non-linear in nature. In fact, at any  $n$ -valence level, the eq. (7.3.18) involves quadratic powers of  $S^{(n)}$ . Sinha et al/135/ exploited the non-linear nature of the equations to generate multiple solutions. The main peaks corresponded to those solution sets which could be reached from the starting iterates obtained from the linearized equation. The satellite peaks are obtained when the starting iterates for the  $S^{(1)}$  amplitudes are estimated from the eigenvectors of a small CI problem involving the model space and the dominant shake-up amplitudes. This finding bolsters the view that a CC-approach does transcend an MBPT in both compactness and information content. In a RS MBPT/3-9/, the number of roots obtained will always be just  $M$ . The Bloch-Horowitz/2/ theory, however, can furnish more than  $M$  roots but at the expense of the loss of size-extensivity with respect to  $N_v$ .

Very recently Sinha et al/136/ have analytically shown that the CC-based LRT for IP is algebraically equivalent to the CC formalism of Haque and Mukherjee/69/ for one-hole valence problem. This indicates that the CC equations for determining the IP's can be cast into an equivalent eigenvalue equation. In fact Sinha et al /136/ have demonstrated that the relation between the CC-LRT for IP and the corresponding CC theory is the same as that between a CI problem and the associated partitioned problem as obtained by the Soliverrez transformation/137/. For open-shells containing more than one valence, the correspondence between the CC-LRT and the CC equations no longer holds, but Sinha et al showed the CC



equations can still be cast into an equivalent eigenvalue problem/137/. This is computationally very useful, since the associated eigenvalue problems are usually numerically easier to solve and techniques exist which allow eigenvectors of a given structure to be reached by suitable root-homing procedures. Using this strategy, Sinha *et al* have computed the Auger spectra of  $F_2$ ,  $H_2O$  and  $HF/135/$ , including satellite peaks/136/ for the first time using an open-shell CC theory. Main peaks for  $N_2$ ,  $CO$  and  $H_2O$  have also been computed by Pal *et al*/139,140/.

We have shown as typical examples of constructing the CC equations, the diagrams contributing to open parts of  $Z^{(2)}$ ,  $Z^{(2)}$  in Figs. 9(a) to (f). Those contributing to  $\theta^{(2)}$  are shown in Figs.9(g) to (h). The valence orbitals are particle type. The diagrams contributing to  $H_{eff}^{(1)}$  and  $H_{eff}^{(2)}$  are shown in Figs.10(a), (b) respectively. As another example, the CC equations for  $S^{(1)}$  amplitudes for IP are shown in Figs.11.

#### 7.4 Variational CC Theory In Fock Space :

Pal *et al*/73(a)/ formulated a variational CC theory for energy differences in Fock space, which generates a hermitian  $H_{eff}$ . They employed the ansatz, eq. (7.2.22) for  $\Omega$  :

$$\Omega = \{\exp(S+X)\} \quad (7.4.1)$$

with  $S$  of the form of eq. (7.2.11), and  $X$  are closed operators causing scatterings with the model space.  $\Omega$  is assumed, as in eq.(7.2.23), to satisfy the 'isometry' condition :

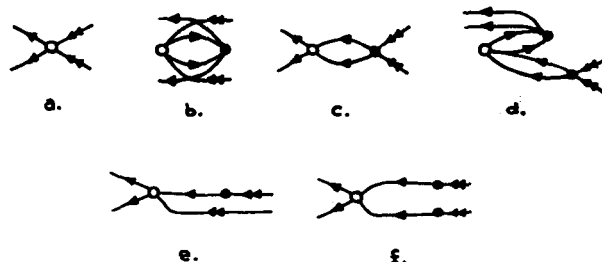
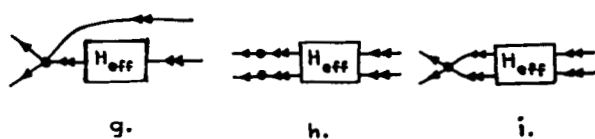
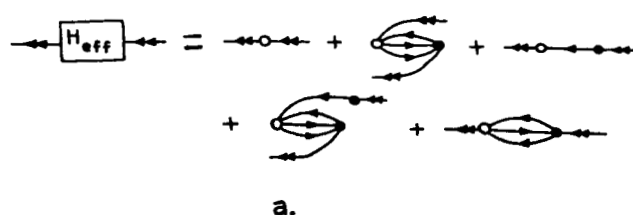
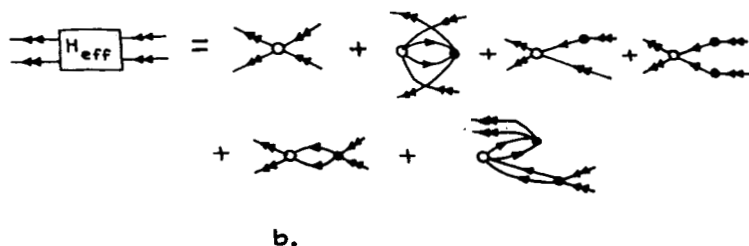


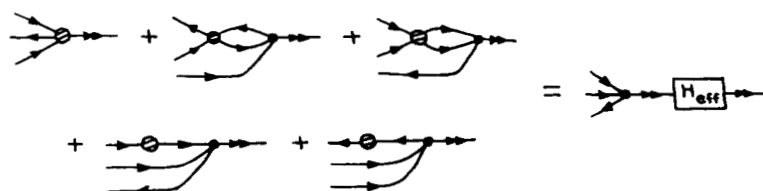
Fig.9 (a)-(f), Diagrams for  $Z_{op}^{(2)}$ .

Fig.9 (g)-(i), Diagrams for  $\theta_{op}^{(2)}$ .

a.



b.

Fig.10 Typical diagrams for  $H_{eff}^{(1)}$  and  $H_{eff}^{(2)}$ .Fig.11 CC equations for  $S^{(1)}$  for the one hole problem.

$$P^{(n)}\{\Omega^\dagger\Omega\}P^{(n)} \equiv P^{(n)}\{\exp(M)\}P^{(n)} = P^{(n)} \quad (7.4.2)$$

Eq. (7.4.2) guarantees that  $H_{\text{eff}}$  will be hermitian.  $M$  is a hermitian closed operator.

In their formulation, Pal et al determined the closed-shell cluster amplitudes  $S^{(0)}$  by varying the expectation value of the ground state energy  $E_0$  with respect to the cluster amplitudes/73(b)/ :

$$\partial E_0 / \partial s_i^{(0)\dagger} = 0 \quad \forall i \quad (7.4.3)$$

$$E_0 = \langle \Phi_0 | \exp(S^{(0)\dagger}) H \exp(S^{(0)}) | \Phi_0 \rangle_C \quad (7.4.4)$$

$\langle \Phi_0 | \dots | \Phi_0 \rangle_C$  stands for connected quantities.

The cluster-amplitudes of  $S^{(1)}$  for the one-valence problems (IP or EA) were determined by minimizing the expectation values  $E_I$  :

$$E_I = \langle \Phi_I | \{ \exp(S^{(0)} + S^{(1)} + X^{(1)}) \}^\dagger H \{ \exp(S^{(0)} + S^{(1)} + X^{(1)}) \} | \Phi_I \rangle_C \quad (7.4.5)$$

Owing to normal ordering,  $E_I$  is at the most quadratic in  $S^{(1)}$  and  $X^{(1)}$  and their adjoints. Minimization of  $E_I$ , subject to the constraint eq. (7.4.2) for  $P^{(1)}$ , leads to equations of the form

$$\partial E_I / \partial s_i^{(1)\dagger} - \sum_{JK} \lambda_{JK}^I \partial M_{JK}^{(1)} / \partial s_i^{(1)\dagger} = 0 \quad (7.4.6a)$$

$$\partial E_I / \partial x_i^{(1)\dagger} - \sum_{JK} \lambda_{JK}^I \partial M_{JK}^{(1)} / \partial x_i^{(1)\dagger} = 0 \quad (7.4.6b)$$

where  $\lambda$ 's are the Lagrange's multipliers.

Solution of eq. (7.4.6), along with the isometry condition

$$M_{JK}^{(1)} = 0, \quad \forall J, K \quad (7.4.7)$$

which is equivalent to eq. (7.4.2), furnishes us with the  $S^{(1)}$  and  $X^{(1)}$  amplitudes. We now write energy  $E_k^{(1)}$  for each  $\Psi_k^{(1)}$  as

$$E_k^{(1)} = \langle \Psi_k^{O(1)} | \Omega^\dagger H \Omega | \Psi_k^{O(1)} \rangle \times [\langle \Psi_k^{O(1)} | \Omega^\dagger \Omega | \Psi_k^{O(1)} \rangle]^{-1} \quad (7.4.8)$$

where

$$\Psi_k^{(1)} = \Omega \Psi_k^{O(1)} \equiv \Omega \sum_I C_{Ik} \Phi_I \quad (7.4.9)$$

Minimization of  $E_k^{(1)}$  with respect to  $C_{Ik}$ 's leads to a set of equations of the form

$$\sum_J \tilde{H}_{IJ}^{(1)} C_{Jk} = E_k^{(1)} C_{Ik} \quad (7.4.10)$$

where  $\tilde{H}_{IJ}^{(1)}$  is given by

$$\tilde{H}_{IJ}^{(1)} = \langle \Phi_I | \Omega^\dagger H \Omega | \Phi_J \rangle_C \quad (7.4.11)$$

The matrix  $\tilde{H}$  is hermitian. By dropping the vacuum diagrams from  $H$ , we may also obtain the energy differences  $\Delta E = E_k - E_0$  directly.

For the two-valence problem, one may proceed similarly. In the two-valence sector, the one valence amplitudes will be rigorously frozen. Pal *et al*/73(a)/ discussed the general  $n$ -valence case in detail.

Haque and Kaldor/132(b)/ applied this formalism to compute IP's of  $N_2$  and  $H_2O$ .

## 8. SIZE EXTENSIVE FORMULATIONS WITH INCOMPLETE MODEL SPACE :

### 8.1 Preliminaries :

There are mainly two inter-related reasons why one would like to formulate a size-extensive theory involving incomplete model spaces: (i) to bypass intruders, (ii) for calculating the low-lying EE's of a closed-shell ground state, one feels that a set of hole-particle determinants would suffice as a choice for a reasonable model space, which involves valence holes as well as valence particles. This is an IMS.

It came to be generally acknowledged /20,44,66,90/ that the theories using IMS generate disconnected

diagrams in  $H_{\text{eff}}$ . In the quasi-complete model space theory of Lindgren/90/, of which the h-p model space is a special case, disconnected diagrams appear in  $H_{\text{eff}}$ . Appearance of disconnected diagrams in  $H_{\text{eff}}$  spells a breakdown of the size-extensivity of the calculated energies/146,147/.

This situation has changed very recently with the discovery of Mukherjee/91-93/ that the real reason behind getting a disconnected  $H_{\text{eff}}$  for an IMS lies in (i) not using a Fock-space valence-universal  $\Omega$  and (ii) not abandoning the intermediate normalization (IN) for  $\Omega$ , which is not size-extensive for IMS. If the assumption of IN is abandoned, it is relatively straight-forward to prove the linked cluster theorem (LCT) for incomplete model space. Abandoning IN, Mukherjee/91/ initially proved LCT for incomplete model spaces having n-hole n-particle determinants, showing also at the same time the validity of the core-valence separation. The corresponding open-shell perturbation theory of Brandow/20/ for such cases leads to unlinked terms and a breakdown of the core-valence separation, which used IN for  $\Omega$ . Mukherjee emphasized that it is essential to have a valence-universal wave operator  $\Omega$  within a Fock space formulation/91/ such that it also correlates the subdued valence sectors. Later on, the proof of LCT was shown for more general model spaces/92/. The Fock space approach is particularly important for this, with the additional conditions of abandonment of the conventional IN. Mukherjee used a normal ordered ansatz/91-93/ for the wave operator  $\Omega$  in order that the solution for the different valence sectors will be decoupled (SEC, as in /69/ ). This is particularly useful computationally and leads to significant saving of the computer time.

## 8.2 General Formulation

In what follows, it will be useful to classify the parent n-valence IMS belonging to the space  $P$ , the complementary active space, whose union with  $P$  forms the CMS, as  $R$  and the 'true' virtual space containing inactive labels as  $Q$ . The cluster operators inducing the transition  $P \rightarrow R$  are all labelled by valence lines, and will henceforth be denoted as 'quasi-open'. The operators making transitions  $P \rightarrow Q$  will be called 'open'. The operators connecting model spaces only will be called 'closed'. In general, for a valence-universal  $\Omega$ , it so happens that products of quasi-open

operators are closed. Thus, for example, consider an IMS with h-p determinants. Since both the vacuum  $\Phi_0$  and the doubly excited states are in the space  $R$ , there will be both de-excitation and excitation types of quasi-open  $S$  operators in  $\Omega$ , as depicted in Figs. 12(a) and (b). Their product, however, is a closed operator, as shown in Fig.12(c).

Thus, although we may choose a cluster operator  $S$  as containing quasi-open and open operators only, a wave-operator of the form  $\exp(S)$  or  $\langle \exp(S) \rangle$  will generate closed operators stemming from the expansion involving products of quasi-open operators. The intermediate normalization is thus incompatible with the above choice of  $S$ .

We show below, following Mukherjee/91-94/, how a connected  $H_{\text{eff}}$  can be obtained using a Valence-universal  $\Omega_{\text{eff}}$  which uses a normalization convention different from intermediate normalization. We start from the Fock-space Bloch Equation:

$$H\Omega P^{(k)} = \Omega H_{\text{eff}} P^{(k)} \quad \forall 0 \leq k \leq n, \quad (8.2.1)$$

$$\Omega = \langle \exp(S) \rangle, \quad (8.2.2)$$

$$S = S_{q-op} + S_{op} \quad (8.2.3)$$

Using the arguments similar to those used in Sec 7.3, we arrive at

$$\langle Z \rangle^{(k)} = \langle \theta \rangle^{(k)} \quad \forall 0 \leq k \leq n, \quad (8.2.4)$$

$$\langle Z \rangle = \langle H\Omega \rangle \quad (8.2.5a)$$

$$\langle \theta \rangle = \langle \Omega_{\text{eff}} \rangle \quad (8.2.5b)$$

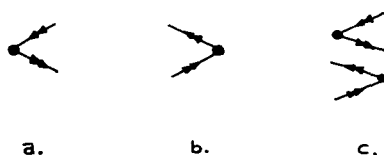


Fig.12 (a) De-excitation and (b) excitation q-op operators in an IMS with h-p determinants. (c) Demonstration that their product is closed.

Since  $H_{\text{eff}}$  is a closed operator, we have

$$\{Z\}_{q\text{-op}}^{(k)} = \{\theta\}_{q\text{-op}}^{(k)} = \{\overline{\Omega_{q\text{-op}} H_{\text{eff}}}\}^{(k)} \quad (8.2.6)$$

$$\{Z\}_{\text{op}}^{(k)} = \{\theta\}_{\text{op}}^{(k)} = \{\overline{\Omega_{\text{op}} H_{\text{eff}}}\}^{(k)} \quad (8.2.7)$$

$$\{Z\}_{\text{cl}}^{(k)} = \{\theta\}_{\text{cl}}^{(k)} = \{\overline{\Omega_{\text{cl}} H_{\text{eff}}}\}^{(k)} \quad (8.2.8)$$

The eqs. (8.2.6) and (8.2.7) are the defining equations for  $S_{q\text{-op}}$  and  $S_{\text{op}}$ , and eq. (8.2.8) defines  $H_{\text{eff}}$ . The structure of the equations shows that  $S$  is a connected operator, and so is  $H_{\text{eff}}$ . The most conspicuous difference between IMS and CMS lies in the appearance of  $\Omega_{\text{cl}}$  in the former, as in eq. (8.2.8).

Since eqs. (8.2.6) and (8.2.7) are quite similar in structure to those for a CMS, we need not discuss their construction in detail. In fact, the left hand side of eqs. (8.2.6) and (8.2.7) are entirely analogous to those for a CMS. On their right hand side, they will have contractions of powers of  $S$  with  $H_{\text{eff}}$ , and since  $H_{\text{eff}}$ , from eq. (8.2.8), has a different structure compared to that in a CMS, we shall discuss this part in some detail.

Writing  $\Omega_{\text{cl}}$  as

$$\Omega_{\text{cl}} = 1 + \kappa_{\text{cl}} \quad (8.2.9)$$

we may write  $H_{\text{eff}}^{(k)}$  as

$$H_{\text{eff}}^{(k)} = \{Z\}_{\text{cl}}^{(k)} - \{\overline{\kappa_{\text{cl}} H_{\text{eff}}}\}^{(k)} \quad (8.2.10)$$

which shows a way to recursively generate  $H_{\text{eff}}^{(k)}$ .

As  $H_{\text{eff}}$  is usually a small matrix, the extra complication of its structure in an IMS does not pose any computational disadvantage.

Kutzelnigg et al/95/ have discussed in detail the Fock-space classification of operators for quasi-complete model space. They also introduced a new type of IMS, called the isolated incomplete model space (IIMS). In IIMS, products of q-open operators are all q-open, never closed. As a result  $\Omega_{\text{cl}} = 1_{\text{cl}}$ , just as in a CMS. The resulting CC equations for IIMS have thus exactly the same structure as in the CMS. Mukherjee et al/96/ have discussed the Fock-space

classification scheme of the operators for a general model space. They also considered the unitary cluster ansatz for  $\Omega$

$$\Omega = \exp(\sigma) , \quad (8.2.11)$$

$$\sigma = S - S^\dagger , \quad (8.2.12)$$

$$S = S_{q-op} + S_{op} \quad (8.2.13)$$

The choice (8.2.12) for  $\Omega$  leads to a hermitian  $H_{eff}$ . Mukherjee/93/ has discussed the 'isometric normalization' for  $\Omega$  in IMS, which also leads to a hermitian  $H_{eff}$ .

Chaudhuri et al have recently formulated an open-shell MBPT for IMS, abandoning intermediate normalization, by choosing  $\Omega_1$  at each order in a way which ensures the connectedness of  $H_{eff}$ /150/.

### 8.3 The Case of IMS Having Valence-holes and -particles

As has been stressed in Sec. 8.1 the incomplete model space is both a computational and conceptual necessity for quasidegenerate systems. Since it is formally possible to have a completely connected cluster expansion theory based on a general incomplete model space it has been possible to apply such methods to the problem of excitation energies (EE). Because of the hierarchical generation of terms, the process of obtaining EE also leads to ionization potentials (IP) and electron affinities (EA). We will discuss the theory in relation to generating such quantities in this section. For the low-lying excited states of a closed-shell  $N$ -electron ground state one chooses the Hartree-Fock function  $\Phi_0$  as the vacuum and uses the particle-hole (p-h) excited determinants as spanning the model space for excited states. The p-h excited determinants are a reasonable choice for the model space on the energetic consideration of the orbitals. One chooses a set of 'active' holes and 'active' particles to constitute the model space. This choice, according to our experience, avoids the intruder state problems for most cases of study if the active orbitals are chosen properly. As emphasized previously, the h-p excited determinants constitute an incomplete model space, even if we exhaust all the



active holes and active particles in forming the determinants. Sinha et al formulated the MRCC theory for EE assuming all holes and particles to be active/138/. However, for the sake of computational convenience it is usual to treat only a few of the particle and hole orbitals as active, namely those near the Fermi level. Pal et al, arguing on similar lines as Sinha et al/138/ developed a similar formulation for the case of active-inactive subdivision/139-140/. Since the IMS under consideration has both valence holes and valence particles, it is better to classify the operators according to their p-h valence ranks.

In Fock space we designate by  $\{\Phi_i^{(k,l)}\}$  a model space of kp-lh determinants. In the present case  $k$  and  $l$  range from 0 to 1. An operator  $A$  which can destroy  $m_p$  active particles and  $m_h$  active holes is denoted as  $A_{p^{(m_p)}, h^{(m_h)}}$ . Each of these is a sum of several one, two ..body operators.

In the spirit of earlier discussions the valence universal wave operator  $\Omega$  may be written as,

$$\Omega = \{\exp(S^{(k,l)})\}, \quad (8.3.1)$$

$$S^{(k,l)} = \sum_{m=0} \sum_{n=0} S^{(m,n)} \quad (8.3.2)$$

Such a wave operator  $\Omega$  is capable of correlating all the lower valence model spaces. Thus, for the IMS  $\Psi_i^{(k,l)}$  with  $k=l=1$ , the lower valence model spaces are  $\Psi_i^{(1,0)}$ ,  $\Psi_i^{(0,1)}$ , and  $\Psi_i^{(0,0)}$ , corresponding to  $(N+1)$ -electron,  $(N-1)$ -electron and  $N$ -electron ground state problems.

One can set up the equations for the  $k=l=1$  valence- problem in a hierarchical manner. The valence sectors are defined in terms of particle-hole rank. Thus the Fock space Bloch equation would be

$$H\Omega P^{(k,l)} = \Omega H_{\text{eff}} P^{(k,l)} \quad 0 \leq k \leq 1; \quad 0 \leq l \leq 1 \quad (8.3.3)$$

Using arguments similar to those given in Sec 8.2, we find

$$\{Z^{(k,l)}\} P^{(k,l)} = \{\theta^{(k,l)}\} P^{(k,l)} \quad (8.3.4a)$$

where

$$\{Z^{(k,l)}\} = \{\prod \Omega\}^{(k,l)} \quad (8.3.4b)$$

$$\{\theta^{(k,l)}\} = \{\Omega_{\text{eff}}\}^{(k,l)} \quad (8.3.4c)$$

For the solution one starts from the lowest hole-particle sector (0,0) for  $S^{(0,0)}$  and then solves the (0,1) and (1,0) sectors finally reaching the target (1,1) sector. The equation for the (0,0) sector reduces to that for the closed shells. The model spaces for (0,1) and (1,0) sectors are always complete by definition. The energies of these two sectors are those of (N-1) electron states and (N+1) electron states respectively. If we collectively call the q-open and open operators as external, then the equations for  $S^{(k,l)}$  are given by

$$\{Z^{(k,l)}\}_{\text{ext}} = \{\theta_{\text{ext}}^{(k,l)}\} = \{\Omega_{\text{eff}}\}_{\text{ext}}^{(k,l)} \quad (8.3.5)$$

The model space projection yields  $H_{\text{eff}}^{(k,l)}$  :

$$\{Z^{(k,l)}\}_{\text{cl}} = \{\theta_{\text{cl}}^{(k,l)}\} = \{\Omega_{\text{eff}}\}_{\text{cl}}^{(k,l)} \quad (8.3.6)$$

For complete model spaces, since IN is valid, we have

$$\Omega_{\text{cl}}^{(k,l)} = 1_{\text{cl}} \quad \text{for } k+l = 1 \quad (8.3.7)$$

i.e., eqs (8.3.5) and (8.3.6) yield the complete model space equations as in the case of (N+1) and (N-1) electron states. However, for the excited states  $k+l=1$ ,  $\Omega_{\text{cl}}^{(k,l)} \neq 1_{\text{cl}}$ , and eq. (8.3.7) is not valid. Using eq. (8.2.7),  $H_{\text{eff}}^{(k,l)}$  may be defined recursively as

$$H_{\text{eff}}^{(k,l)} = \{Z_{\text{cl}}^{(k,l)}\} - \{\overline{x_{\text{cl}}} H_{\text{eff}}\}^{(k,l)} \quad (8.3.8)$$

In recent calculations by Kaldor et al /133,134/, use was made of p-h excited determinants with IN for  $\Omega$  ignoring the fact that such a choice causes the generation of unlinked diagrams in  $H_{\text{eff}}$ . What Kaldor and co-workers did was essentially to ignore the de-excitation operators of the type shown in Fig. 12(a), and consequently impose IN on  $\Omega$ . At first sight, it then appears that their ignoring the disconnected terms in  $H_{\text{eff}}$  is a result of the cancellation of errors: (1) use of IN for  $\Omega$ ; (2) neglect of de-excitation  $S$ . However, Sinha et al /138/ have shown that, when all the particles and holes are active, it is impossible

to connect  $\chi_{c1}$  to get a connected  $H^{(1,1)}$ . The reason for this lies in the fact that the operator  $S_2^{(0,0)}$ , of Fig. 12(b) has no lines on its right, and hence cannot be contracted to  $H_{eff}^{(1,1)}$  on its right. Thus although IN does not hold good,  $\theta^{(1,1)}$  is still equal to  $H_{eff}^{(1,1)}$ , exactly as in a CMS theory with IN for  $\Omega$ . They also noted that the  $S_2^{(1,1)}$  operator, which can de-excite the p-h model space to  $\Phi$  lying outside the model space, can contribute only to the  $S_2^{(1,1)}$  determining equations only and the calculation of all other excitation operators  $S^{(1,1)}$  as well as  $H^{(1,1)}$  do not depend on the  $S_2^{(1,1)}$  operator. Hence for computing the energy of the excited states one does not have to involve  $S_2^{(1,1)}$  operators at all. If the vacuum diagrams are dropped in the construction of  $H_{eff}$ , the results are directly the difference energies. The same observation was found to be true even when only some holes and particles are active, as shown by Pal *et al* /139,140/. See /94(a), 138-140/ for a more extensive treatment of this aspect.

The above formulation is quite general and applies equally well to quasi-complete model spaces having m holes and n particles. When there are several p-h valence ranks in the parent model space, the situation is fairly complicated. The subduced model spaces in this case may belong to the parent model space itself. The valence-universality of  $\Omega$  in such a situation implies that  $\Omega$  is the wave-operator for all the subduced model spaces, in addition to those which have same number of electrons as in the parent model space. It appears that a more convenient route to solve this problem is to redefine the core in such a way that holes for the problem become particles and treat it as an IMS involving valence particles only.

The general IMS containing 0h-0p, 1h-1p, ..., nh-np has recently been considered by Stolarczyk and Monkhorst /85/.

#### 8.4. Applications of IMS-based CC Theories

The formalisms described in Secs. 8.2 and 8.3 have found several recent applications. Apart from the IP results /99,100,132/, there have also been calculations of some atomic electron affinities/134/ all of which accompany EE computations. There are a few molecular EE results of  $N_2$ , CO,  $CH_2N_2$ ,  $CH_2CO$ , CH by Pal *et al* /139,140/, Rittby *et al* /141/ and Mattie *et al* /141/. The results so far have been obtained mostly

in the singles and doubles approximations. Kaldor has reported some IP results including approximate triples/134/. There have also been IP results with approximate triples in the group of Bartlett/142/. According to the experience gained, it is computationally convenient to construct  $\exp(-S^{(0,0)})H \exp(S^{(0,0)})$ , which is a necessary building block, separately and store it as advocated by Mukhopadhyay et al/99/. Most of the results for IP are reasonably good. However, there are certain third-order terms which can be accounted for by triples. Initial IP calculation with triples in a basis including polarization functions yield almost chemically accurate or near-chemically accurate results. The quality of EA results in the same basis is not usually as good because one needs to have a reasonably large basis set for EA calculations. The low-lying EE calculations with modest basis sets of systems like  $N_2$ , CO,  $CH_2CO$  etc., yield results comparable with experiments and those obtained by other methods. At double zeta and one polarization function basis level the discrepancies of some of the EE's for the systems studied from the experimental results are usually much less than 0.1 eV. The discrepancy for most others is from 0.1-0.5eV and are often lesser than other methods. Koch used the IMS-based CC theory to compute the  $\pm P$  states of He and the excited states of  $C^{+3}$  to  $F^{+6}$ /149/.

In low-lying excited state calculations one has so far been able to get rid of the intruder state problem. However, it is still not established that in studying potential surfaces even such a general incomplete model space-based MRCC is capable of handling intruder state problems. Although, in a very recent pilot potential energy calculation, Koch and Mukherjee/143/ could get most regions of the potential curves for  $He_2^{+2}$ , and Kaldor/144/ and Ben-Shlomo and Kaldor/145/ got curves for LiH, further studies are clearly needed.

## 8.5 Size-extensivity of Energies with IMS

It may appear that the connectedness of  $H_{eff}$  is not by itself a guarantee for the size-extensivity of energies, since disconnected terms may appear during the diagonalization. Chaudhuri et al /146/ have demonstrated recently that such is not the case because of the special structure of  $H_{eff}$ .

One may follow the diagonalization process in orders of perturbation. In a general CI matrix with IMS, the disconnected terms arising from the norm-correction may be cast as products of closed diagrams, with intermediates belonging to the P space or to the R space. There are no compensating diagrams with R space intermediates coming from the general sum, since R space functions are not present in the CI matrix. This is the reason why there are disconnected diagrams in a CI with IMS. In the  $H_{\text{eff}}$  matrix, however, the intermediates in the norm term can never have R space denominators, since all quasi open L operators leading to  $P \rightarrow R$  transitions are by construction zero. The disconnected terms with P space intermediates are exactly cancelled by the contributions from the general sum with P space intermediates.

The situation can be illustrated using, as an example, an IMS with  $\Phi$  and singly excited states  $\{a, a\Phi\}$ . The operators of the type  $A_{\alpha}^{\dagger} a_{\alpha}$  are quasi-open type. Thus  $L_{\alpha\alpha}$  matrix-elements are zero. As a result  $\Phi$  does not couple with  $\{a, a\Phi\}$  (Fig.13(a)). The ground state energy is then just  $\langle \Phi | H_{\text{eff}} | \Phi \rangle$ , and is size-extensive. It can also be shown that the other roots are also size-extensive. In a CI matrix with the same functions, however, the situation is qualitatively different. Here  $H_{\alpha\alpha}$  is not zero (Fig.13(b)), and as a result we have disconnected terms in the ground state energy from the norm-correction having doubly, triply... R space intermediates which remain uncanceled. This is shown in Figs.14(a),(b). For more extensive discussions see e.g. /146/.

It is tempting to force IN on  $\bar{\Omega}$ , once a connected formalism is developed. Thus, it may appear that one can define a different  $\bar{\Omega}$  satisfying IN for the parent model space  $P^{(N_v)}$

$$\bar{\Omega} = \bar{\Omega} \bar{\Omega}_{c1}^{-1} \quad (8.5.1)$$

which will define an  $\bar{H}_{\text{eff}}$  related to  $H_{\text{eff}}$  by a similarity transformation and thus will generate a size-extensive theory. This is not so in practice, since for a truncated calculation it is impossible to find out an  $\bar{\Omega}$  with an exponential structure which generates an  $\bar{H}_{\text{eff}}$  similar to an  $H_{\text{eff}}$ . A simple way to look at the transformation (8.5.1) is to include in  $\bar{\Omega} = \{\exp(\bar{S})\}$  some closed operator  $X_{c1}$ , such that  $\bar{\Omega}_{c1} = 1$ . For the h-p model space, for example, a two-body  $X_{c1}$  can be used to eliminate the disconnected closed terms, as shown in Fig. 15. This, however, generates a

$$\begin{aligned}
 & \left[ \begin{array}{ccc} \langle \Phi_0 | H_{\text{eff}} | \Phi_0 \rangle & & \text{diagram} \\ 0 & \text{diagram} & + \text{diagram} - \text{diagram} \end{array} \right] \\
 & \left[ \begin{array}{ccc} \langle \Phi_0 | H & | \Phi_0 \rangle & \\ \text{diagram} & \text{diagram} & + \text{diagram} - \text{diagram} \end{array} \right]
 \end{aligned}$$

Fig.13 (a) The structure of  $H_{\text{eff}}$  matrix. Note the null-block in the lower diagonal.  
 (b) The structure of  $H$ -matrix in the same IMS.

$$\begin{aligned}
 & \text{diagram} = \text{diagram} \\
 & \text{diagram} = \text{diagram} \\
 & \text{diagram} = \text{diagram}
 \end{aligned}$$

a. b.

Fig.14 (a) The norm-correction term from Fig.13(b) at second order. (b) This can be converted to diagrams with doubly excited intermediates.

$$\left[ \text{diagram} + \text{diagram} \right] = 0$$

Fig.15 Generation of a disconnected  $X_{c1}$  after the imposition of IN on  $\Omega$ .

disconnected  $X_{c1}$  operator which will yield a disconnected  $H_{eff}$  from the expression

$$\bar{H}_{eff} = \{H\bar{\Omega}\}_{c1} \quad (8.5.2)$$

For an extensive discussion of this aspect, see the recent paper by Chaudhuri et al/146/, who also demonstrate the explicit break-down of size-extensivity of energy if IN is enforced on  $\bar{\Omega}$ .

## 9. CONCLUDING REMARKS :

We may summarize our conclusions as follows:

(a) For a size-extensive theory for both core- and valence-electrons, a Fock space formulation offers the best insight as regards the connectivity of the cluster operator  $S$  and  $H_{eff}$  for a CMS. In particular, the Fock space strategy can be used to generate energy differences with respect to a closed shell ground state.

(b) For an IMS, the connectivity of  $S$  and  $H$  can be proved only in Fock space. This requires the use of normalization for a valence universal  $\bar{\Omega}$  that is size-extensive.

(c) When  $N_C \gg N_V$ , and  $N_V$  is kept fixed, core-extensive theories using semi-cluster expansion may be a viable strategy. The quality of calculation will be increasingly poorer with the increase in  $N_V$ .

(d) From the initial trends of producing accurate results, rigorous maintenance of size-extensivity and the potentiality to obviate the intruder-state problem, it appears that the formulation covered in Sec. 8 will turn out to be one of the most versatile and viable tools in the current status of energy difference calculations (such as IP, EA and EE) in particular, and electronic structure calculations in general.

## ACKNOWLEDGEMENTS

The authors thank W.Kutzelnigg, I.Lindgren, U.Kaldor, R.J.Bartlett and S.Koch for many stimulating

discussions and for explaining their viewpoints about the open-shell cluster expansion theories, and P.O. Löwdin for his kind patience with us as the editor of this volume. Special thanks are due to S.Mukhopadhyay, D.Mukhopadhyay and A.Majumdar for their kind help in bringing the manuscript in final form. The financial support of the Department of Science and Technology (New Delhi) is also gratefully acknowledged.

## REFERENCES

1. See, e.g., (a) J. Paldus, and J. Cizek, *Adv. Quantum Chem.* 9, 305 (1975).  
(b) P. Jorgensen, and J. Simons, *Second-quantization Based Methods In Quantum Chemistry* (Acad. Press, N.Y., 1981) for pedagogical surveys.
2. C.Bloch and J.Horowitz, *Nucl. Phys.*, 8, 91 (1958).
3. B.Brandow, *Rev. Mod. Phys.*, 39, 771 (1967).  
*Adv. Quantum Chem.*, 10, 187 (1977).
4. I.Lindgren, *J. Phys.* B7, 2441 (1971).  
I.Lindgren and J.Morrison, *Atomic Many Body Theory* (Springer Verlag, Heidelberg, 1981).
5. V.Kvasnicka, *Adv. Chem. Phys.*, 36, 345 (1977).
6. G.Oberlechner, F.Owono-N-Guema and J.Richert, *Nuovo Cim.*, B68, 23 (1970).
7. T.T.S.Kuo, S.Y.Lee and K.F.Ratcliffe, *Nucl. Phys.*, A176, 65 (1971).
8. M.B.Johnson and M.Baranger, *Ann. Phys.*, 62, 172 (1971).
9. A.Banerjee, D.Mukherjee and J.Simons, *J. Chem. Phys.*, 62, 1979, 1995 (1982).
10. J.Linderberg and Y.Ohrn, *Propagators In Quantum Chemistry* (Acad. Press, N.Y., 1973).  
G.Csanak, H.S.Taylor and R.Yaris, *Adv. At. Mol. Phys.*, 7, 288 (1971).
11. L.S.Cederbaum and W.Domcke, *Adv. Chem. Phys.*, 36, 206 (1977).  
W von Niessen, L.S.Cederbaum, W.Domcke and J. Schirmer, in *Computational Methods In Chemistry* (Ed: J.Bargon, Plenum Press, N.Y., 1980).
12. (a) Y.Ohrn and G.Born, *Adv. Quantum Chem.*, 13, 1 (1981).  
(b) J.Oddershede, *Adv. Quantum Chem.*, 11, 275 (1976).
13. J.Simons, *Ann. Rev. Phys. Chem.*, 28, 15 (1975).



14. J.Paldus and J.Cizek, J. Chem. Phys., 60, 49 (1974)
15. M.F.Herman, K.F.Freed and D.Y.Yeager, Adv. Mole. Phys., 48, 1 (1981).
16. T.H.Dunning and V.Mckoy, J. Chem. Phys., 47, 1735 (1967).  
T.Shibuya and V.Mckoy, Phys. Rev., A2, 2208 (1970).
17. R.J.Bartlett and G.D.Purvis, Phys.Scripta, 21,255 (1980).  
R.J.Bartlett and G.D.Puvis, Int.J. Quantum Chem., 14, 561 (1978).
18. R.J.Bartlett, Ann. Rev. Phys. Chem., 32, 359 (1981)
19. J.Paldus, in New Horizons Of Quantum Chemistey (Ed: P.Lowdin and B.Pullman, Reidel, Dordrecht, 1983)
20. B.Brandow, in New Horizons Of Quantum Chemistry (Ed: P.Lowdin and B.Pullman, Reidel, Dordrecht, 1983)
21. (a) R.J.Bartlett, C.E.Dykstra and J.Paldus, in Advanced and Computational Approaches to the Electronic Structure of Molecules (Ed: C.E.Dykstra, Reidel,Dordrecht, 1984)  
(b) K.Jankowski, in Methods in Computational Chemistry, Vol 1 (Ed: S.Wilson, Plenum Press, N.Y., 1987)
22. V.Kvasnicka, V.Laurinc and S.Biskupic, Phys. Rep., 90, 159 (1982)  
V.Kvasnicka, V.Laurinc, S.Biskupic and M.Haring, Adv. Chem. Phys., 52, 181 (1983)
- 23 B.Brandow, Int. J. Quantum Chem., 15, 207 (1979)
24. J.A.Pople, J.S.Binkley and R.Seeger, Int. J. Quantum Chem. Symp.,10, 1 (1976)
25. K.A.Bruckner, Phys. Rev., 100, 36 (1955)
26. J.Goldstone, Proc. Roy. Soc., A239, 267 (1957)
27. J.Hubbard, Proc. Roy. Soc., A240, 539 (1957)  
N.M.Hugenholtz, Physica, 23, 481 (1957)
28. O.Sinanoglu, Adv. Chem. Phys., 6, 315 (1964)
29. H.Primas, in Modern Quantum Chemistry; Istambul Lectures (Ed: O.Sinanoglu, Acad. Press, N.Y., 1965)
30. F.Coester, Nucl. Phys., 1, 421 (1958)
31. F.Coester and H.Kumel, Nucl. Phys., 17, 477 (1960)  
H.Kumel, Nucl. Phys., 22, 177 (1969)
32. J.Cizek, J. Chem. Phys., 45, 4256 (1966)  
Adv. Chem. Phys., 14, 35 (1969)

33. J.Paldus, J.Cizek and I.Shavitt, *Phys. Rev.*, A5, 50 (1972)
34. See, e.g., refs 18,19,21,22 for extensive references
35. (a)S.Koch and W.Kutzelnigg, *Theo. Chim. Acta.*, 59, 387 (1981).  
(b)M.R.Hoffmann, H.F.Schaefer, *Adv. Quantum. Chem.*, 18, 207 (1986).
36. P.R.Taylor, G.B.Bacskay, N.S.Hush and A.C.Hurley, *Chem. Phys. Lett.*, 41, 444 (1976).
37. R.Ahlrichs, *Comput. Phys. Comm.*, 17, 31 (1979).
38. J.Paldus and J.Cizek, *Physica Scripta*, 21, 251 (1980).
39. J.W.Kenney, J.Simons, G.D.Purvis and R.J.Bartlett, *J. Am. Chem. Soc.*, 100, 6930 (1978).
40. Y.Lee, S.A.Kucharski and R.J.Bartlett, *J. Chem. Phys.*, 81, 5906 (1984).  
M.Urban, J.Noga, S.J.Cole and R.J.Bartlett, *J. Chem. Phys.*, 83, 4041 (1985).
41. S.Wilson, K.Jankowski and J.Paldus, *Int. J. Quantum Chem.*, 23, 1781 (1983)
42. G.F.Adams, G.D.Bent, G.D.Purvis and R.J.Bartlett, *Chem. Phys.*, L461 (1981).  
G.F.Adams, G.D.Bent, R.J.Bartlett and G.D.Purvis, in *Potential Energy Surfaces and Dynamics Calculations* (Ed: D.G.Truhlar, Plenum Press, N.Y., 1981).
43. P.E.Siegbahn, *Int. J. Quantum Chem.*, 18, 1229 (1980).
44. G.Hose and U.Kaldor, *J. Phys.*, B12, 3827 (1979). *Phys. Scripta*, 21, 357 (1980).  
*Chem. Phys.*, 62, 419 (1981).  
*J. Phys. Chem.*, 86, 2133 (1982).
45. M.Shepard, *J. Chem. Phys.*, 80, 1225 (1984).
46. For a thorough and modern treatment, see, e.g., W.Kutzelnigg in *Modern Theoretical Chemistry Vol 3* (Ed: H.F.Schaefer, Plenum Press, 1977).
47. W.Meyer, *J. Chem. Phys.*, 58, 1017 (1973).
48. H.Silverstone and O.Sinanoglu, *J. Chem. Phys.*, 44, 1899,3608 (1966).
49. K.Roby, *Int. J. Quantum Chem.*, 6, 101 (1972).
50. H.Nakatsuji, *Chem. Phys. Lett.*, 59, 362 (1978).  
*Chem. Phys. Lett.*, 67, 329 (1979).  
H.Nakatsuji and K.Hirao, *J. Chem. Phys.*, 68, 2053 (1978).  
*J. Chem. Phys.*, 68, 4279 (1978).

51. H.Nakatsuji and T.Yonezawa, Chem. Phys. Lett., 87 426 (1982).
52. H.Nakatsuji, Int. J. Quantum Chem. Symp., 17, 241 (1983) and references therein
53. K.Hirao, J. Chem. Phys., 79, 5000 (1983)
54. J.Paldus, J.Cizek, M.Saute and L.Laforgue, Phys. Rev., A17, 805 (1978)
55. M.Saute, L.Laforgue, J.Paldus and J.Cizek, Int. J. Quantum Chem., 15, 463 (1979)
56. H.J.Monkhorst, Int. J. Quantum Chem., S11, 421 (1977).
57. E.Dalgaard and H.J.Monkhorst, Phys. Rev., A28, 1217 (1983).
58. D.Mukherjee and P.K.Mukherjee, Chem. Phys., 37, 325 (1979).  
S.Ghosh, D.Mukherjee and S.N.Bhattacharyay, Chem. Phys., 72 961 (1982).
59. S.Ghosh, D.Mukherjee and S.N.Bhattacharyay, Mol. Phys., 43, 173 (1981).
60. S.Ghosh and D.Mukherjee, Proc. Ind. Acad. Sci., 93, 947 (1984).
61. H.Sekino and R.J.Bartlett, Int. J. Quantum Chem. S18, 255 (1984).
62. K.Emrich, Nucl. Phys., A351, 379 (1981).
63. M.D.Prasad, S.Pal and D.Mukherjee, Phys. Rev., A31, 1287 (1985).
64. A.Banerjee and J.Simons, Int. J. Quantum Chem., 19, 207 (1981).
65. A.Banerjee and J.Simons, J.Chem. Phys., 76, 4548 (1982).  
Chem. Phys., 81, 297 (1983).  
Chem. Phys., 87, 215 (1984).
66. B.Jeziorski and H.J.Monkhorst, Phys. Rev., A24, 1668 (1981).
67. D.Mukherjee, R.K.moitra and A.Mukhopadhyay, Pramana, 4, 247 (1975).  
Mol. Phys., 30, 1861 (1975).
68. (a)D.Mukherjee, R.K.Moitra and A.Mukhopadhyay, Mol. Phys., 33, 955 (1977).  
(b)D.Mukherjee, A.Mukhopadhyay and R.K.Moitra , Z.Naturforsch , 33a ,1549 (1978)
69. D.Mukherjee, Pramana, 12, 203 (1979).  
M.Haque and D.Mukherjee, J.Chem. Phys., 80, 5058 (1984).
70. R.Offermann, W.Ey and H.Kummel, Nucl. Phys., A273, 349 (1976).  
R.Offermann, Nucl. Phys., A273 , 368 (1976).  
W.Ey, Nucl. Phys., A296, 189 (1978).
71. I.Lindgren, Int. J. Quantum Chem. Symp., 12, 33 (1978).

72. V.Kvasnicka, Chem. Phys. Lett., 79, 89 (1981), Adv. Chem. Phys., 52, 181 (1982).
73. (a) S.Pal, M.D.Prasad and D.Mukherjee, Theoret. Chim. Acta., 66, 311 (1984), Theoret. Chim. Acta., 62, 523 (1983).
74. M.Haque and D.Mukherjee, Proc. 5th ICQC, Montreal(1985).
75. P.Westhaus, Int.J.Quantum Chem., 75, 463 (1973)
76. W.Kutzelnigg, J.Chem. Phys., 77, 3081 (1981) J. Chem. Phys., 80, 822 (1984).
77. W.Kutzelnigg and S.Koch, J. Chem. Phys., 79, 4315 (1983).
78. H.Reitz and W.Kutzelnigg, Chem. Phys. Lett., 66, 11 (1979).
79. P.des Cloizeaux, Phys.Rev.,135A, 685,698 (1964)
80. B.Lukman and O.Goscinski, Chem. Phys. Lett., 7, 573 (1970).
81. B.Brandow, Ann. Phys., 64, 21 (1971).
82. M.A.Robb, D.Hegarty and S.Prime, in Excited States in Quantum Chemistry (Ed: C.A.Nicolaides and D.R.Beek,Reidel).
83. T.T.S.Kuo and E.M.Krenciglowa, Nucl. Phys., A342, 454 (1980).
84. D.Mukherjee and W.Kutzelnigg, in Many Body Methods in Quantum Chemistry (Ed: U.Kaldor, Springer Verlag, Heidelberg, 1988).
85. L.Z.Stolarczyk and H.J.Monkhorst, Phys. Rev., A32, 725,743 (1986). Phys. Rev. A37,1906,1926 (1988).
86. T.H.Schucan and H.A.Weidenmuller, Ann. Phys., 73, 108 (1972). Ann. Phys., 76, 483 (1973).
87. D.Hegarty and M.A.Robb, Mol. Phys., 37, 1455 (1979). H.Baker, M.A.Robb and Z.Slattery, Mol. Phys., 44, 1035 (1981).
88. M.Haque and D.Mukherjee, Pramana, 23, 651(1984).
89. I.Lindgren, Phys. Scripta, 32, 291 (1985).
90. I.Lindgren, Phys. Scripta, 32, 611 (1985).
91. D.Mukherjee, Proc. Ind. Acad. Sci., 96, 145 (1986).
92. D.Mukherjee, Chem. Phys. Lett., 125, 207 (1986).
93. D.Mukherjee, Int. J. Quantum Chem. Symp., 20, 409 (1986).
94. (a) I.Lindgren and D.Mukherjee, Phys. Rep., 151, 93 (1987). (b) D.Mukherjee, in Condensed Matter Theories, Vol 3 (Ed: J.Arponen, R.F.Bishop and M.Monnenen, Plenum Press, N.Y., 1988).

95. W.Kutzelnigg, D.Mukherjee and S.Koch, J. Chem. Phys., 87, 5902 (1987).
96. D.Mukherjee, W.Kutzelnigg and S.Koch, J. Chem. Phys., 87, 5911 (1987).
97. S.S.Z.Adnan, S.N.Bhattacharyay and D.Mukherjee, Chem. Phys. Lett., 85, 204 (1982)
98. S.S.Z.Adnan, S.N.Bhattacharyay and D.Mukherjee, Mol. Phys., 39, 519 (1982).
99. S.K.Mukhopadhyay, D.Sinha, M.D.Prasad and D.Mukherjee, Chem. Phys. Lett., 117, 437 (1986).
100. S.Roy, S.Sengupta, D.Mukherjee and P.K.Mukherjee, Int. J. Quantum Chem., 29, 205 (1986).
101. In the collection of the various expressions for  $\bar{F}$  and  $\bar{V}$ , a particular type of integral was unfortunately dropped. In actual applications these terms are, however, present.
102. K.Hirao and H.Nakatsuji, J. Comput. Phys., 45, 246 (1982).  
See also, S.Rettrup, J. Comput. Phys., 45, 100 (1982).
103. E.R.Davidson, J.Comput. Phys., 17, 87 (1975)  
J. Phys., A13, L179 (1980).
104. P.W.Langhoff, S.T.Epstein and M.Karplus, Rev. Mod. Phys., 45, 602 (1972).
105. H.Takahashi and J.Paldus, J. Chem. Phys., 85, 1486 (1986).
106. J.Geertsen and J.Oddershede, J. Chem. Phys., 85, 2112, (1986).
107. S.Ghosh, S.K.Mukhopadhyay, R.Chaudhuri and D.Mukherjee, to be published.
108. P.O.Löwdin, Phys. Rev., 97, 1474 (1955).
109. See, e.g., P.O.Löwdin, J. Math. Phys., 3, 1171 (1962) for an early discussion of this aspect. Also, P.O.Löwdin, in Horizons of Quantum Chemistry (Ed: K.Fukui and B.Pulmann, Reidel, 1980).
110. O.Sinanoglu and I.Oksuz, Phys. Rev.Lett., 21, 507 (1988), Phys. Rev., 181, 42 (1969).  
O.Sinannoglu and B.J.Skutnik, J.Chem. Phys., 61, 3670 (1974).
111. R.J.Bruna, S.D.Peyerimhoff and W.Butscher, Mol. Phys., 35, 771 (1978).  
R.J.Bruckner and S.Peyerimhoff, in New Horizons of Quantum Chemistry, (Ed: P.O.Lowdin and B.Pulmann, Reidel, Dordrecht, 1983).
112. B.Jonsson, B.O.Ross, P.R.Taylor and P.E.M.Siegbagn, J. Chem. Phys., 74, 4566 (1981).

- H.Lischka, R.Shepard, F.B.Brown and I.Shavitt, *Int. J. Quantum Chem. Symp.*, 15, 91 (1981).
113. J. Paldus, in *Theoretical Chemistry : Advances and Perspectives Vol 2*, (Ed: H.Eyring and D.Henderson, Acad Press, N.Y., 1976).
114. I.Shavitt, *Int. J. Quantum Chem.*, S11, 131 (1977) .  
*Int. J. Quantum Chem.*, S12, 5 (1978).
115. (a) W.D.Laidig, P.Saxe and R.J.Bartlett, *J. Chem. Phys.*, 86, 887 (1987).  
(b) M.Hoffmann and J.Simons, *J. Chem. Phys.*, 88, 993 (1988).  
(c) H.Baker and M.A.Robb, *Mol. Phys.*, 50, 20 (1983).  
(d) R.Tanaka and H.Terashima, *Chem. Phys. Letts.*, 106, 558 (1984).
116. C.Bloch, *Nucl. Phys.*, 6, 329 (1958).
117. P.O.Löwdin, *Adv. Chem. Phys.*, 2, 207 (1959).
118. (a) W.D.Laidig and R.J.Bartlett, *Chem. Phys. Letts.*, 104, 424 (1984).  
(b) B.Jeziorski and J.Paldus, *J. Chem. Phys.*, 88, 5673 (1988).
119. D.Sinha, R.Chaudhuri, S.Sengupta and D.Mukherjee, to be published .
120. D.Mukherjee, R.K.Moitra and A.Mukhopadhyay, *Pramana*, 9, 6 (1977) .
121. P.Westhaus, *Int. J. Quantum Chem.*, 75, 463 (1973).  
P.Westhaus and E.G.Bradford, *J. Chem. Phys.*, 63, 5416 (1975).  
P.Westhaus, E.G.Bradford and D.Hall, *ibid*, 62, 1067 (1975) .
122. D.Mukherjee, R.K.Moitra and A.Mukhopadhyay, *Ind. J. Pure and Appl. Phys.*, 15, 613 (1977)  
A.Mukhopadhyay, R.K.Moitra and D.Mukherjee, *J. Phys*, 12, 1.
123. M.Haque, Ph. D thesis, University of Calcutta (1983).
124. S.Koch, Ph. D. thesis, Ruhr University of Bochum (1983) .
125. W.Kutzelnigg, H.Durmaz, H.Reitz and S.Koch, *Proc. Ind. Acad. Sciences*, 96, 177 (1986).
126. M.Haque and D.Mukherjee, 5th ICQC Proceedings Montreal (1985).
127. F.Jorgensen, *Mol. Phys.*, 29, 1137 (1975).
128. F.Coester, in *Lectures in Theoretical Physics*, Vol 11B (Ed: K.T.Mahanthappa, and W.E.Brittin, Gordon and Breach, N.Y., 1969) .

129. I.Lindgren, Phys. Rev., A31, 1273 (1985).
130. S.Salomonsen, Z. Phys., A316, 135 (1984).  
S.Salomonsen and A.M.Martensson-Pendrill,  
Phys. Rev., A30, 712 (1984).
131. S.Salomonsen, I.Lindgren and A.Martensson,  
Physica Scripta, 21, 351 (1981).
132. (a) A.Haque and U.Kaldor, Chem. Phys. Lett.,  
117, 347 (1985).  
ibid, 120, 261 (1985).  
ibid, 128, 45 (1986).  
(b) A.Haque and U.Kaldor, Int. J. Quantum  
Chem., 29, 425 (1986) .
133. U.Kaldor, J. Comput. Phys., 8, 448 (1987).
134. U.Kaldor, J. Chem. Phys., 87, 467 (1987).
135. D.Sinha, S.K.Mukhopadhyay, M.D.Prasad and  
D.Mukherjee, Chem. Phys. Lett., 125,  
213 (1986) .
136. D.Sinha, S.K.Mukhopadhyay, R.Chaudhuri and  
D.Mukherjee, Chem.Phys.Letts,154,544(1989).
137. E.Soliverez, Phys. Rev. A24, 4 (1981).
138. D.Sinha, S.Mukhopadhyay and D.Mukherjee,  
Chem. Phys. Letts., 127, 369 (1986).
139. S.Pal, M.Rittby, R.J.Bartlett, D.Sinha and  
D.Mukherjee, Chem. Phys. Letts, 137,  
273 (1987).
140. S.Pal, M.Rittby, R.J.Bartlett, D.Sinha and  
D.Mukherjee, J. Chem. Phys., 88,  
4357 (1988).
141. (a) M.Rittby, S.Pal and R.J.Bartlett, J.  
Chem. Phys., (in press).  
(b) R.Mattie, M.Rittby, R.J.Bartlett and  
S.Pal, in Aspects of Many Body Effects in  
Molecules and Extended Systems  
(Ed: D.Mukherjee, Springer Verlag,  
Heidelberg, in press).
142. S.Pal, R.Mattie, M.Rittby and R.J.Bartlett,  
to be published .
143. S.Koch and D.Mukherjee, Chem. Phys. Letts.,  
145, 321 (1988).
144. U.Kaldor, in Condensed Matter Theories, Vol 3  
(Ed: J.Arponen, R.F.Bishop and M.Manninen,  
Plenum Press, N.Y., 1988).
145. S.Ben-Shlomo and U.Kaldor, J. Chem. Phys.,  
89, 956 (1988).
146. R.Chaudhuri, S.Guha and D.Mukherjee, in Many  
Body Methods in Quantum Chemistry  
(Ed: U.Kaldor, Springer Verlag, Heidelberg,  
1989, in press).
147. R.Chaudhuri, D.Mukherjee and M.D.Prasad,  
in Aspects of Many Body Effects in Molecules  
and Extended Systems, (Ed: D.Mukherjee,

- Springer Verlag, Heidelberg, 1989, in press)
148. U.Kaldor, S.Roszak, P.C.Hariharan and J.J.Kaufman, J. Chem. Phys., (1988), in press.
149. S. Koch, in Aspects of Many Body Effects in Molecules and Extended Systems (Ed : D. Mukherjee, Springer Verlag, 1989, in press).
150. R.Chaudhuri, D.Mukhopadhyay and D.Mukherjee, in Aspects of Many Body Effects in Molecules and Extended Systems, (Ed: D.Mukherjee, Springer Verlag, Heidelberg, 1989, in press)

## ADDENDUM :

There have been some very recent works which we quote for completeness :

- (a) L. Meissner, K. Jankowski and J. Wasilewski, Int. J. Quantum Chem. 34, 535 (1988).
- (b) R.F. Bishop, J. Arponen and E. Pajanne, in Aspects of Many-Body Effects in Molecules and Extended Systems (Ed : D. Mukherjee, Springer-Verlag, 1989, in press).
- (c) B.Jeziorski and J.Paldus, in Many Body Methods In Quantum Chemistry (Ed :U.Kaldor, Springer Verlag, Heidelberg, 1989, in press)
- (d) U.Kaldor, in Condensed Matter Theories, Vol 4 (Ed :J.Keller, Plenum Press, 1989, in press)
- (e) S.Guha, R.Chaudhuri and D.Mukherjee, in Condensed Matter Theories, Vol 4 (Ed.:J.Keller, Plenum Press, 1989, in press)



This Page Intentionally Left Blank

JACOBI ROTATIONS: A GENERAL PROCEDURE  
FOR ELECTRONIC ENERGY OPTIMIZATION\*

Ramon Carbó  
Joan Miró

Unitat de Química,  
Col·legi Universitari de Girona,  
17071-Girona , SPAIN

Llorenç Domingo

Seccio de Química Quantica, Dept. de Quimiometria,  
Institut Quimic de Sarria,  
08017-Barcelona, SPAIN

Juan J. Novoa

Dept. de Química Física, Facultat de Química,  
Univ. de Barcelona,  
08028-Barcelona, SPAIN

\* A contribution from the "Grup de Química Quàntica del  
Institut d'Estudis Catalans"

## TABLE OF CONTENTS

1. Introduction
  2. Elementary Jacobi Rotations
  3. Variation of Quantum Mechanical Objects in the MO framework via Elementary Jacobi Rotations
  4. Some application examples
    - 4.1. EJR on monoconfigurational density matrices
    - 4.2. EJR on CI energy
    - 4.3. EJR on two electron molecular integrals
  5. Monoconfigurational energy expressions
    - 5.1. General variation structure
    - 5.2. The Roothaan-Bagus Atomic Energy as a particular case
  6. Multiconfigurational Energy Expression
    - 6.1. Variational structure
    - 6.2. Structure of MC Elementary Jacobi Rotations
    - 6.3. The CI-MO computation algorithms
    - 6.4. Results and discussion
    - 6.5 Comparison with the exponential MCSCF methodology
  7. Optimization of rotation sine
    - 7.1. An exact solution
    - 7.2. Performance improving procedures
    - 7.3. Localization procedures and Jacobi rotations
- Acknowledgements
- Appendix I
- Appendix II
- References

## 1. INTRODUCTION

Some time ago, we started /1/ the development of the Elementary Jacobi Rotation (EJR) algorithms which were known from old times /2, 3, 4/. As a consequence of this development some intermediate work has been published /5,6/. The present paper corresponds, essentially, to the application of our EJR experience to multiconfigurational calculations.

The current work presented here must be necessarily considered as a new step towards an easy, comprehensive and cheap way to obtain atomic and molecular wavefunctions. In the same path we are attempting here to show that the Jacobi Rotation techniques may be considered viable alternative procedures to the usual SCF and Unitary Transformation algorithms.

We present here some developments which can compete with other procedures found in the current literature and program lore. In this sense, we deal with the current state of the art in the calculation of the minimal rotation steps leading to the optimal energy and wavefunction.

After this, perhaps, the reader will be indulgent with the structure of this paper. First we will present a broad survey of the ideas which are necessary to build up the Jacobi Rotation techniques. A second part will study various energy forms, mono- and multiconfigurational, and some results will be given to illustrate the way optimal orthonormality constrained energies are obtained. A final part describes some algorithms used, which are important for efficient

implementation of the rotations as well as some relationships between energy optimization and localization procedures.

## 2. ELEMENTARY JACOBI ROTATIONS

1) An Elementary Jacobi Rotation (EJR) over  $n$ -dimensional space can be constructed, in order to fulfill the needs of the present work, as an orthogonal matrix  $\mathbf{J}_{ij}(\alpha) = \{J_{ij,pq}(\alpha)\}$  with the following prescription:

a) Chosen an active pair of indices  $\{i,j\}$  and a rotation angle, then:

$$b) J_{ij,ii}(\alpha) = J_{ij,jj}(\alpha) = \cos \alpha$$

$$c) J_{ij,ij}(\alpha) = -J_{ij,ji}(\alpha) = -\sin \alpha$$

$$d) J_{ij,pq}(\alpha) = \delta_{pq}; \quad p, q \neq i, j$$

It is easy to find that

$$i) \mathbf{J}_{ij}^+(\alpha) = \mathbf{J}_{ij}^{-1}(\alpha)$$

$$ii) \mathbf{J}_{ij}(-\alpha) = \mathbf{J}_{ij}^+(\alpha)$$

$$iii) \mathbf{J}_{ij}(\alpha) = \mathbf{J}_{ij}^+(\alpha)$$

2) The action of  $\mathbf{J}_{ij}$  over a  $n$ -dimensional column vector  $\mathbf{c} = \{c_p\}$  produces a new vector such as

$$\mathbf{J}_{ij} \mathbf{c} = \mathbf{r} \quad (2.1)$$

whose components are:

$$\begin{aligned} r_i &= c_i \cos \alpha - c_j \sin \alpha \\ r_j &= c_i \sin \alpha + c_j \cos \alpha \\ r_p &= c_p \quad \forall p \neq i, j \end{aligned} \quad (2.2)$$

This is sufficient to ensure norm invariance for the involved vectors:

$$\mathbf{r}^+ \mathbf{r} = \mathbf{c}^+ \mathbf{J}_{ij}^+ \mathbf{J}_{ij} \mathbf{c} = \mathbf{c}^+ \mathbf{c} \quad (2.3)$$

3) From a set of orthormalized column vectors  $\mathbf{C}=(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m)$ , that is

$$\mathbf{C}^+ \mathbf{C} = \mathbf{I} \quad (2.4)$$

any EJR over  $\mathbf{C}$  leaves the induced metric over the set invariant, that is, if:

$$\mathbf{R} = \mathbf{J}_{ij} \mathbf{C} \quad (2.5)$$

then

$$\mathbf{R}^+ \mathbf{R} = \mathbf{C}^+ \mathbf{J}_{ij}^+ \mathbf{J}_{ij} \mathbf{C} = \mathbf{C}^+ \mathbf{C} = \mathbf{I} \quad (2.6)$$

The same property is obtained if  $\mathbf{R}=\mathbf{C} \mathbf{J}_{ij}$ .

4) It is well known that the product of a set of unitary matrices gives as a result a unitary matrix. Therefore, if  $S=\{i,j\}$  is some set of active index pairs, and  $\mathfrak{R}=\{\mathbf{J}_{ij}\}$  and attached set of EJR matrices, the product

$$U = \prod_{(i,j) \in S} J_{ij} \quad (2.7)$$

is a unitary matrix.

5) A usual representation of a unitary matrix is the exponential form:

$$U = e^{iA} \quad (2.8)$$

where  $A$  is hermitean that is  $A^+ = A$ . But one can also write

$$e^{iA} = \cos A + i \sin A \quad (2.9)$$

or using the well known representation of complex matrices, it can be written:

$$\cos A + i \sin A = \begin{bmatrix} \cos A & -\sin A \\ \sin A & \cos A \end{bmatrix} \quad (2.10)$$

Thus the EJR can be viewed as the simplest form of an exponential unitary transformation matrix, expressed in a real space.

### 3. VARIATION OF QUANTUM MECHANICAL OBJECTS IN THE MO FRAMEWORK VIA ELEMENTARY JACOBI ROTATIONS

The use of EJR as a tool for Quantum Chemical purposes may be based on the simple transformation of a couple of MO

functions.

Let us suppose that  $\{|1\rangle, |2\rangle, \dots, |n\rangle\}$  are a set of monoelectronic functions, then choosing an active pair of indices  $\{i, j\}$  there can be defined an EJR of the MO pair  $\{|i\rangle, |j\rangle\}$  as:

$$\begin{aligned} |i\rangle &\longrightarrow c |i\rangle - s |j\rangle \\ |j\rangle &\longrightarrow s |i\rangle + c |j\rangle \end{aligned} \quad (3.1)$$

taking into account that  $c^2 + s^2 = 1$ .

The above transformation can be put as a variation of the MO pair, rewriting simply:

$$\begin{aligned} \delta |i\rangle &= (c-1) |i\rangle - s |j\rangle \\ \delta |j\rangle &= s |i\rangle + (c-1) |j\rangle \end{aligned} \quad (3.2)$$

Practical use of this transformation or variation scheme can be greatly simplified if in the MO framework one redefines the functions and integrals involved as quantum mechanical objects of first-, second-, ..., n-th order following that they depend on one, two, ..., n active indices prone to be transformed through an EJR.

Suppose that one is facing the transformation of any object bearing the active index pair  $\{i, j\}$ . Then, for example, the MO's  $\{|i\rangle, |j\rangle\}$  behave as first-order objects, the monoelectronic integrals  $\{h_{ii}, h_{jj}, h_{ij}\}$  or the projectors  $\{|i\rangle\langle i|, |i\rangle\langle j|, |j\rangle\langle i|, |j\rangle\langle j|\}$  as second-order objects and so on.

The interesting fact about this idea appears to be included in the transformation mechanics under EJR of objects of any order, which can be deduced from recursive transformation of lower order ones, without any further need to derive a new expression for each MO structure.

The general transformation algorithm can therefore be



described in the following form:

a) Given an active index pair  $\{i,j\}$  and an EJR sine  $s$ , as well as a set of first order objects  $\{[p]\}$ , then

b) First order objects transformation is defined as:

$$\begin{aligned} [i] &\text{-----} \rightarrow c [i] - s [j] \\ [j] &\text{-----} \rightarrow s [i] + c [j] \end{aligned} \quad (3.3)$$

c) Second order objects transform by means of the algorithm:

$$[ii] = [i] \otimes [i] \text{-----} \rightarrow (c [i] - s [j]) \otimes (c [i] - s [j]) \quad (3.4)$$

that is

$$[ii] \text{-----} \rightarrow c^2 [ii] - cs ([ij] + [ji]) + s^2 [jj] \quad (3.5)$$

which can be written as a variation:

$$\delta[ii] = (c^2-1) [ii] - cs([ij] + [ji]) + s^2 [jj] \quad (3.6)$$

and taking into account the relationships between sine and cosine, the expression is further simplified:

$$\delta[ii] = s^2 ([jj] - [ii]) - cs ([ij] + [ji]) \quad (3.7)$$

Now, calling:

$$\begin{aligned} A &= [jj] - [ii] \\ B &= [ij] + [ji] \end{aligned} \quad (3.8)$$

one arrives, using the same procedure, at

$$\delta[ij] = \delta[ji] = - (s^2 B + cs A) \quad (3.9)$$

and also:

$$\delta[jj] = -s^2 A + cs B \quad (3.10)$$

d) Any other n-th-order object can be manipulated in the same way, as, for example, if an object is written as  $[\theta_1, \theta_2, \dots, \theta_n]$ , with  $\theta_p = \{i \text{ or } j\}$ , then the EJR transformation can be deduced from any one of the direct products

$$\begin{aligned} [\theta_1, \theta_2, \dots, \theta_n] &= [\theta_1, \theta_2, \dots, \theta_{n-1}] \otimes [\theta_n] = \\ &= [\theta_1, \theta_2, \dots, \theta_{n-2}] \otimes [\theta_{n-1}, \theta_n] = \\ &= \dots \\ &= [\theta_1] \otimes [\theta_2] \otimes \dots \otimes [\theta_n] \end{aligned} \quad (3.11)$$

#### 4. SOME APPLICATION EXAMPLES

As application examples of the EJR algorithm let us first show the variation of functions of second order objects as Density matrices and CI energy expressions.

##### 4.1. EJR on Monoconfigurational Density Matrices

Using the expression

$$D = \sum_p \omega_p |p\rangle \langle p| \quad (4.1)$$

for the first order density matrix it is easy to prepare the

expression for the EJR over an active pair  $\{i,j\}$ ,

$$D = \sum_{p \neq i,j} \omega_p |p\rangle\langle p| + \omega_i |i\rangle\langle i| + \omega_j |j\rangle\langle j| \quad (4.2)$$

and using the previous results, for the variation of second-order objects to find that:

$$\delta D = (\omega_i - \omega_j) (s^2 A - c s B) \quad (4.3)$$

where  $A$  and  $B$  are defined as in the formulae (3.8) substituting the second order objects by the appropriate projectors.

#### 4.2. EJR on CI Energy

Another application example can be described in respect of the CI energy expression

$$E = \sum_P \sum_Q C_P C_Q H_{PQ} \quad (4.4)$$

where  $\{C_P\}$  are variational coefficients and  $\{H_{PQ}\}$  the hamiltonian matrix representation over the configuration state functions. EJR here can be performed in the active index pair  $\{I,J\}$  on the set  $\{C_P\}$ . This is a case of mixed first- and second-order objects, as one can write, supposing real coefficients and symmetric matrix elements:

$$\begin{aligned}
 E = & \sum_{P \neq I, J} \sum_{Q \neq I, J} C_P C_Q H_{PQ} + 2 \sum_{P \neq I, J} C_P (C_I H_{PI} + C_J H_{PJ}) + \\
 & + C_I^2 H_{II} + C_J^2 H_{JJ} + 2 C_I C_J H_{IJ}
 \end{aligned} \quad (4.5)$$

Then, after some manipulation

$$\delta E = (c-1) E_{10} + s E_{01} + s c E_{11} + s^2 E_{02} \quad (4.6)$$

where

$$\begin{aligned}
 E_{01} &= 2 \sum_{P \neq I, J} C_P (C_I H_{PI} + C_J H_{PJ}) = C_I K_I + C_J L_J \\
 E_{10} &= 2 \sum_{P \neq I, J} C_P (C_I H_{PJ} - C_J H_{PI}) = C_I L_J - C_J L_I \\
 E_{11} &= 2 (H_{JJ} - H_{II}) C_I C_J + 2 H_{IJ} (C_I^2 - C_J^2) \\
 E_{02} &= (H_{JJ} - H_{II}) (C_I^2 - C_J^2) - 4 H_{IJ} C_I C_J
 \end{aligned} \quad (4.7)$$

if the gradient of  $\delta E$  with respect to the rotation sine is found there is obtained

$$\begin{aligned}
 \frac{\partial \delta E}{\partial s} &= -t E_{10} + E_{01} + c E_{11} - s t E_{11} + 2 s E_{02} = \\
 &= E_{01} + c E_{11} + 2 s E_{02} - t (E_{10} - s E_{11})
 \end{aligned} \quad (4.8)$$

where  $t=s/c$ . So, if  $s \rightarrow 0$ ,  $c \rightarrow 1$  and thus  $t \rightarrow 0$

$$\left(\frac{\partial \delta E}{\partial s}\right)_{s=0} = E_{01} + E_{11} \quad (4.9)$$

then at the extremum

$$E_{01} + E_{11} = 0 \quad (4.10)$$

This is the same as to obtain

$$[C^+ H]_J C_I - [C^+ H]_I C_J = 0 \quad (4.11)$$

which can be considered as a CI form of Brillouin theorem.

Furthermore, this development is sufficiently general as to be considered a way to obtain the minimal eigenvalue and eigenvector of a symmetric matrix.

#### 4.3. EJR on Two-electron Molecular Integrals

Two-electron integrals of the form

$$(pq | rs) = \iint \phi_p^*(1) \phi_q(1) r_{12}^{-1} \phi_r^*(2) \phi_s(2) d\tau_1 d\tau_2 \quad (4.12)$$

transform under EJR as first, second, third, and fourth-order objects /1/. Formulae for these transformations were deduced and published some time ago /1/ and will not be repeated here, but some properties may be useful to the interested reader in order to implement the present formalism.

### 4.3.1. Coulomb and Exchange integrals.

4.3.1.1: Considering the usual Coulomb and exchange integrals  $J_{pi}$ ,  $J_{pj}$ ,  $K_{pi}$  and  $K_{pj}$  it is interesting to note that

$$\delta J_{pj} = -\delta J_{pi} \quad (4.13)$$

and

$$\delta K_{pj} = -\delta K_{pi} \quad (4.14)$$

4.3.1.2: Considering the integrals  $J_{ii}$ ,  $J_{jj}$ ,  $K_{ii}$  and  $K_{jj}$  then it is easy to show, since  $J_{ii} = K_{ii}$  and  $J_{jj} = K_{jj}$ , that  $\delta J_{ii} = \delta K_{ii}$  and  $\delta J_{jj} = \delta K_{jj}$ .

Concerning also  $J_{ij}$  and  $K_{ij}$  one can obtain a similar relationship, that is

$$\delta J_{ij} = \delta K_{ij} \quad (4.15)$$

### 4.3.2. Invariants and relationships

4.3.2.1: The following integral expressions are invariant under an EJR:

$$a) \delta(J_{ii} + J_{jj} + 2J_{ij}) = 0 \quad (4.16)$$

$$b) \delta(J_{ii} + J_{jj} - 2J_{ij} + 4K_{ij}) = 0 \quad (4.17)$$

## 4.3.2.2: Calling

$$\begin{aligned}
 P &= J_{ii} - J_{jj} \\
 Q &= (i i | i j) + (i j | j j)
 \end{aligned}
 \tag{4.18}$$

and using

$$\begin{aligned}
 \gamma &= c^4 - s^4 \\
 \sigma &= s c^3 + s^3 c
 \end{aligned}
 \tag{4.19}$$

then

$$\begin{aligned}
 P &\longrightarrow \gamma P - \sigma Q \\
 Q &\longrightarrow \gamma Q + \sigma P
 \end{aligned}
 \tag{4.20}$$

## 4.3.2.3: If now

$$\begin{aligned}
 P &= J_{ii} + J_{jj} - 2 J_{ij} - 4 K_{ij} \\
 Q &= 4 [(i i | i j) - (i j | j j)]
 \end{aligned}
 \tag{4.21}$$

and

$$\begin{aligned}
 \gamma &= c^4 - 6 s^2 c^2 + s^4 \\
 \sigma &= 4 c s (c^2 - s^2)
 \end{aligned}
 \tag{4.22}$$

then

$$\begin{aligned}
 P &\longrightarrow \gamma P - \sigma Q \\
 Q &\longrightarrow \sigma P + \gamma Q
 \end{aligned}
 \tag{4.23}$$

The theoretical framework to obtain EJR variation formulae from any MO expression has been developed so far, the following sections will deal with useful computational cases.

## 5. MONOCONFIGURATIONAL ENERGY EXPRESSIONS

### 5.1. General Variation Structure

Let us write the monoconfigurational energy in terms of the MO integrals as

$$E = \sum_p \omega_p h_{pp} + \sum_p \sum_q (\alpha_{pq} J_{pq} - \beta_{pq} K_{pq})
 \tag{5.1}$$

where  $\{\omega_p\}$ ,  $\{\alpha_{pq}\}$  and  $\{\beta_{pq}\}$  are parameters depending on the state under consideration,  $\{h_{pq}\}$  the monoelectronic hamiltonian integrals,  $\{J_{pq}\}$  the two-electron Coulomb integrals and  $\{K_{pq}\}$  the exchange ones.

Upon choosing an active orbital index pair  $\{i,j\}$  and a rotation sine, variation of the energy through a EJR can be obtained from the preceding section. The net result is a fourth order sine polynomial

$$\delta E = s c E_1 + s^2 E_2 + s^3 c E_3 + s^4 E_4
 \tag{5.2}$$



for the energy variation, with the additional complication of the presence of a cosine factor in the odd terms. The polynomial coefficients  $\{E_k\}$  are complicated functions of the MO integrals and can be written using

$$\begin{aligned} a_{ij} &= \alpha_{ii} - 2\alpha_{ij} + \alpha_{jj} \\ b_{ij} &= \beta_{ii} - 2\beta_{ij} + \beta_{jj} \end{aligned} \quad (5.3)$$

as

$$\begin{aligned} E_1 &= 4 \langle i | F_j - F_i | j \rangle \\ E_2 &= 2 \{ \langle i | F_j - F_i | i \rangle - \langle j | F_j - F_i | j \rangle + 2 a_{ij} K_{ij} - b_{ij} (J_{ij} + K_{ij}) \} \\ E_3 &= 4 \langle i | J_i - J_j | j \rangle (a_{ij} - b_{ij}) \\ E_4 &= (J_{ii} + J_{jj} - 2J_{ij} - 4K_{ij}) (a_{ij} - b_{ij}) \end{aligned} \quad (5.4)$$

where  $\{F_i, F_j\}$  are the Fock operators for the orbitals  $\{i, j\}$ , that is, for example:

$$F_i = \frac{1}{2} \omega_i h_i + \sum_p (\alpha_{ip} J_p - \beta_{ip} K_p) \quad (5.5)$$

being  $\{J_p\}$  and  $\{K_p\}$  the usual Coulomb and exchange operators respectively.

The interesting features of these coefficients can be summarized in the following points:

a)  $E_1$  is no more than the Brillouin matrix element for orbitals  $\{i, j\}$ . Thus, at convergence, one must have  $E_1=0$ . This feature will be further discussed when describing sine optimization procedures, in section 7.1.4.

b)  $E_4$  contains an invariant integral structure as has been previously obtained.

c)  $E_3, E_4$  are related by a quadruple rotation as described in section 4.3.2.

d)  $E_2$  is the term which drives the sign of the hessian of the energy with respect to the sine, thus governing the structure of the energy minimum.

Finally, Fock operators have been used in order to obtain compact expressions for the coefficients but they are not necessary, obviously, to the implementation of the algorithm, which can be entirely written over MO integrals.

Application examples to various molecular structures will be shown in the multiconfigurational section, where the monoconfigurational MO space has been used as starting point for every calculation reported.

## 5.2. The Roothaan-Bagus Atomic Energy as a Particular Case

According to Roothaan and Bagus /7/, the energy of some atomic states can be written as a variational functional of two density matrices: the usual total density matrix containing all the occupied orbitals with their corresponding well known occupation numbers, and a second one, in which the occupation numbers for the doubly occupied orbitals are set to zero, namely, the open shell density matrix.

In this way, two different sets of occupation numbers may be defined:  $\omega_T$ , the classical ones, and the open shell occupation numbers:  $\omega_\theta$  which fulfill

$$\omega_\theta = \delta(i \in \theta) \omega_T \quad (5.6)$$

The particular electronic state being computed is defined through a set of parameters which are included in the construction of the two-electron contribution hypermatrices called P and Q. These hypermatrices correspond to the total and open shell density matrices, respectively, and are built up in such a way that they behave as Coulomb terms in front of MO variation.

In this way, the energy expression is written as

$$E = H:D_T + \frac{1}{2} [D_T:P:D_T - D_\theta:Q:D_\theta] \quad (5.7)$$

where  $D_T$  and  $D_\theta$  are the density matrices above mentioned, and H, P and Q are the one-electron hamiltonian and two-electron integral contribution hypermatrices.

Any Jacobi rotation between two MO's will simply modify both density matrices in the way described in section 4.1. Once these density variation expressions, see formula (4.3), are applied on the Roothaan-Bagius energy and after collecting terms of similar order the usual four-parameter expression for the energy increment (5.2) is obtained. The parameters are now defined by the formulae:

$$\begin{aligned} E_1 &= F:B \\ E_2 &= F:A + \frac{1}{2}Y:B \\ E_3 &= X:B \\ E_4 &= \frac{1}{2}(X:A - Y:B) \end{aligned} \quad (5.8)$$

where F is the Fock matrix:

$$F = \Delta\omega_T H + \Delta\omega_T D_T : P - \Delta\omega_0 D_0 : Q \quad (5.9)$$

and X, Y two auxiliary matrices defined by:

$$\begin{aligned} X &= (\Delta\omega_T^2 P - \Delta\omega_0^2 Q) : A \\ Y &= (\Delta\omega_T^2 P - \Delta\omega_0^2 Q) : B \end{aligned} \quad (5.10)$$

being A, B, and  $\Delta\omega_T$ ,  $\Delta\omega_0$  defined as in formulae (3.8) and (4.3).

When symmetry relationships are introduced in the calculation, it is important to realize that only the submatrices corresponding to the rotating orbital symmetry have to be computed for both sets of matrices A, B and X, Y.

If optimization is carried out in the MO basis, the variational terms are the integrals instead of the density matrices, which become diagonal and their values equal to the occupation numbers. The Roothaan-Bagus energy expression can therefore be written as:

$$\begin{aligned} E = & \sum_{i \in C} \omega_i h_{ii} + \sum_{p \in O} \omega_p h_{pp} + \frac{1}{2} \left( \sum_{i \in C} \sum_{j \in C} \omega_i \omega_j P(i i | j j) + \right. \\ & \left. 2 \sum_{i \in C} \sum_{p \in O} \omega_i \omega_p P(i i | p p) + \sum_{p \in O} \sum_{q \in O} \omega_p \omega_q (P(p p | q q) - Q(p p | q q)) \right) \end{aligned} \quad (5.11)$$

with the same definitions as above for the  $\omega$ , h, P and Q, but now, these last three must be considered calculated on the MO basis.

The variation of the integrals under a single Jacobi rotation over an active pair of orbitals  $\{i,j\}$  has already been defined in section 4.3, and noting that both  $P$  and  $Q$  hypermatrices transform in the same way, these expressions alone will suffice, through substitution in the energy expression, to obtain the formulae for the definition of the energy variation parameters:

$$\begin{aligned}
 E_1 &= 2 \Delta\omega_T [h_{ij} + \sum_k \omega_k P(ij|kk)] - 2\Delta\omega_\theta \sum_{p \in O} \omega_p Q(ij|pp) \\
 E_2 &= 2\Delta\omega_T [(h_{ii} - h_{jj}) + \sum_k \omega_k (P(i i | kk) - P(j j | kk))] + \\
 &\quad + 2 \Delta\omega_T^2 P(ij|ij) - 2\Delta\omega_\theta [\sum_{p \in O} \omega_p (Q(i i | pp) - Q(j j | pp))] - \\
 &\quad - 2 \Delta\omega_\theta^2 Q(ij|ij) \\
 E_3 &= 2\Delta\omega_T^2 [P(i i | ij) - P(i j | j j)] - 2\Delta\omega_\theta^2 [Q(i i | ij) - Q(i j | j j)] \\
 E_4 &= \frac{1}{2} \Delta\omega_T^2 [P(i i | i i) + P(j j | j j) - 2P(i i | j j) - 4P(i j | i j)] - \\
 &\quad - \frac{1}{2} \Delta\omega_\theta^2 [Q(i i | i i) + Q(j j | j j) - 2Q(i i | j j) - 4Q(i j | i j)]
 \end{aligned} \tag{5.12}$$

This expression together with the ones that perform the rotation of the  $h$ ,  $P$  and  $Q$  integral sets, provide the algorithm to implement the Jacobi rotation method on the MO basis using the Roothaan-Bagius energy expression.

Both algorithms were implemented on a highly modified version of Roos, Salez, Veillard and Clementi atomic program /8/ and tested for a very large number of the calculations reported by Clementi and Roetti /9/.

## 6. MULTICONFIGURATIONAL ENERGY EXPRESSION

In the initial work /1/ a general review on SCF methodology was presented, but as time has passed by we think necessary an updating of the state of the art in this field. For this purpose it will be presented a broad analysis of the development of multiconfigurational techniques covering the 1980-1987 period.

During this years some reviews have appeared. The first one is due to Hinze /10/. Inside a report edited by Dupuis /11/ one can find some other revision articles, as the one due to Dietrich and Wahl. Another excellent review is due to Olsen et al. /12/. Finally, Almlof and Taylor reviewed the implementation on vector-oriented hardware, while some of the chapters in the two volume series were dedicated to the MCSCF methodology.

Other contributions in this field deal on some specific subjects as follows:

Unitary transformations had been studied by Polezzo and Fantucci /15, 16, 17/, Matsen and Nelling /18/, and Fantucci et al. /19/.

Stability and initial vectors had been discussed by Levy /20/, general theoretical and optimization problems had been studied by Mukherjee /21/, McCullough /22/, Igawa et al. /23/, Lengsfeld /24/, Polezzo and Fantucci /25/, Lengsfeld and Liu /26/, Polezzo /27/, Werner and Meyer /28/, Ishikawa /29/, Siegbahn et al. /30/, Brooks et al. /31/.

General convergence problems and other aspects had been treated by Yurtsever /32/, Das /33/, Camp and King /34/.

Baushlicher et al. /35/, and Werner and Knowles /36/.

During this period, Joergensen's group has made a large contribution to this subject, as the interested reader can find in references /37/ to /47/.

After facing the cumbersome but simple problem of integral energy transformation in EJR techniques, and having solved and used it in atomic and molecular monoconfigurational frameworks, we managed to construct a multiconfigurational code, ARIADNE, which has been evolved through the various hardware facilities available to us. The following theoretical discussion was tested and results were obtained by means of a VAX-11 version of the ARIADNE program.

### 6.1. Variational Structure

The use of unitary transformations in MCSCF procedures is a common practice in Quantum Chemistry today. The most extended type of unitary transformations are the exponential ones, introduced by Levy /48/, defined using the following equation:

$$U=\exp(X), \quad X^T=-X \quad (6.1)$$

The implementation of this type of transformation in the multiconfigurational problem has been very successful when associated with energy direct minimization. Newton-Raphson or related techniques have been used for this task /44/. The problem in these implementations is the truncation normally done on the exponential transformation during the computational process, either by the use of the Taylor expansion of the exponential

$$U = \exp(X) = 1 + X + (1/2) X^*X + \dots \quad (6.2)$$

or by the use of the Hausdorff transformation

$$H' = \exp(-X) H \exp(X) = H + [H, X] + 0.5 * [[H, X], X] + \dots \quad (6.3)$$

In both cases, the energy increment in each transformation is given by the approximate formula truncated at the second order

$$E' - E \cong g^*X + 0.5 * X^* G * X \quad (6.4)$$

where  $g$  is the gradient and  $G$  the hessian matrix of the energy with respect to the variational parameters  $X$ . The direct consequence of such a truncation is that these methods are no longer exact and, therefore, one needs to use a Newton-Raphson iterative procedure in order to find the solution. Very efficient techniques have been described in the literature in order to optimize the performances of the process. An excellent review of some of these techniques can be found in the work of Olsen, Yeager, and Joergensen /14/.

The Elementary Jacobi Rotations can be taken as an alternative form for the unitary matrix transformation of the multiconfiguration energy. The exponential and the EJR forms of the unitary matrix can be related through the expressions, already discussed in section 2.5, because the EJR are the minimal parts in which the exponential transformations can be divided.

Based on the EJR technique it is possible to develop a new procedure for MC energy optimization with no truncation in the energy increment expressions. Therefore, this formalism may be more powerful than those based on the exponential



transformations. In the following sections we will study the general properties of the MC method based on EJR, comparing some of our results with those obtained from the exponential transformations.

## 6.2. Structure of MC Elementary Jacobi Rotations

Using the same conventions previously introduced one can distinguish between two types of rotations: The MO rotations and the CI coefficients rotations of the MC vector, which appear in a CI wavefunction

$$\Psi = \sum_K C_K |K\rangle \quad (6.5)$$

where  $|K\rangle$  are the chosen polyelectronic basis functions.

Each MO EJR transformation is defined as in the monoconfigurational case, see formulae (3.1), whilst the CI rotations are similarly defined on the coefficients of the MC vector as in formulae (2.2).

Application of such transformations to the electronic energy expression:

$$\begin{aligned} E &= \sum_{K,L} C_K C_L \langle K | H | L \rangle = \\ &= \sum_{i,j} \omega_{ij} \langle i | h | j \rangle + \sum_{i,j,k,l} \gamma_{ijkl} \langle i j | r_{12}^{-1} | k l \rangle \end{aligned} \quad (6.6)$$

generates the energy increment for each rotation. In order

to simplify the problem one can suppose that the CI and MO rotations are mutually independent. Thus, a two-step iteration MC formalism is obtained, which can be resumed in the equations:

**CI:**

$$\Delta E = (\sigma - 1) \xi_{10} + \sigma \xi_{01} + \sigma \gamma \xi_{11} + \sigma^2 \xi_{02} \quad (6.7)$$

where  $\sigma$ ,  $\gamma$  are the CI rotation sine and cosine respectively.

**MO:**

$$\begin{aligned} \Delta E = E_{10} (c-1) + (E_{01} + E_{11} c)s + (E_{02} + E_{12} c) s^2 + (E_{03} + E_{13} c)s^3 + \\ + E_{04} s^4 \end{aligned} \quad (6.8)$$

where  $s, c$  are the MO rotation sine and cosine respectively.

Both polynomials define completely the structure of the EJR MC procedure by means of the values of the coefficients  $\xi_{ab}$  and  $E_{ab}$ . The equations used to compute the MO coefficients can be found in Appendix 1.

In order to find the optimal rotation one only needs to compute the absolute minimum of each polynomial with respect to the MO and CI rotational sines. This can be done using a Newton-Raphson or any related technique, in a very simple way, as will be shown in the section 7. Once the computed optimal sines are known, one can obtain the value of the electronic energy increment for each rotation and compute a new set of values for MO and CI coefficients.

### 6.3. The CI-MO Computation Algorithms

Using the equations described in section 6.2, it is possible to compute the optimal rotation for each pair of MO or CI coefficients until convergence. As in the monoconfigurational case, the process can be simplified for the MO rotations if one takes into account the fact that rotations between orbitals of the same shell or of different symmetry species keep invariant the electronic energy. Similarly one can apply the same symmetry considerations to the CI rotations. A scheme of the whole MC computational process is described in Appendix 2.

The MC procedure, as given in Appendix 2, will converge in one iteration if no coupling between the MO's and CI coefficients is present, and the MO rotations are fully independent one from the other. However, this is an ideal situation, because in practice the rotation of a pair of MO's changes the remaining orbital structure. This is not surprising if one considers that each MO rotation is a step forward in the diagonalization of the generalized Brillouin matrix, whose elements are defined as:  $\langle i | F_{ij} - F_{ji} | j \rangle$ ; where  $F_{ab}$  are generalized Fock operators, and it is well known that in Jacobi matrix diagonalization each off-diagonal element annihilation produces some changes in the remaining matrix elements. The considerations made above are a direct consequence of the structure of the coefficients of the MO energy increment polynomial, as it is shown in Appendix 1, where it can be seen that the molecular integrals are present everywhere.

Due to the MO coupling, the procedure given in Appendix 2 has an iterative form in which the rotations in the MO space are repeated until convergence.

Because of the couplings within the CI coefficients rotation scheme, a similar iterative procedure is necessary for this part.

Finally, the coupling between MO and CI coefficients, although small, as can be seen in Tables III and IV, forces repetition of the two MO and CI rotation schemes.

Depending on the case and precision, each partial process converges in few iterations.

An alternative procedure, hereafter identified as algorithm MC-B in order to distinguish it from the previous one which will be called algorithm MC-A, can be constructed from the one described in Appendix 2. It will be sufficient for this purpose to skip the MO and CI internal iterative steps, keeping the external loop, until convergence is reached.

## 6.4. Results and Discussion

### 6.4.1. Global convergence tests

In order to study the behaviour of the EJR algorithms the computations of Wood and Veillard on the H<sub>2</sub> molecule /49/ have been repeated, owing to the fact this is one of the tests normally used in the exponential MC methods. The results are included in Table I.

As can be seen in Table I, algorithm MC-A does not converge in one iteration but almost quadratically in the same number of iterations as the full Newton-Raphson exponential technique. This is due, as it is shown in Table II, to the existence of a small coupling between the MO and CI rotations after the first iteration which is significative at the degree of precision used ( $1.0 \times 10^{-8}$  a.u.). It must be noticed that for this

TABLE I. Total energy (a.u.) computed by different MC methods for the H<sub>2</sub> molecule in its ground state. The basis set and geometry are those of Wood and Veillard /49/.

METHOD				
ITERATION	Exact exponential <sup>a</sup>	Algorithm MC-A (this work)	Algorithm MC-B (this work)	First order MC method <sup>b</sup>
0	-1.13411 <sup>c</sup>	-1.13268666 <sup>d</sup>	-1.13268788 <sup>d</sup>	-1.13410 <sup>c</sup>
1	-1.15089	-1.15129176	-1.15121878	-1.13956
2	-1.15130	-1.15130474	-1.15130498	-1.14883
3	-1.15131	-1.15130503	-1.15130503	-1.15087
4				-1.15122
5				-1.15129
6				-1.15130
7				-1.15130
8				-1.15131

(a) Results taken from reference 61.

(b) Results from reference 49. A good starting point or level shifting used. Otherwise the process converges to the SCF value.

(c) After a CI computation (see ref. 49).

(d) After the SCF result. The result obtained in a complete CI with this basis set is -1.16838 a.u. , the exact energy being -1.1745 a.u. /49/.

molecule algorithms MC-A and MC-B give the same final result in the same number of global iterations. Furthermore, the results of both algorithms are invariant upon the choice of starting CI and MO coefficients.

In order to complete the study we have also tested the convergence of both algorithms for the Li<sub>2</sub> system ground state, taking as reference the MC results of Lengsfeld /50/. The

**TABLE II.** Structure of the electronic energy increment (a.u.) using algorithm MC-A on the H<sub>2</sub> molecule in its ground state and at the equilibrium geometry ( $r(\text{H-H})=1.4$  a.u.). The basis set used is the uncontracted 5s of Huzinaga (ref. 59) plus a set of p orbitals with exponents 0.9 .

ITERATION	ELECTRONIC ENERGY	ENERGY INCREMENT	
		CI	MO
0	-1.84697237	-0.01297389	-0.00563118
1	-1.86557746	-0.00001132	-0.00000166
2	-1.86559045	-0.00000024	-0.00000003
3	-1.86559072	-0.00000001	-0.00000000
4	-1.86559073		
TOTAL ENERGY =		-1.15130503	

**TABLE III.** Variation of the electronic energy (a.u.) during the MC computation using the algorithm MC-A on the ground state of the Li<sub>2</sub> molecule. The internuclear distance is set to 2.5 a.u. and the basis set is given in reference 50.

ITERATION	ELECTRONIC ENERGY	ENERGY INCREMENT	
		CI	MO
0	-18.31803889	-0.00281133	-0.00252618
1	-18.32337641	-0.00643254	-0.00053385
2	-18.33034282	-0.00006846	-0.00000526
3	-18.33041655	-0.00000021	-0.00000012
4	-18.33041688		
TOTAL EJR ENERGY =		-14.73041687	
TOTAL ENERGY <sup>a</sup> =		-14.73041770	

(a) Result obtained by Lengsfeld (ref. 50) using a Full MC exponential method after 4 iterations.

**TABLE IV.** Variation of the electronic energy (a.u.) during the MC computation using algorithm MC-B on the ground state of the  $\text{Li}_2$  molecule. The internuclear distance is set to 2.5 a.u. and the basis set is given in reference 50.

ITERATIONS	ELECTRONIC ENERGY	ENERGY INCREMENT	
		CI	MO
0	-1.831818370	-0.00275609	-0.00223826
1	-1.832317806	-0.00666303	-0.00054629
2	-1.833038739	-0.00002246	-0.00000361
3	-1.833041347	-0.00000000	-0.00000182
4	-1.833041529	-0.00000000	-0.00000086
5	-1.833041615	-0.00000000	-0.00000041
6	-1.833041656	-0.00000000	-0.00000019
7	-1.833041676	-0.00000000	-0.00000009
8	-1.833041685	-0.00000000	-0.00000004
9	-1.833041689	-0.00000000	-0.00000002
10	-1.833041692	-0.00000000	-0.00000001
11	-1.833041693	-0.00000000	-0.00000000
TOTAL ENERGY =		-14.73041693	

results for this molecule, which are shown in Table III and Table IV for algorithms MC-A and MC-B respectively, are similar to those reported for the  $\text{H}_2$  molecule; namely, the convergence when using algorithm MC-A is found in 4 iterations and not in only one, due to the existence of a coupling between MO's and CI coefficients. Algorithm MC-B needs more iterations in order to reach the same final result, and has after the first three iterations a linear convergence behaviour instead of the quadratic shown by the algorithm MC-A.

#### 6.4.2. Structure of each EJR rotation.

As was stated before, the structure of each EJR rotation is fully defined by the values of the characteristic parameters  $\xi_{ab}$  and  $E_{ab}$  for the CI and MO rotations, respectively. The energy parameters seem arbitrary but, in practice, their values present some relationships which simplify notably the form of the polynomial. The values of those parameters for the first iteration, included in the Table V for the  $\text{Li}_2$  molecule computation, are a good example of this fact. The values found show that, for the MO rotations, the  $E_{03}$ ,  $E_{12}$ ,  $E_{01}$ , and  $E_{10}$  coefficients are null. Therefore, the energy variation polynomial shape is defined by the remaining four parameters, and, in particular, by the dominant one, the  $E_{02}$  parameter. Only when this parameter is small enough the optimal sine has a value which is significantly different from zero.

Some systematic study has shown that the polynomial shape is only affected by the relative weight of the  $E_{04}$  term against the  $E_{02}$  one. In the polynomial structure, one can appreciate three different parts: when  $E_{02} > 0$  where only one minimum at  $s=0$  is present, next when  $E_{02} < 0$  then two minima at values of the sine  $+1$  and  $-1$  appear, and finally when  $E_{02} = 0$  in which case the situation depends on the particular problem and there can be more than one minima. Therefore, one needs good optimization algorithms in order to locate the absolute minimal energy sine value.

Finally, again for  $\text{Li}_2$ , Table V also illustrates that at the end of the iterative process, at sweep 7, the sum of  $E_{11}$  and  $E_{01}$  must be null, a consequence of the Generalized Brillouin theorem.

In the case of the CI rotations, as seen in Table VI the



**TABLE V.** Values (in a.u.) of the non-null  $E_{ij}$  coefficients of the MO energy increment for the first MC iteration of the  $\text{Li}_2$  molecule given in Table IV. The optimum sine ( $s$ ) and the computed energy increment with this sine (in a.u.) are also given for the first two and the last EJR sweep whose number is also indicated. Only those rotations giving energy increments greater than  $1.0 \times 10^{-12}$  a.u. are included in the table.

ROTATING							ENERGY
SWEEP	MO's	$E_{04}$	$E_{13}$	$E_{02}$	$E_{11}$	$s$	INCREMENT
1	3 - 1	0.0065	0.0003	0.0431	-0.0120	0.1348	-0.00082209
	4 - 2	0.3162	0.4902	0.0526	-0.0036	0.0003	-0.00000064
	5 - 1	-0.1258	-0.4356	6.7793	0.0012	-0.0001	-0.00000006
	5 - 3	-0.0209	-0.1655	2.1898	0.0141	-0.0032	-0.00002294
	6 - 2	0.6857	-0.2948	5.4391	-0.0000	-0.0000	-0.00000000
	6 - 4	0.0004	0.0007	0.0081	-0.0172	0.5285	-0.00533516
	7 - 1	0.0896	0.3382	7.9691	0.0003	-0.0000	-0.00000000
	7 - 3	0.0041	0.0783	3.3117	0.0131	-0.0020	-0.00001296
	8 - 2	-0.1898	-0.3898	9.2903	-0.0009	0.0001	-0.00000002
	8 - 4	0.0008	-0.0019	0.0528	-0.0128	0.1195	-0.00077124
2	3 - 1	0.0054	0.0036	0.0562	0.0037	-0.0333	-0.00006238
	4 - 2	0.6998	0.2366	4.6245	0.0019	-0.0002	-0.00000019
	5 - 1	-0.1226	-0.4529	6.7888	0.0001	-0.0000	-0.00000000
	5 - 3	-0.0127	-0.1588	2.1915	-0.0046	0.0010	-0.00000244
	6 - 2	0.3289	-0.4692	6.0079	-0.0022	0.0002	-0.00000020
	6 - 4	-0.0004	-0.0000	0.0201	0.0046	-0.1127	-0.00026236
	7 - 1	0.0935	0.3481	7.9747	0.0006	-0.0001	-0.00000001
	7 - 3	0.0114	0.0752	3.3067	-0.0023	0.0003	-0.00000040
	8 - 2	-0.2337	-0.3889	9.3618	0.0003	-0.0000	-0.00000000
	8 - 4	0.0013	-0.0018	0.0548	-0.0046	0.0421	-0.00009747
7	3 - 1	0.0058	0.0029	0.0557	-0.0000	0.0000	-0.00000000
	4 - 2	-0.0125	-0.1588	4.6468	-0.0000	0.0000	-0.00000000
	5 - 3	-0.0125	-0.1588	2.1916	0.0007	-0.0002	-0.00000005
	6 - 4	-0.0001	0.0002	0.0205	0.0000	-0.0007	-0.00000001
	7 - 3	0.0112	0.0754	3.3071	0.0007	-0.0001	-0.00000003
	8 - 4	0.0015	-0.0017	0.0549	-0.0000	0.0003	-0.00000000

TABLE VI. Values (in a.u.) of the non-null  $\xi_{ij}$  coefficients of the CI energy increment for the first MC iteration of the  $\text{Li}_2$  molecule computed with the same basis set as in Table IV with a CI space of 20 configurations. The optimum computed energy increment (in a.u.) is also given. For the first two sweeps all the rotations giving energy increments greater than  $1.0 \times 10^{-6}$  a.u. are included. Only the 5-1 rotation is shown in the last sweeps.

SWEEP	ROTATING CI COEFFICIENTS	$\xi_{10}$	$\xi_{01}$	$\xi_{11}$	$\xi_{02}$	ENERGY INCREMENT
1	4-1	0.	0.	-0.015	4.970	-0.000012
	5-1	0.	-0.004	-0.017	2.967	-0.000035
	6-1	0.	-0.007	-0.100	0.910	-0.003141
	6-4	0.006	-0.001	0.002	0.014	-0.000002
	6-5	0.006	-0.001	-0.007	0.008	-0.001274
	13-1	0.006	0.006	0.	7.142	-0.000001
	15-1	0.006	0.	0.204	9.069	-0.001149
	16-1	0.008	-0.016	0.185	7.365	-0.000964
	16-5	0.006	-0.002	0.002	0.001	-0.000003
	17-1	0.010	0.017	-0.089	5.601	-0.000234
2	20-1	0.010	0.017	0.	9.735	-0.000007
	5-1	0.013	-0.137	0.110	2.956	-0.000062
	5-4	0.003	0.001	0.	0.002	-0.000005
	6-1	0.008	-0.057	-0.001	0.915	-0.000893
	6-5	0.009	0.	-0.006	0.024	-0.000341
	15-4	0.002	0.	-0.001	0.000	-0.000013
	15-5	0.009	0.004	-0.003	0.007	-0.000003
3	20-1	0.014	0.025	-0.017	9.690	-0.000002
	5-1	0.020	-0.206	0.188	2.940	-0.000026
4	5-1	0.025	-0.244	0.234	2.928	-0.000009
5	5-1	0.028	-0.266	0.260	2.921	-0.000003
6	5-1	0.029	-0.278	0.275	2.916	-0.000001
...		...	...	...	...	...
11	5-1	0.032	-0.293	0.293	2.910	-0.000000

structure of the EJR rotations is very simple. At convergence, Generalized Brillouin Theorem holds and therefore imposes that the sum  $\xi_{01} + \xi_{11}$  becomes zero. As a consequence, the magnitude of the rotation sine will be strongly related with the evolving values of  $\xi_{01}$  and  $\xi_{11}$ . The convergence is normally achieved in few iterations, when the number of configurations is small.

#### 6.4.3. Extension to other molecular systems

In order to test the general validity of the previous results, first we have applied algorithm MC-A to a set of diatomic systems, shown in Table VII. In all the cases, we have used SCF MO's as starting point. The "split valence" basis set of Dunning and Hay /51/ was used for each atom and the interatomic distances were taken from the work of Snyder and Basch /52/ or, if not quoted there, as the experimental reported value. Almost quadratic convergence is obtained in all cases. The number of MO iterations needed in each global iteration varies with the number of orbitals present, as expected. Furthermore, we have also verified the invariance of the final result from the starting point.

Results for some polyatomic molecules, given in Table VIII using Algorithm MC-B, show a similar behaviour.

#### 6.4.4. On choosing the CI space.

In order to obtain information on the best CI space we have carried out a systematic search with a different number of configurations on the helium atom using a STO triple zeta quality basis set. The results are given in Table IX and show that the best convergence with optimum correlation

**TABLE VII.** Total energy for the ground state of various diatomic molecular structures studied using the Algorithm MC-A

MOLECULE	num. orb.	num. el.	E(SCF) (a.u.)	num. conf.	E(MC) (a.u.)	num. iter.	C(SCF)
H <sub>2</sub>	16	2	-1.13268787	2	-1.15130503	3	0.994
LiH	11	4	-7.98243979	2	-7.98349017	6	0.999
Li <sub>2</sub>	8	6	-14.71814758	2	-14.73041693	3	0.968
BeH	11	5	-15.14360491	2	-15.15449758	3	0.996
CH <sup>+</sup>	11	6	-37.88516611	2	-37.90448079	4	0.994
CH	11	7	-38.25735724	2	-38.27569839	5	0.994
NH	11	8	-54.94899181	2	-54.97034529	4	0.993
OH <sup>+</sup>	11	8	-74.96433590	2	-74.98674456	4	0.994
OH	11	9	-75.38088047	2	-75.40549505	5	0.993
OH <sup>-</sup>	11	10	-75.34782807	2	-75.37201130	5	0.992
C <sub>2</sub>	18	12	-75.35644086	3	-75.41475156	4	0.964
BeO	18	12	-89.39400793	2	-89.40884632	8	0.995
CN	18	13	-92.14880317	3	-92.21242918	4	0.965
FLi	18	12	-106.96395611	2	-106.97651066	5	0.997
N <sub>2</sub> <sup>+</sup>	18	13	-108.29770309	3	-108.30286184	5	0.999
N <sub>2</sub>	18	14	-108.87813657	3	-108.94795882	4	0.966
CO	18	14	-112.68483960	3	-112.73475610	5	0.979
BF	18	14	-124.08140340	3	-124.10727997	4	0.997
NO	18	15	-129.19383614	2	-129.23964110	4	0.975
O <sub>2</sub>	18	16	-149.57112871	3	-149.60159868	4	0.992
F <sub>2</sub>	18	18	-198.70749705	2	-198.78989444	5	0.960
F <sub>2</sub> <sup>-</sup>	18	19	-198.81385465	2	-198.81689740	4	0.999
AlH	20	14	-242.42947043	2	-242.44235399	4	0.999
SiH	20	15	-289.23178590	2	-289.38999376	4	0.999
PH	20	16	-341.24176813	2	-341.25451937	4	0.999
SH	20	17	-398.03741786	2	-398.05039714	4	0.999
HCl	20	18	-460.03030032	2	-460.04303296	4	0.999

**TABLE VIII.** Total energy (in a.u.) with the selected configurations for the ground state of some polyatomic molecules. Snyder and Basch geometry (ref. 52), the basis set given by Dunning and Hay (ref. 51), and Algorithm MC-B are used.

<u>Molecule</u>	<u>Symmetry</u>	<u>State</u>	<u>Number of state functions</u>	<u>Total Energy</u>	<u>Iterations</u>
CH <sub>2</sub>	C <sub>2v</sub>	<sup>1</sup> A <sub>1</sub>	1	-38.848867	8
			95	-38.901893	
N <sub>2</sub> H <sub>2</sub>	C <sub>2h</sub>	<sup>1</sup> A <sub>g</sub>	1	-109.697171	8
			8	-109.713451	
H <sub>2</sub> CO	C <sub>2v</sub>	<sup>1</sup> A <sub>1</sub>	1	-113.827063	4
			12	-113.877026	
CH <sub>3</sub> OH	C <sub>1h</sub>	<sup>1</sup> A'	1	-115.004535	6
			12	-115.033926	

energy is obtained when the CI space is formed by the reference configuration plus all the double excitations from that one, or, alternatively, using a complete active space for the selected AO's. The conclusion that for a two electron system the best MCSCF function consists of the reference function and all the double excited ones, is a reformulation of the fact that, using natural orbitals, the full CI function may be written in terms of configurations consisting of a doubly occupied orbital only /53, 54/. In the cases where the configurational space was highly insufficiently described the process failed to converge in a reasonable number of iterations using either Algorithms MC-A or MC-B.

**TABLE IX.** Total energy (in a.u.) computed by different configurational spaces for the Helium atom using a STO of triple zeta quality basis set. The configurations included in each case are indicated by the values of the coefficients when the process converges, otherwise are not shown. The criterion of convergence is set to  $1.0 \times 10^{-6}$  a.u.

Configurations selected						Total Energy	Global Iterations
(1s,1s)	(1s,2s)	(2s,2s)	(1s,3s)	(2s,3s)	(3s,3s)		
0.9708	-0.2367	-0.0358	-0.0066	-0.0067	-0.0048	-2.878417	1
0.9705	-0.2384	-0.0360	-0.0059	-0.0002		-2.877850	27
0.9706	-0.2377	-0.0362	-0.0069		-0.0043	-2.878417	2
0.9706	-0.2381	-0.0356		-0.0072	-0.0048	-2.878417	2
0.9396	-0.3413		-0.0176	-0.0186	-0.0088	-2.878417	22
0.9979		-0.0623	-0.0134	-0.0115	-0.0054	-2.878417	2
	0.3573	-0.9330	-0.0080	0.0415	0.0035	-2.878417	24
0.9705	-0.2382	-0.0361	-0.0072			-2.877846	2
0.9703	-0.2393	-0.0358		-0.0017		-2.877846	16
0.9349	-0.3547		0.0035	-0.0021		-2.877846	23
0.9978		-0.0634	-0.0166	-0.0008		-2.877882	20
0.9703	-0.2391	-0.0359			-0.0043	-2.878417	3
0.9979		-0.0634	-0.0138		-0.0043	-2.878417	3
	0.3575	-0.9338	-0.0030		0.0044	-2.878417	26
0.9380	-0.3460			-0.0184	-0.0072	-2.878417	23
0.9979		-0.0622		-0.0121	-0.0056	-2.878417	2
	0.3576	-0.9338		0.0066	0.0044	-2.878417	24
0.9980			-0.0084	-0.0011	-0.0633	-2.877846	27
		0.9977	-0.0012	-0.0232	-0.0632	-2.877846	20
0.9702	-0.2396	-0.0357				-2.877846	3
0.9978		-0.0634	-0.0145			-2.877885	3
	0.3573	-0.9340	-0.0006			-2.877846	26
0.9351	-0.3544			-0.0017		-2.877846	24
0.9980		-0.0635		-0.0012		-2.877846	17
	0.3576	-0.9339		0.0070		-2.877846	24
	0.9702		-0.1712	0.1713		-1.349630	49
		0.9351	-0.0019	-0.3545		-2.877846	53
0.9980		-0.0634			-0.0043	-2.878417	4
	0.3578	-0.9338			0.0043	-2.878417	25

TABLE IX (Continued)

Configurations selected						Total	Global
(1s,1s)	(1s,2s)	(2s,2s)	(1s,3s)	(2s,3s)	(3s,3s)	Energy	Iterations
0.9979			-0.0128		-0.0633	-2.877846	5
	0.9994		-0.0333		-0.0026	-1.301915	3
		0.9980	-0.0012		-0.0635	-2.877846	21
0.9980				-0.0012	-0.0633	-2.877846	27
	0.9998			-0.0177	-0.0025	-1.301915	3
		0.9978		-0.0155	-0.0633	-2.877846	5
			0.3168	-0.1664	-0.9338	-2.877846	22
0.9980		-0.0634				-2.877846	3
	0.3578	-0.9337				-2.877846	25
	0.9994		-0.0333			-1.301722	2
		0.9999	-0.0097			-2.863280	3
0.9999				-0.0097		-2.863280	2
	0.9998			-0.0182		-1.301722	2
		0.9352		-0.3539		-2.877846	86
			0.9232	-0.3843		-1.301722	2
0.9980					-0.0634	-2.877846	5
	0.9999				-0.0025	-1.301908	3
		0.9979			-0.0634	-2.877846	5
			0.3577		-0.9338	-2.877846	25
				0.3576	-0.9338	-2.877846	25
1.0000						-2.861671	1
	1.0000					-1.301585	1

6.4.5. On choosing the active orbital space.

In light of the previous results on helium, in the construction of state functions we used a limited number of active MO, and on them we performed a complete set of excitations. The procedure is related to a similarity measure between density functions attached to the occupied MO. The

algorithm can be described with the following scheme:

- a) A set of occupied MO's is chosen  $\{li_1>, li_2>, \dots li_n>\}$ .
- b) A set of virtual MO's is chosen  $\{lv_1>, lv_2>, \dots lv_n>\}$  in such a way as the pairs  $\{li_1>, lv_1>\}$ ,  $\{li_2>, lv_2>\}$ ,  $\dots \{li_n>, lv_n>\}$  are most similar according to the measure

$$r(i,v) = (i \text{ ilvv}) [(i \text{ ili } i)(vv \text{ lvv})]^{-1/2}$$

where  $(i \text{ ilvv})$  are Coulomb integrals involving the occupied-virtual MO pair  $\{li>, lv>\}$ . The similarity measure  $r(i,v)$  is nothing more than a correlation coefficient between the density functions of the MO pair.

The meaning of  $r(i,v)$  is such as  $li>$  is more similar to  $lv>$  then  $r(i,v) \rightarrow 1$ . Therefore, given any occupied MO  $li>$ , the virtual pair  $lv>$  is chosen as to fulfill  $\max\{v'\}\{r(i,v')\}$ .

- c) A complete excitation set of state functions is built from  $\{li_1>, li_2>, \dots li_n>\}$  and  $\{lv_1>, lv_2>, \dots lv_n>\}$ .

The underlying idea of this kind of virtual MO selection consists of maintaining as constant as possible the density of the basic configuration when substituting the MO  $li>$  by the most similar virtual  $lv>$ , in an attempt to provide in this manner the highest interaction between the original determinant and the new ones.

Following the previous procedure, we have performed some computations on the ground state of the water molecule, which can be found in Table X, in order to have some information on how many and which symmetry type the occupied MO, selected as active, must have. Optimal results were obtained when every symmetry species was represented in the occupied active space.



**TABLE X.** Dependence of the total energy (in a.u.) with the selected configurations for the ground state of water. Snyder and Basch geometry (ref. 52), the basis set given by Dunning and Hay (ref. 51), and algorithm MC-B are used. The selected configurations are a complete active space defined by the active orbitals whose ordinal numbers are shown. The SCF reference energy, is -76.000945 a.u.

Active orbitals					Number of state functions	Total energy	Iterations
1a <sub>1</sub>	2a <sub>1</sub>	1b <sub>1</sub>	3a <sub>1</sub>	2b <sub>2</sub>			
1					3	-76.003430	17
	2				3	-76.014813	6
		3			3	-76.011140	2
			4		3	-76.014813	19
				5	3	-76.016043	3
1	2				12	-76.017983	7
1		3			8	-76.013830	21
1			4		12	-76.017982	15
1				5	12	-76.054142	7
	2	3			12	-76.040079	8
	2		4		20	-76.029336	4
	2			5	12	-76.054142	6
		3	4		12	-76.040784	20
		3		5	12	-76.040649	3
			4	5	12	-76.054142	11
1	2		4		100	-76.032476	4
1		3		5	55	-76.088017	8
	2	3		5	55	-76.088017	7
	2		4	5	95	-76.078172	5
		3	4	5	55	-76.088017	14

**TABLE XI.** Total energy (in a.u.) and optimum geometry (bohrs and degrees) computed using mono- and multiconfigurational EJR methods for various molecules in the given state. The number of state functions used as basis set is also given.

Molecule	State	State functions	Total Energy	Geometry	
				r(A-H)	$\Theta(\text{HAH})$
BH <sub>2</sub>	<sup>2</sup> A <sub>1</sub>	1	-25.739455	2.27	128
		23	-25.753884	2.28	128
CH <sub>2</sub> <sup>-</sup>	<sup>2</sup> A <sub>1</sub>	1	-38.810618	2.07	132
		59	-38.830790	2.08	133
CH <sub>2</sub>	<sup>3</sup> B <sub>1</sub>	1	-38.909755	2.05	130
		28	-38.929805	2.07	128
	<sup>3</sup> Σ <sub>g</sub> <sup>-</sup>	1	-38.891992	2.03	180
		10	-38.906385	2.04	180
	<sup>1</sup> A <sub>1</sub>	1	-38.859876	2.11	106
		33	-38.901139	2.13	103
		37	-38.905188	2.14	105
	<sup>2</sup> A <sub>1</sub>	1	-38.552996	2.08	139
CH <sub>2</sub> <sup>+</sup>	<sup>2</sup> A <sub>1</sub>	23	-38.571390	2.10	138
		1	-38.545720	2.07	180
	<sup>2</sup> Π	14	-38.563882	2.09	180
		1	-55.479463	1.98	103
NH <sub>2</sub> <sup>-</sup>	<sup>1</sup> A <sub>1</sub>	56	-55.531094	2.05	101
NH <sub>2</sub>	<sup>2</sup> A <sub>1</sub>	1	-55.498149	1.89	143
		59	-55.522335	1.91	143
	<sup>2</sup> B <sub>1</sub>	1	-55.539397	1.95	108
		49	-55.592828	2.01	106
	<sup>2</sup> Π	1	-55.492833	1.88	180
		75	-55.510087	1.90	180

TABLE XI (Continued)

Molecule	State	State functions	Total Energy	Geometry	
				r(A-H)	$\Theta(\text{HAH})$
NH <sub>2</sub> <sup>+</sup>	3Σ <sub>g</sub> <sup>-</sup>	1	-55.191655	1.96	180
		3	-55.209092	1.97	180
	1A <sub>1</sub>	1	-55.111118	1.99	115
		37	-55.161953	2.02	114
H <sub>2</sub> O	1A <sub>1</sub>	1	-76.002172	1.82	111
		55	-76.090631	1.88	107
		56	-76.061319	1.88	108
H <sub>2</sub> O <sup>+</sup>	2B <sub>1</sub>	1	-75.600820	1.90	119
		49	-75.659075	1.96	115
	2Π	1	-75.579178	1.89	180
		27	-75.604880	1.92	180
	2A <sub>1</sub>	1	-75.578866	1.89	167
		59	-75.605722	1.91	180
FH <sub>2</sub> <sup>+</sup>	1A <sub>1</sub>	1	-99.829222	1.85	125
		56	-99.888494	1.90	119

#### 6.4.6. Applications on the geometry optimization of AH<sub>2</sub> systems

With the above described techniques, a general mono- and multiconfigurational geometry optimization procedure has been implemented. The optimization algorithm is the numerical Newton-Raphson procedure described by Payne /55/. Results on triatomic AH<sub>2</sub> systems are presented in Table XI for various states of the neutral molecules and ions. As expected, the inclusion of correlation gives rise to some modifications in the geometry computed at the

monoconfigurational level. In all the cases reported in Table XI, no convergence problems were found.

#### 6.5.- Comparison with the exponential MCSCF methodology.

A comparison of the present algorithm with every one of the various MCSCF procedures used at present in literature is a task beyond the scope of this paper. But some kind of test must be done at least with one of the best procedures described so far. According to Werner /14/ the Augmented Hessian (AH) procedure of Lengsfeld /62/ constitutes one of the best algorithms available today. Therefore, we have compared the present EJR procedure with the AH using the version programmed by Brooks, Laidig and Saxe contained in HONDO7 /63/. But before making the comparison one must discuss some details concerning the exponential MCSCF methodology.

It must be taken into account that second order MCSCF procedures, as the AH and other exact or approximate second order methods, converge quadratically when close to the final solution, but with a very small radius of convergence. More than this, when the MO-CI coupling is not included one finds linear convergence even with second order methods. For example, see Table II of Werner's paper /14/.

A possible way to solve the convergence problems consists in using high order energy derivatives in the Taylor energy expansion. The drawback is that a higher derivative calculation is really expensive. Consequently some authors had included these terms in some approximate way / 12,14, 28/. This, combined, for instance, with the Direct CI method of Knowles and Handy, using Slater determinants instead of CSF's, overcomes some problems.

The advantage of the EJR procedure can be based on the fact that the present method includes all the energy derivatives up to any order in the energy expansion through the MO transformation, making the convergence radius infinite, as numerous computations had shown up to date /65/. These computations range from very simple molecules and basis sets, like some of those discussed in this paper, to large transition metal clusters. Consequently, the Jacobi Rotation procedure is truly independent of the starting MO's and opens the way to the construction of a Black Box MCSCF procedure.

The claim of independent convergence is not a common feature of the usual MCSCF methodology. Indeed, MCSCF quadratic procedures may fail to converge even in simple cases. Let us take, for example, the case of the CH<sub>2</sub> molecule computed with an STO-3G basis set and a CI space formed by the <sup>1</sup>A<sub>1</sub> ground state configuration and the first biexcitation. Table XII shows the results for MCSCF computations using the fully quadratic AH method including MO-CI couplings, the two-step AH method (where the MO-CI coupling is not considered) and the EJR procedure, where MO-CI coupling is not, at present, explicitly taken into account. The starting orbitals for all the computations are the SCF MO's. It is evident from these results that the fully quadratic AH procedure does not converge and one has to switch to the two-step AH which presents linear convergence. EJR converge in a similar number of iterations like the two-step AH method. It is possible to see that the linear convergence of the EJR procedure is due to the non inclusion of the MO-CI coupling in the energy variation of the MO transformations, as the convergence is achieved by optimizing only the MO rotations when the CI vector is set equal to the final CI values.

The previous behaviour does not depend on the number of

configurations in the CI expansion as can be inferred from TABLE XIII results, where the final MCSCF energy and the number of iterations for various CI spaces of the CAS (Complete Active Space) type is presented. In all the cases the two-step AH method gives the same energy and converges in a similar number of iterations than the EJR. For non-CAS CI spaces, see the results shown in Table IX where one can see how EJR convergence is fulfilled even in cases where no convergence is found with the two-step or full quadratic procedures.

**TABLE XII.** Total energy (in a.u.) for a biconfigurational MCSCF computation of  $\text{CH}_2$  in the  $^1\text{A}_1$  state with various algorithms. The STO-3G basis set was employed at the following geometry (in a.u.): C (0., 0., 0.); H<sub>1</sub> (-1.8411892, 0., 1.0531792); H<sub>2</sub> (1.8411892, 0., 1.0531792).

Aumented Hessian method		
With MO-CI coupling	Without MO-CI coupling	Elementary Jacobi Rotations
-36.85564239	-36.85564239	-36.85564239
-38.21989305	-38.22103444	-38.28935957
-38.37692871	-38.37701021	-38.37694611
-38.36733483	-38.37750608	-38.37750796
-38.37036663	-38.37756460	-38.37756543
-38.63733740	-38.37757648	-38.37757663
-37.51216343	-38.37757222	-38.37757927
-37.46986509	-38.37757988	-38.37757983
-37.46993122	-38.37758005	-38.37757006
-35.90110428	-38.37758009	-38.37758009
-33.62845154	-38.37758010	-38.37758010
-23.71655960		
-20.31578866		
-15.86176625		
-16.87542415		
-13.16446563		

**TABLE XIII.** Total energy (in a.u.) for a MCSCF computation of  $\text{CH}_2$  in the  $^1\text{A}_1$  state with a STO-3G basis set for various CAS expansions. The geometry (in a.u.) used was the following: C (0., 0., 0.);  $\text{H}_1$  (-1.8411892, 0., 1.0531792);  $\text{H}_2$  (1.8411892, 0., 1.0531792). In all cases the initial orbitals feed into the MCSCF program were SCF orbitals.

Number of configurations	Total energy	Number of Iterations	
		AH	EJR
152 (full CI)	-38.41935901	1	1
19	-38.39282679	3	5
8	-38.38255372	11	11
2	-38.37758010	10	10

Similar results have been obtained for the  $\text{AlH}$ , the  $\text{AlO}$ , or  $\text{AlO}_2$  systems and, therefore, will not be presented here.

Previous results show that the EJR method is very promising to solve the convergence problem in MCSCF computations. The only snag is the number of integral transformations to be done in the exact formulation. However, this problem can be obviated if instead of the "exact" EJR algorithm outlined before, some kind of accumulated EJR is used /1/, where the integrals are only transformed after the full set of rotations is done, in the way this problem is treated in the exponential transformations. Preliminary work in this direction shows that the efficiency of the method is kept while the amount of time in the integral transformations is lowered.

## 7. OPTIMIZATION OF ROTATION SINE

### 7.1. An Exact Solution

Finding the sine of the angle that minimizes the energy when two orbitals are transformed according to EJR is a univariate minimization problem and as such may be solved using any of the classical methods available [56].

However, in this particular case a straightforward solution may be obtained after a close inspection of the equations attached to each EJR particular case.

#### 7.1.1. A simple procedure

In most usual situations, when the energy variation polynomial coefficients are computed, those multiplying the higher powers of the sine may be considered negligible in front of the one multiplying the lower terms. Therefore the energy may be approximated by the expression:

$$E = E_0 + E_1 s + E_2 s^2 \quad (7.1)$$

which will be very useful to start our analysis about the behaviour of the family of polynomial functions associated with the EJR variation problem.

The derivatives with respect to the sine of the approximate polynomial may be written:



$$\begin{aligned}\partial E / \partial s &= -c (E_1 t^2 - 2 E_2 t - E_1) \\ \partial^2 E / \partial s^2 &= 2 E_2 - E_1 t (3+t^2)\end{aligned}\quad (7.2)$$

where  $t=s/c$  is the tangent of the rotation angle.

Being quadratic in nature, the first derivative may have two real solutions which produces a vanishing value, that is: the energy extremal. These two solutions may be compactly written that:

$$t = (E_2 \pm \sqrt{E_2^2 + E_1^2}) / E_1 = (1/a) \pm \sqrt{1 + (1/a^2)} \quad (7.3)$$

where  $a=E_1/E_2$ .

The single parameter  $a$  may be used to further simplify all the previously written equations, in such a manner that:

$$\begin{aligned}E &= E_0 + E_2 (a s c + s^2) \\ \partial^2 E / \partial s^2 &= E_2 (2 - a t (3+t^2))\end{aligned}\quad (7.4)$$

Introducing the energy extrema previously found into these latter equations, it is easy to see, that the sign of the root that corresponds to a minimum depends on the sign of  $E_2$ , namely if  $E_2 > 0$  then the root has to be taken as negative, and viceversa.

Furthermore, the solutions found in this way always correspond to energy decreasing steps.

When  $E_2 > 0$ , it is convenient, in order to avoid roundoff errors, to use a modification of the classical quadratic equation solution, such as:

$$t = E_1 / (E_2 + \sqrt{E_1^2 + E_2^2}) \quad (7.5)$$

which prevents subtractions between magnitudes of similar order .

The results, obtained with this simplified formula, can be used as starting points to solve the cases that arise from both mono- and multiconfigurational energy variation.

### 7.1.2. Monoconfigurational energy

The first derivative of the monoconfigurational energy variation (5.2) can be written as

$$\partial E / \partial s = -c[(E_1 + E_3 s^2)t^2 - 2(E_2 + 2E_2 s^2) - (E_1 + 3E_3 s^2)] \quad (7.6)$$

where  $t$  is again the tangent of the rotation angle.

One can see that, although having its parameters dependent on the sine, this first derivative is again quadratic in the tangent.

A very simple, and fast convergent, recursive method may be devised from this formula, leaving the rotation sine as a constant and finding, at each iteration step, the optimal tangent value, which will lead to a new sine value for the next iteration.

Therefore, calling

$$\begin{aligned} A &= E_1 + E_3 s \\ B &= E_2 + 2E_4 s \\ C &= E_1 + 3E_3 s \end{aligned} \quad (7.7)$$

the solution has the form

$$t = (B + \sqrt{B^2 + AC}) / A = A / (\sqrt{B^2 + AC} - B) \quad (7.8)$$

where the sign of the square root was taken according to the considerations made in the previous section.

The expression on the left is used when  $B > 0$ , while the one on the right for the case  $B < 0$  to avoid roundoff errors.

These formulae normally lead to small angle values, and this is going to be the most usual case, as one expects to obtain rotation values significantly far from zero only in the initial iterations.

### 7.1.3 Multiconfigurational energy

A very similar reasoning may be used for the multiconfigurational case where the energy is written as in (6.8).

The expression for the first derivative is:

$$\begin{aligned} \partial E / \partial s = & -c[t^2 (E_{11} + E_{21} s + E_{31} s^2) - t(2E_{20} + 3E_{30} s + 4E_{40} s^2 - E_{01}/c) \\ & - (E_{11} + 2E_{21}s + 3E_{31} s^2 + E_{10}/c)] \end{aligned} \quad (7.9)$$

which leads again to a quadratic form when using

$$\begin{aligned} A &= E_{11} + E_{21} s + E_{31} s^2 \\ B &= E_{20} + 3 E_{30} s/2 + 2 E_{40} s - E_{01}/c \\ C &= E_{11} + 2 E_{21} s + 3 E_{31} s^2 + E_{10}/c \end{aligned} \quad (7.10)$$

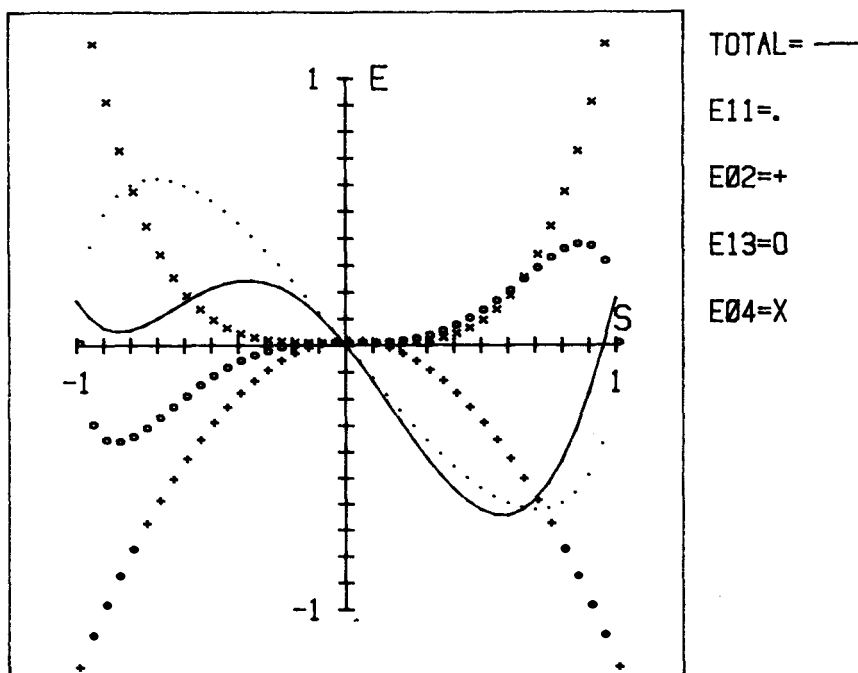


FIGURE 1.- Form adopted by the energy increment curve and its most important components for a particular case in a helium atom computation.

In this case the convergence surely will be a little worse as the multiconfigurational coefficients depend on smaller powers of the sine and will have a larger contribution. Figure 1 shows the value of energy increment versus the rotation sine (solid line) for a particularly complicated case found when optimizing the helium atom, together with the weight of the most important terms, identified by its  $E_{ij}$  component. In any case, the convergence is still fast enough to make the algorithm suitable but one has to be careful with the two or more than two

minima cases.

#### 7.1.4. Brillouin theorems

Brillouin theorems (see for example, reference /57/ ) can be deduced in both mono- or multiconfigurational cases using the same technique as in the CI expression of section 4.2. Then making  $s=0$  in the first derivative formulae and forcing the gradient to nullity one easily obtains:

##### a) Monoconfigurational

$$(\partial \delta E / \partial s)_{s=0} = -E_1 = 0 \quad (7.11)$$

This general expression takes the form

$$\langle i | \Delta\omega_T h + \Delta\omega_T D_T : P - \Delta\omega_0 D_0 : Q | j \rangle = 0 \quad (7.12)$$

for the Roothaan-Bagus energy expression given in section 5.2.

##### b) Multiconfigurational

$$(\partial \delta E / \partial s)_{s=0} = - (E_{11} + E_{10}) = 0 \quad (7.13)$$

It seems in this way that throughout EJR formalism one is able to find some sort of generalized Brillouin theorem for any arbitrary expression involving Quantum Mechanical objects.

In fact, any Quantum Mechanical object may be written in a general way as:

$$Z = f(|i\rangle, |j\rangle) \quad (7.14)$$

Then, a EJR of indices  $\{i, j\}$  will give rise to some structure which at least will contain terms in either  $s$  or  $sc$ , or both,

$$\partial Z = Z_{01} s + Z_{11} s c + O(s^2) \quad (7.15)$$

thus, the first derivative may be written as:

$$\partial \delta Z / \partial s = Z_{01} + Z_{11} c + O(s^2) \quad (7.16)$$

and the extremum condition at  $s=0$  gives

$$(\partial \delta Z / \partial s)_{s=0} = Z_{01} + Z_{11} = 0 \quad (7.17)$$

which constitutes a Brillouin Theorem valid for any Quantum Mechanical object expression.

## 7.2. Performance Improving Procedures

It was mentioned before that as soon as the EJR iterative process is settled and has begun to converge, the rotations will be fairly small, - in fact, computational experience shows that this happens after the fourth sweep, as in classical Jacobi diagonalization - so that approximate solutions of the energy polynomial, which discard terms with high powers of the rotation sine, may be accurate enough to guarantee the total process convergence.

It is also quite usual that from the whole set of active rotations, very few of them have energy decreasing contributions several orders in magnitude higher than the

remaining ones. These rotations that lay behind in the iterative process should be done more often than the others, if one wants the process to converge in a reasonable number of rotations.

These considerations are important for an efficient implementation of a Jacobi Rotation Procedure.

### 7.2.1. Approximate optimal solutions

When only the terms up to second order in the sine are considered in the EJR energy variation expressions, simple and non-iterative expressions are found for the optimal rotation angle and its corresponding energy variation.

In the monoconfigurational case, one obtains an expression similar to the one described in section 7.1.1.,

$$E = E_0 + E_1 s + E_2 s^2 \quad (7.18)$$

which has trivial solutions:

$$s_{\text{opt}} = -E_1/2E_2 \quad (7.19)$$

and

$$E = -E_1^2/4E_2 \quad (7.20)$$

The multiconfigurational energy can be approximated to

$$\Delta E = E_{00} + E_{01} s^2/2 + E_{11} s + E_{21} s^2 + E_{10} s + E_{20} s \quad (7.21)$$

where the well known approximation to the cosine  $c=1-s/2$  has been used. The solutions are therefore:

$$s_{opt} = - (E_{11} + E_{10}) / (E_{01} + 2E_{21} + E_{20})$$

$$\Delta E = - (E_{11} + E_{10})^2 / 2(E_{01} + 2E_{21} + E_{20}) \quad (7.22)$$

### 7.2.2. Rotation selection and filtering

Between the collection of computational details which may be commented here, the strategy of performing the minor number of possible rotations, through a given energy minimization process, can help to cut the necessary computer time by a significant amount.

The selection of rotations which can produce an energy lowering within a given precision range, avoiding the remaining meaningless ones, can be performed using energy increments estimated by means of the formulae given in the previous section, comparing the obtained results with a predetermined precision value. Although this precision threshold may be arbitrarily set, computational experience suggests that the best order of magnitude will be given by the arithmetic mean of the estimated energy lowerings for all the active rotations.

Such a strategy requires the computation of the low order energy parameters for all the active orbital pairs before each sweep, but a significant amount of time is spent in this previous filtering.

The alternative of estimating the energy decrements based upon the ones obtained in the previous sweep, although saving filtering time, do not give as good results as using the whole set of available rotations.

Furthermore, a significant performance improvement is



found if, at each sweep, the rotation that leads to the largest energy lowering is performed first, followed by a normal sweep; this sweep may even include the rotation already performed if its newly estimated energy decrement happens to be larger than the average value.

The main steps in each iteration for the best actual strategy are:

- a) Compute the estimated energy lowerings for every active index pair rotation:  $\Delta E_{ij}$ .
- b) Find the minimum of  $\Delta E_{ij}$  and perform the corresponding rotation.
- c) Compute a threshold value as:

$$T = \frac{1}{n_{\text{rot}}} \sum \Delta E_{ij} \quad (7.23)$$

where  $n_{\text{rot}}$  is the number of rotations.

- d) For each rotation: estimate the energy lowering  $\Delta E_{ij}$  and perform the rotation if  $\Delta E_{ij} < T$ .

### 7.3. Localization Procedures and Jacobi Rotations

EJR can be used for purposes other than energy minimization. As an example of the present scheme flexibility a short discussion on OM localization follows.

The Edmiston-Ruedenberg localization procedure /58/ consists of an intrashell transformation of selfrepulsion energy until the sum

$$L = \sum_p (pp | pp) \quad (7.24)$$

attains its maximal value.

Let us consider this expression more generally as a sum of norms of the density functions  $\{\rho_p = |p\rangle\langle p|\}$  attached to a given MO set  $\{|p\rangle\}$ .

A generalized norm can then be defined as

$$(\rho_p | \rho_p) = \int \rho_p \Theta \rho_p dV = (pp | \Theta | pp) \quad (7.25)$$

where  $\Theta$  must be some positive definite operator. In the case of repulsion integrals  $\Theta = 1/r_{12}$ . Thus, a general localization function may be defined as

$$L_\Theta = \sum_p (pp | \Theta | pp) \quad (7.26)$$

The localization function  $L_\Theta$  may be considered from the point of view of EJR as a sum of fourth-order objects. Thus given an active pair  $\{i, j\}$  and a rotation sine one obtains

$$\begin{aligned} \delta \Sigma &= \delta [(ii | \Theta | ii) + (jj | \Theta | jj)] = \\ &= -2sc[(ii | \Theta | ij) - (ij | \Theta | jj)] - \\ &\quad - (1/2)s^2[(ii | \Theta | ii) + (jj | \Theta | jj) - 2(ii | \Theta | jj) - 4(ij | \Theta | ij)] \end{aligned} \quad (7.27)$$

where  $S = \sin 2\alpha$ ,  $C = \cos 2\alpha$ .

But the interesting fact is that the above expression becomes the same as the variation of the square euclidean distance between functions  $\rho_i$  and  $\rho_j$ :

$$D^2 = (i|i| \Theta |i|i) + (j|j| \Theta |j|j) - 2(i|i| \Theta |j|j) \quad (7.28)$$

which properly varied produces

$$\delta D^2 = 2 \delta \Sigma \quad (7.29)$$

Thus, maximization of  $L_\Theta$  in a general Edmiston-Ruedenberg scheme becomes the same as maximization of distances between pairs of intrashell orbitals.

Another interesting feature is the possibility of using exactly the EJR scheme to obtain localized functions. That is, localized orbitals can be produced in the same general procedure as the energy minimization one, providing the way to obtain simultaneously with the best orbitals some orbitals which bear some kind of canonical localized form with respect to a positive definite operator.

The best operator which fits both minimization and localization schemes is, of course, the electron repulsion  $1/r_{12}$ .

In this manner as optimal energy gives a minimal non-definite measure between orbitals belonging to different shells, localization provides a maximal distance between orbitals belonging to the same shell.

Brillouin theorem here looks as

$$\partial \delta \Sigma / \partial s = -2[(i|i| \Theta |j|j) - (i|j| \Theta |j|j)] = 0 \quad (7.30)$$

that is, at the maximal distance one will have the relationship

$$(i|i|j) = (i|j|j) \quad (7.31)$$

between repulsion integrals if this measure has been chosen.

Let us finally say that maximizing a distance measure becomes something analogous to minimize the similarity measure  $r(i,j)$  defined in section 6.4.5. Thus localization in this context becomes the same as to obtain within a shell the least similar set of functions.

#### ACKNOWLEDGEMENTS

Computer time has been made available through a grant CACYT-0657/81. One of us also (J.J.N.) wishes to acknowledge the support given by the "Caixa d'Estalvis de Barcelona" and the British Council. The authors wish also to express their gratitude to Mr. J. Peris and Mr. J. Moreno for the helpful work done during early stages of the EJР theoretical and computational development.

#### REFERENCES

1. Carbo, R., Domingo, Ll., and Peris, J. J.,(1982) *Adv. Quantum Chem.* **15**, 215.
2. Miller, K. J., Ruedenberg, K., (1968) *J. Chem. Phys.*, **48**, 3414.
3. Raffenetti, R. C., Ruedenberg, K., (1970) *Int. J. Quantum Chem.*, **3**, 625.
4. Hoffman, D. K., Raffenetti, R. C., Ruedenberg, K., (1972) *J. Math. Phys.*, **13**, 528.

5. Carbo, R., Domingo, Ll., Peris J. J., and Novoa J. J., (1983) *J. Mol. Struct.*, **93**, 15.
6. Carbo, R., Domingo Ll., and Novoa J. J., (1985) *J. Mol. Struct.*, **120**, 357.
7. Roothaan, C. C. J., and Bagus, P. S., (1963) *Methods in Computational Physics*, **2**, 47.
8. Roos, B., Salez, C., Veillard, A., and Clementi, E., (1968) "A general program for calculation of Atomic SCF Orbitals by the expansion method". IBM Research. **RJ 518**.
9. Clementi, E. and Roetti, C., (1974) *Atomic Data and Nuclear Data Tables*, **14**, 185.
10. Hinze, J., (1981) *Int. J. Quantum Chem., Quantum Chem. Symp.*, **15**, 69.
11. Dupuis, M., (ed) (1981) NRCC Proceedings: "Proceedings of the Workshop on Recent Developments and Applications of Multi-Configurational Hartree-Fock Methods".
12. Olsen, J., Yeager, D. L. and Joergensen, P., (1983) *Adv. Chem. Phys.*, **54**, 1.
13. Almlof, J and Taylor, P. R., (1984) *NATO ASI Ser., Ser. C*, **133**, 107.
14. Werner, H. J., (1987) *Adv. Chem. Phys.*, **69** (part II), 1; Shepard, R., *Adv. Chem. Phys.*, **69** (part II), 63.
15. Polezzo, S., and Fantucci, P., (1980) *Mol. Phys.*, **39**, 1527.
16. Polezzo, S., and Fantucci, P., (1980) *Mol. Phys.*, **40**, 759.
17. Polezzo, S., and Fantucci, P., (1981) *Gazz. Chim. Ital.*, **111**, 49.
18. Matsen, F., and Nelin, C. J., (1981) *Int. J. Quantum Chem.*, **20**, 861.
19. Fantucci, P., Polezzo, S., Trombetta, L., (1981) *Int. J. Quantum Chem.*, **19**, 493.
20. Levy, B., (1981) Lawrence Berkeley Lab., [Rep.] LBL-12157, 134.
21. Mukherjee, N. G., (1980) *Int. J. Quantum Chem.*, **18**, 1045.
22. McCullough, E. A., (1982) *J. Phys. Chem.*, **86**, 2178.
23. Igawa, A., Yeager, D. L., and Fukutome, H., (1982) *J. Chem. Phys.*, **76**, 5388.
24. Lengsfeld, B. H., (1982) *J. Chem. Phys.*, **77**, 4073.
25. Polezzo, S., and Fantucci, P., (1982) *Stud. Phys. Theor. Chem.* **21**, 181.

26. Lengsfeld, B. H., and Liu, B., (1981) *J. Chem. Phys.*, **75**, 478.
27. Polezzo, S., (1980) *Chem. Phys. Lett.*, **75**, 303.
28. Werner, H. J., and Meyer, W., (1980) *J. Chem. Phys.*, **73**, 2342; (1981) *J. Chem. Phys.*, **74**, 5794.
29. Ishikawa, Y. (1980) *Z. Naturforsch.*, **35**, 408.
30. Siegbahn, P., Heiberg, A., Roos, B., and Levy, B., (1980) *Phys. Scr.*, **21**, 323.
31. Brooks, B. R., Laiding, W. D., Saxe, P., Goddard, J. D., Schaeffer, H. F., (1981) *Lect. Notes Chem.*, **22**, 158.
32. Yurtsever, E., and Shillady, D., (1983) *Chem. Phys. Lett.*, **94**, 316.
33. Das, G., (1981) *J. Chem. Phys.*, **74**, 5775.
34. Camp, R. N., King, H. F., (1982) *J. Chem. Phys.*, **77**, 3056.
35. Bauschlicher, C. W., Bagus, P. S., Yarkony, D. R., and Lengsfeld, B. H., (1981) *J. Chem. Phys.*, **74**, 3965.
36. Werner, H. J., and Knowles, P. J., (1985) *J. Chem. Phys.*, **82**, 5053; Knowles, P. J. and Werner, H. J., (1985) *Chem. Phys. Lett.*, **115**, 259.
37. Yeager, D. L., and Joergensen, P., (1980) *Mol. Phys.*, **39**, 587.
38. Yeager, D. L., Albertsen, P., and Joergensen, P., (1980) *J. Chem. Phys.*, **73**, 2811.
39. Joergensen, P., Albertsen, P., and Yeager, D. L., (1980) *J. Chem. Phys.*, **72**, 6466.
40. Joergensen, P., Olsen, J., and Yeager, D. L., (1981) *J. Chem. Phys.*, **75**, 5802.
41. Olsen, J., and Joergensen, P., (1982) *J. Chem. Phys.*, **77**, 6109.
42. Yeager, D. L., Lynch, D., Nichols, J., Joergensen, P., and Olsen, J., (1982) *J. Phys. Chem.*, **86**, 2140.
43. Joergensen, P., Swannstroem, P., and Yeager, D. L., (1983) *J. Chem. Phys.*, **78**, 347.
44. Olsen, J., Joergensen, P., and Yeager, D. L., (1983) *Int. J. Quantum Chem.*, **24**, 25.
45. Jensen, H. J. A. and Joergensen, P. (1984) *J. Chem. Phys.*, **80**, 1204.
46. Joergensen, P., and Simons, J., (1983) *J. Chem. Phys.*, **79**, 334.
47. Golab, J. T., Yeager, D. L., Joergensen, P., (1985) *Chem.*

- Phys.*, **93**, 83.
48. Levy, B., (1970) *Int. J. Quantum Chem.*, **4**, 297.
49. Wood, M. H. and Veillard, A., (1973) *Mol. Phys.*, **26**, 595.
50. Lengsfeld, B. H., (1980) *J. Chem. Phys.*, **73**, 382.
51. Dunning, T. H. and Hay, P. J., (1977) *Mod. Theoret. Chem.*, **3**, 1.
52. Snyder, L. C. and Basch, H. B., (1972) "Molecular Wave Functions and Properties", John Wiley and Sons, New York.
53. Lowdin, P. O., (1955) *Phys. Rev.*, **97**, 1509.
54. Lowdin, P. O. and Shull, H., (1956) *Phys. Rev.*, **101**, 1730.
55. Payne, P. W. (1976) *J. Chem. Phys.*, **65**, 1920.
56. Beveridge, G. S. G. and Schechter, R. S., (1970) "Optimization, Theory and Practice", McGraw-Hill Kogakusha, Tokyo.
57. Levy B., and Berthier G., (1968) *Int. J. Quantum Chem.*, **2**, 307.
58. Edmiston, C. E. and Ruedenberg, K., (1963) *Rev. Mod. Phys.*, **35**, 457.
59. Huzinaga, S., (1965) *J. Chem. Phys.*, **42**, 1293.
60. Iwata, S., (1981) *Chem. Phys. Lett.*, **83**, 134.
61. Kendrick, J. and Hillier, I. H., (1976) *Chem. Phys. Lett.*, **41**, 283.
62. Lengsfeld, B. H., (1980) *J. Chem. Phys.*, **73**, 382.
63. HONDO7, Dupuis, M., Watts, J. D., Villar, H. O., Hurst, G. J. B., (1988) IBM Scientific and Engineering Department, Kingston, NY.
64. Rubio, J., Novoa, J. J., Illas, F., (1986) *Chem. Phys. Lett.*, **126**, 98; Illas, F., Merchan, M., Pelissier M., and Malrieu, J. P., (1986) *Chem. Phys.*, **107**, 361; Illas, F., Rubio, J., Ricart, J. M., (1988) *J. Chem. Phys.*, **88**, 260; Rubio, J., Ricart, J. M., Illas, F., (1988) *J. Comp. Chem.*, **9**, 836; Illas, F., Rubio, J., Ricart, J. M., *J. Chem. Phys.*, (1988) **89**, 6376.

## APPENDIX I

Expression of MO multiconfigurational energy  
variational coefficients

The multiconfigurational energy may be written in a general way as:

$$E = \sum_{ij} \omega_{ij} h_{ij} + \sum_{ijkl} \gamma_{ijkl} (ij|kl)$$

where the sums run over the indices with the following constraints:  $j \leq i$ ,  $l \leq k \leq i$  and  $l \leq j$  when  $i=k$ , or, with packed indices, this is the same as to consider  $[kl] \leq [ij]$ .

The energy variation is given by the formula (6.8) in the text where the parameters are defined by:

$$E_{04} = (\gamma_{pppp} + \gamma_{qqqq} - \gamma_{ppqq} - \gamma_{qqpp}) [(pp|pp) + (qq|qq) - 2(pp|qq) - 4(pq|pq)] + \\ + 4(\gamma_{pppq} - \gamma_{ppqp}) [(pp|pq) - (pq|qq)]$$

$$E_{03} = \sum_{i \neq pq} [(\gamma_{iipq} - \gamma_{ipiq} - \gamma_{ipqp}) [(ip|pp) - 2(iq|pq) - (ip|qq)] - \\ - (\gamma_{ippp} - \gamma_{ipqp} - \gamma_{ipqq}) [(iq|qq) - 2(ip|pq) - (iq|pp)] ]$$



$$\begin{aligned}
E_{02} = & (\omega_{pp} - \omega_{qq}) (h_{qq} - h_{pp}) - 2 \omega_{pq} h_{pq} + \\
& + (\gamma_{ppqq} + \gamma_{ppqq}) [(pp|pp) + (qq|qq) - 2(pp|qq) - 4(pq|pq)] + \\
& - 2\gamma_{pppp} [(pp|pp) - (pp|qq) - 2(pq|pq)] + \\
& + 2\gamma_{qqqq} [(qq|qq) - (pp|qq) - 2(pq|pq)] + \\
& - 4(\gamma_{ppqq} - \gamma_{ppqq}) [(pp|pq) - (pq|qq)] - \\
& - (\gamma_{ppqq} + \gamma_{ppqq}) [(pp|pq) + (pq|qq)] + \\
& + \sum_{ij \neq pq} [(\gamma_{ijpp} - \gamma_{ijqq})[(i j | qq) - (i j | pp)] - 2\gamma_{ijpq} (i j | pq) + \\
& + (\gamma_{ijqp} - \gamma_{ijqp}) [(iq|jq) - (ip|jp)] ] - \\
& - \sum_{i \neq pq} \sum_{j \neq pq} \gamma_{ijkl} [(ip|jq) + (iq|jp)]
\end{aligned}$$

$$\begin{aligned}
E_{01} = & \sum_{i \neq pq} [\omega_p h_{ip} - \omega_q h_{iq} + \gamma_{iqqi} [(2(iq|pq) + (ip|qq))] - \\
& - \gamma_{iipq} [(iq|qq) - (iq|pp) - (ip|pq)] + \gamma_{iipp} [(ip|pp) - 2(iq|pq)] - \\
& - \gamma_{iipq} [2(ip|pq) + (iq|pp)] + \gamma_{iipq} [(ip|pp) - (ip|qq) - (iq|pq)] - \\
& - \gamma_{iipq} [(iq|qq) - 2(ip|pq)] + \sum_{k,l \neq p,q} [\gamma_{ijkl} (ip|kl) - \gamma_{ijkl} (iq|kl)] ]
\end{aligned}$$

$$\begin{aligned}
E_{13} = & 4 (\gamma_{pppp} + \gamma_{qqqq} - \gamma_{ppqq} - \gamma_{ppqq}) [(pp|pq) - (pq|qq)] - \\
& - (\gamma_{pppp} - \gamma_{ppqq}) [(pp|pp) + (qq|qq) - 2(pp|qq) - 4(pq|pq)]
\end{aligned}$$

$$E_{12} = \sum_{i \neq p q} [ (\gamma_{i p q} + \gamma_{i q p} - \gamma_{i p p}) [(i p | p p) - 2(i q | p q) - (i p | q q)] + \\ + (\gamma_{i p q} + \gamma_{i q p} - \gamma_{i q q}) [(i q | q q) - 2(i p | p q) - (i q | p p)] ]$$

$$E_{11} = 2(\omega_{q i} - \omega_{p p}) h_{p i} - \omega_{p i} (h_{q i} - h_{p p}) + 4[\gamma_{q i i} (p q | q q) - \gamma_{p p p} (p p | p q)] + \\ + 2[\gamma_{p p i} + \gamma_{p i p}] [(p p | p q) - (p q | q q)] + \gamma_{p p i} [(p p | p p) - (p p | q q) - \\ - 2(p q | p q)] + \sum_{i, j \neq p q} [ 2(\gamma_{i j i} - \gamma_{i j p}) (i j | p q) - \gamma_{i j p} [(i j | q q) - (i j | p p)] + \\ + (\gamma_{i j q} - \gamma_{i p p}) [(i p | j q) + (i q | j p)] - \gamma_{i p q} [(i q | j q) - (i p | j p)] ]$$

$$E_{10} = \sum_{i \neq p q} [ \omega_p h_{p i} + \omega_i h_{i q} + \gamma_{p p p} (i p | p p) + \gamma_{i p p} (i p | p q) + \gamma_{p q i} (i p | q q) + \\ + \gamma_{i p p} (i q | p p) + \gamma_{i p q} (i q | p q) + \gamma_{i q i} (i q | q q) ] + \\ + \sum_{i, k, l \neq p q} [ \gamma_{i p k l} (i p | k l) + \gamma_{i q k l} (i q | k l) ]$$

## APPENDIX II

Computational procedure followed in Algorithm MC-A

- a) Compute AO integrals
- b) Compute initial molecular orbitals
- c) Transform atomic to molecular integrals
- d) Define initial CI vector
- e) Compute first and second order density matrices
- f) Jacobi rotations:

global-iterations = 0

do while energy-increment > global-precision

MO rotations:

MO-sweep = 0

do while MO-energy-increment > MO-precision

loop over {i,j} active MO pairs

\* compute the MO-energy polynomial coefficients

\* compute optimum sine

\* rotate the MO's, and the molecular integrals

end loop {i,j}

MO-sweep = MO-sweep + 1

enddo MO-energy

CI rotations:

CI-sweep = 0

do while CI-energy-increment > CI-precision

loop over {I,J} active CI coefficient pairs

```
        *   compute the energy polynomial coefficients
        *   compute optimum sine
        *   rotate CI vector
            *   update the first and second order density matrices
    end loop {I,J}
    CI-sweep = CI-sweep + 1
enddo CI-energy
```

```
global-iterations = global-iterations + 1
enddo energy-increment
```

g)End algorithm MC-A

This Page Intentionally Left Blank

## A

*Ab initio* quantum chemical methods, 101–112  
 computational cost, 104  
 embedding theory, 105–106  
 error bound method, 104–106  
 generalized valence bond-CI method, 108–110  
 generalized valence bond method, 106–108  
 localization scheme, 105  
 modified effective potential, 102–103  
 multiconfiguration electron state, 102  
 oxygen, binding site on Si(100), 103–104  
 self-consistent field Hartree-Fock method, 101–102  
 self-consistent field studies, 102  
 vibrational frequency shift for oxygen on Ni(100), 102–103  
 wave function for closed shell singlet state, 101

Acetylene, chemisorption binding energies, 87–90

Active orbital space, choosing, multiconfigurational energy expression, 412–414

Adatom-adatom forces, 113–114

Adatom-substrate binding forces, 114

Adjoint operators, stability problem, 190–198  
 bi-orthogonality theorem, 191  
 block-diagonalized, 194  
 classical canonical form, 191  
 complex symmetric operator, 193–194  
 general theory, 190–194  
 Hartree-Fock scheme, 195–198  
 Jordan blocks, 192–193  
 ket-bra operator, 190  
 Segré characteristics, 192

Adsorbate  
 on metal surfaces, surface structure, 144–146  
 surface coverage, 114

Adsorbate-adsorbate interaction, 115

AH<sub>2</sub> systems, geometry optimization, multiconfigurational energy expression, 415–417

Alkylidyne, 120

Alloys, surface structure, 131–134

Aluminum, chemisorption on GaAs (110), 112

Angle-resolved auger electron spectroscopy  
 amplitudes of outgoing partial waves, 72  
 surface structure, 56

Angle-resolved photoelectron emission spectroscopy, surface structure, 51

Angle-resolved photoemission fine-structure, 67  
 relative intensity modulations, 68  
 surface structure, 28–29, 52

Angle-resolved ultraviolet photoemission spectroscopy, surface structure, 51

Angle-resolved x-ray photoelectron spectroscopy, surface structure, 51–52

Angle-resolved x-ray photoemission, surface structure, 51

Anonymous parentage approximation scheme, 325

Atomic adsorption, semiconductor surfaces, 165

Atomic beam diffraction, surface structure, 33–34

Atom-superposition electron-delocalization molecular orbital theory, 84

Auger electron emission, 71–72

Augmented Hessian procedure, 417–418

## B

Bi-orthogonality theorem, 191

Bloch equation, 328  
 Fock-space, 334, 347, 355, 358  
 zero valence problem, 345

Bloch equation based Fock space theory, 345–350  
 IP calculations, 349  
 satellite peaks, 349

- subsystem embedding condition, 348
- valence-universal wave-operator, 347
- wave-operator, 345
- Bloch-Horowitz theory, 349
- Bloch wave-operator, 345
- $\pi$ -Bonded model, 96–97
- Bond length, contractions and reconstructions, 117–118
- Born-Oppenheimer approximation, 220
- Bragg peaks, LEED, 77
- Brillouin theorem, 432
  - CI form, 386
  - rotation sine optimization, 426–427
- Broyde-Fanno-Goldfarb-Shanno method compared with Murtagh-Sargent method, 280
- optimization
  - hydrogen peroxide, 271
  - NH<sub>3</sub>, 275, 277, 282–283

## C

- Carbon monoxide
  - binding energy, 85
  - bonding geometry, 120
    - adsorbate, 95
  - bonding to platinum (100) surface, 97
  - chemisorption on metals
    - extended Hückel theory, 85–87
    - Ni(100), 99, 108–109
    - surface structure, 157–158
  - as molecular adsorbate, 119–120
- CAS-SCF function, 325
- C–C bond, 120
- CH<sub>2</sub>, total energy, biconfigurational MCSCF method, 418–420
- Chalcogen, chemisorption on metals
  - Ni surfaces, 100
  - surface structure, 136–139
- Chemical composition, surface structure, 36
- CI energy
  - elementary Jacobi rotations, 384–386
  - increment, non-null  $\xi_{ij}$  coefficients, 407–408
- CI-MO computation algorithms, 400–401
- CI space, choosing, multiconfigurational energy expression, 408, 410–412
- CI wave function, multiconfigurational energy expression, 398
- Clean metal structures, surface structure, 122–126
- Clebsch-Gordon coefficients, 63
- Cluster expansion methods, 292–293
  - incomplete model space, 357
  - size-extensivity, 298–302
    - core-extensive theories, 301
    - linked clusters, 300–301
    - open-shell cluster expansion theories, 301
    - pair-correlation, 299
    - particle correlations, 299
    - size-extensive theory, 301–302
    - wave-function, 298–300
- Complementary active space, 354
- Complete active space, 297
- Complete model space, 298
- Complete model space theory, 360
- Complex scaling method, 188, 222–230
  - dilatation operator, 189, 217–218
  - e-Be scattering problems, 223, 230
  - nearest scattering root, 227–228
  - resonant orbital energy, 222–226, 229
  - SCF resonances, 223
  - shape resonance, 222
  - unbounded similarity transformation, 219
- Conjugate gradients, line search strategy, 262–263
- Core-extensive theories, 301
- Core-hole decay, 30
- Coulombic many-particle Hamiltonian, 218–220
- Coulomb integral, elementary Jacobi rotations, 387
- Coupled-cluster equations
  - constructing, 350–351
  - IP calculations, 349–350
- Coupled-cluster linear response theory, 315–319
  - hermitian version, 321
  - interaction Hamiltonian, 315, 321
  - IP calculations, 349
  - nonhermitian matrix, 317
  - numerical applications, 318–319
  - time-dependent version, 323
  - working equations, 316–318
- Coupled-cluster theory
  - closed shell, 295
  - incomplete model space, applications, 360–361
  - open-shell, 345
  - time-dependent version, 321–323
  - variational
    - cluster amplitudes, 352
    - energy, 353
    - Fock-space, 350, 352–353
    - isometry condition, 352

Coupled perturbed Hartree-Fock equations,  
255–258, 287  
Coupled perturbed Hartree-Fock theory,  
241–242  
CuCO clusters, 95  
Cu<sub>5</sub>CO clusters, 95  
Cu<sub>30</sub>H<sub>2</sub> cluster, 106

## D

*d*-band, hybridization with *s*-band, 97  
Debye-Waller attenuation factor, 60–61  
DFP, update optimization of NH<sub>3</sub>, 275, 278  
Diffuse low energy electron diffraction,  
surface structure, 27–28  
Dilatation operator, 189, 217–218  
Di-nitrogen, chemisorption on metals,  
surface structure, 157–158  
Dipole scattering, 74  
Discrete variational linear combinations of  
atomic orbitals, 98  
Discrete variation self-consistent charge, 98

## E

*e*-Be scattering problems, 223, 230  
Edmiston-Ruedenberg localization  
procedure, 430–432  
Electron affinities, 357  
Electron appearance potential fine structure  
spectroscopy, surface structure, 32  
Electron-atom scattering, 59–61  
Electron diffraction  
four-step description, 39–41  
surface structure, 5, 25–29  
Electron energy loss near-edge structure,  
surface structure, 55–56  
Electronic energy  
increment structure, 401, 403  
variation, 402–404  
Electronic spectroscopy, surface structure, 36  
Electron loss spectroscopy, surface  
structure, 54  
Electron microscopy, surface structure, 53  
Electron-photon scattering, 74–76  
Electron propagation, 58–59  
Electron scattering  
multiple, *see* Multiple electron scattering  
single versus multiple, 76–77  
Electron-stimulated desorption, surface  
structure, 37–38

Elementary Jacobi rotations, 377–380  
advantage, 418  
CI energy, 384–386  
comparison with exponential MSCSF  
methodology, 417–420  
Coulomb integral, 387  
exchange integral, 387  
invariants and relationships, 387–389  
localization procedures, 430–433  
MO framework, 380–383  
monoconfigurational density matrix,  
383–384  
monoconfigurational energy, *see*  
Monoconfigurational energy  
MO pair, 381  
multiconfiguration  
energy, *see* Multiconfigurational energy  
expression  
structure, 398–399  
optimization of rotation sine, *see*  
Rotation sine optimization  
quantum mechanical objects, 381  
rotation structure, 405–408  
two-electron molecular integrals, 386–389  
unitary matrix, 379–380  
Eley-Rideal mechanism, 87  
Energy hypersurface, characterization of  
minima, 243–245  
Equation of motion methods, 304–305  
Error bound method, 104–106  
Ethylidyne, chemisorption, 89, 91  
Exchange integral, elementary Jacobi  
rotations, 387  
Excitation energies, 357  
Excitation operator, 312, 329  
Extended appearance potential fine  
structure  
detection modes, 52–53  
surface structure, 52–53  
Extended fine structure techniques, 31  
Extended Hückel theory, 83–89  
acetylene chemisorption binding energies,  
87–90  
advantage, 84  
application, 85–89  
atom-superposition electron-delocalization  
molecular orbital theory, 84  
binding energies, 85  
chemisorption of hydrocarbon molecules, 87  
CO chemisorption, 85–87  
CO-Pt, 86  
electron delocalization energies, 84



- ethynidyne, chemisorption, 89, 91
- molecular orbitals, 83
- total energy, 83–84
- Extended-surface model, 82
- Extended x-ray absorption fine structure
  - phase shifts, 69
  - surface structure, 31–32, 49
- Extended x-ray energy loss fine structure,
  - surface structure, 55

## F

- Field ion microscopy, surface structure, 33
- Fine-structure techniques, *see also* specific techniques
  - range of multiple scattering, 77
  - surface structure, 29–32
- Flecher-Reeves algorithm, 262
- Fock-Dirac density matrix, 196
- Fock-Dirac density operators, 206
- Fock-Dirac operator, 187–188
- Fock matrix, 393
- Fock operator, 390–391
- Fock-space equation, 348
- Fock space theory, *see* Bloch equation
  - based Fock space theory; Full cluster expansion theories
- Full cluster expansion theories
  - Fock space, 332–353
    - advantage, 335
    - Bloch equation, 334
    - Bloch equation based theory, 345–350
    - flexibility in wave operator, 333
    - generating coupled-cluster equations, 334
    - similarity transformation-based theories, *see* Similarity transformation-based Fock-space theories
    - valence universal wave-operator, 333
    - variational CC theory, 350, 352–353
- Hilbert space, 324–323
  - Bloch equation, 328
  - CAS-SCF function, 325
  - cluster operators, 324
  - cluster structure, 324–325
  - effective Hamiltonian, 328
  - multireference CI approach, 325–326
  - multi-root formalisms, 331
  - Silverstone-Sinanoglu strategy, 327
  - single root formalisms, 328–331
- Full-valence model space, 297

## G

- GaAs, zincblende structure, 108, 110–111
- Generalized valence bond theory, 106–108
  - via configuration interaction, 108–110
- Geometry optimization, *see also* NH<sub>3</sub>,
  - AH<sub>2</sub> systems, multiconfigurational energy expression, 415–417
  - hydrogen peroxide, 271
- Global convergence tests, multiconfigurational energy expression, 401–404
- Golden-rule expression, 72
- Greenstadt method, optimization of NH<sub>3</sub>, 279

## H

- Hamiltonian, *see also* Many-particle Hamiltonian
  - dressed molecular, 318
  - effective, 328
    - disconnected, 364
    - energy-dependent, 303
    - generating hermitian, 344
    - hermitian, 350
    - matrix, 362–363
    - one-particle, 208
    - valence universal, 355
    - valence universal nonhermitian, 338
  - hermitian valence-shell, 339
  - interaction, 321
  - many-electron, transformed, 187–188
  - skeletons, 310
  - transformed Fock space, 343
- Hartree-Fock equations
  - canonical, 197, 213
  - coupled perturbed, 255–258
  - solving, 215
- Hartree-Fock function, 357
  - canonical, 208, 215
  - closed shell single determinant, 295
  - ground state, 313
  - transformation, 231
- Hartree-Fock method, *see also* Complex scaling method; Tarfala theorem
  - application to many-particle Hamiltonian, 221
  - bi-variational expression, 195–196
  - canonical Hartree-Fock equations, 197
  - complex conjugation, 211
  - complex symmetric, 189, 215–216, 231
  - operator, 205–209
  - real operator, 209–211

- dilated SCF equations, 231
  - effective one-electron operator, 196
  - effective one-particle Hamiltonian, 208
  - Fock-Dirac density matrix, 196
  - Fock-Dirac density operators, 206
  - iterative SCF-procedure, 210–211
  - one-particle projector, 206, 208–209
  - purpose, 198
  - restricted, 210
  - self-consistent, 101–102
  - similarity transformations, 201–205
  - Slater determinants, 195
  - stability problem of adjoint operators, 195–198
  - symmetry dilemma, 210
  - Hartree-Fock-Slater method, 94, 98–101
    - chemisorption on Ni surface
      - carbon monoxide, 99
      - chalcogen atoms, 100
    - determination of energetically favorable adsorption site, 99–100
    - discrete variational linear combinations of atomic orbitals, 98
    - discrete variation self-consistent charge, 98
    - molecular ionization energies, 99
    - total energy calculations, 100–101
  - Hartree-Fock theory, coupled perturbed, 255–258, 287
  - Hartree-Fock total energy, 84
  - Hausdorff transformation, 397
  - HCN, augmented Hessian, 285
  - H<sub>2</sub>CO, exact Hessian optimization, 273–274
  - Helium
    - energy increment curve, 425
    - total energy, configurational spaces, 410–412
  - Hermitian matrices, 313–314
  - Hessian matrix, 243–245, *see also* Quasi-Newton methods
    - approximate, 282–283
    - BFGS optimization cycles required, 281–282
    - estimating, coupled perturbed Hartree-Fock equations, 287
    - H<sub>2</sub>CO, 273–274
    - inverse, 252–254
    - NH<sub>3</sub>, 273–274
    - resolving for analysis, 263
  - Hessian method
    - approximate analytic, 255–258
    - augmented, 260–262, 284–285
  - High-resolution electron energy loss spectroscopy, 119
  - surface structure, 37, 54
  - Hilbert space, *see also* Full cluster expansion theories
    - one-electron, 189
  - HREELS, scattering, 74–75
  - Hydrocarbon, chemisorption systems, 87
  - Hydrogen, chemisorption, 106
  - Hydrogen peroxide
    - coordinates and gradients, 267–268
    - energy descent, 270
    - geometry optimization, 271
    - Newton iterations, 267, 269
- I**
- Impact scattering, 74
  - Incomplete model space, *see* Size extensive formulations, incomplete model space
  - Inelastic electron-electron scattering, 72–73
  - Inelastic low-energy electron diffraction, surface structure, 54
  - Infra-red reflectance-adsorption spectroscopy, surface structure, 37
  - Insulators, surface structure, 160–163
  - Invariance theorem, 214
  - Ion scattering, surface structure, 35
- J**
- Jordan blocks, 192–193
- L**
- Langmuir-Hinshelwood mechanism, 87
  - Lattice-gas chemisorption systems, 28
  - LEED
    - beam-set neglect, 67, 79
    - Bragg peaks, 77
    - defects, 78–79
    - diffuse theory, 66
    - disorder effects, 80
    - inverse, 65
    - long- versus short-range order, 78–79
    - range of multiple scattering, 77–78
    - renormalized forward scattering, 66
    - reverse scattering perturbation, 66
    - two-dimensional surface periodicity, 65
  - Level-shift function approach, 306
  - Li<sub>2</sub> molecule
    - electronic energy variation, 402–404
    - non-null  $\xi_{ij}$  coefficients, 405–408
  - Linear response theory, 306

- Line search strategy, 245–250
    - approximate analytic Hessians, 255–258
    - augmented Hessian method, 260–262
    - conjugate gradients, 262–263
    - cosine test, 248
    - coupled perturbed Hartree-Fock
      - equations, 255–258
    - descent condition, 246–247
    - downhill search, 260
    - extrapolation formula, 250
    - interpolation formula, 250
    - Newton's method, 251–252
    - quadratics surface, 259
    - quasi-Newton methods, 252–255
    - restricted step methods, 258–260
    - search direction methods analysis, 263–265
    - steepest descent, 251
    - tests for right and left extremes, 248
  - Linked cluster theorem, 301, 313–314
    - incomplete model space, 354
  - Localization procedures, elementary Jacobi rotations, 430–433
  - Low energy electron diffraction
    - advantage, 26
    - high Miller-index surfaces, 26
    - surface structure, 3, 25–27, 53–54
- M**
- Many-body perturbation theory
    - core-extensive open-shell, 302–303
    - incomplete model space version, 308
    - open-shell, 292, 326, 354
    - Rayleigh-Schrödinger, 303–304
  - Many-particle Hamiltonian
    - under complex scaling, 217–221
      - application of Hartree-Fock method, 221
      - Born-Oppenheimer approximation, 220
      - dilatation operator, 217–218
      - matrix property, 219
    - Coulombic, 218–220
    - stability problem, 188
    - symmetry property, 188
  - Many-particle operator, 187
  - similarity transformation, 199–216
    - associated kernels, 203
    - eigenfunction, 200–201
    - Hilbert space, 199
    - invariance of complex symmetry, 211–216
    - invariance property, 201
    - invariance theorem, 214
    - many-particle projector, 202–203
    - operator is self-adjoint and real, 214
    - real complex symmetric operator, 209–211
    - restricted, 211–216
    - Tarfala theorem, 201–205
    - transformation formula, 203
    - unbounded operator, 199
  - Many-particle projector, 202–203
  - MCSCF method
    - biconfigurational, total energy of  $\text{CH}_2$ , 418–420
    - comparison with elementary Jacobi rotations, 417–420
  - Medium-energy electron diffraction, surface structure, 53
  - MEED, chain method, 67
  - Method of complex scaling, *see* Complex scaling method
  - Minima, characterization, 243–245
  - MO energy increment, non-null  $E_{ij}$  coefficients, 405–406
  - MO functions, elementary Jacobi rotations, 380–383
  - MO integrals, monoconfigurational energy expressions, in terms of, 389
  - Molecular adsorption, 119–121
    - surface structure, 168–171
  - Molecular conformation, prediction through models, 241
  - Molecular integrals, two-electron, elementary Jacobi rotations, 386–389
  - Monoconfigurational density matrix, elementary Jacobi rotations, 383–384
  - Monoconfigurational energy expressions, 389–394
    - energy variation parameters, 394
    - Fock matrix, 393
    - Fock operators, 390–391
    - polynomial coefficients, 390
    - Roothaan-Bagus atomic energy, 391–394
    - rotation sine optimization, 423–424
    - sets of occupation numbers, 391
    - two-electron contribution
      - hypermatrices, 392
      - variation structure, 389–391
  - MO pair, elementary Jacobi rotations, 381
  - MO rotations, multiconfigurational energy expression, 398–399

Muffin-tin model, 58  
     multiple electron scattering, 62  
 Mulliken population analysis, 98–99  
 Multi-commutator Haussdorff formula, 336  
 Multiconfigurational energy, 395–420  
     choosing active orbital space, 412–414  
     choosing CI space, 408, 410–412  
     CI-MO computation algorithms, 400–401  
     CI wave function, 398  
     computational procedure, 440–441  
     electronic energy increment structure, 401, 403  
     electronic energy variation, 402–404  
     elementary Jacobi rotation, 387  
     extension to other molecular systems, 408–410  
     geometry optimization of  $AH_2$  systems, 415–417  
     global convergence tests, 401–404  
     Haussdorf transformation, 397  
     MO rotations, 398–399  
     rotation sine optimization, 424–426  
     total energy, 401–402  
         ground state, 408–409  
         selected configurations, 408, 410  
     unitary transformations, 396–397  
     variational coefficients, 437–439  
     variational structure, 396–398  
 Multiple electron scattering, 61–65  
     approximations, 66–69  
     forward, 76  
     high-energy fine-structure techniques, 67–68  
     at higher energies, 62  
     higher-order events, 79  
     incident electron beam, 65  
     intensity modulation, 68  
     at low energies, 61–62  
     matrix dimension, 64  
     muffin-tin model, 62  
     plane-wave expansion, 64–65  
     point-group symmetries, 64  
     range, 77–78  
     scattering amplitude, 63  
     versus single, 76–77  
 Multiple-scattering method, 89  
 Multireference CI approach, 325–326  
 Multi-root Hilbert space formalisms, 331–332  
 Murtagh-Sargent method, compared with BFGS, 280

## N

Near-edge fine structure techniques, surface structure, 30  
 Near-edge x-ray absorption fine structure, surface structure, 30, 50  
 Newton iterations, hydrogen peroxide, 267, 269  
 Newton's method  
     line search strategy, 251–252  
     optimizations, 271–274  
         Hessian eigenvalues, 273–274  
         number of cycles required, 272  
 $NH_3$   
     BFGS optimization, 275, 277, 282–283  
     DFP update optimization, 275, 278  
     exact Hessian optimization, 273–274  
     Greenstadt update in optimization, 279  
     Rank 1 update optimization, 275–276  
 $Ni_3O$  cluster, 99, 102  
 $Ni_{25}O_5$  cluster, 102–103  
 Nitric oxide, chemisorption on metals, surface structure, 157–158

## O

One-particle subspaces, stable under complex conjugation, 233–236  
 Open-shell coupled cluster theory, wave-operator, 345  
 Open-shell cluster expansion approach, 301, 305–309  
     core-excitation, 306–307  
     Fock-space strategy, 305–397  
     Hilbert space, 305–306  
     intermediate normalization, 308–309  
     occupation number representation, 307  
     quasi-complete model space, 308  
     virtual space determinants, 308  
 Open-shell many body formalism, 292, 302–310, 328  
     arrow conventions, 310–311  
     classification of orbitals into holes and particles, 310–311  
     earlier perturbative developments, 302–304  
     energy-dependent effective Hamiltonian, 303  
     Hugenholtz convention, 309–310  
     linked valence expansion, 303–304  
     open-shell cluster expansion approach, 305–309  
     propagator methods, 304–305

Open-shell many-body perturbation theory,  
292, 326, 354  
Optical techniques, surface structure, 37  
Optimization algorithms, 245  
Oxygen  
  binding site on Si (100), 103–104  
  chemisorption  
    GaAs (110), 108, 110–111  
    metals, 136–138  
    Ni(100), ordering, 114  
  molecular ionization energies, 99

## P

Pair-correlation model, 299  
Phase shifts, 59  
  EXAFS, 69  
  temperature-dependent, 60–61  
Photoelectron diffraction, surface structure,  
  28–29  
Photoelectron emission, 69  
Photoemission matrix elements, 70–71  
Photon propagation, 69–70  
Physisorption, 115  
 $\pi$ -back-bonding, 97  
Plane waves, 57–58  
Potential energy searches, 241–242  
  computational procedures, 266–267  
  summary of optimizations, 284  
Potential energy surface  
  characterization of minima, 243–245  
  line search strategy, *see* Line search strategy  
  optimization algorithms, 245  
Projection operator, antisymmetric, 201–202  
Propagator methods, 304–305  
Pseudo-selection rule, 73

## Q

Quantum mechanical objects, elementary  
  Jacobi rotations, 381  
Quasi-complete model space, 308  
  classification of operators, 356–357  
Quasi-complete model space theory, 354  
Quasi-Newton methods  
  line search strategy, 252–255  
  with non-unit starting Hessians, 281–283  
  optimization cycles required, 272  
  optimization of hydrogen peroxide, 271  
  with unit Hessian, 274–281  
  BFGS optimization, 275, 277

comparison of MS and BFGS  
  procedures, 280  
DFP update optimization, 275, 278  
Rank 1 update optimization, 275–276

## R

Rayleigh-Schrödinger many-body perturba-  
  tion theory, 303–304  
Reconstructed metals, surface structure,  
  131–134  
Reflection high-energy electron diffraction,  
  surface structure, 53  
Resonant orbital energy, 223–226, 229  
Restricted Hartree-Fock scheme, 210  
Restricted step methods, line search  
  strategy, 258–260  
RHEED, chain method, 67  
Roothaan-Bagus atomic energy, 391–394  
Rotation angle, 219  
Rotation sine optimization, 421–433  
  derivatives, 421–422  
  exact solution, 421–427  
    Brillouin theorems, 426–427  
    monoconfigurational energy, 423–424  
    multiconfigurational energy, 424–426  
    simple procedure, 421–423  
  modification of quadratic equation, 422–423  
  performance improving procedures, 427–430  
  approximate optimal solutions, 428–429  
  rotation selection and filtering, 429–430

## S

*s*-band, hybridization with *d*-band, 97  
Scale factor, 217, 219, 221  
Scanning tunneling microscopy, surface  
  structure, 34  
Scattered wave, 60  
Scattering amplitude, 60  
Schrödinger equations, 327  
Secondary-ion mass spectrometry, 36  
Second harmonic generation, surface  
  structure, 37  
Segré characteristics, 192  
Selenium, chemisorption on metals, 139  
Self-consistent field Hartree-Fock method,  
  101–102  
Self-consistent-field- $X\alpha$ -scattered wave  
  method, 89, 92  
  adsorption of ethylene on Ni, 96–97

- advantages and disadvantages, 94, 98
- calculations, 92–95
- carbon monoxide
  - adsorbate bonding geometries, 95
  - bonding to platinum (100) surface, 97
- core and valence photoemission spectra, 95–96
- core-level satellite structure, 96
- electronic wave functions, 92
- exchange potential, 92
- molecular orbital study, 96
- muffin-tin approximation to potential, 98
- orbital energies, 94
- $\pi$ -bonded model, 96–97
- potentials, 93–94
- Self-consistent-field- $X\alpha$ -scattered wave
  - method, 89, 92
  - results, 95–98
- Semiconductors
  - atomic adsorption on, 165
  - surface structure, 149–155
- Semi-cluster expansion theories, 310–323
  - anonymous parentage approximation scheme, 325
  - CC-LRT, 315–319
  - computation of transition moments, 323
  - core-correlation, 311–312
  - eigenvalue equation, 319
  - excitation operator, 312
  - hermitian matrices, 313–314
  - h-h matrix elements, 319–320
  - IP and EE computations, 319–321
  - linked cluster theorem, 313–314
  - non-variational methods, 315–323
  - physical considerations, 310–313
  - spin-adapted version, 323
  - time-dependent version, 321–323
  - vacuum term, 322
  - variational methods, 313–315
- Shape resonance, 222
- Similarity transformation-based Fock-space theories, 335–344
  - classification of operators, 342
  - closed-shell cluster amplitudes, 339–340
  - cluster wave-operator, 344
  - correlated core energy, 338
  - decoupling conditions, 342–343
  - eigenvalue equations, 340
  - hermitian valence-shell Hamiltonian, 339
  - multi-commutator expansion, 336–337
  - one-valence block, 337–338
  - passive scattering, 337–338
  - passive valence lines, 336
  - S operators, 341
  - spectator scattering, 337–338
  - two valence blocks, 338–339
  - unitary cluster operator, 343
  - valence universal cluster operator, 335
  - valence universal wave-operator, 339–340
  - zero-valence core problem, 334, 336–338
- Single root Hilbert space formalisms, 328–331
  - core excitations, 330
  - energy, 330
  - excitation operator, 329
  - wave operator, 329
- Size extensive formulations, 301–302
  - incomplete model space, 353–364
    - applications, 360–361
    - Bloch equation, 358
    - complementary active space, 354
    - construction zero, 362
    - defining equations for S operators, 356
    - disconnected effective Hamiltonian, 354, 364
    - effective Hamiltonian matrix, 362–363
    - electron affinities, 357
    - excitation energies, 357
    - general formulation, 354–357
    - having valence-holes and -particles, 357–360
    - IP calculations, 361
    - linked cluster theorem, 354
    - model space projection, 359
    - parent model space, 362
    - p-h excited determinants, 357–359
    - products of quasi-open operators, 354–355
    - quasi-open S operators, 355
    - size-extensivity of energies, 361–364
    - valence-universal effective Hamiltonian, 355
    - valence universal wave operator, 358
  - quasi-complete model space, classification of operators, 356–357
- Slater's transition-state theory, 94
- Slater type orbitals, 98
- Small-angle inelastic scattering, 74
- S operators, 341
  - defining equations, 356
  - quasi-open, 355
- Spherical-wave, 58
  - outgoing, amplitude, 71

- Substrates, atomic penetration into, 118–119  
Subsystem embedding condition, 348  
Sulfur, chemisorption on metals, 138–139  
Superlattice, 116  
Surface chemistry, 81–82, *see also Ab initio*  
    quantum chemical methods; Self-consistent-field- $X\alpha$ -scattered wave method  
    extended Hückel theory, 83–89  
    Hartree-Fock-Slater method, 98–101  
    multiple-scattering method, 89  
Surface crystallography, 112–121  
    adsorbate surface coverage, 114  
    Miller indices, 115  
    mutual atomic interaction, 113–114  
    notation for surface lattices, 115–116  
    ordering principles, 113–115  
    Wood notation, 116  
Surface EXAFS  
    inner potential, 76–77  
    surface structure, 32, 49–50  
Surface extended energy loss fine structure, surface structure, 55  
Surface strain, 110  
Surface structure, 2–4, *see also* Multiple scattering  
    angle integrated mode, 41  
    angle-resolved auger electron spectroscopy, 56  
    angle resolved mode, 41  
    angle-resolved photoelectron emission spectroscopy, 51  
    angle-resolved photoemission fine-structure, 28–29, 52  
    angle-resolved ultraviolet photoemission spectroscopy, 51  
    angle-resolved x-ray photoelectron spectroscopy, 51–52  
    angle-resolved x-ray photoemission, 51  
    atomic adsorbates on metal surfaces, 144–146  
    atomic beam diffraction, 33–34  
    atomic penetration into substrates, 118–119  
    bond angles versus bond directions, 81  
    bond length contractions and reconstructions, 117–118  
    chalcogen chemisorption on metals, 136–139  
    chemical composition, 36  
    clean metal structures, 122–126  
    cluster model, 82  
    CO, di-nitrogen and nitric oxide chemisorption on metals, 157–158  
    complementary techniques, 35–38  
    determination, 4–5  
    diffuse low energy electron diffraction, 27–28  
    electron diffraction, 5, 25–29, 39–41  
    electron energy loss near-edge structure, 55–56  
    electronic spectroscopy, 36  
    electron loss spectroscopy, 54  
    electron microscopy, 53  
    electron-stimulated desorption, 37–38  
    energy versus time diagrams, 42–46  
    extended appearance potential fine structure, 52–53  
    extended fine-structure techniques, 31  
    extended-surface model, 82  
    extended x-ray absorption fine structure, 49  
    extended x-ray energy loss fine structure, 55  
    field ion microscopy, 33  
    fine structure techniques, 29–32  
    forms of disorder, 79–80  
    future needs and directions, 173  
    high-resolution electron loss spectroscopy, 54, 119  
    inelastic low-energy electron diffraction, 54  
    insulators, 160–163  
    ion scattering, 35  
    layer spacing versus bond lengths, 80–81  
    low energy electron diffraction, 3, 25–27, 53–54  
    medium-energy electron diffraction, 53  
    molecular adsorption, 119–121, 168–171  
    multiple scattering, 47  
    near-edge fine structure techniques, 30  
    near-edge x-ray absorption fine structure, 50  
    numbers and types, 117  
    optical techniques, 37  
    ordering, 113  
    photoelectron diffraction, 28–29  
    process formalisms, 56–57  
    reconstructed metals, 131–134  
    reflection high-energy electron diffraction, 53  
    scanning tunneling microscopy, 34  
    semiconductors, 149–155  
    SEXAFS, 32, 49–50  
    single versus multiple scattering, 76–77  
    structural information, 41, 47  
        electron propagation, 58–59

surface extended energy loss fine structure, 55  
surface topography techniques, 32–35  
techniques, 6–24  
unsaturated hydrocarbons, 120–121  
vibrational spectroscopy, 37  
x-ray absorption near-edge structure, 50  
Surface topography techniques, surface structure, 32–35  
Symmetry dilemma, 210

## T

Tarfala theorem, 189, 221, 231  
many-particle operator, 201–205  
Tellurium, chemisorption on metals, 139  
Theoretical model, 293–298  
all-electron extensive models, 296  
core-extensive models, 296  
core-valence extensive models, 296–297  
criteria, 294–296  
size-consistent, 295, 297  
size-extensivity, 294, 297  
Thermal desorption spectroscopy, 36  
Two-electron molecular integrals, elementary  
Jacobi rotations, 386–389

## U

Unitary transformations, 396–397  
Unsaturated hydrocarbons, surface structure, 120–121

## V

Valence lines, passive, 336  
Valence universal cluster operator, 335  
Valence universal wave-operator, 333, 335, 347, 358  
Bloch, 345  
Variable metric methods, line search strategy, 252–255  
Vibrational spectroscopy, surface structure, 37

## W

Water, ground state, total energy, 413–414  
Wick's theorem, 346  
Wolfsberg-Helmholz formula, 83

## X

X-ray absorption near-edge structure, surface structure, 30, 50



This Page Intentionally Left Blank