

LOCAL QUANTUM CHEMISTRY

by
Hugo J Bohórquez

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

AT

DALHOUSIE UNIVERSITY
HALIFAX, NOVA SCOTIA
FEBRUARY 2011

© Copyright by Hugo J Bohórquez, 2011

DALHOUSIE UNIVERSITY

DEPARTMENT OF CHEMISTRY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “**Local Quantum Chemistry**” by **Hugo J Bohórquez** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: February 18 2011

External Examiner:

Research Supervisors:

Examining Committee:

Departmental Representative:

DALHOUSIE UNIVERSITY

Date: **February 18 2011**

Author: **Hugo J Bohórquez**

Title: **Local Quantum Chemistry**

Department: **Department of Chemistry**

Degree: **Ph.D.**

Convocation: **May**

Year: **2011**

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions. I understand that my thesis will be electronically available to the public.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

The author attests that permission has been obtained for the use of any copyrighted material appearing in the thesis (other than brief excerpts requiring only proper acknowledgement in scholarly writing) and that all such use is clearly acknowledged.

Signature of Author

In memory of my grandmother, Josefina.

Table of Contents

List of Tables	ix
List of Figures	x
Abstract	xiii
List of Abbreviations and Symbols Used	xiv
Acknowledgements	xviii
Chapter 1 Introduction	1
Chapter 2 Theoretical Background	6
2.1 The Electron Density as the Input	7
2.1.1 Atomic Properties and Group Additivity	10
Chapter 3 On the Local Representation of the Electronic Momentum Operator in Atomic Systems	13
3.1 Introduction	13
3.2 Quantum-Statistical Estimation of Local Observables	14
3.3 Local Estimate of Momentum Operator in Coordinate Representation	15
3.3.1 Local Kinetic Energy Theorem (LKET)	17
3.4 Density Functionals	19
3.5 Results	19
3.5.1 Bohr Atomic Model	20
3.5.2 The Local Momentum at the Origin is Given by Kato'S Cusp Condition	20
3.5.3 The Local Momentum for Valence Electrons Gives the Ionization Potential	22
3.5.4 The Local Momentum in the Middle Range Describes the Shell Structure of Atoms	23
3.6 Conclusions	30

Chapter 4	Lower Estimate to Relativistic Electron-Mass Corrections	32
4.1	Introduction	32
4.2	A Local Expression for the Relativistic Corrections	33
Chapter 5	Is The Size of an Atom Determined by its Ionization Energy?	38
5.1	Introduction	38
5.2	Methodology	38
5.3	Results and Discussion	42
5.4	Conclusions	45
Chapter 6	A Localized-Electrons Detector for Atomic and Molecular Systems	47
6.1	Introduction	47
6.2	Theoretical Considerations	49
6.2.1	The Single Particle Kinetic Energy Density and QTAIM	53
6.3	Localized-Electrons in Molecules	56
6.3.1	Single, Double and Triple Bonded Atoms	57
6.3.2	Hydrogen Bonded Systems	59
6.4	Conclusions	61
Chapter 7	The Localized Electrons Detector as an <i>Ab Initio</i> Representation of Molecular Structures	62
7.1	Introduction	62
7.2	Theory	63
7.2.1	The Single Particle Kinetic Energy and the Electron Localization	65
7.2.2	A Localized Electrons Detector (LED)	66
7.2.3	Electron Density Symmetry Around the Critical Points	67
7.3	Results	68
7.4	Conclusions	71
Chapter 8	Taxonomy of Chemical Bonds	73
8.1	Introduction	73

8.2	Theory	74
8.2.1	Symmetry of the Electron Density at the Stationary Points . . .	76
8.3	Covalent Bonds	79
8.4	Weak Interactions	81
8.5	Metallic Bonding	88
8.6	Discussion and Conclusions	91
Chapter 9	Quantum Molecular Similarity in Terms of Physical Observables	94
9.1	Introduction	94
9.2	Quantum Molecular Similarity	95
9.3	On the CPU Cost of Molecular Similarity	96
9.4	From Quantum Similarity to Observable Comparison	97
9.5	A Maximum Overlapping Implies Molecular Structure Similarity	99
9.6	Discussion and Conclusions	101
Chapter 10	Methods in Biocomputational Chemistry: a Lesson From the Amino Acids	103
10.1	Conformers, Rotamers and Physicochemical Variables	104
10.2	Side Chain Polarizations and the Theoretical Classification of Amino Acids	109
10.3	Quantum Mechanical Studies of Peptide-Host Interactions	115
10.4	Conclusions	119
Chapter 11	<i>Ab Initio</i> Study of the Structure, Energy and Polarizability of Amino Acids	122
11.1	Introduction	122
11.2	Methods	124
11.3	Optimized Structures	125
11.4	Energetic Results	125
11.5	Polarizability of the Amino Acids	129
11.6	Conclusions	132

Chapter 12	Final Remarks and Conclusions	133
12.1	Outlook	136
Bibliography		137
Appendix A	Copyright Permission Letters	146

List of Tables

3.1	Valence local kinetic energies \tilde{K}_Z^v , ionization energies I_Z and experimental ionization energies I_{exp}	24
3.2	Shell radii and electron populations for closed shell systems.	25
4.1	Relativistic correction to the electron mass for the fourth and fifth rows of the periodic table.	34
4.2	Periodic table with the relativistic correction to the inner shells as obtained from the mass correction term (Eq. 4.3).	37
5.1	Linear fit between five sets of theoretical radii and the experimentally-based atomic radius.	42
5.3	Experimentally-based atomic radius, $r_o(\text{au})$, for the first 104 elements.	46
7.1	Rank three critical points classification according to the sign of the Hessian eigenvalues.	67
7.3	Curvatures of benzene. All quantities in atomic units.	70
7.4	Curvatures of water dimer.	70
7.5	Critical points classification according to the magnitude of eigenvalues.	71
10.1	Properties selected for the representation of the amino acids.	107
11.1	Energy and dipole moment for the residues and peptides.	128
11.2	Polarizability of the amino acids in au.	130

List of Figures

1.1	Electron densities for Arg and Trp	2
3.1	Ionization potential: local theory, Koopmans' theorem and experimental.	21
3.2	Radial distribution function and local momentum for H and He.	27
3.3	Local electron momentum and the radial distribution function for Rb and Xe.	28
3.4	The shell structure of the calcium atom as given by the local momentum.	29
3.5	Local electron momentum for closed shells.	30
4.1	Relativistic electron mass correction.	35
4.2	Effective relativistic core for Cu.	36
5.1	Strontium (1S_0) local momentum and radial distribution function.	41
5.2	Plots of five sets of theoretical radii versus the experimentally-based radii proposed herein.	43
5.3	Scaled atomic radii versus the atomic number for the first five rows of the periodic table.	44
6.1	Variables involved in ELF for the ground state of F_2	50
6.2	LED for the ground state and the anti-bonding state of H_2 . . .	56
6.3	\tilde{P} isocontours for C_2H_2 , C_2H_4 , and C_2H_6	58
6.4	LED isocontours for the 20-electron hydrogen-bonded isoelectronic series.	60
7.1	LED for benzene.	69
7.2	Interatomic surfaces and LED isocontours for the water dimer. .	71
8.1	Electronic regions for propane.	74
8.2	BCPs classification according to the local symmetry.	78
8.3	LED for ethene and benzene.	78
8.4	LED for fluorine molecule.	80

8.5	LED for the formaldehyde-water complex.	81
8.6	QTAIM and LED for benzene dimer.	83
8.7	LED for phenanthrene.	85
8.8	Relative density and steric pressure at hydrogen atoms in phenanthrene.	85
8.9	LED for He@adamantane.	87
8.10	LED for ferrocene.	88
8.11	LED for clusters.	90
10.1	Capped amino acid residues used in this study.	105
10.2	Amino acids structure.	106
10.3	Theoretical variables responsible for the amino acids' main structural propensities.	108
10.4	Amino acids classification based on 8 principal components.	109
10.5	Energy magnitude versus mass, as provided by QTAIM, for the genetically-encoded amino acid side chains.	110
10.6	Andrews plots for the 10 QTAIM variables on 40 amino acid side chains.	111
10.7	Similar side chains as revealed by their corresponding Andrews plots.	111
10.8	Quantum theoretical classification of the genetically-encoded amino acids.	113
10.9	Peptide binding region (PBR) of the Major Histocompatibility Complex Class II (MHCII-P).	116
10.10	Diagram of the quantum study of MHC-P complexes.	118
10.11	Diagram of the quantum study of MHC-P complexes using the Mulliken-based multipole method.	119
10.12	Diagram of the quantum study of MHC-P complexes using the electrostatic potential.	120
10.13	Diagram of the quantum study of MHC-P complexes using a wavefunction analysis.	121
11.1	Optimized geometry of acetyl(Ala) ₇ NH ₂ at the PCM-B97D/D95(d,p) level of theory.	126

11.2	Comparison of the energetic effects of the mutation Ala→X for the 4th residue in acetyl(Ala) ₇ NH ₂ (11.2).	129
11.3	Theoretical polarizability of the amino acids.	131

Abstract

The single-particle momentum is studied as a tool for the visualization of the electronic regions in atoms and molecules. The limiting values of this function correctly obey two fundamental theorems: Kato's cusp condition and the Hoffmann-Ostenhof and Hoffmann-Ostenhof exponential decay. The local momentum also depicts the electron shell structure in atoms as given by its local maxima and inflection points. The integration of the electron density in a shell gives electron populations that are in agreement with the ones expected from the periodic table of the elements. The shell structure obtained is in agreement with higher level of theory computations. The average of the local kinetic energy associated with the local momentum is the Weizsäcker kinetic energy. It is shown that this quantity provides an estimate of steric interactions in molecules.

The single-particle momentum is a practical tool for the exploration of new stabilizing interactions for all kinds of molecular systems. It provides a three-dimensional representation of the molecular structure and depicts the polarizability regions, a feature not available with other continuous analyses.

A general definition of the radius of an atom in terms of its ionization energy is found. A relationship between these two fundamental properties is derived from the radial distribution function and the local momentum for the valence electrons. Strong correlations with well-known atomic radii suggest that this is a universally valid definition of the atomic radius.

The stability of peptides in the α -helix conformation upon replacement of the central amino acid is studied. These systems were optimized with a continuous solvent model and a recently developed DFT functional with empirical terms accounting for dispersion interactions. Both, the dispersion terms and the solvent model are directly related to the polarizability of the involved atoms. A new formula for an ab initio computation of the polarizability is introduced and tested for the amino acids.

List of Abbreviations and Symbols Used

Symbols

α	Polarizability
ϵ	Transverse ellipticity of a BCP
η	Longitudinal ellipticity of a BCP
γ	Lorentz factor
\hat{H}	Molecular Hamiltonian
λ_i	i^{th} eigenvalue of the Hessian of the electron density
$\langle A \rangle_\psi$	Expectation value of operator A
$\tilde{\mathbf{P}}$	Local momentum vector
\mathbf{J}	Vector current density
$\mathbf{p}(\mathbf{r})$	Statistical average of the single particle momentum
\mathbf{v}	Vector velocity
\mathcal{V}	Potential energy density
$\nabla\rho$	Gradient of the electron density
$\nabla^2\rho$	Laplacian of the electron density
$\bar{A}(\mathbf{r})$	Local value of observable A
ψ	Wavefunction
ρ	Electron density
ρ'	Derivative of the spherically-averaged electron density
τ	Kinetic energy density

τ_w	Weizsäcker kinetic energy density
\tilde{p}	Magnitude of the local momentum vector
$\tilde{A}(\mathbf{r})$	Local spread of observable A
G	Positive defined kinetic energy density
K	Kinetic energy
K_w	Expectation value of the Weizsäcker kinetic energy density
L	Laplacian of the electron density
m_e	Electron mass
r_o	Experimentally-derived atomic radius
Z	Atomic number

Abbreviations

BCP	Bond critical point
BQP	Bohm quantum potential
CADD	Computer-aided drug design
CD	Circular dichroism (spectroscopy)
DFT	Density functional theory
DRF	Direct reaction field
ECP	Effective core potential
ELF	Electron localization function
FIDIC	Fundacion Instituto de Immunologia de Colombia (Immunology Institute of Colombia)
GGA	Generalized gradient approximation

HF Hartree-Fock

HKT Hohenberg-Kohn theorem

KEF Kinetic energy functional

KSKE Kohn-Sham kinetic energy density

LED Localized-electrons detector

LKET Local kinetic energy theorem

MHCII-P Major histocompatibility complex class II

MM Molecular mechanics

NMR Nuclear magnetic resonance (spectroscopy)

PBR Peptide binding region

PCA Principal component analysis

PCM Polarizable continuum model

POS Proper open system

QM Quantum mechanics

QQ-SAR Quantum quantitative structure-activity relationship

QS Quantum similarity

QSAR Quantitative structure-activity relationship

QTAIM Quantum theory of atoms in molecules

RCP Ring critical points

SRF Solvent reaction field

TFE trifluoroethanol

vdW van der Waals

VSCC Valence shell charge concentration

VSEPR Valence shell electron pair repulsion

Acknowledgements

I would like to express my deepest and sincere gratitude to my thesis advisers, Prof. Russell J. Boyd and Prof. Chérif F. Matta for their great support and inspiring instruction in the development and completion of this study.

Thanks to the very sharp and assertive participation of the members of my committee, Profs. Axel Becke, Joe Zwanziger and Donald Weaver.

Special thanks to my friends and colleagues at Dalhousie for their time to proofread my work, Laura Albrecht, Gavin Heverly-Coulson, Felix Kannemann and Erin Martin. Special thanks to former members of the group whose contribution with discussion and comments was relevant, in particular Erin Johnson.

Thanks to the Natural Sciences and Engineering Research Council of Canada for financial support. Thanks to the computational facilities provided by ACEnet, the regional high performance computing consortium for universities in Atlantic Canada. ACEnet is funded by the Canada Foundation for Innovation (CFI), the Atlantic Canada Opportunities Agency (ACOA), and the provinces of Newfoundland and Labrador, Nova Scotia, and New Brunswick.

And thanks especially to my family and friends for their continuous support.

Chapter 1

Introduction

The electronic structure of molecules can be studied from many different perspectives. In crystals, for instance, the continuous electronic states are described in the reciprocal space by the band structure. Observable properties related to the electrical conductivity can be addressed using this representation of the electronic states, which helps in obtaining simple models useful in the simulation and design of electronic semiconductor devices, for example. Unfortunately, for non periodic systems there is no analysis equivalent to that of the band structure which can depict the electronic structure of molecules in the three-dimensional (3D) coordinate space.

Several proposals for a 3D representation of the electronic structure in chemistry have flourished within the past three decades. Each of them deals with terms such as “the electron pair” or “the chemical bond” which do not unambiguously correspond to directly measurable physical quantities. Nevertheless, due to their conceptual utility, clear and rigorous physical definitions of these and related terms are indispensable.

In the present thesis a new tool for the analysis of the electronic regions in atoms, molecules and crystals is developed. It depends explicitly on the electron density and can be derived formally from the local quantum theory. The main variable studied is closely related to other quantum chemical theories, such as the density functional theory (DFT) and the quantum theory of atoms in molecules (QTAIM).

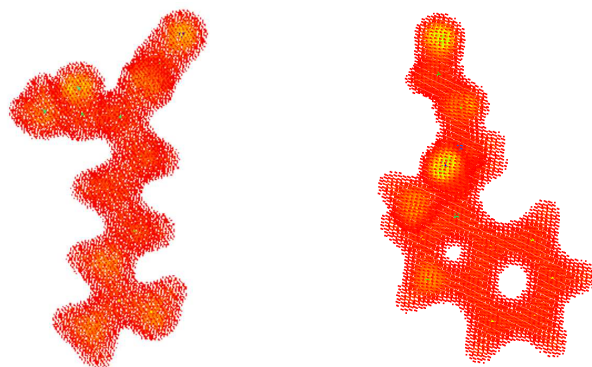
The main motivation for this work was the idea that an electron density analysis must provide complete information about chemically-relevant electron regions in atoms and molecules. This assumption is supported by the Hohenberg-Kohn theorems (HKT), which state that the external potential is uniquely defined by the electron density ρ ; and that any trial density leads to an energy that is greater than or equal to the energy of the ground state.¹ According to the HKTs, the electron density completely characterizes the ground state of any molecular system. The electron density also provides a theoretical bridge between the Hilbert space where the molecular wave function ψ is represented, and the three-dimensional world.² These theorems clearly suggest that the physics and the chemistry of any molecular system can be completely

studied in the 3D coordinate space, similar to how the classical systems are studied.³

However, a direct analysis of the electron density is impractical for the identification of electronic regions in the 3D space. Specifically, ρ changes by several orders of magnitude within a few angstroms, making direct 3D analysis difficult. It is necessary to apply precise transformations to the electron density in order to visualize chemically-relevant regions.

The original motivation for the present thesis was the development of a theoretical tool for the study of the electronic regions in molecules. In a preliminary attempt, each coordinate point in the 3D space is given by a pixel that is coloured according to a sigmoid function applied to ρ , and the transparency and intensity of each pixel is calibrated with the density values. An example of this application appears in Figure 1.1, which reveals certain chemical features like the core regions in oxygen atoms and the presence of bonding interactions. This is a practical tool for a quick assessment of certain regions of interest in molecular systems, but any improvement required additional research.

Figure 1.1: Electron densities for Arg and Trp at HF/6-31G level with the software SOLID, FIDIC 2000© by H. J. Bohórquez and M. Obregón, (C++ & Open Inventor libraries, SGI). The original black background was digitally changed to white.



The object of this thesis is thus to develop a theoretical tool capable of providing an accurate description of the electronic regions in atoms and molecules by using the electron density as the main variable. Here it is demonstrated, through a series of papers (published or to be submitted), that the single-particle local momentum is especially well suited for the localization of chemical bonds and paired electrons in the three-dimensional coordinate space.

The method developed here is entirely based on the electron density and accomplishes a set of requisites:

1. It is derivable from first principles.
2. It provides a direct and practical visualization of different electronic regions in the coordinate space.
3. It is universal in the sense that it is equally applicable to atoms, molecules, molecular complexes, clusters and crystals.
4. It is consistent with other electron density analyses such as QTAIM.
5. It provides a visualization consistent with both the physics and the chemistry of the system.
6. It is independent of the level of theory used.

Several quantum-chemical analyses found in the literature already accomplish the items of this list. For instance, the electron localization function (ELF) of Edgecombe and Becke⁴ fulfils all requirements except item 4, which may be a consequence of its DFT conceptual origins.

In this thesis, the single-particle momentum has been systematically studied as a density-based analysis tool. The single-particle momentum can be derived from the local quantum mechanics of a single particle. As it has been shown in a series of publications, the single-particle momentum is able to identify the electron density regions in atoms, molecules, complexes and crystals. With this analysis, electronic sub-regions such as covalent bonds or atomic shells can be readily located in the 3D space, and therefore it is called the “localized electrons detector” LED. The method has been extensively tested for a variety of systems in order to validate its applicability. It has been demonstrated that LED is fully compatible with QTAIM theorems. For instance, the central variable for the topology of the electron density is the gradient function which is the *current density* associated with the single-particle local momentum. In addition, LED is conceptually related to DFT. The kinetic energy density associated with the single-particle momentum is the Weizsäcker term, τ_w , which appears as a “gradient correction” in generalized gradient approximations (GGAs). Moreover, in conceptual DFT τ_w leads to an estimate of steric effects, as recently shown by Liu.⁵

During the development of this theoretical tool several important findings have been achieved. For instance, a topological criterion for the precise definition of electron shells in atoms was found (Chapter 3). These ideas suggest that it is possible to determine the effective cores used in relativistic approximations in a rational way, as is sketched in Chapter 4.

The electronic structure depicted by LED also indicates the existence of an observable-dependent confinement region for the atom. This idea leads to the proposal of an experimentally-based atomic radii table (Chapter 5). It is also linearly correlated ($R > 0.9$) to five different sets of theoretical atomic radii.

For molecules (Chapters 6-8), LED reveals a wealth of chemically-relevant information. It depicts the core electrons in a spherically-symmetrical shape and typical covalent bonds as rods connecting the involved nuclei. This 3D picture alone constitutes an ab initio molecular structure that agrees with the chemical intuition. In addition, and perhaps more importantly, LED reveals the local electronic polarization of the valence electrons. This feature helps in understanding challenging bonding situations that contradict chemical intuition. For instance, it helps to explain how certain types of hydrogen-hydrogen interactions can actually stabilize the phenanthrene molecule, contrary to what might be expected. A quantitative estimate of this effect is given by the Weizsäcker kinetic energy density. In addition, a new set of topological properties for the stationary points of the electron density have been found that may be added to the existing properties as described by QTAIM. These findings indicate that LED provides a clear graphical representation of all kinds of interatomic interactions, from covalent to metallic bonds.

One of the main goals of this project was to develop tools for the theoretical assessment of electronic features in molecules, with an aim for the characterization of organic macromolecules. Despite recent developments in QTAIM algorithms, the level of theory available is still too prohibitive to be practical in fields such as drug design, where the classical mechanics approach still dominates.

The electronic polarizability is responsible for many fundamental properties of biomolecules. In fact, the electronic multipole polarizations of the side chains of amino acids lead to a quantum classification resembling their experimental biochemical classification. Unfortunately, such polarizability is absent in most classical mechanics models.

The final part of this thesis is concerned with the molecular similarity problem

(Chapters 9-11). In addition, a study of the stability of peptides in α -helix conformation upon replacement of the central amino acid is introduced. These systems were optimized with a continuous solvent model and a recently developed DFT functional with empirical terms accounting for dispersion interactions. Both, the dispersion terms and the solvent model are directly related to the polarizability of the involved atoms. A formula for an ab initio computation of the polarizability is introduced and tested for the amino acids.

Chapter 2

Theoretical Background

“Today, we celebrate the fact that mathematics has invaded chemistry, that by means of theoretical calculations we can predict a large variety of chemical phenomena...

An atom consists of a nucleus and electrons. The motion of the electrons is described by the laws of quantum mechanics. When these laws were formulated more than 70 years ago, researchers realized immediately that in them was contained the explanation of the chemical bond.”

Professor Björn Roos delivering the Presentation Speech for the 1998 Nobel Prize in Chemistry at the Stockholm Concert Hall.

Such an enthusiastic speech makes partial account of the famous prediction made by Paul Dirac regarding his views on chemistry and physics, seven decades before:

“The underlying laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that exact applications of these laws lead to equations which are too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.”

P.M.A. Dirac, Proceedings of the Royal Society A 123, 714 (1929)

Interestingly, the first part of the quotation is usually the most cited, but it sounds rather pessimistic and has been the target of critics because of its implicit “reductionism” of chemistry to physics.¹ The second part of Dirac’s phrase adds a more realistic view, because it clearly states that *any* practical quantum model of chemical systems

¹A notable adherent of this view is Roald Hoffmann (Nobel Laureate in Chemistry, 1981), who says: “only the wild dreams of theoreticians of the Dirac school make nature simple”, within a philosophical paper exposing several of his “concerns” regarding the contemporary role of theoreticians in chemistry.⁶

will be merely an *approximate* one. Such is the character of every computational solution of a molecular system.

A precise simulation of the physicochemical properties of matter requires precise computations of the electronic properties of molecules. But a system involving more than one electron is technically considered a many-body system. Therefore, any chemical system of interest falls into this category. This fact, together with the inherently probabilistic nature of the wave function, makes the solution of the Schrödinger equation a formidably complex computational task, even for the simplest molecule. Such level of complexity demanded by molecular computations is commonly referred to as the *exponential wall* of quantum mechanics. As Walter Kohn admits, overcoming these limitations was one of the main motivations for developing the density functional theory (DFT).³ Another reason, Kohn points out, is the fundamental understanding gained by a simplified, but equivalent, formulation of the quantum theory in terms of the electron density.

DFT transforms a many-body system into a system of non-interacting particles that move within an *effective potential*, similar to other mean field theories. Walter Kohn’s contribution to DFT was awarded with the 1998 Nobel Prize in Chemistry, which he shared with John Pople, one of the developers of modern computational techniques for the numerical solution of molecular wavefunctions.

In this section, some of the fundamentals that are used in this thesis are briefly outlined. In particular, some of the basic ideas of QTAIM are mentioned after a digression on the interpretation of the electron density.

2.1 The Electron Density as the Input

Richard Feynman published in 1939 a short paper entitled “Forces in Molecules” that greatly influenced the theoretical chemistry language afterwards, mainly through the quantum-mechanical principle known as the Hellmann-Feynman theorem.⁷ This theorem expresses in electrostatic terms a way to correlate the quantum nature of molecules with classical observables. For example, the average electrostatic forces are given by

$$F_{X_\gamma} = - \left\langle \frac{\partial \hat{H}}{\partial X_\gamma} \right\rangle \quad (2.1)$$

where F_{X_γ} is the force acting on the X_γ -component of a given nucleus, and \hat{H} is the molecular Hamiltonian. After replacing the Hamiltonian, the force F_{X_γ} is given in terms of the single-particle electron density ρ

$$F_{X_\gamma} = -Z_\gamma \left(\int d\mathbf{r} \rho(\mathbf{r}) \frac{x - X_\gamma}{|\mathbf{r} - \mathbf{R}_\gamma|} - \sum_{\alpha \neq \gamma} Z_\alpha \frac{X_\alpha - X_\gamma}{|\mathbf{R}_\alpha - \mathbf{R}_\gamma|} \right) \quad (2.2)$$

where the first term is the electron-nucleus interaction and the second one the nucleus-nucleus.

In the same paper, Feynman proposes a hypothetical atomic partitioning of the electron density, $\rho = \sum_i \rho_i$. He also mentions a detailed description of the electrostatic nature of bonding interactions in molecules; and an explanation to the van der Waals interactions in terms of the electronic polarizability.

The single-particle electron density ρ is formally defined as the diagonal contraction of the spin-free reduced density matrix, i.e. the trace of the density operator

$$\hat{\rho}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}_0) \quad (2.3)$$

where N is the number of electrons⁸

$$\rho(\mathbf{r}) = Tr \hat{\rho}(\mathbf{r}) \quad (2.4)$$

The density operator is a Hermitian operator that describes the statistical state of a quantum system. Hence the expectation value of an observable \hat{A} is

$$\langle \hat{A}(\mathbf{r}) \rangle = Tr [\hat{\rho}(\mathbf{r}) \hat{A}(\mathbf{r})] \quad (2.5)$$

Expectation values of observables are a primary goal in quantum mechanics; however, for chemistry a more attractive idea is the decomposition of an observable into atomic contributions and transferable fragments. Such an idea led Richard Bader to develop the *quantum theory of atoms in molecules*⁹ (QTAIM). He demonstrated that the same boundary conditions that apply to a molecular system are found for atomic regions. These subsystems are called proper open systems (POS), because they can exchange electronic population, energy and volume.¹⁰

Bader introduced the concept of a *proper open system* as the quantum chemical definition of the region that pertains to an atom inside a molecule.¹⁰ The boundary

conditions of a POS are equivalent to the boundary conditions of the whole molecule, which is mathematically expressed by the zero-flux condition integral

$$-\frac{\hbar^2}{4m} \oint dS \nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad (2.6)$$

that leads to the local zero-flux condition

$$\nabla \rho(\mathbf{r}_S) \cdot \mathbf{n}(\mathbf{r}_S) = 0 \quad (2.7)$$

that defines the inter-atomic surfaces (IAS) as the 2D sub-regions $\mathbf{r}_S \in \mathcal{S}(\Omega)$ which are embedded inside the 3D coordinate space. The zero-flux equation is a local boundary condition that defines the limits between atoms. The 3D space is divided into disjoint Ω_i volumes, where i identifies each atom in the molecule. Thus, each atom has an exclusive portion of space where its physical properties can be evaluated.

Bader introduced the *dynamical systems analysis* for the study of molecular stability,¹¹ giving rise to the topologically-defined molecular structure theory. The main quantity in this formulation is the gradient of the electron density $\nabla \rho$. The stationary points of the electron density are the critical points, i.e. those points where the gradient vanishes. The molecular structure is given in terms of gradient paths linking the nuclei through saddle points generically called *bond critical points* (BCP). The gradient vector field also provides an exhaustive partitioning of the 3D space into *atomic basins*ⁱⁱ, i.e. atomic sub-regions where the virial theorem holds.¹²

It is known from experiments that atoms and functional groups of atoms can be transferable to a high degree. When this occurs, one can determine the atomic or group contribution to the total properties for a system. Therefore, any transferability probe necessarily implies the evaluation of sub-molecular properties. Building quantum chemical representations of molecules by combining molecular fragments has been the subject of many studies,^{13,14,15,16,17} and QTAIM is one of the most popular theories for the study of chemical groups in terms of its fundamental components.

QTAIM language greatly permeated the conventional analysis of molecular systems.¹⁸ Bader recognizes the source of his inspiration in the works of four Nobel Laureates in Physics: "... I started off as an organic chemist, and I had to learn all of this physics. In my work I quote Schrödinger, Schwinger, Dirac and Feynman because they are the only references I need. My theory is based on the physics".¹⁹

ⁱⁱHowever, some non atomic basins can be topologically possible, like in the case of Li₂.

Bader’s theory is not limited to the topology of the electron density, although it is the *practical* approach followed by most of the theoretical chemists citing his work. Its contribution to the visualization of molecular interactions and atomic regions is equally pervasive in quantum chemistry and solid state physics.¹⁸

There is a close link between the topology of the electron density and the Nobel laureates mentioned by Bader. His interpretation of the electron density seems to follow Schrödinger’s interpretation of the wavefunction as the *charge density* spread out over the infinite space. Remarkably enough, Kohn most certainly acquired his familiarity with variational methods, which are key for DFT formulation, during the years he spent as a PhD student at Harvard University under Julian Schwinger’s supervision. In 1965 Schwinger shared the Nobel prize in physics with Sin-Itiro Tomonaga and Richard Feynman “for their fundamental work in quantum electrodynamics (QED), with deep-ploughing consequences for the physics of elementary particles”, as stated in the official announcement.²⁰ Therefore, at the very core of the fundamental works on the electron density most widely used in theoretical chemistry, there is a resilient appearance of the ideas proposed initially by just a few of the developers of quantum mechanics and quantum electrodynamics.

This brief historical recall of events attempts to illustrate just how convergent the philosophical views underneath the electron density studies in chemistry developed during the past five decades, from DFT to QTAIM have been. While other ideas and developments based on the very same quantum mechanical postulates have been explored since the early days of the theory, not many transcended successfully through the theoretical chemistry community.

2.1.1 Atomic Properties and Group Additivity

An atomic property corresponding to the i^{th} atom is defined as being the volume integral of the property density, $\mathbf{G}(\mathbf{r})$, over the atomic region Ω_i :

$$G(\Omega_i) = \int_{\Omega_i} \mathbf{G}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \quad (2.8)$$

QTAIM preserves the additivity of physical properties. That is, the molecular expectation value of the property \mathbf{G} is

$$\langle \mathbf{G} \rangle = \sum_i^N G(\Omega_i) \quad (2.9)$$

where N is the total number of participating atoms in the 3D volume Ω for property $\mathbf{G}(\mathbf{r})$.

The atomic population, $N(\Omega_i)$, can be considered as the zeroth charge density moment,

$$N(\Omega_i) = -e \int_{\Omega_i} \rho(\mathbf{r}) d\mathbf{r}. \quad (2.10)$$

Thus, the atom's net charge is:

$$q(\Omega_i) = Z_i e + N(\Omega_i) \quad (2.11)$$

where Z_i is the atomic number of i^{th} atom.

The energy of the i^{th} atom in a molecule, $E_e(\Omega_i)$, is given by

$$E_e(\Omega_i) = T(\Omega_i) + \mathcal{V}(\Omega_i) \quad (2.12)$$

where the kinetic energy $T(\Omega_i)$ and the potential energy $\mathcal{V}(\Omega_i)$ are related by the atomic virial theorem,⁹ and therefore

$$E_e(\Omega_i) = -T(\Omega_i) = \frac{1}{2} \mathcal{V}(\Omega_i) \quad (2.13)$$

The first moment of charge density, $\mathbf{M}(\Omega_i)$, provides a measure of the extent and direction of the atom's charge density dipolar polarization by determining the displacement of the atom's centre of negative charge from the position of its nucleus:

$$\mathbf{M}(\Omega_i) = -e \int_{\Omega_i} \mathbf{r}_{\Omega_i} \rho(\mathbf{r}) d\mathbf{r}. \quad (2.14)$$

The second moment, $\mathbf{Q}(\Omega_i)$, gives information on planar distributions of charge, which is particularly prevalent in aromatic groups. This is the quadrupolar polarization of an atomic density. When it is measured with respect to one of the three orthogonal planes (z), its expression is given by:

$$Q_{zz}(\Omega_i) = -\frac{e}{2} \int_{\Omega_i} (3z_{\Omega_i}^2 - r_{\Omega_i}^2) \rho(\mathbf{r}) d\mathbf{r} \quad (2.15)$$

and an off-diagonal element involving the coordinates x and y is

$$Q_{xy}(\Omega_i) = -\frac{3e}{2} \int_{\Omega_i} x_{\Omega_i} y_{\Omega_i} \rho(\mathbf{r}) d\mathbf{r} \quad (2.16)$$

QTAIM is remarkably insensitive to the choice of the basis set for the solution of the molecular wavefunction.^{21,22} Hence this theory might be a useful tool for rational drug design where interesting molecules include hundreds of atoms, which conditions the selection of basis sets.¹⁸

From the chemical perspective it is relevant to ask about the electron-pair density and how this information is reflected in the electron density ρ alone. In a recent study by Lobayan et al, the electron density is decomposed into the *paired* and the *effectively unpaired* contributions, $\rho = \rho_p + \rho_u$, via the density operator formalism.²³ Their results show that unpaired electron density concentrations, as accounted by the Laplacian, occur outside the topological bonding regions, whereas the paired electron densities exhibit accumulations inside those regions. In addition, the paired density contribution to the total electron density is a few orders of magnitude bigger than the unpaired one, i.e. $\rho_p \gg \rho_u$, and hence $\rho \approx \rho_p$. Therefore, for practical purposes the topology of the electron density alone provides as much information as the paired density itself, especially in those regions where electron pairs are usually found, like inside the core shells or in the covalent-bonding regions. These results and previous topological studies on the electron pairs²⁴ prove that the quantum chemical topology has a firm theoretical support.

Chapter 3

On the Local Representation of the Electronic Momentum Operator in Atomic Systems

Reprinted with kind permission from the American Institute of Physics.

H. J. Bohórquez and R. J. Boyd, *J. Chem. Phys.*, **129**, 024110 (2008)

3.1 Introduction

The single-particle electron density $\rho(\mathbf{r})$ contains a wealth of chemical information. It also provides the bridge between the abstract Hilbert space, where the state wave function ψ is represented, and the three-dimensional world.² As stated by the Hohenberg-Kohn theorem (HKT), $\rho(\mathbf{r})$ completely characterizes the ground state of any molecular system.¹ Such statements about the electron density lay down the hypothesis that the physics of a molecular system can be *completely* studied in the real, three-dimensional coordinate space, just as the classical systems usually are.³ Despite proven success for describing the physics of molecular systems in a comprehensive and practical manner,^{25,26,18} the validity of single-electron analysis is still an active matter of debate in quantum chemistry, largely centred on the differences between *topological* density-based and molecular orbital-based approaches.²⁷ Herein is described an alternative approach to the physical properties of chemical systems based on results extracted from a quantum local statistics theory.²⁸

The *local quantum theory* provides a comprehensive understanding of electron-density-derived quantities,²⁸ including a representation of the electron momentum in terms of the single-electron density. The extreme values of this property are confirmed by two paramount theorems in quantum chemistry: Kato's cusp condition,²⁹ and the exponential decay of the electron density found by Hoffmann-Ostenhof and Hoffmann-Ostenhof.³⁰ Additionally, it is shown how this single physical variable is able to adequately depict the atomic electron shells. The shell structure is given by the critical points of the local momentum as obtained from the spherically-averaged electron

density at Hartree-Fock (HF) level. The electrons contained within the shells coincide with the values expected on the basis of the periodic table of the elements.

These results suggest that the local quantum theory provides reliable statistical values of the electronic momentum in atomic systems in a formal and unified way. The approach to the local quantum theory illustrated here is an attempt to show its practical use as well as its interpretative power in quantum chemistry. Previous discussions were devoted to the formalism itself rather than to practical numerical results.^{31,32,33,34,28,35,36,37}

In the next section the minimum aspects of Luo’s local theory are reviewed.²⁸ In the rest of the chapter, the physical properties of the local momentum are discussed. The potential use of the local quantum theory for the further study of other local variables in quantum chemistry is also discussed.

3.2 Quantum-Statistical Estimation of Local Observables

Luo derived local values of quantum observables in the 3D coordinate space, whose results are in perfect agreement with other recent studies on local representation of observables.^{32,38} Here the main aspects of local observables theory are merely outlined. A more comprehensive derivation is found in the original references.^{28,39,40,32,41,34}

According to Luo, “given a quantum mechanical observable A and a quantum system at a definite state ψ , one can construct a *classical observable*, that is, a real function in the configuration space $A(\mathbf{r})$, such that it is the optimal estimate of the quantum observable, in the sense of minimum variance”.²⁸ The Heisenberg uncertainty principle prohibits the possibility of simultaneously assigning exact values to conjugate observables. However, it is possible to assign an exact value to one observable which serves as a reference observable, and only a statistical average to the conjugate one. The local quantum theory formally defines the proper way to do it.^{41,40,39} In the following expressions it is assumed that \mathbf{r} is the reference variable, and the momentum \mathbf{p} is described by a statistical average, $\mathbf{p}(\mathbf{r})$.

Within the local quantum theory, any physical observable A can be decomposed into two parts: a real part $\bar{A}(\mathbf{r})$ (the *local value*) that corresponds to a classical approximation of A , and an imaginary part $\tilde{A}(\mathbf{r})$ (the *local spread*) which is a quantum fluctuation of A . The expectation value of A , $\langle A \rangle_\psi$, is therefore given by

$$\langle A \rangle_\psi = \int A(\mathbf{r}) |\psi|^2 d\mathbf{r} \quad (3.1)$$

The classical estimate of A when the quantum system is in the state ψ , is the real function $\bar{A} = \bar{A}(\mathbf{r})$ on \mathbb{R}^3 (rather than a quantum operator on L^2), such that the variance $Var_\psi(\tilde{A})$ is minimized subject to the condition $\langle \bar{A} \rangle_\psi = \langle A \rangle_\psi$, i.e. $\langle \tilde{A} \rangle_\psi = 0$. This condition guarantees that the real function $A(\mathbf{r})$ is the optimal estimate of the quantum observable A . Lou shows that the unique solution for this optimization problem is

$$\bar{A}(\mathbf{r}) = Re\left(\frac{A\psi}{\psi}\right)(\mathbf{r}) \quad (3.2)$$

$$\tilde{A}(\mathbf{r}) = Im\left(\frac{A\psi}{\psi}\right)(\mathbf{r}) \quad (3.3)$$

and

$$Var_\psi(A - \bar{A}) = \int (\tilde{A}(\mathbf{r}))^2 |\psi|^2 d\mathbf{r} \quad (3.4)$$

As demonstrated by Luo,²⁸ the quantum fluctuation is essentially the Fisher information.⁴² Summarizing, any quantum observable A can be decomposed into two local contributions: an average real-valued function $\bar{A}(\mathbf{r})$ and an imaginary fluctuation $i\tilde{A}(\mathbf{r})$,

$$A(\mathbf{r}) \equiv \bar{A}(\mathbf{r}) + i\tilde{A}(\mathbf{r}) \quad (3.5)$$

3.3 Local Estimate of Momentum Operator in Coordinate Representation

The single-electron complex wave function $\psi(\mathbf{r})$ can be expressed in polar form as $\psi(\mathbf{r}) = \rho(\mathbf{r})^{1/2} e^{i\frac{\hbar}{\hbar}S(\mathbf{r})}$, where $\rho(\mathbf{r})$ is the ground state electron density. Using $\psi(\mathbf{r})$ and the momentum operator in coordinate representation $\mathbf{P} = -i\hbar\nabla$ in Eq. 3.5, the following local representation of the momentum operator is obtained

$$\mathbf{P}(\mathbf{r}) = \nabla S(\mathbf{r}) - i\frac{\hbar}{2} \frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})} \quad (3.6)$$

Equation 3.6 states that the wave function phase $\nabla S(\mathbf{r})$ is responsible for the *classical* part of the momentum average $\bar{\mathbf{P}}(\mathbf{r}) = \nabla S(\mathbf{r})$, while the quantum *spread* is accounted for by the term $\tilde{\mathbf{P}}(\mathbf{r}) = -\frac{\hbar}{2} \frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})}$.

This local representation is equivalent to the recently proposed *deformed* momentum operator by Mosna *et al.*³² which, in turn, is similar to the *velocity field operator*

for spinning particles by Salesi and Recami.³⁸ Additionally, as derived from Bohmian quantum mechanics,^{43,44} Hirschfelder *et al.* used a similar velocity concept to obtain *imaginary and real streamlines* for the study of hydrodynamical aspects of quantum mechanics.⁴⁵ Ghosh and Deb also studied general properties of this local momentum expression within their *quantum fluid dynamics*.³⁵ In summary, the formal aspects of the local representation of the momentum operator have already been discussed in the literature in different contexts. In the present work those concepts are used for exploring the electron-density based properties of atoms.

From Eq. 3.6 it is easy to derive a local expression for the velocity vector field, as the local momentum $\mathbf{P}(\mathbf{r})$ divided by the electron mass, m_e . In quantum chemistry the notion of electron velocity is rather uncommon while a closely-related quantity, the current density, is a very well studied observable. These two properties are related via the equation $\mathbf{J} = \rho\mathbf{v}$.⁴³ Following this relationship the local current density can be defined as $\mathbf{J}(\mathbf{r}) = \frac{1}{m_e}\mathbf{P}(\mathbf{r})\rho(\mathbf{r})$, which after replacing Eq. 3.6 leads to

$$\mathbf{J}(\mathbf{r}) = \frac{1}{m_e}\rho(\mathbf{r})\nabla S(\mathbf{r}) - i\frac{\hbar}{2m_e}\nabla\rho(\mathbf{r}) \quad (3.7)$$

It is easy to prove that the first term in this equation is exactly the quantum mechanical vector current density,⁹ which is defined in terms of the single-particle wave function as $\frac{\hbar}{2m_e i}(\psi^*\nabla\psi - \psi\nabla\psi^*)$. The *current local value* is then $\bar{\mathbf{J}}(\mathbf{r}) = \frac{1}{m_e}\rho\nabla S$, and the *current fluctuation* is given by $\tilde{\mathbf{J}}(\mathbf{r}) = -\frac{\hbar}{2m_e}\nabla\rho(\mathbf{r})$. Clearly, in the absence of an external field, $\bar{\mathbf{J}}(\mathbf{r}) = 0$, i.e. $\nabla S(\mathbf{r}) = 0$.

According to Bader,¹⁰ the gradient of the electron density brings the physical definition of the molecular structure and the boundaries of an atom within a molecule.⁹ This quantity, $\nabla\rho(\mathbf{r})$, is the cornerstone of the quantum theory of atoms in molecules (QTAIM),⁹ and still is considered a mathematical vector field rather than a physical property, but according to previous equations a physical interpretation is feasible. In fact, Delle Site suggested a possible extension of Bader's concepts to atoms and molecules in condensed phases using a local current expression such as Eq. 3.7.^{46,47}

A local velocity vector field $\mathbf{v}(\mathbf{r})$ can be defined via the identity $\mathbf{J}(\mathbf{r}) = \rho(\mathbf{r})\mathbf{v}(\mathbf{r})$, which from Eq. 3.7 gives

$$\mathbf{v}(\mathbf{r}) = \frac{1}{m_e}\nabla S(\mathbf{r}) - i\frac{\hbar}{2m_e}\frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})} \quad (3.8)$$

Kan and Griffin consider this velocity field as a mathematical object due to its

non-observability in contrast to the proven observability of the current density.³⁷ On the other hand, according to Bohm's interpretation of quantum mechanics a quantum particle can have a definite velocity (or momentum) in a coordinate representation.⁴³

As demonstrated by Bohm, spherically symmetrical systems (as atoms) have a constant phase factor $S(\mathbf{r})$ which leads to a null value for the classical part of the momentum operator, $\nabla S(\mathbf{r}) = 0$,⁴³ which is consistent with the fact that in the absence of an external field $\bar{\mathbf{J}}(\mathbf{r}) = 0$.⁹ The local representation of the momentum operator given by Eq. 3.6 is therefore reduced to its quantum spread. i.e.,

$$\mathbf{P}(\mathbf{r}) \equiv \tilde{\mathbf{P}}(\mathbf{r}) = -\frac{\hbar}{2} \frac{\nabla \rho(\mathbf{r})}{\rho(\mathbf{r})} \quad (3.9)$$

In Nelson's terminology,³⁴ this imaginary part of the momentum operator is called the *osmotic* (current) *velocity*, and can be expressed alternatively as

$$\tilde{\mathbf{P}}(\mathbf{r}) = -\hbar \nabla \ln \left(\sqrt{\rho(\mathbf{r})} \right) \quad (3.10)$$

This equation shows the conservative (irrotational) character of the local momentum spread, i.e. its curl vanishes, $\nabla \times \tilde{\mathbf{P}}(\mathbf{r}) = 0$.

Real-valued wave functions are commonly used in computational chemistry for the study of electronic states,⁸ and therefore the definition of local momentum as provided by the *quantum spread*, $\tilde{\mathbf{P}}(\mathbf{r})$, can be studied by using numerically-solved electron densities.

The next section deals with the kinetic energy operator as given by the local quantum representation.

3.3.1 Local Kinetic Energy Theorem (LKET)

By using a similar approach to that used in the previous section for the momentum operator, a local representation of the kinetic energy operator, $K = -\frac{\hbar^2}{2m_e} \nabla^2$, is obtained

$$K(\mathbf{r}) = \bar{K}(\mathbf{r}) + i\tilde{K}(\mathbf{r}) \quad (3.11)$$

and the real part the Bohm's quantum potential (BQP)

$$\bar{K}(\mathbf{r}) = -\frac{\hbar^2}{2m_e} \frac{\nabla^2 R(\mathbf{r})}{R(\mathbf{r})} \quad (3.12)$$

and the imaginary term is given by the expression

$$\tilde{K}(\mathbf{r}) = -\frac{\hbar}{2m_e} (\nabla R(\mathbf{r}) + R(\mathbf{r}) \nabla) \cdot \nabla S(\mathbf{r}) \quad (3.13)$$

with $R(\mathbf{r}) = \rho(\mathbf{r})^{1/2}$. As previously stated, the average value of the imaginary part is null, $\langle \tilde{K} \rangle_\psi = 0$, and therefore the average value is given only by the real part, which in terms of the electron density is

$$\bar{K}(\mathbf{r}) = -\frac{\hbar^2}{2m_e} \left(\frac{\nabla^2 \rho}{2\rho} - \frac{\nabla \rho \cdot \nabla \rho}{4\rho^2} \right) \quad (3.14)$$

Taking the expectation value of the kinetic energy $\langle K(\mathbf{r}) \rangle_\psi$, the integral of Eq. 3.14 is

$$\langle K(\mathbf{r}) \rangle_\psi = \frac{\hbar^2}{8m_e} \int \frac{\nabla \rho(\mathbf{r}) \cdot \nabla \rho(\mathbf{r})}{\rho(\mathbf{r})} d\mathbf{r} \quad (3.15)$$

which is the well-known functional, $\tau_w[\rho]$. It can be proved that $\tau_w[\rho]$ is also proportional to the variance of the momentum operator in the local representation, $Var_\psi(\tilde{\mathbf{P}}(\mathbf{r}))$ (by using Eq. 3.4). These results can be stated as the *local kinetic energy theorem*, and can be summarized by the double identity

$$\tau_w[\rho] = \frac{1}{2m_e} Var_\psi(P - \bar{P}) = \langle K(\mathbf{r}) \rangle_\psi \quad (3.16)$$

Provided that the real part of the local representation of the kinetic energy is the only one responsible for the average value, this local kinetic energy can be expressed in terms of the local variables momentum $\tilde{\mathbf{P}}(\mathbf{r})$ and current $\tilde{\mathbf{J}}(\mathbf{r})$, as

$$\bar{K}(\mathbf{r}) = \frac{1}{2m_e} \tilde{\mathbf{P}}(\mathbf{r}) \cdot \tilde{\mathbf{P}}(\mathbf{r}) + \frac{\hbar}{2\rho} \nabla \cdot \tilde{\mathbf{J}}(\mathbf{r})$$

This equation states that the local value of the kinetic energy arises from the square of the local fluctuation of the momentum operator plus the divergence of the fluctuation of the local current density. This is an alternative way to express the BQP in terms of local variables. It is evident from it that the integral of the second term vanishes, because the net flux of the current density is equal to zero, $\int \nabla \cdot \tilde{\mathbf{J}}(\mathbf{r}) dv = 0$. The study of the local representation of kinetic energy functionals in the context of the density functional theory (DFT) has been recently discussed by Hamilton, Mosna and Delle-Site.³³

A local kinetic energy functional in terms of the electron density is not uniquely defined, but is an important issue in density functional theory. Nevertheless, the relevance of the Weizsäcker term is beyond doubt,⁴⁸ being even considered a leading term in any kinetic energy density functional (see e.g. Jaramillo, Scuseria and Ernzerhof,⁴⁹ or Kaupp, Bahmann and Arbuznikov, and references therein).⁵⁰ For a detailed analysis of the relevance of this kinetic energy term in DFT the reader can check the recent papers from García-Aldea and Alvarillos who investigated kinetic density functionals using Weizsäcker-like semilocal and nonlocal terms.^{51,52}

3.4 Density Functionals

It must be noticed that the integral expression 3.4 can be interpreted as the expectation value of the term $\left(\tilde{A}(\mathbf{r})\right)^2$. It can therefore be stated that the variance $Var_\psi(A - \bar{A})$ is the expectation value of the square of the local fluctuation, i.e.

$$Var_\psi(A - \bar{A}) = \left\langle \left(\tilde{A}(\mathbf{r})\right)^2 \right\rangle_\psi \quad (3.17)$$

This statement better explains the relevance of the local representation of quantum observables: the statistical variance of the difference between the observable A and its classical estimate \bar{A} is equivalent to the average of the square of the fluctuation, $\left(\tilde{A}(\mathbf{r})\right)^2$. This quantity can be adequately considered as the the density of the square of the observable $A^2(\mathbf{r}) \equiv \left(\tilde{A}(\mathbf{r})\right)^2$.

Additionally, by considering that the expression 3.17 always involves the probability distribution $P_\psi = |\psi|^2$, and a density $A^2(\mathbf{r})$, the result is a functional of the form:

$$f_A[P_\psi] = \left\langle \left(\tilde{A}(\mathbf{r})\right)^2 \right\rangle_\psi \quad (3.18)$$

Therefore, it is possible to obtain observable densities in a very straightforward way.

3.5 Results

The single-particle density $\rho(r)$ is the main property used in the present study. Hunter studied the use of a single-particle model in atomic and molecular systems.^{25,53,54} The

local momentum (Eq. 3.9) was computed for the first 54 elements of the periodic table at the numerical Hartree-Fock level with the NUMOL package from Becke's group.^{55,56} The following results are based on those computations.

3.5.1 Bohr Atomic Model

In the hydrogen atom the radial distribution function $g(r) = \rho(r)r^2$ exhibits a maximum at the atomic radius a_0 , which is the characteristic length for Bohr's atomic model. After substituting the ground state electron density of the hydrogen atom, $\rho(r) = \frac{1}{\pi a_0^3} e^{-2r/a_0}$, into Eq. 3.9 the atomic unit of momentum is obtained, $p_o = \frac{\hbar}{a_0}$ which yields the atomic unit of velocity $v_0 = \frac{\hbar}{a_0 m_e} = 2.1877 \cdot 10^6 \text{m/s}$. This is exactly the classical value of the velocity for the first *orbit* in Bohr's model of the hydrogen atom. The derivation of this velocity unit in the context of local quantum mechanics is very straightforward, while conventional derivations require additional assumptions. It should be emphasized that the electron velocity mentioned here is purely a quantum-statistical estimate, while in Bohr's model the electron remains in a classical circular trajectory around the nucleus with certain *dynamical* velocity, clearly a very different picture.

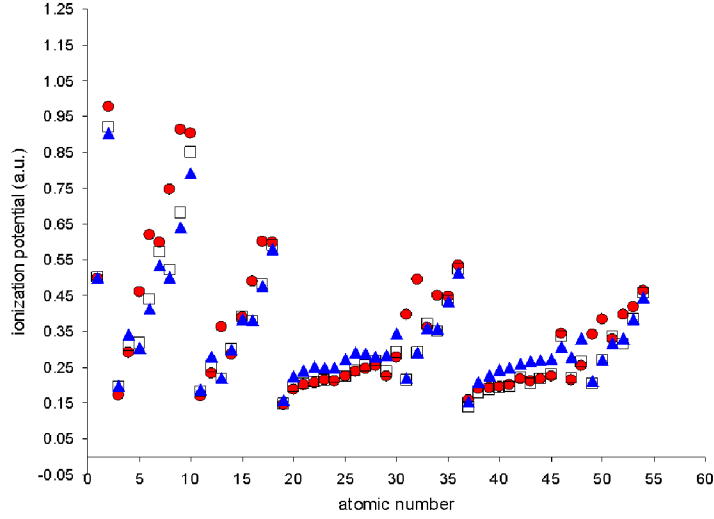
The expectation value of the kinetic energy in the ground state, $\langle K(\mathbf{r}) \rangle_{\psi}$, can be obtained from Eq. 3.15 giving $\langle K(\mathbf{r}) \rangle_{\psi_{1s}} = \frac{p_0^2}{2m_e} = \frac{m_e e^4}{2\hbar^2}$, as expected.

The local velocity for the hydrogen atom is the constant v_0 , which means that the single electron of the hydrogen atom is closer (physically speaking) to the free-particle model than any other atom, in the sense that it has zero electron-electron interaction effects, including correlation and therefore pair-density effects. The radial distribution function and the local momentum for the hydrogen atom are shown in Fig. 3.2 a) . The local velocity in the helium atom drops from $2v_0$ to $1.4v_0$, as shown in Fig. 3.2 b). This is a change in momentum of $\sim 30\%$ that shows the overall effect of mutual repulsion between the two electrons located in the same electron shell. Hence, electron-electron interactions are reflected as an averaged effect in the local momentum.

3.5.2 The Local Momentum at the Origin is Given by Kato's Cusp Condition

As derived from the theorem stated by Kato,²⁹ the value of the electron density at the nucleus, $\rho(0)$, and the average of the derivative of $\rho(\mathbf{r})$ are related according to⁵⁷

Figure 3.1: Ionization potential as derived from local theory (\bullet), from Koopmans' theorem (\square) and experimental (\blacktriangle).



$$\lim_{|\mathbf{r}| \rightarrow 0} [\nabla \rho(\mathbf{r})]_{av} = (-2Z/a_0)\rho(0) \quad (3.19)$$

Therefore, from Eq. 3.9, the local momentum for the atom with atomic number Z , $\tilde{v}_Z(\mathbf{r})$, at the origin has the limit

$$\tilde{p}_Z^o \equiv \lim_{|\mathbf{r}| \rightarrow 0} \tilde{p}_Z(\mathbf{r}) = p_0 Z \quad (3.20)$$

where p_0 is the atomic unit of momentum. Thus Kato's cusp condition can be interpreted as the upper limit value for the electron speed within the atom. The theorem of Kato in atomic units can be simply stated as

$$\tilde{p}_Z^o = Z \quad (3.21)$$

The atomic number Z and the electronic behaviour near the nucleus are connected by the average value of the electron momentum of the core electrons. The corresponding velocity value is used, in fact, for the estimation of the relativistic corrections for atoms beyond the third row, where these corrections become significant.^{58,59}

3.5.3 The Local Momentum for Valence Electrons Gives the Ionization Potential

According to Kohout *et al*, it is also expected that the ratio $\frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})}$ would be asymptotically convergent.⁶⁰ The asymptotic behaviour of the ratio $\frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})}$ was examined in more detail by Nagy and March.⁶¹ Using the Hoffmann-Ostenhof results,³⁰ they showed that this ratio at large r exhibits the limit value of (in SI units)

$$\lim_{|\mathbf{r}| \rightarrow \infty} \left| \frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})} \right| = \frac{2}{a_0} \left(\frac{2I_Z}{e^2/(4\pi\epsilon_0 a_0)} \right)^{1/2} \quad (3.22)$$

where I_Z is the first ionization potential of the atom with atomic number Z . Using this result the local momentum corresponding to valence electrons, \tilde{P}_Z^v , is obtained as

$$\tilde{P}_Z^v \equiv \lim_{|\mathbf{r}| \rightarrow \infty} \left| \tilde{\mathbf{P}}(\mathbf{r}) \right| \quad (3.23)$$

which can be simplified to

$$\tilde{P}_Z^v = p_0 \sqrt{\frac{2I_Z}{E_0}} \quad (3.24)$$

where $E_0 = \hbar^2/m_e a_0^2 = 4.35974 \cdot 10^{-18} \text{J}$ is the Hartree unit of energy. Therefore Eq. 3.24 in atomic units can be reduced to

$$\tilde{P}_Z^v = \sqrt{2I_Z} \quad (3.25)$$

This equation shows the relevance of the asymptotic behaviour of the local momentum: according to Eq. 3.25 the external electrons or valence electrons have a finite momentum \tilde{P}_Z^v given exclusively in terms of the square root of two times the ionization energy. It is also easy to see that the ionization potential is equivalent to the kinetic energy of the valence electrons, from Eq. 3.24

$$K_Z^v = \frac{\left(\tilde{P}_Z^v \right)^2}{2m_e} = I_Z \quad (3.26)$$

Alternatively, this equation can be verified directly from the local expression of the kinetic energy, Eq. 3.14, giving as a resultⁱ

ⁱFrom the results in ref.⁶¹ the expressions for the second derivative limit can also be obtained, which are required for the use of Eq. 3.14.

$$\lim_{|\mathbf{r}| \rightarrow 0} |K_Z(\mathbf{r})| = I_Z \quad (3.27)$$

In summary, from the local theory it is deduced that in order to remove an electron from the outermost atomic shell, the energy transferred to the atom must be *at least* equal to the local kinetic energy of the valence electron, K_Z^v , which is exactly the ionization potential I_Z .

Table 3.1 shows the numerical values for the valence electron kinetic energies as derived from Eq. 3.24 and the ionization potential values from both, theory (from Koopmans' theorem) and experiments for the first 54 elements.⁶² In Hartree-Fock theory, Koopmans' theorem states that the atomic ionization energies are equal to the negatives of the eigenvalues of the Fock operator associated with the occupied molecular orbitals (if orbital relaxation is neglected).

The local kinetic energies derived from the local theory are in good agreement with the ionization values shown in Fig. 3.1, as provided by equation 3.27. The values included in the table are the local kinetic energies as obtained at a distance from the nucleus of ~ 10 a.u..

From these kinetic energies it follows that all the valence electron velocities lies within the $\sim 1-3 \cdot 10^6$ m/s range for the first three rows and within $\sim 1-2 \cdot 10^6$ m/s for the fourth and fifth rows. This narrow range of velocities for the valence electrons has been suggested to be responsible for the persistent presence of a maximum in the stopping cross section for protons of ~ 100 keV/amu for most targets.⁶³ The values reported by Cabrera *et al* for the valence electron velocities contributing to the stopping section span a wider range of $\sim 1-6 \cdot 10^6$ m/s. This wider range can be explained by interactions with electrons having higher kinetic energies, i.e. located inside inner shells.

3.5.4 The Local Momentum in the Middle Range Describes the Shell Structure of Atoms

In the traditional atomic model proposed by Bohr, electrons are located in concentric orbits around the nucleus, labelled K, L, M, etc. The shell structure in atoms is a consequence of the Pauli exclusion principle, and the number of allowed electrons inside each shell is given by the formula $2n^2$ ($n = 1, 2, 3 \dots$).⁶⁷ Nevertheless, the

Table 3.1: Valence local kinetic energies \tilde{K}_Z^v , and the ionization energies from the Koopmans' theorem, I_Z , and experimental ionization energies I_{exp} .⁶²

Atom	\tilde{K}_Z^v	I_Z	I_{exp}	Atom	\tilde{K}_Z^v	I_Z	I_{exp}	Atom	\tilde{K}_Z^v	I_Z	I_{exp}
¹ H	0.499	0.500	0.500	¹⁹ K	0.146	0.148	0.160	³⁷ Rb	0.159	0.138	0.154
² He	0.976	0.918	0.904	²⁰ Ca	0.189	0.195	0.225	³⁸ Sr	0.190	0.178	0.209
³ Li	0.172	0.196	0.198	²¹ Sc	0.203	0.203	0.241	³⁹ Y	0.193	0.188	0.228
⁴ Be	0.291	0.309	0.343	²² Ti	0.208	0.210	0.251	⁴⁰ Zr	0.196	0.192	0.244
⁵ B	0.462	0.319	0.305	²³ V	0.214	0.216	0.248	⁴¹ Nb	0.200	0.197	0.248
⁶ C	0.619	0.439	0.414	²⁴ Cr	0.213	0.222	0.249	⁴² Mo	0.216	0.223	0.261
⁷ N	0.599	0.571	0.534	²⁵ Mn	0.224	0.226	0.273	⁴³ Tc	0.213	0.203	0.268
⁸ O	0.746	0.522	0.500	²⁶ Fe	0.238	0.242	0.290	⁴⁴ Ru	0.217	0.219	0.270
⁹ F	0.913	0.680	0.640	²⁷ Co	0.247	0.256	0.290	⁴⁵ Rh	0.226	0.231	0.274
¹⁰ Ne	0.903	0.850	0.792	²⁸ Ni	0.256	0.268	0.281	⁴⁶ Pd	0.345	0.336	0.306
¹¹ Na	0.169	0.182	0.189	²⁹ Cu	0.227	0.240	0.284	⁴⁷ Ag	0.214	0.221	0.278
¹² Mg	0.234	0.253	0.281	³⁰ Zn	0.278	0.292	0.345	⁴⁸ Cd	0.255	0.265	0.331
¹³ Al	0.363	0.218	0.220	³¹ Ga	0.398	0.216	0.220	⁴⁹ In	0.343	0.205	0.213
¹⁴ Si	0.287	0.301	0.300	³² Ge	0.496	0.291	0.290	⁵⁰ Sn	0.383	0.269	0.270
¹⁵ P	0.390	0.392	0.385	³³ As	0.361	0.370	0.361	⁵¹ Sb	0.327	0.336	0.318
¹⁶ S	0.491	0.380	0.381	³⁴ Se	0.450	0.349	0.358	⁵² Te	0.397	0.315	0.331
¹⁷ Cl	0.602	0.481	0.477	³⁵ Br	0.446	0.434	0.434	⁵³ I	0.420	0.385	0.384
¹⁸ Ar	0.600	0.591	0.579	³⁶ Kr	0.535	0.524	0.514	⁵⁴ Xe	0.463	0.457	0.446

Table 3.2: Shell radii and electron populations at the HF level as given by the inflection points of the local momentum as given by Eq. 3.9 (first row), from the local minima in electron localization function (second row),⁶⁴ and Kohn-Sham kinetic energy density by Navarrete-Lopez *et al.* (third row),⁶⁵ and by Schmider and Becke (fourth row).⁶⁶ Q_o^{10} is the electron density enclosed at a distance of 10a.u.

Atom	r^K	q^{Kii}	r^L	q^L	r^M	q^M	r^N	q^N	r^O	q^O	Q_o^{10}
⁴ Be	0.94	1.99	6.91	2.01							4.00
	1.02	2.00	–	2.00							
	0.93	2.00	–	2.00							
	0.92	1.97	–	2.03							
¹⁰ Ne	0.26	1.96	1.59	7.25							10.00
	0.30	2.20	–	7.80							
	0.24	1.80	–	8.20							
	0.26	1.98	–	8.02							
¹² Mg	0.20	1.94	1.56	8.03	7.71	2.02					12.00
	0.24	2.20	1.69	7.90	–	1.90					
	0.18	1.80	1.54	8.10	–	2.10					
	0.21	1.97	1.55	7.96	–	2.07					
¹⁸ Ar	0.12	1.88	0.69	8.04	3.14	7.90					18.00
	0.14	2.20	0.74	7.90	–	7.90					
	0.11	1.70	0.66	8.00	–	8.30					
	0.13	1.96	0.69	7.88	–	8.17					
²⁰ Ca	0.11	1.89	0.58	8.02	2.34	8.07	9.83	2.02			20.00
	0.13	2.20	0.62	7.90	2.55	8.00	–	1.90			
	0.10	1.70	0.56	8.00	2.33	8.20	–	2.10			
	0.11	1.95	0.58	7.86	2.37	8.11	–	2.07			
³⁶ Kr	0.05	1.85	0.25	8.38	0.96	17.35	8.08	8.44			36.00
	0.07	2.20	0.27	8.50	1.02	17.10	–	8.10			
	0.05	1.60	0.22	7.60	0.91	17.80	–	9.00			
	0.06	1.95	0.25	8.30	0.95	17.15	–	8.60			
³⁸ Sr	0.05	1.89	0.23	8.39	0.84	17.37	2.72	8.43	8.55	1.92	38.00
	0.06	2.20	0.25	8.60	0.89	17.10	2.95	8.30	–	1.80	
	0.05	1.60	0.20	7.60	0.81	17.80	2.70	8.90	–	2.10	
	0.06	1.95	0.23	8.33	0.84	17.08	2.76	8.57	–	2.07	
⁵⁴ Xe	0.03	1.81	0.14	8.60	0.45	17.37	1.28	17.70	4.67	8.47	54.00
	–	–	–	–	–	–	–	–	–	–	
	0.03	1.60	0.13	7.50	0.42	17.80	1.25	18.00	–	9.10	
	0.04	1.94	0.15	8.6	0.44	17.02	1.29	17.69	–	8.75	

electron shells are not uniquely defined. The radial distribution function $g(r) = \rho(r)r^2$, for example, is usually employed for the study of the electron shells. However this definition fails to reproduce the expected number of shells for atoms in the fourth and fifth periods.⁶⁸

In Fig. 3.3 the local momentum and the radial distribution function for Rb and Xe are shown. The radial distribution of Rb in Fig. 3.3a shows only three peaks, while the local momentum exhibits five local maxima, corresponding to the five electronic shells. A similar situation is shown for the Xe atom in Fig. 3.3b, where only four peaks are visible for the radial distribution function, whereas the electron momentum exhibits five maxima. The local momentum provides a complete description of the electron shells as anticipated by Kohout *et al.* in their step-function analysis of the ratio $\frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})}$.⁶⁰ They used relativistic wave functions to compare the shell structures predicted by $\nabla^2\rho$, $\nabla^2\frac{\sqrt{\rho}}{\sqrt{\rho}}$ and the ratio $\frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})}$.⁶⁰ They found no difference between relativistic and non-relativistic results in the definition of shell structures. The Laplacian of the electron density, $\nabla^2\rho$, depicts the shell structure in atoms, but is limited to s-block and most p-block atoms.^{69,70,71} Hunter also studied the one-electron BQP (Eq. 3.14), in atoms and molecules, and mentioned a possible connection between the local momentum and QTAIM.⁵⁴

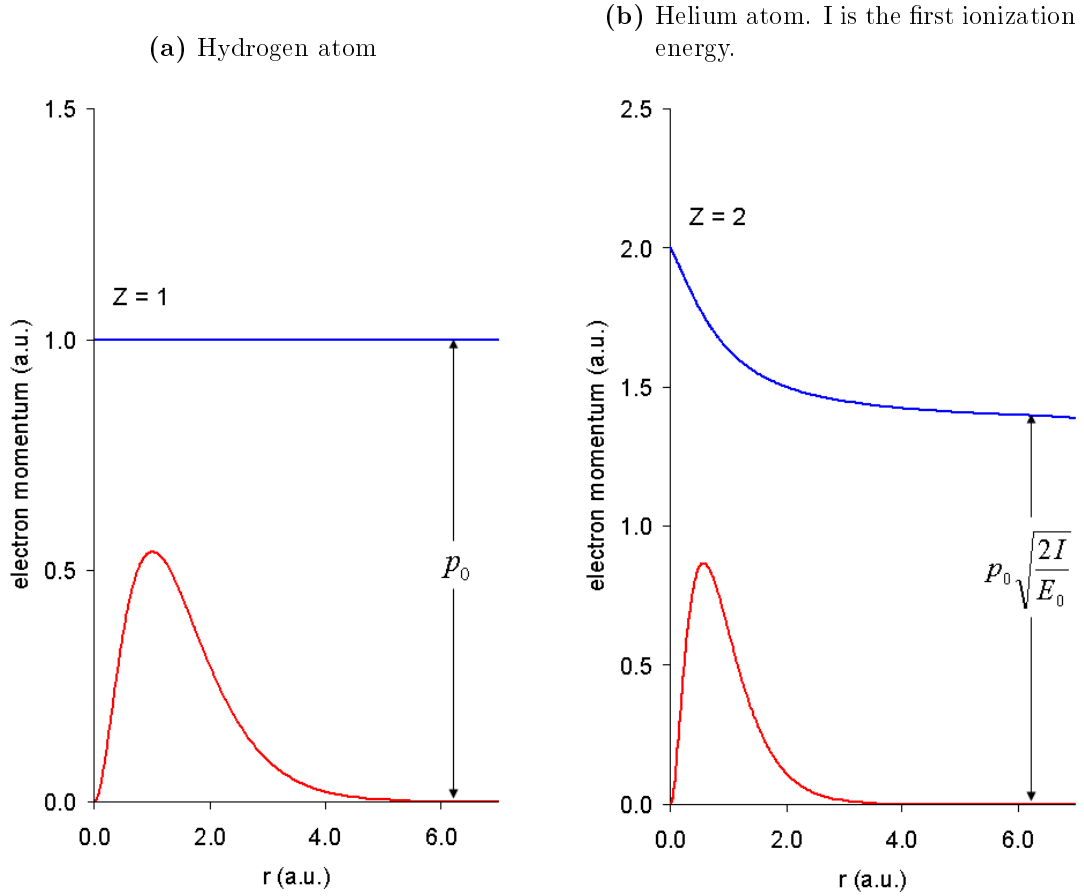
The following definition for shell structure is introduced: the shells are characterized by the local maxima in the momentum (Eq. 3.9), and their boundaries are the *inflection points* located after each local maximum. The inflection points are zeros in the second derivative of the momentum field. This definition uniquely identifies critical points of the local momentum and assigns definite electron populations with spatial regions.

As an illustration, the local momentum for the calcium atom is shown in Fig. 3.4. The electronic populations for the four atomic shells correspond to what is expected from the *Aufbau* principle.

The local momentum correctly provides the atomic shell structure for the atoms studied, i.e. for the closed shell atoms and noble gases. As shown in Table 3.2 the number of electrons per shell and the corresponding shell radius agrees with the expected shell occupancy.

In order to compare present results with other approaches, the table includes shell-structure values reported by Kohout and Savin,⁶⁴ who used the electron localization function (ELF) from Becke and Edgecombe, which is based on the kinetic energy functional $\tau_\sigma = \frac{1}{2} \sum_{i=1}^{n_\sigma} |\nabla\psi_i^\sigma|^2 - \tau_W$.⁴ In the same table are included the recent

Figure 3.2: Radial distribution function (red) and local momentum (blue) for H and He at numerical Hartree-Fock level.

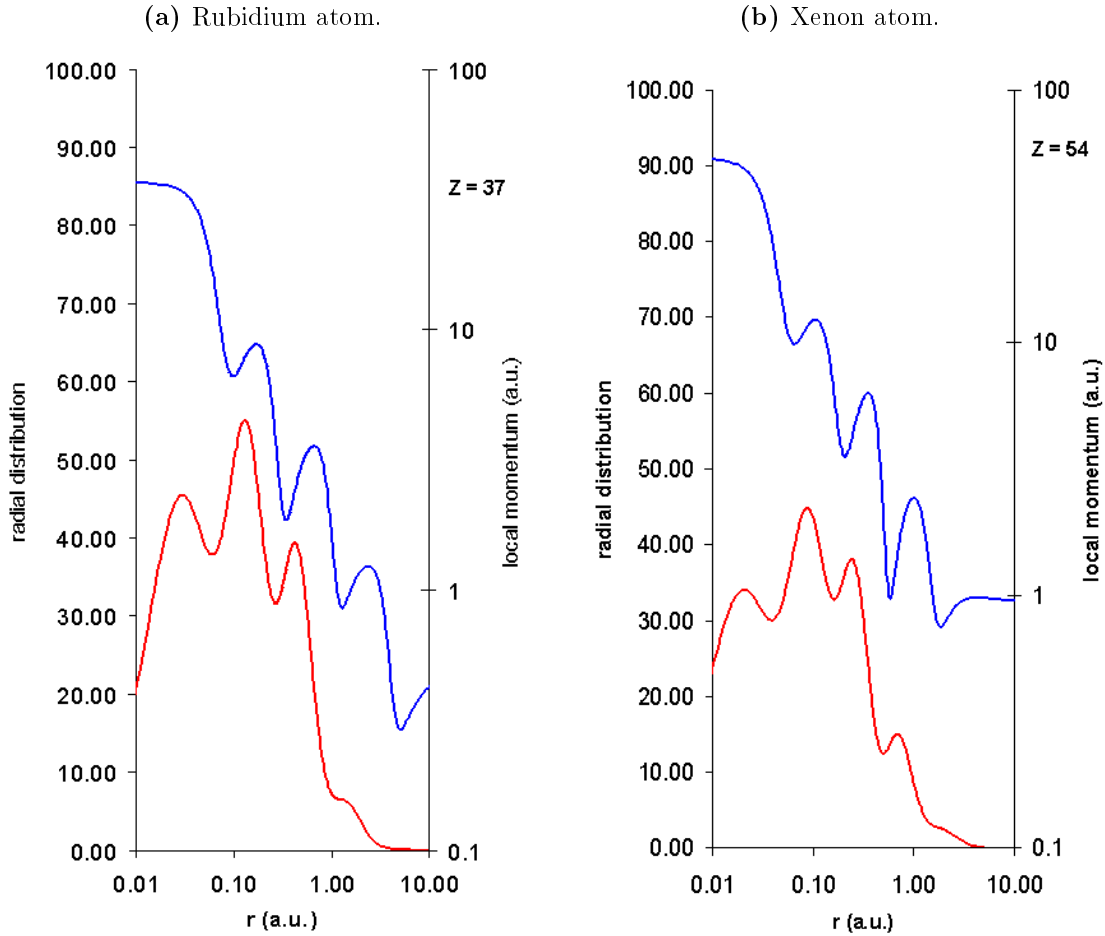


results by Navarrete-Lopez et al., who used the Kohn-Sham kinetic energy density (KSKE) $\tau_{KS} = \frac{1}{2} \sum_{i=1}^{n_\sigma} \nabla \psi_i^{*\sigma} \left(-\frac{1}{2} \nabla^2\right) \psi_i^\sigma$.⁶⁵ This approach is very similar to the shell structures reported by Schmider and Becke, also included in the table.⁶⁶ In these studies the quantities used are Lorentzian forms of the normalized kinetic energies, using as normalization factor the Thomas-Fermi kinetic energy density functional.

ELF and KSKE require a knowledge of orbital-based kinetic energy densities, whereas the present approach does not require the explicit use of orbitals because it is based exclusively on the spherically-averaged electron density and its gradient (Eq. 3.9). This fact makes it possible to compute this quantity at different levels of theory and even from experimentally-obtained electron densities.

The results in Table 3.2 show that the atomic shells' radii (r^K, r^L, r^M, \dots) are virtually the same among the three different approaches. In most cases, the shell radii

Figure 3.3: Local electron momentum (logarithmic scale, blue) versus the radial distribution function (linear scale, red) for the initial and final elements of the fifth row of the periodic table.

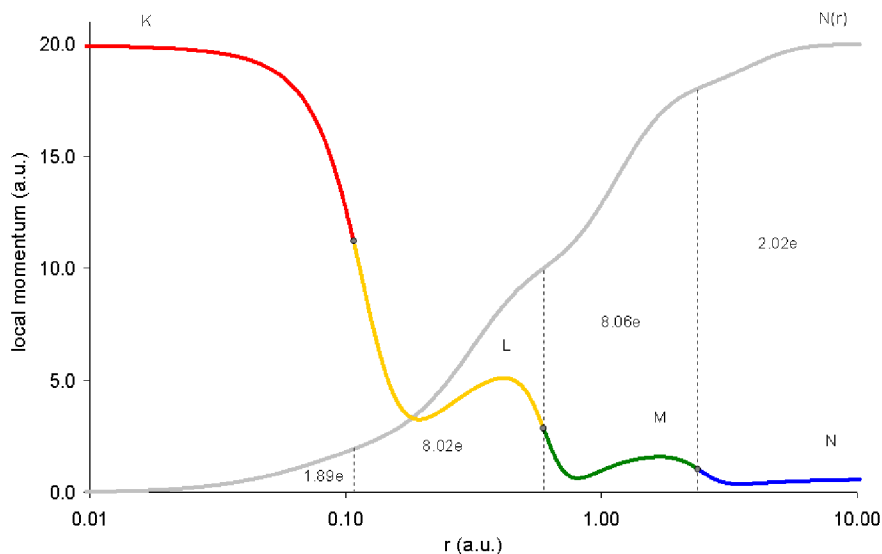


obtained from the local momentum are closer to those from the KSKE by Schmider and Becke than to the ELF ones.

The core shell given by r^K clearly shows the contraction effect with the increasing atomic number. The corresponding populations given by ELF computations overestimate the number of electrons closer to the nucleus, predicting an increment on the $1s$ electron population ($2.2e$ for Xe), while by other hand, KSKE predicts a decrement in the electron population of $1.6e$ for Xe. The local momentum predicts a value between these two of $1.8e$.

The outer shells' radius defined as the inflection point, encloses $\sim 99\%$ of the total electron population for most of the cases. In the other studies it is assumed that these

Figure 3.4: The shell structure of the calcium atom as given by the local momentum. The shells are denoted by capital letters and the corresponding number of electrons are given below the electronic population curve, $N(r)$ (in grey). The inflection points that separates each shell are given by the circles.



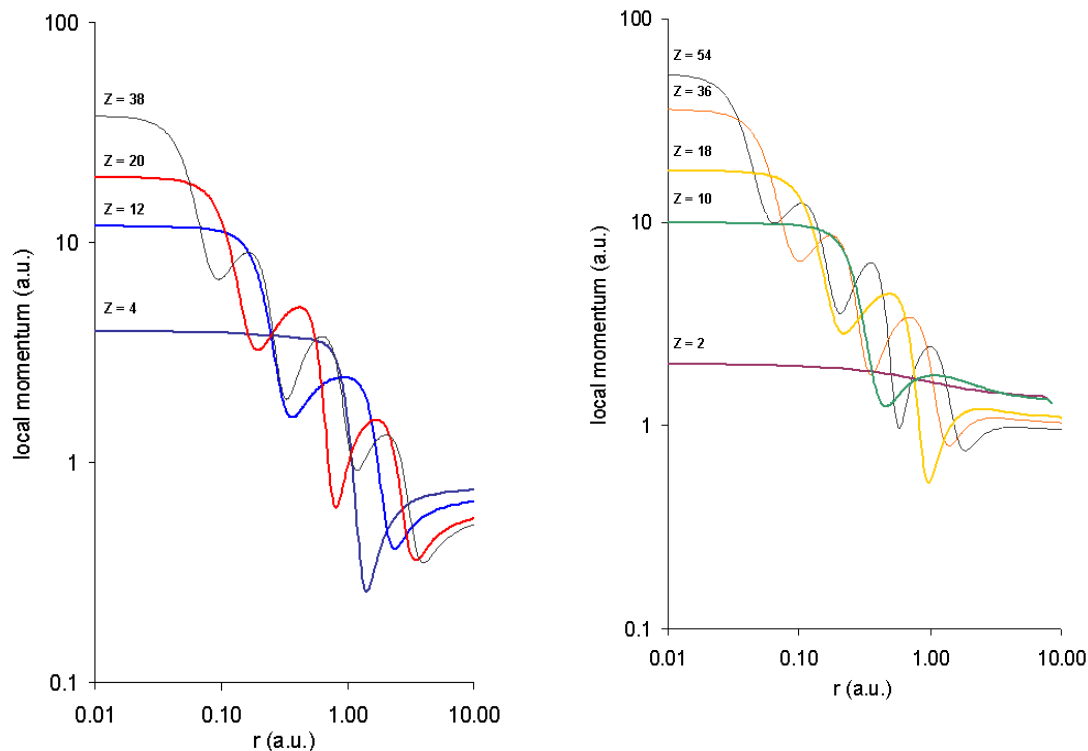
radii are located at the infinity.

While the shell radii obtained from the local momentum lie between the ELF and KSKE values, the KSKE valence shell electron population exhibit major discrepancies. For example, in the N shell of Kr, KSKE gives an electron population of $9e$. Similarly, KSKE yields an electron population of $9.1e$ for the O shell of Xe.

Fig. 3.5 shows the electronic shells as depicted by the local maxima in the local momentum for the closed shell atoms and noble gases. The highest momentum value occurs at the origin and increases with the atomic number as a consequence of Kato's theorem, as discussed previously. Near the nucleus the local momentum changes by almost an order of magnitude, whereas from He to Xe, the values corresponding to the outer shells are located within a narrow range $\sim [0.3, 1.4]$ a.u., as previously discussed. Therefore, the core electrons exhibit greater relativistic effects than those in outer shells (as shown in the next Chapter).⁵⁹

In brief, the local momentum provides a description of atomic shell structure that is consistent with the results obtained with ELF or KSKE. This may suggest that

Figure 3.5: Local electron momentum for closed shells.
 (a) Noble gases. (b) Elements of the IIA group.



using ELF or KSKE adds as much information as the local momentum for describing the shell structure of atoms.

3.6 Conclusions

The local momentum definition given by equation 3.9 yields information about the local kinetic energy of the electrons in atoms. The limits of this property are determined by two paramount theorems in quantum chemistry: the electron density at the origin is given by Kato's cusp condition,²⁹ and at long distances the electron density decays exponentially according to the theorem by Hoffmann-Ostenhof and Hoffmann-Ostenhof.³⁰ The first limit yields information about relativistic effects in atoms, while the second is related to the first ionization energy.

The integral of the local kinetic energy (Eq. 3.14) is the Weizsäcker kinetic energy functional $\tau_w[\rho]$. This functional is non-negative and yields a lower bound to the true kinetic energy.⁷² First introduced as a correction to the Thomas-Fermi kinetic

energy functional (KEF), it has been recognized as being an essential component of any KEF proposal. The local theory makes $\tau_w[\rho]$ to be interpreted as the variance of the momentum operator P , minus its classical estimate \bar{P} as given by Eq. 7.1. In that sense, $\tau_w[\rho]$ is a statistical estimate of the kinetic energy, depending exclusively on the inherent quantum fluctuations of the momentum operator, as given by $\tilde{\mathbf{P}}(\mathbf{r})$.

The local momentum depicts atomic shells better than the radial distribution function or the Laplacian of the electron density. The shell structure of closed shell systems as given by the corresponding radii are in agreement with previous works based on the ELF,⁶⁴ and the Kohn-Sham kinetic energy density.^{65,66} The local momentum has the potential advantage that it doesn't depend on orbital computations, but exclusively on the electron density.

In conclusion, Weizsäcker electron kinetic energy, the ionization potential of an atom, Kato's cusp condition and the shell structure of atoms have been unified in this discussion by means of the local quantum theory. Thus, a single plot of the electron momentum (in atomic units) yields the atomic number of the element, the atomic shell structure and the first ionization energy of that atom (see e.g. Fig. 3.4). Such simplification is only possible thanks to the power of the representation exerted by the local quantum theory.

Chapter 4

Lower Estimate to Relativistic Electron-Mass Corrections

4.1 Introduction

Accurate electronic energy calculations for molecules involving heavy elements require the inclusion of relativistic corrections.⁵⁸ In noble gas chemistry, for instance, compounds containing heavy noble gas atoms should, in principle, be treated by relativistic methods. Recent results show that in order to obtain the correct pattern of the ionization spectrum in XeF₂ relativistic corrections must be included in the calculations.⁷³

The increasing interest in the evaluation of relativistic effects in molecules, motivated by, for example, organometallic chemistry, faces several practical challenges. For instance, the difficulty associated with solving the Dirac equation for large systems dictates the use of approximate models. The idea that only the valence electrons of an atom determine its chemical behaviour is the basis for the effective core potential (ECP) approach. This approach uses a valence-only Hamiltonian and attempts to model the valence properties of atoms and molecules as accurately as the corresponding all-electron results. The main advantage of effective core potentials is the ease with which relativistic effects can be included in the calculations. While this approach is in line with the chemist's view that the valence electrons of an element determine its chemical behaviour, from a quantum mechanical point of view the partitioning of a many-electron system into electronic subsystems is forbidden, since electrons as elementary particles are indistinguishable. However, in the framework of effective one-particle states approximations like Hartree-Fock or Dirac-Hartree-Fock theory, a definition of core and valence orbitals/shells is possible on the basis of either energetic or spatial arguments. If the core shells of a system are determined for the free atoms, and then transferred to the molecule, this is the frozen-core or frozen-orbital approximation. This approach underlies all valence-only schemes. However, the chemist's qualitative view of partitioning core and valence shells is usually not

suitable for quantitative calculations. In fact, the treatment of Ti ([Ar] 3d² 4s²) or Ce ([Xe] 4f¹ 5d¹ 6s²) as four-valence electron systems leads to poor results, whereas it works very well for C ([He] 2s² 2p²). The reason is the presence of partially occupied valence shells which have the same or even lower main quantum number as the fully occupied core shells. An orbital-based separation between core and valence shells may seem reasonable, but it may be incorrect from a spatial point of view: the Ti 3d shell has its maximum density close to the one of the 3s and 3p shells, and the Ce 4f shell has its maximum density even closer to the nucleus than the 5s and 5p shells.⁷⁴ An electron density determination of the the relevant valence electrons required for a given electronic computation might lead to an improvement of the study of relativistic molecules, in particular if their effective core shells can be systematically assessed.

In the present work a new method for the theoretical identification of relativistic spatial extent in atoms is introduced. The single-particle local momentum expression leads to a simple local form of the Lorentz factor that provides the spatial extent of relativistic electronic effects.

4.2 A Local Expression for the Relativistic Corrections

This single-particle local momentum provides the correct limiting values for the electronic velocity within the valence and core regions, as was discussed in the previous chapter.

$$\tilde{\mathbf{P}}(\mathbf{r}) = -\frac{\hbar}{2} \frac{\nabla \rho(\mathbf{r})}{\rho(\mathbf{r})}$$

It depends exclusively on the electron density, and hence it can provide information about the spatial extent of electronic relativistic effects in atoms. The local momentum expression provides a lower estimate of the relativistic corrections by defining $\beta = \tilde{p}/m_e c$, that leads to a local expression for the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - \tilde{p}^2(\mathbf{r})/(m_e^2 c^2)}} \quad (4.1)$$

In turn, a local expression for the relativistic mass correction is obtained

$$m_e^*(\mathbf{r}) = \frac{m_e}{\sqrt{1 - \left(\frac{\hbar}{2cm_e}\right)^2 \left(\frac{\nabla \rho(\mathbf{r})}{\rho(\mathbf{r})}\right)^2}} \quad (4.2)$$

An electron in the valence shell typically exhibits velocities around $10^6 m/s$, which carries little relativistic effects, as the ratio $\beta = v/c$ is below 0.01. At the origin the electron reaches a velocity peak, given by Kato's cusp condition as $\tilde{p}(0)/m_e = v_o Z$, where the atomic unit of velocity $v_0 = 2.187691 \times 10^6 m/s$ is used, and Z is the atomic number.

Hence, after replacing the limiting value of $\tilde{p}(0)/m_e$ into 4.2 a theoretical limit to the relativistic corrections for the electron mass is obtained,

$$m_e^{Z*} = m_e \gamma_Z = \frac{m_e}{\sqrt{1 - v_0^2 Z^2 / c^2}} \quad (4.3)$$

Fig. 4.1 shows the electron-mass correction as given by the latter expression (continuous line) up to $Z = 104$, and the numerical values obtained for the first 54 elements (dots).

Table 4.1: Relativistic correction to the electron mass for the fourth and fifth rows of the periodic table.

	$m_e^{Z*(a)}$	$m_e^{Z*(b)}$		$m_e^{Z*(a)}$	$m_e^{Z*(b)}$
K	1.0097	1.0098	Rb	1.0386	1.0386
Ca	1.0108	1.0108	Sr	1.0408	1.0408
Sc	1.0119	1.0120	Y	1.0431	1.0431
Ti	1.0131	1.0131	Zr	1.0455	1.0455
V	1.0144	1.0144	Nb	1.0480	1.0480
Cr	1.0157	1.0157	Mo	1.0505	1.0506
Mn	1.0171	1.0171	Tc	1.0532	1.0532
Fe	1.0185	1.0185	Ru	1.0559	1.0559
Co	1.0200	1.0200	Rh	1.0587	1.0587
Ni	1.0215	1.0216	Pd	1.0616	1.0616
Cu	1.0232	1.0232	Ag	1.0646	1.0646
Zn	1.0249	1.0249	Cd	1.0676	1.0676
Ga	1.0266	1.0266	In	1.0708	1.0708
Ge	1.0284	1.0284	Sn	1.0740	1.0740
As	1.0303	1.0303	Sb	1.0774	1.0774
Se	1.0323	1.0323	Te	1.0808	1.0808
Br	1.0343	1.0343	I	1.0844	1.0844
Kr	1.0364	1.0364	Xe	1.0880	1.0880

^aValues obtained from the equation 4.3.

^bValues obtained from the local velocity at the nucleus $v_Z(0)$.

In Table 4.1 the numerical values of the relativistic mass correction for the first

54 elements are included.

Another known relativistic effect is the shrinking of the radial distribution function, which, according to Pyykko, is about 23% for Hg.⁵⁸ As can be seen better in Table 4.2, this effect is below 1% for atoms with atomic number lower than 19, i.e. basically the first three rows of the periodic table. For the fourth row this relativistic effect is between 1 – 3.6%, which can be considered as significant especially for the case of the elements within the *d-block* or transition metals. From these results it is clear that the the relativistic correction to the electron mass will be higher than 3% for elements heavier than Ge ($Z = 32$), as is already known.⁵⁸

Figure 4.1: Relativistic electron mass correction.

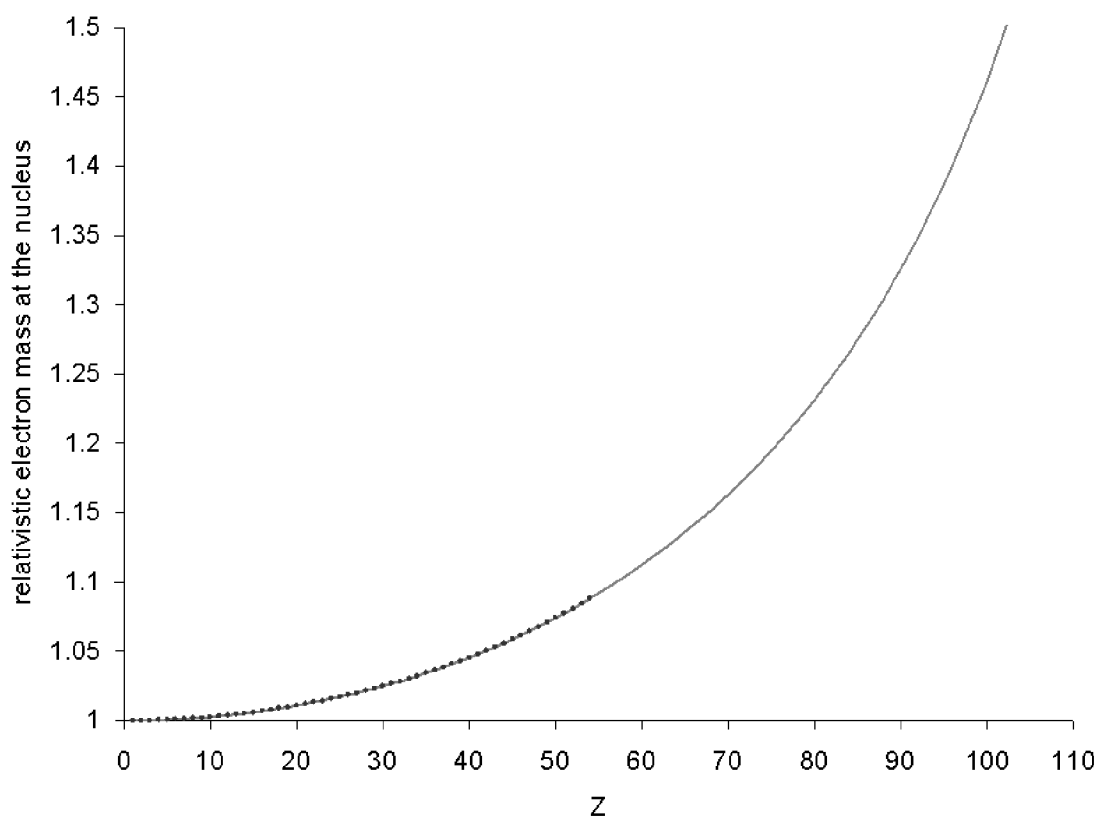
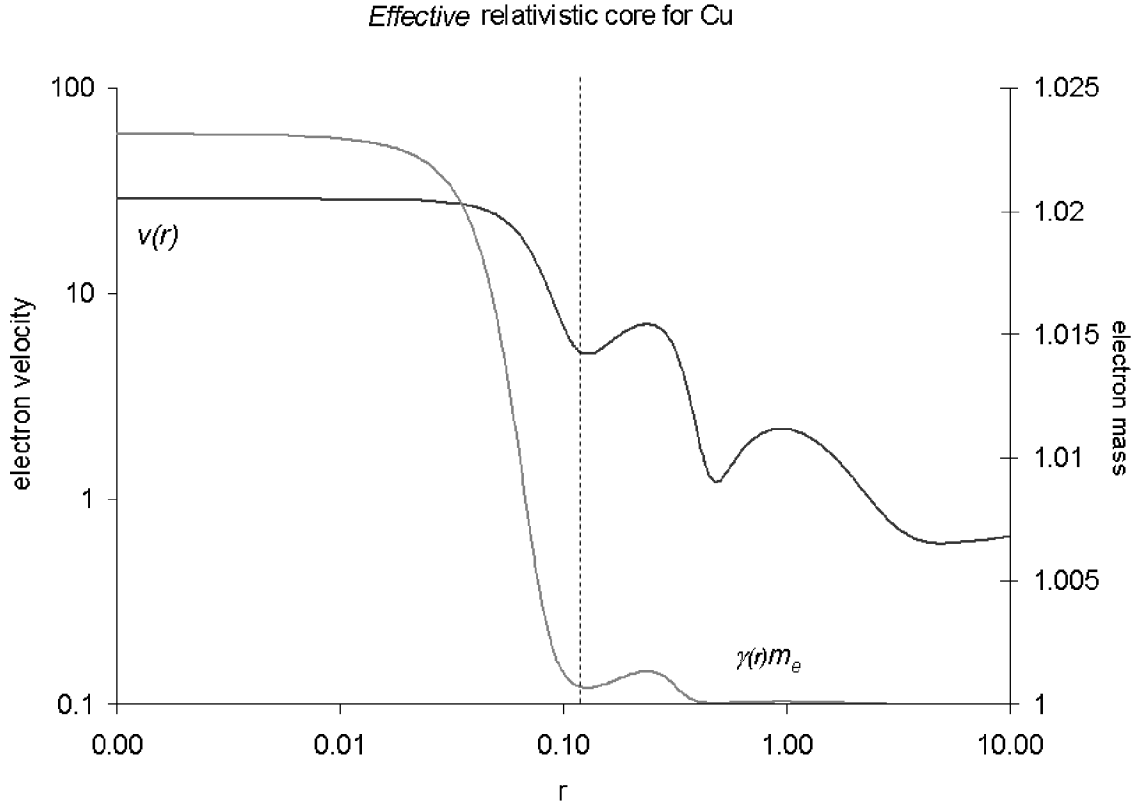


Figure 4.2: Effective relativistic core for Cu.

A spatial identification of core and valence shells may be possible via equation 4.2. As an example, in Fig.4.2 the local velocity and the local relativistic mass, $m_e^*(\mathbf{r})$, are shown for the Cu atom. As can be seen, the inner shell is the one that is more affected by relativistic effects by $\sim 2.3\%$, as indicated by the dashed line. While the second shell is marginally affected (less than 0.15%), the external shells are almost unaffected by this relativistic correction. From Fig. 4.2 an *effective* radius involving the majority of relativistic effects of the atom can be determined. By following the methods developed in the previous chapter, the precise number of electrons included inside the effective core can be calculated.

A more detailed approach to the design of effective cores may determine the validity of the present approach to the spatial determination of relativistic effects in atoms. The local expression of the Lorentz factor is a promising direction for future developments on local theories, such as density functional theory.

Table 4.2: Periodic table with the relativistic correction to the inner shells as obtained from the mass correction term (Eq. 4.3).

		Relativistic contraction of the inner shell (%)																			
		IA			IIA			IIIB			IIB			IVB			VIB			VIIB	
1	1	H	0.00	3	Li	0.02	11	Na	0.32	19	K	0.98	37	Rb	3.86	55	Cs	9.18	2	He	0.01
				4	Be	0.04	12	Mg	0.38	20	Ca	1.08	38	Sr	4.08	56	Ba	9.57	9	F	0.22
2																			10	Ne	0.27
3																			13	Al	0.45
																			14	Si	0.53
																			15	P	0.60
																			16	S	0.69
																			17	Cl	0.78
																			18	Ar	0.87
																			31	Ga	2.66
																			32	Ge	2.84
																			33	As	3.03
																			34	Se	3.23
																			49	In	7.08
																			50	Sn	7.40
																			51	Sb	7.74
																			52	Te	8.08
																			53	I	8.44
																			81	Tl	23.98
																			82	Pb	24.81
																			83	Bi	25.67
																			84	Po	26.57
																			85	At	27.49
																			86	Rn	28.44
																			29	Cu	2.32
																			28	Ni	2.16
																			27	Co	2.00
																			26	Fe	1.85
																			25	Mn	1.71
																			24	Cr	1.57
																			23	V	1.44
																			22	Ti	1.31
																			21	Sc	1.20
																			44	Ru	5.87
																			45	Rh	6.16
																			46	Pd	6.46
																			47	Ag	6.76
																			48	Cd	7.08
																			77	Ir	20.89
																			76	Os	21.62
																			75	Re	22.38
																			74	W	23.17
																			73	Ta	24.81
																			72	Hf	25.67
																			17	Yb	16.32
																			18	Lu	16.92
																			69	Tm	46.26
																			70	Yb	47.96
																			100	Fm	44.63
																			99	Es	43.07
																			98	Cf	41.57
																			97	Bk	40.13
																			96	Cm	38.75
																			63	Eu	12.61
																			64	Gd	13.09
																			65	Tb	13.59
																			66	Dy	14.11
																			67	Ho	14.64
																			68	Er	15.18
																			69	Tm	15.74
																			70	Yb	16.32
																			95	Am	37.43
																			96	Cm	38.75
																			94	Pu	36.15
																			93	Np	34.93
																			92	U	33.75
																			11.23	Nd	11.67
																			10.79	Pr	10.37
																			91	Pa	32.61
																			90	Th	31.51
																			89	Ac	30.45
																			60	Pm	20.18
																			59	Pr	18.81
																			58	Ce	17.53
																			57	La	9.96
6																			61	Pm	20.18
																			62	Sm	20.89
																			63	Eu	21.62
																			64	Gd	22.38
																			65	Tb	23.17
																			66	Dy	23.98
																			67	Ho	24.81
																			68	Er	25.67
																			69	Tm	26.57
																			70	Yb	27.49
																			71	Lu	28.44
																			101	Md	102
																			102	No	103
																			103	Lr	103
																			104	Lr	103

Chapter 5

Is The Size of an Atom Determined by its Ionization Energy?

Reprinted with kind permission from Elsevier Science Ltd.

H. J. Bohórquez and R. J. Boyd, *Chem. Phys. Lett.*, **480**, 127-131 (2009)

5.1 Introduction

In the absence of an external field, the time-averaged electron density of a neutral atom exhibits spherical symmetry. The vast majority of the electron cloud is confined to a region within a few angstroms of the nucleus, which suggests that there should be an effective radius r associated with the atomic size. This radius characterizes the extensive properties of the electron distribution.

Many properties, including atomic polarizability,⁷⁵ electronegativity,⁷⁶ chemical hardness,⁷⁷ diamagnetic susceptibility⁷⁸ and atomic capacitance,⁷⁹ among others, are associated with the atomic size. Whereas the mass of an atom is clearly defined and readily measured, the volume of an atom is not uniquely defined. There is no consensus about how to estimate atomic sizes. Several quantum-based sets of atomic radii, based mainly on wave-function averages rather than in terms of experimentally-accessible observables, have been reported.^{80, 81, 82, 83, 84, 85, 86, 87}

This chapter considers whether or not the atomic radius can be determined in terms of the physically-observable variables of the system. Is derived a theoretical relationship between the atomic radius and the first ionization energy using results from a local quantum theory. An absolute scale for atomic radii emerges that is defined in terms of the intrinsic properties of the isolated atoms. A strong correlation between the resultant set of atomic radii and several other definitions is observed.

5.2 Methodology

The size of an atom is related to the probability of finding an electron at a given distance from the nucleus as measured, e.g., by the radial distribution function, $D(r)$.⁸⁸

The radial distribution function of the ground state of the hydrogen atom provides a standard definition of the atomic length unit, a_0 . A generalization of this approach for obtaining an atomic radius definition valid for any neutral atom is considered in this chapter. The maximum value in the radial distribution function measures the statistical spread in the radial distances from the nucleus. Assuming that the spherically-symmetrical electron distribution for the ground state $\rho(r)$ is known, the radial distribution is $D(r) = 4\pi r^2 \rho(r)$, and its extreme value is given by the condition $D'(r) = 0$, which leads to the equation

$$\frac{1}{r} + \frac{\rho'(r)}{2\rho(r)} = 0 \quad (5.1)$$

The ratio of the density derivative to the electron density that appears in equation 5.1 was studied by Nagy and March as a local wave-number that characterizes the ground state of atoms and molecules.⁶¹ Kohout, Savin and Preuss also studied this quantity, showing that it provides the shell structure of atoms.⁶⁰ It also appeared as a local velocity term in the hydrodynamical approach to local quantum chemistry proposed by Ghosh.^{35,36} Within the context of the *statistics of local values* proposed by Cohen⁴¹ and later developed by Luo,^{28,89} this quantity is the imaginary part of a local representation of the momentum operator,⁹⁰ as independently proposed recently by Hamilton, Mosna and Delle Site.³³ It has been shown previously that the inflexion points of this quantity provides the boundaries for the electron shells, whose electronic populations correctly obey the Aufbau principle, as computed at the Hartree-Fock level.⁹⁰

On the basis of these previous works, the second term in equation 5.1 can be identified as a component of the local representation of the momentum operator,

$$\tilde{p}(r) = -\frac{\hbar \rho'(r)}{2 \rho(r)} \quad (5.2)$$

whose limiting values are precisely known. In fact, at the origin the local momentum is given by Kato's cusp condition,²⁹ which leads to $\tilde{p}_{core} = Z p_0$, where Z is the atomic number and $p_0 = \hbar/a_0$ is the atomic unit of momentum. This local value of the momentum at the nucleus is used for the estimation of relativistic effects in core electrons.⁵⁸ The kinetic energy associated with $\tilde{p}(r)$ is the Weizsäcker local kinetic energy term $K_W(r)$, which is a very well-known lower limit for the electronic local kinetic energy functional, $K(r)$. These two local kinetic energy functionals coincide at the limiting distances from the nucleus: $K(r)|_{r=0,\infty} = K_W(r)|_{r=0,\infty}$. Hence, equation

5.2 provides a correct estimate of the momentum for core ($r \sim 0$) and valence electrons ($r \gg 0$).

As previously shown (Table I in ref.⁹⁰), within the valence region, the local momentum approaches asymptotically to a constant limit, \tilde{p}_{val} , according to which

$$\tilde{p}_{val} = \sqrt{2m_e I} \quad (5.3)$$

where I is the ionization energy and m_e is the electron mass. This equation is supported by the exponential decay of the electron density as established by a theorem due to Hoffmann-Ostenhoff and Hoffmann-Ostenhoff³⁰ and therefore is generally valid for any neutral atom. Substitution of this expression into equation 5.2 yields

$$r\tilde{p}_{val} = \hbar \quad (5.4)$$

This equation assigns to every value \tilde{p}_{val} a corresponding characteristic length r . Hence, the characteristic lengths associated with \tilde{p}_{val} , via equation 5.4, define a set of atomic radii, hereafter noted r_o , given exclusively in terms of the first ionization energies, I , by

$$r_o = a_0 \sqrt{\frac{E_0}{2I}} \quad (5.5)$$

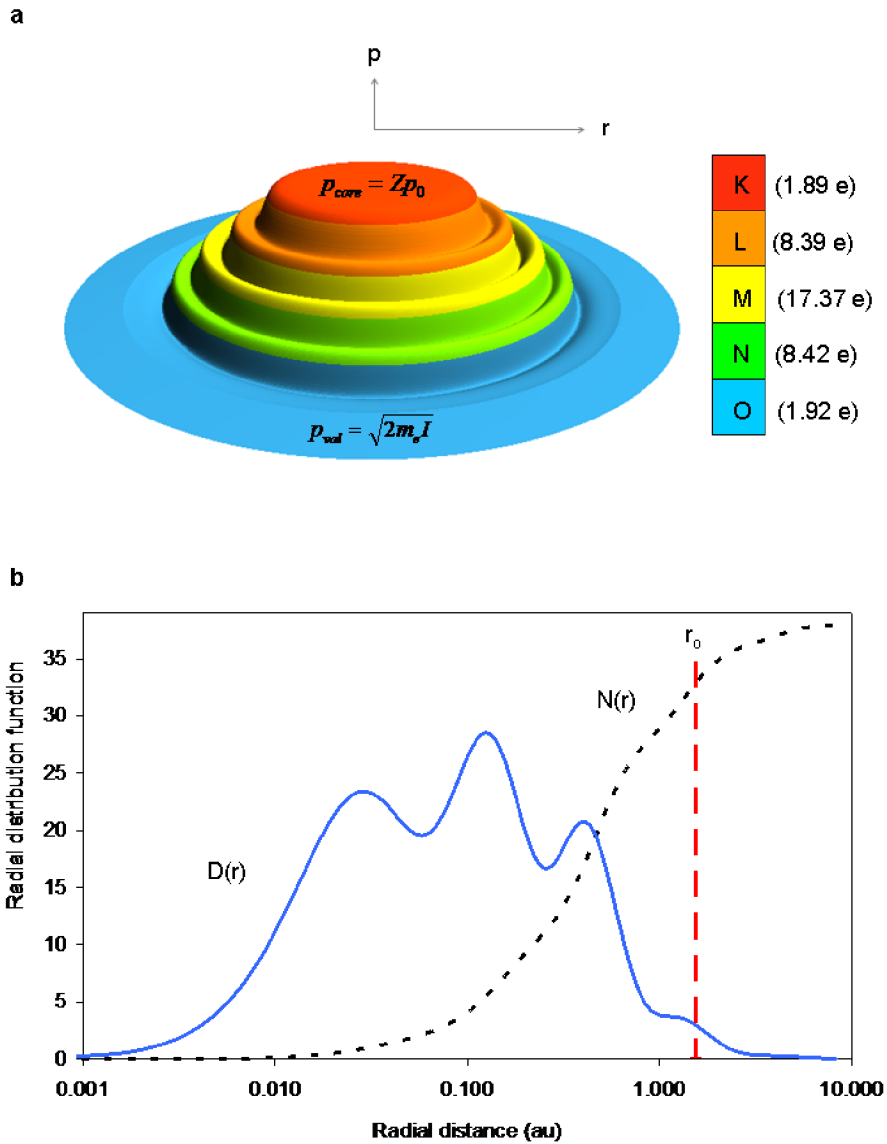
Considering that the atomic unit of energy, E_0 , is twice the absolute value of the ionization energy of the electron in the ground state of the hydrogen atom, I_H , equation 5.5 can be reduced to the simpler expression

$$r_o = a_0 \sqrt{\frac{I_H}{I}} \quad (5.6)$$

Consequently, the radius of the hydrogen atom is the Bohr radius, a_0 , as expected. Alternatively, equation 5.5 can be derived by using the quantization of angular momentum postulated by Bohr in conjunction with the local momentum of the valence electron \tilde{p}_{val} (equation 5.3).

As an example, the local momentum term given by equation 5.2 is shown in Fig. 5.1a for the strontium atom ($Z = 38$). The colours show the five electron shells of this atom, as provided by the inflection points in \tilde{p} , i.e. those points where the second derivative changes sign, $\tilde{p}'' = 0$. For the core electrons, it is found that $\tilde{p}_{core} = 37.995p_0$, in accordance with Kato's cusp condition. Fig. 5.1b shows the

Figure 5.1: Strontium (1S_0) local momentum and radial distribution function. a) Local momentum term (Eq. 5.2) for Sr, at the numerical Hartree-Fock level of theory (log-log plot), with $r \in [7.4 \times 10^{-4}, 16.74]$ au. The atomic shells are given by the inflection points of $\tilde{p}(r)$; the colour code is accompanied by the electronic population obtained for each shell. $Z = 38$ and $I = 5.695$ eV for strontium.⁹¹ b) Radial distribution function of the strontium (solid line) and cumulative electron population $N(r)$ (dotted line), with the radial distance in logarithmic scale, as obtained at the same level of theory. Within the radial distance $r_o = 1.545$ au (purple vertical dashed line), about 87% of the electronic population is enclosed ($N(r_o) \simeq 33e$).



radial distribution function $D(r)$, the cumulative electronic population $N(r)$ and the location of the experimental atomic radius r_o (equation 5.6) for the same atom. The computations were done at the numerical Hartree Fock level with the NUMOL package from Becke and Dickson.⁵⁶

Table 5.1: Linear fit ($r_{theor} = \alpha r_o + r_b$) between five sets of theoretical radii and the experimentally-based atomic radius, $r_o = a_0\sqrt{E_0/2I}$. The radii are given in angstroms (\AA) and R denotes the correlation coefficient.

Method	α	$r_b(\text{\AA})$	R
Froese ⁸¹	2.2173	-1.4182	0.958
Clementi et al. ⁸³	2.0445	-1.4007	0.902
Waber and Cramer ⁸²	1.7134	- 1.1543	0.948 ^a
Zhang et al. ⁸⁵	2.4354	- 1.1854	0.963
Boyd ⁸⁴	2.2528	0.1542	0.924

^aFor this linear fit the value corresponding to Pd was excluded.

5.3 Results and Discussion

In order to compare the atomic radii given by equation 5.6 with other sets of atomic radii, two categories of non-covalent quantum-based radii were selected: a) those based on the maximum electron density in the outermost orbital (Froese,⁸¹ Clementi et al.⁸³ and Waber & Cramer⁸²), and b) those based on the total electron density (Zhang et al.⁸⁵ and Boyd⁸⁴). The experimental ionization energies I for computing the atomic radii r_o were taken from the NIST database.⁹¹

It is apparent from Fig. 5.3 that all data sets show similar periodic trends. In order to facilitate a comparison of the data sets, all sets were scaled to map the range $[0, 1]$. Thus, the respective radii were transformed according to the linear relation $r_{std} = (r - r_{He}) / (r_{Rb} - r_{He})$, where r_{He} and r_{Rb} are the minimum and maximum radii, respectively, for $Z \leq 54$. Interestingly, the radii provided by equation 5.6 are closer to the values of Zhang et al. (open circles), which involve indirectly the ionization energy in its computation. This similarity is confirmed by the higher correlation coefficient listed in Table 5.1. All sets of atomic radii exhibit a correlation $R > 0.9$ with respect to r_o . This means that these atomic radii can be approximated by the linear fits shown in Table 5.1.

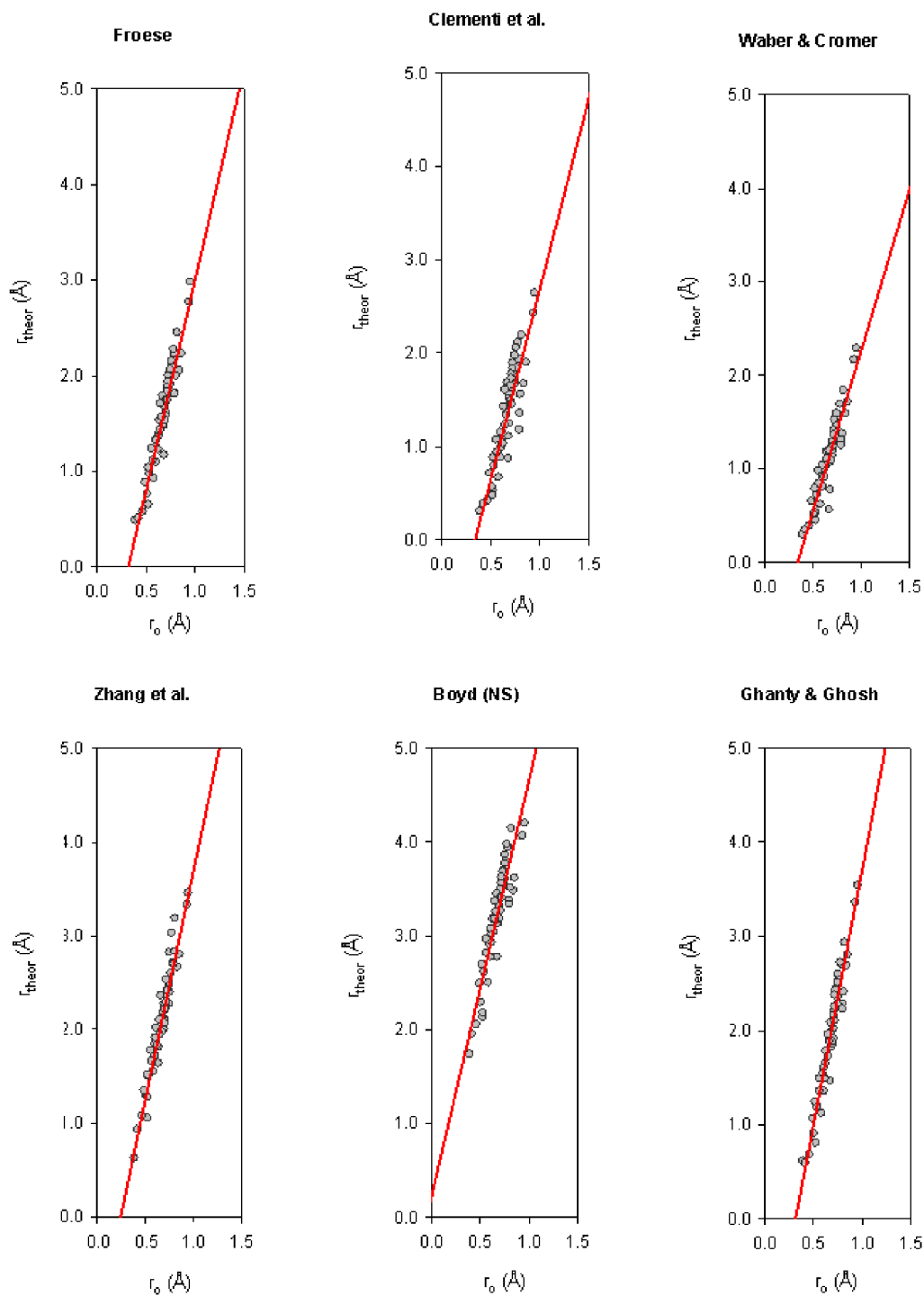
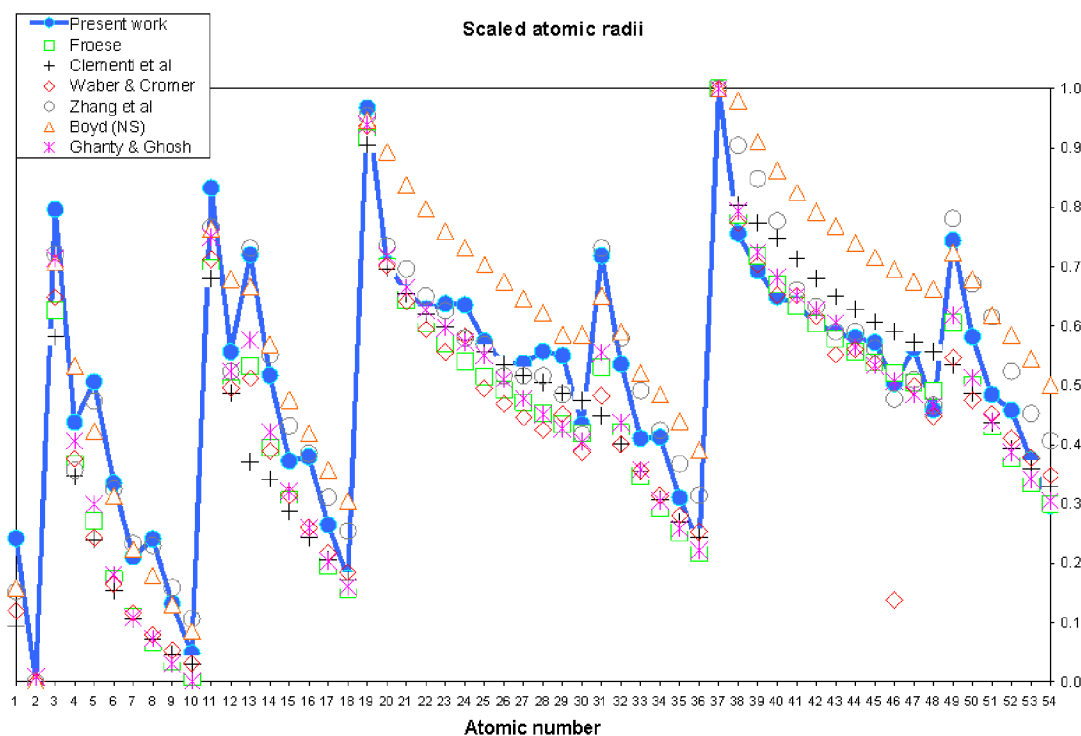


Figure 5.2: Plots of five sets of theoretical radii versus the experimentally-based radii proposed herein, r_o . From left to right: Froese,⁸¹ Clementi et al.,⁸³ Waber and Cramer,⁸² Zhang et al.⁸⁵ and Boyd (non-scaled).⁸⁴ The corresponding linear fits are plotted as red lines; their respective parameters appear in Table 5.1.

The five sets of theoretical radii are plotted against experimentally-based radii in Fig. 5.2. The least-squares linear fits (see Table 5.1) are shown in each case. The Pd radius reported by Waber and Cramer⁸² is clearly an outlier, and is omitted from the fit. The experimentally-based atomic radii as obtained from the equation 5.6 are shown in Table 5.3.

In addition to the strong correlations noted above, it should be emphasized that the method introduced herein is universally valid and can be applied to both experimental and theoretical ionization energies.

Figure 5.3: Scaled atomic radii versus the atomic number for the first five rows of the periodic table. Five different theoretical definitions are plotted (open symbols): Froese⁸¹ (squares), Clementi et al.⁸³ (crosses), Waber and Cramer⁸² (diamonds), Zhang et al.⁸⁵ (circles), Boyd⁸⁴ (triangles), and the present work (solid circles connected by lines).



5.4 Conclusions

Early reports of the relationship between the ionization energy I and the atomic radius r are based on Bohr's model of the hydrogen atom,^{92,93} which is often invoked to explain experimental observations. In these pre-quantum approaches the ionization energy is considered to be proportional to the electron-nucleus point charge interaction and therefore $I \propto 1/r$. The validity of this model is limited to single-electron systems, and its theoretical extension to other neutral atoms is not clearly justified. Nevertheless, this classical approach underlies subsequent atomic radius-ionization energy studies.^{94,95,96} Application of the local quantum theory to atomic systems, however, supports the concise relationship $I/I_H = a_0^2/r_o^2$, as given by equation 5.5.

The present results support the utility of local quantum theory for the study of atomic systems because the momentum of the valence electrons given by equation 5.3 is derived from the *statistics of local values*^{41,28} applied to atomic systems.⁹⁰ For the derivation of the atomic radius introduced here, no explicit assumptions about trajectories or orbital-like behaviour of the valence electrons are required. In this sense, the radius r_o provided by equation 5.5 is a characteristic length associated with the ground-state electron density of an atom. Given that it was derived from a maximal condition on the radial distribution, it is an optimal statistical estimate of the atomic extensive properties, and hence a universally valid definition of atomic radius.

Table 5.3: Experimentally-based atomic radius, $r_o(\text{au})$, for the first 104 elements as obtained from the ionization energies reported at NIST.⁹¹

H	1.00																	He	0.74																
Li	1.59	Be	1.21																	F	0.88	Ne	0.79												
Na	1.63	Mg	1.33																	S	1.02	Ar	0.93												
K	1.77	Ca	1.49	Sc	1.44	Ti	1.41	V	1.42	Cr	1.42	Mn	1.35	Fe	1.31	Co	1.31	Ni	1.33	Cu	1.33	Zn	1.20	Ga	1.51	Ge	1.31	As	1.18	Se	1.18	Br	1.07	Kr	0.99
Rb	1.80	Sr	1.55	Y	1.48	Zr	1.43	Nb	1.42	Mo	1.38	Tc	1.37	Ru	1.36	Rh	1.35	Pd	1.28	Ag	1.34	Cd	1.23	In	1.53	Sn	1.36	Sb	1.26	Te	1.23	I	1.14	Xe	1.06
Cs	1.87	Ba	1.62	La	1.56	Hf	1.41	Ta	1.34	W	1.31	Re	1.32	Os	1.27	Ir	1.23	Pt	1.23	Au	1.21	Hg	1.14	Tl	1.49	Pb	1.35	Bi	1.37	Po	1.27	At	1.21	Rn	1.12
Fr	1.83	Ra	1.61	Ac	1.62	Rf	1.51																												
Ce	1.57	Pr	1.58	Nd	1.57	Pm	1.56	Sm	1.55	Eu	1.55	Gd	1.49	Tb	1.52	Dy	1.51	Ho	1.50	Er	1.49	Tm	1.48	Yb	1.47	Lu	1.58								
Th	1.47	Pa	1.52	U	1.48	Np	1.47	Pu	1.50	Am	1.51	Cm	1.51	Bk	1.48	Cf	1.47	Es	1.46	Fm	1.45	Md	1.44	No	1.43	Lr	1.67								

Chapter 6

A Localized-Electrons Detector for Atomic and Molecular Systems

Reprinted with kind permission from Springer Science and Business Media.

H. J. Bohórquez and R. J. Boyd, *Theor. Chem. Acc.*, **127** (4) 393-400 (2010)

6.1 Introduction

"Sometimes it seems to me that a bond between two atoms has become so real, so tangible, so friendly, that I can almost see it. Then I awake with a little shock, for a chemical bond is not a real thing. It does not exist. No one has ever seen one. No one ever can. It is a figment of our own imagination." Charles A. Coulson

Electronic bonding interactions are not directly observable, as Coulson asserts, but our intuitive perception of molecular phenomena in the three-dimensional space demands such a representation. With a similar mind Lewis conceived the idea of electron pairs.⁹⁷ It is reasonable to think that an adequate representation of chemical bonding should be given by a physical observable defined in coordinate space. The electron density is the best choice because it is a local function defined within the exact many-body theory, and it is also an experimentally-accessible scalar field. Its paramount role in the description of many-body problems is supported by the Hohenberg-Kohn theorem.¹

Although the Hohenberg-Kohn theorem guarantees that all the molecular information is encoded in the electron density, the physical description of chemical systems requires additional postulates for extracting observable information in terms of atomic contributions. This is achieved by the quantum theory of atoms in molecules (QTAIM) introduced by Bader.⁹ The *proper open system* concept provides a quantum

topological partitioning of the molecular space into chemically-transferable molecular fragments for which the energy and all other measurable properties can be precisely defined.¹⁰

The localization of electron pairs is elusive within the electron density topology analysis, because a direct link between the local maxima in the electron density and the electron pairs of the Lewis model has not been established, despite the fact that the Laplacian of the electron density provides some information about electron localization.²⁴ Several attempts to depict electron localization from different perspectives have been proposed in recent years. The variety of proposals for assessing that single task leads to the logical question of why the conventional analysis appears to be insufficient to fully explain electron localization in molecules.

Two main causes are identified in this chapter: one is interpretation, and the other is of a practical nature. The conventional analyses consist mainly of the study of the Laplacian of the electron density,⁷⁰ and the electron localization function (ELF).⁴ Both approaches entail conceptual and practical limitations. The Laplacian provides information about local concentration or depletion of electron density, but its values are not bounded, and it fails to correctly produce atomic shells for atoms beyond the third row.⁶⁹ A direct link between ELF and QTAIM is still missing, whereas an homotopic relationshipⁱ with the Laplacian has been suggested.⁷⁰

ELF, by construction, provides values within the $[0, 1]$ range, and its topological analysis by Savin and Silvi made ELF a preferred tool for the study of electronic bonding interactions.⁹⁸ In spite of its formally-sound derivation from the electronic pair probability, ELF interpretation is not straightforward and the respective plots are far from intuitively evident, a feature to be expected of an ideal representation of chemical bonding interactions. Additionally, ELF fails to provide insight for non-covalent bonding interactions, limiting even more its application for the study of unconventional bonding situations and weak intermolecular interactions.

In the present chapter a variable is introduced that has the ability to detect electron pairs inside an electron density distribution, and overcomes the limitations of the other analyses. The localized-electrons detector (LED) depends exclusively on the electron density and its gradient. Here it is shown how this variable consistently fits within the conventional atoms in molecules analysis, by providing complementary information about the physics of bonding interactions and their local symmetries.

ⁱTwo mathematical objects are said to be homotopic if one can be continuously deformed into the other.

The present work extends to molecular systems a previous investigation of this local variable in atomic systems. It has been shown that LED correctly provides atomic shell structures,⁹⁰ and an atomic radius scale that can be experimentally derived,⁹⁹ among other interesting results. Herein, the application of LED to molecular systems is illustrated.

In the following section we examine the variables involved in ELF are examined in order to show that its key ingredient is the single-particle kinetic energy density, which is connected to theLED by a quantum theorem (see Sect. III A in Ref.⁹⁰). Several important features of this variable, including its bounded character, and its direct connection to QTAIM are discussed in the second part of section II. Section III presents the results of the proposed analysis with several examples, showing how LED identifies the presence of electron shells, and the different bonding interactions adequately. This graphical representation of bonding regions in coordinate space provides distinctive graphical representations for covalent and hydrogen bonding interactions as well.

In summary, LED provides an orbital-free and intuitively interpretable three-dimensional electron-pair localization scalar function that is easy to compute from either theoretically-computed or experimentally-derived electron densities.

6.2 Theoretical Considerations

In this section it is shown that the single-particle kinetic energy local momentum $\tilde{\mathbf{P}}$ is linked to one of ELF's key components, as an alternative to its derivation introduced in Ref.⁹⁰ from the local quantum theory.²⁸

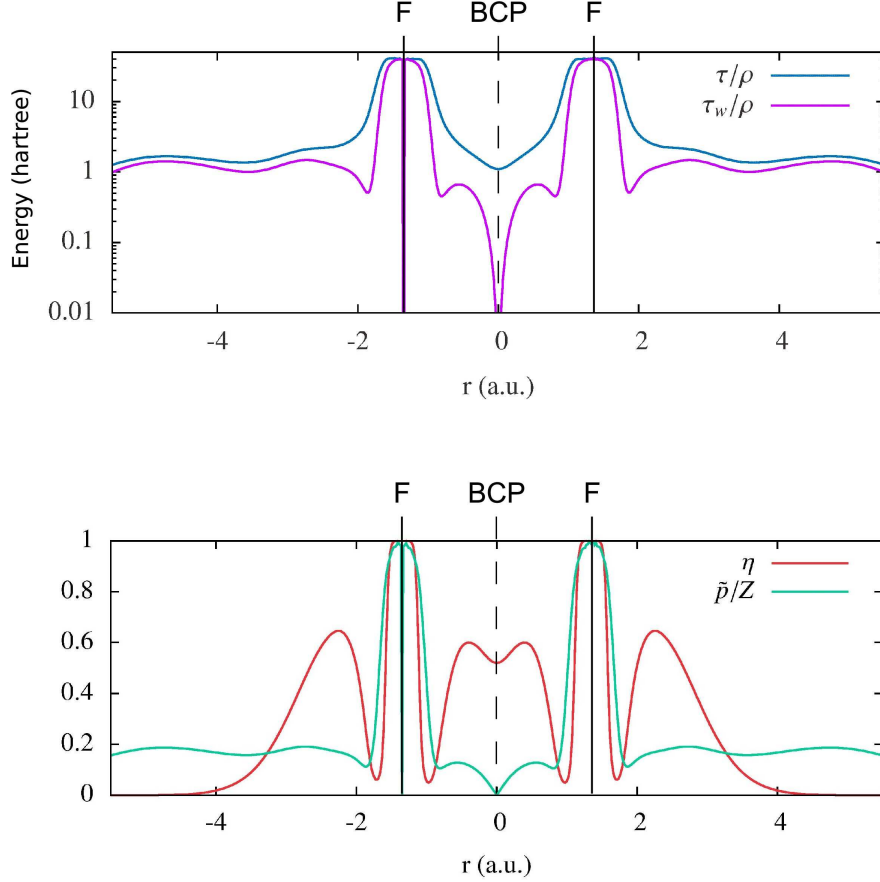
The electron localization function was introduced by Becke and Edgecombe as a "simple measure of electron localization in atomic and molecular systems".⁴ Two terms of a Taylor series expansion of the spherically-averaged conditional same-spin pair probability density provide the main ELF equation, i.e. the kinetic energy density variable

$$\mathcal{D}(\mathbf{r}) = \tau(\mathbf{r}) - \tau_w(\mathbf{r}) \quad (6.1)$$

where the first term is the orbital or positive-definite kinetic energy density

$$\tau(\mathbf{r}) = \frac{\hbar^2}{2m_e} \sum_i |\nabla\psi_i(\mathbf{r})|^2 \quad (6.2)$$

Figure 6.1: Variables involved in ELF for the ground state of F_2 . Top: local values τ/ρ and τ_w/ρ in coordinate space. Bottom: ELF η and LED in coordinate space. The figures are centred at the BCP position.



and the second term is the single particle or Weizsäcker kinetic energy density

$$\tau_w(\mathbf{r}) = \frac{\hbar^2}{8m_e} \frac{\nabla\rho(\mathbf{r}) \cdot \nabla\rho(\mathbf{r})}{\rho(\mathbf{r})} \quad (6.3)$$

\mathcal{D} in Eq. 6.1 is proportional to the Fermi hole mobility function of Luken and Culberston¹⁰⁰ and is related to the curvature of the Fermi hole as shown by Dobson.¹⁰¹ Becke and Edgecombe identify the localization of an electron with the probability density to find an electron in the vicinity of a second same-spin reference electron. The smaller the probability density \mathcal{D} , the higher the localization of the electrons. In order to get values in the range from 0 to 1, ELF is defined as the Lorentzian mapping (\mathcal{L}) of the core variable χ , i.e.

$$\eta(\mathbf{r}) \doteq \mathcal{L}(\chi(\mathbf{r})) = \frac{1}{[1 + \chi^2(\mathbf{r})]} \quad (6.4)$$

with $\chi(\mathbf{r}) = \mathcal{D}(\mathbf{r})/\mathcal{D}_h(\mathbf{r})$, where

$$\mathcal{D}_h(\mathbf{r}) = \frac{3\hbar^2}{5m_e}(6\pi^2)^{2/3}\rho(\mathbf{r})^{5/3} \quad (6.5)$$

is the kinetic energy density of the free electron gas associated with the electron density ρ .

Becke and Edgecombe interpreted the ratio χ as a convenient dimensionless localization index calibrated with respect to the uniform electron gas.ⁱⁱ In this sense, ELF is a local measure of the effect of the Pauli exclusion principle as reflected by the kinetic energy density: in the regions of space where this effect is smaller than the kinetic energy density of a uniform electron gas of identical density, ELF is close to 1, whereas where the local parallel spin pairing is higher, ELF is low.

But the only measure of the electron localization present in ELF is the expression \mathcal{D} , as was recently pointed out by Gatti.¹⁰² However, he adds, ELF cannot yield the value of \mathcal{D} because its dependence on the electron density via the free electron gas term \mathcal{D}_h . In this sense, ELF is a *relative* measure of the electron localization. It is, in fact, a relative measure of the *bosonic character* of the electron density, because \mathcal{D} is the excess kinetic energy electrons have compared to a system of bosons of the same density. This interpretation of ELF was introduced by Savin and Silvi,¹⁰³ who also introduced a generalized kinetic density version of ELF and its topological analysis.⁹⁸

In order to illustrate these observations, the variables involved in ELF are plotted in Fig. 6.1 for the bonding ($\mathbf{r} \in [\mathbf{r}_{BCP}, \mathbf{R}_F]$) and non bonding regions ($\mathbf{r} > \mathbf{R}_F$) of fluorine molecule, where $\mathbf{r}_{BCP} = 0$ is the position of the bond critical point, and the fluorine nuclei are located at a distance $|\mathbf{R}_F - \mathbf{r}_{BCP}| = 1.35$ a.u. from the BCP. In Fig. 6.1 these two points are indicated at the top. Fig. 6.1 a) shows the *local values*ⁱⁱⁱ of the variables involved in \mathcal{D} , τ/ρ (in blue) and τ_w/ρ (in purple). Fig. 6.1 b) shows

ⁱⁱThis is an arbitrarily-chosen reference variable, found also in the Fermi hole mobility function by Luken and Culberson,¹⁰⁰ but instead of a division they made a subtraction, also arbitrarily choosing the uniform electron gas as a reference. While this choice seems physically sound, valence electrons behave as a non-interacting electron gas in crystals mainly and therefore its inclusion in other molecular systems is not entirely justified.

ⁱⁱⁱThese are bounded variables unlike their respective kinetic densities, τ and τ_w , and have units of energy, not energy per volume. For more details on the local values in quantum chemistry, see Ref.⁹⁰ and references therein.

the LED variable (in green), $\tilde{P} = -\hbar |\nabla\rho|/2\rho$ (normalized by Z), and the ELF, η (in red).

While the two electron shells of fluorine are clearly visible with any of these variables, \tilde{P} (and hence τ_w/ρ) also shows the location of the bond critical points, which are relevant for the study of the topological properties in molecules; \tilde{P} reveals the polarization of non bonding regions, which are relevant for intermolecular interactions studies.

Electron pairs are stable groups of electrons that unlike free electrons, have integer spin and therefore can display *bosonic* behaviour. Atomic shells and covalent bonding interactions fit within that description. Recently, the bosonic character of electron pairs was experimentally confirmed.¹⁰⁴ A localized pair behaves as a single particle, and therefore its kinetic energy is given by the Weizsäcker kinetic energy, τ_w . Therefore, its local value (τ_w/ρ) detects those regions of space where the molecule exhibits a marked single-particle character, providing in this way a direct measure of electron pairing. Consequently, this functional is able to identify electron shells and covalent bonding interactions. Hence, this chapter considers the bosonic character of atomic and molecular electron densities as a direct measure of their localized pairs is proposed.

A closer examination of τ_w/ρ depicted in Fig. 6.1 for difluorine reveals the existence of four different regions that can be identified going from the BCP to a long distance from the nucleus (~ 10 a.u.), along the molecular axis:

- The interatomic bonding region, characterized by a continuous approach of $\tau_w/\rho \rightarrow 0$ as $\mathbf{r} \rightarrow \mathbf{r}_{bcp}$, and $\tau \gg \tau_w$. A local maximum value in τ_w/ρ is located between the bond critical point and the core region.
- The core electron region is characterized by a continuous approach of τ_w/ρ to its absolute maximum, $\tau_w/\rho \rightarrow Z^2/2$ (in a.u.) as $\mathbf{r} \rightarrow \mathbf{R}_A$ ($\mathbf{r} \neq \mathbf{R}_A$).^{iv}
- A non-bonding region, characterized by a local maximum in τ_w/ρ for which the outer zone of the valence shell is polarized and where $\tau \approx \tau_w$.
- A molecular boundary region that is characterized by the asymptotic limit $\tau_w \rightarrow \tau$, i.e. τ_w/ρ has a limiting value that depends on the molecular ionization

^{iv}Notice that at the nucleus there is a critical point of the electron density and hence $\kappa(\mathbf{R}_A) = 0$, but the neighbour points obey the limit.

energy.^v

These four distinct regions can be identified in any pair of atoms connected by a gradient path within a molecule. The present discussion is focused mainly on the first three, because their direct relation with the presence of bonding interactions. Although reference is made to the kinetic energy τ , no explicit knowledge of this property is necessary for the bonding analysis provided by the single particle local value τ_w/ρ .

6.2.1 The Single Particle Kinetic Energy Density and QTAIM

In a previous study of the single particle operators in atomic systems,⁹⁰ it was shown that the expectation value of the single particle kinetic energy density, $\langle\tau_w\rangle$, is equivalent to the variance of the momentum operator, $\mathbf{P} = -i\hbar\nabla$, minus its classical estimate $\bar{\mathbf{P}}$, i.e.

$$\tau_w(\mathbf{r}) = \frac{1}{2m_e} \text{Var}_\psi(\mathbf{P} - \bar{\mathbf{P}}) \quad (6.6)$$

where $\tilde{\mathbf{P}} \equiv \mathbf{P} - \bar{\mathbf{P}}$ is the fluctuation of the electronic momentum

$$\tilde{\mathbf{P}}(\mathbf{r}) = -\frac{\hbar}{2} \frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})} \quad (6.7)$$

which leads to

$$\langle\tau_w(\mathbf{r})\rangle = \int \frac{\tilde{\mathbf{P}}(\mathbf{r}) \cdot \tilde{\mathbf{P}}(\mathbf{r})}{2m_e} \rho(\mathbf{r}) d\mathbf{r} \quad (6.8)$$

This equation implies that τ_w/ρ can be expressed in terms of the local momentum component $\tilde{\mathbf{P}}$ by $\tau_w/\rho = \tilde{\mathbf{P}} \cdot \tilde{\mathbf{P}}/2m_e$. This means that $\tilde{\mathbf{P}}$ provides at least the same information as τ_w/ρ . Additional information arises from the fact that $\tilde{\mathbf{P}}$ is a vector variable. For instance, the interatomic surfaces defining the atomic partitioning in QTAIM.⁹

The vector field (Eq. 6.7) $\tilde{\mathbf{P}}$ points in the direction of maximum decrease in the electron density, and its magnitude \tilde{P} is sensitive to local charge depletion. It has been found that atomic shells are limited by the radial distances from the nucleus where a change in the concavity of \tilde{P} occurs, i.e. distances for which the condition $\partial^2\tilde{P}/\partial r^2 = 0$

^vGaussian functions exhibit difficulties for reproducing the exponential decay, and therefore this limiting behaviour is very sensitive to basis set selection.⁶⁰

is satisfied.⁹⁰ The electronic-shell populations are in excellent agreement with those obtained with ELF⁶⁴ and τ ,^{65,66} confirming the robustness of \tilde{P} for depicting electron pairs in atoms.

For graphical analysis purposes, it is also convenient that \tilde{P} is totally bounded by physically meaningful values. The lowest limit occurs at those points where the critical points of the electron density are located, i.e. for all those points that obey $\tilde{P} = 0$. Kato's cusp condition imposes a limit to electron velocities near the nucleus of an atom A ,²⁹ \mathbf{R}_A , making \tilde{P} finite and numerically equal to the atomic number Z_A , when atomic units are used^{vi}.

$$\lim_{\mathbf{r} \rightarrow \mathbf{R}_A} \tilde{P}(\mathbf{r}) = p_0 Z_A \quad (6.9)$$

This velocity is used in the estimation of relativistic corrections for atomic systems.⁵⁸ Valence electron speeds in atoms are limited by the ionization energy I_A ⁹⁰

$$\lim_{|\mathbf{r}| \rightarrow \infty} \tilde{P}(\mathbf{r}) = \sqrt{2m_e I_A} \quad (6.10)$$

This equation determines a natural boundary for an atom that leads to an experimentally-based atomic radii scale.⁹⁹ Given the general validity of the exponential decay, it is expected that molecular regions depicted by $\tilde{P}(\mathbf{r})$ are similarly bounded.

$\tilde{\mathbf{P}}$ exhibits some additional practical advantages for topological analysis. In particular, it has a direct connection with the local variables studied in QTAIM, as anticipated.⁹⁰ This vector field runs anti-parallel to the gradient of the electron density $\nabla\rho$, hence depicting the same electron density gradient paths, but with opposite direction, i.e. they have opposite direction tangent vectors at every given point of the 3D space: $\nabla\rho/|\nabla\rho| = -\tilde{\mathbf{P}}/|\tilde{\mathbf{P}}|$. The gradient paths of the electron density connect the nuclei with BCPs, giving rise to an operative definition of molecular structure.¹¹

The electron density gradient field $\nabla\rho$ can be interpreted as the fluctuation of the current density vector field $\tilde{\mathbf{J}} = \frac{1}{m_e} \tilde{\mathbf{P}}\rho$, or equivalently

$$\tilde{\mathbf{J}}(\mathbf{r}) = -\frac{\hbar}{2m_e} \nabla\rho(\mathbf{r}) \quad (6.11)$$

The local flux of the current density fluctuation $\tilde{\mathbf{J}}(\mathbf{r})$ is given by the divergence operator,

^{vi} $p_0 = \hbar/a_0$ is the atomic unit of momentum.

$$\nabla \cdot \tilde{\mathbf{J}}(\mathbf{r}) = -\frac{\hbar}{2m_e} \nabla^2 \rho(\mathbf{r}) = \frac{2}{\hbar} L(\mathbf{r}) \quad (6.12)$$

This equation indicates that the Laplacian of the electron density, L , effectively identifies local concentrations or depletion of electron density, just as QTAIM prescribes.^{9,18}

There are two forces governing chemical structures, the Feynman force exerted on the nuclei and the Ehrenfest force exerted on the electrons. The virial theorem relates the virial of the Ehrenfest force to the kinetic energy of the electrons, the virial including a contribution from the virial of the Feynman forces acting on the nuclei. The local virial theorem is given in terms of the Laplacian of the electron density by⁹

$$L(\mathbf{r}) = -2G(\mathbf{r}) - \mathcal{V}(\mathbf{r}) \quad (6.13)$$

or equivalently in terms of the current fluctuation (Eq. 6.11),

$$\frac{\hbar}{2} \nabla \cdot \tilde{\mathbf{J}}(\mathbf{r}) = -2G(\mathbf{r}) - \mathcal{V}(\mathbf{r}) \quad (6.14)$$

where $G(\mathbf{r})$ is the positively-defined kinetic energy density, and $\mathcal{V}(\mathbf{r})$ is the virial density.

The integral form of the virial, $-2G = \mathcal{V}$, requires that the net flux of the current density vanish, which is granted by the divergence theorem applied to the current density

$$\int \nabla \cdot \tilde{\mathbf{J}}(\mathbf{r}) d\mathbf{r} = \oint \tilde{\mathbf{J}}(\mathbf{r}) \cdot \mathbf{n} ds = 0 \quad (6.15)$$

In order to make this condition valid for any given atomic basin region, Ω , the local flux of the current density $\tilde{\mathbf{J}}$ over the surface of this region ($\partial\Omega$) must vanish, which leads to the local zero-flux condition at the interatomic surface^{vii}

$$\tilde{\mathbf{J}}(\mathbf{r}_{\partial\Omega}) \cdot \mathbf{n}(\mathbf{r}_{\partial\Omega}) = 0 \quad (6.16)$$

that finally gives, after using Eq. 6.11,

$$\nabla \rho(\mathbf{r}_{\partial\Omega}) \cdot \mathbf{n}(\mathbf{r}_{\partial\Omega}) = 0 \quad (6.17)$$

which defines those points $\mathbf{r}_{\partial\Omega}$ located on the basin surface $\partial\Omega$. The molecular space can be exhaustively partitioned into the atomic basins defined by equation 6.17, as

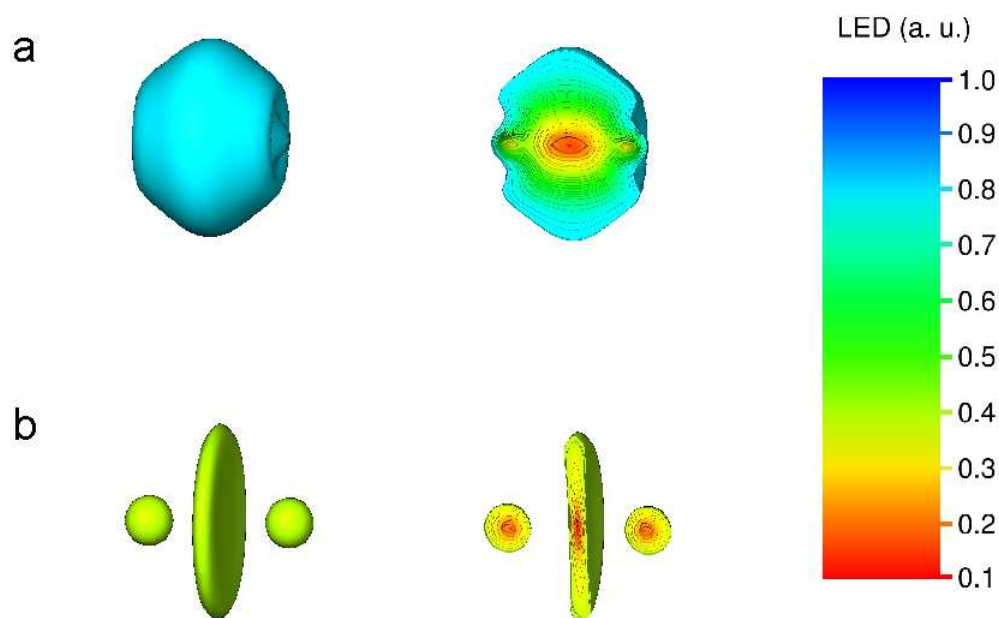
^{vii}A similar derivation is discussed by Delle Site in Ref.⁴⁷

stated by QTAIM.

Equations 6.11 to 6.17 show that $\tilde{\mathbf{P}}$ is consistent with QTAIM definitions of structure in terms of bond paths and that of proper open system. Additionally, and this is the main claim made here, $\tilde{\mathbf{P}}$ also reveals the single-particle character of localized electrons within the electron density by depicting the atomic shell structures and the symmetry of bonding interactions, as illustrated in the following section.

6.3 Localized-Electrons in Molecules

Figure 6.2: LED for the ground state (a) and the anti-bonding state (b) of H_2 at MP2/aug-cc-pVDZ level of theory. The isocontours correspond to $\tilde{P} = 0.98$ a. u.. Cross sections are coloured as indicated.



The study of \tilde{P} can be made for ground and excited states equally,^{viii} as no particular assumptions about the state of the molecule are required for its derivation.⁹⁰ Fig. 6.2 shows the LED graphics for the ground and first excited states of the diatomic hydrogen molecule.^{ix} The isocontour $\tilde{P} = 0.98$ a.u. for the bonding state of H_2 defines

^{viii}All the examples discussed in the present paper were computed at MP2/pVDZ level of theory using Gaussian 09 program,¹⁰⁵ unless otherwise stated.

^{ix}Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco

a *valence region* that encloses both nuclei. In the middle, in red, is located the bond critical point, inside a *covalent region* (orange) given by $\tilde{P} \lesssim 0.3$ a.u. The symmetry of this region indicates that this is a *shared bonding interaction*, that is characterized topologically by a negative Laplacian ($\nabla^2 \rho_{bcP} = -0.7563$, $\rho_{bcP} = 0.248$ a.u).

On the other hand, the same valence region for the ungerade state reveals three disconnected surfaces: the two spherical cores centred at the nuclei, and a flat-shaped surface enclosing the BCP. This covalent region (orange) has the shape of a disc located perpendicularly to the gradient path, which is characteristic of a *closed-shell interaction*. Topologically, it corresponds to a positive Laplacian ($\nabla^2 \rho_{bcP} = 1.7246$), and a drop in the electron density to $\rho = 0.124$ a.u, which is characteristic of a non-bonded interaction.^x

It is important to state clearly once again that the presence of a BCP is not a sufficient condition to declare that there is a chemical bond between two nuclei, as recently revisited by Bader.¹⁰⁷ In both, the gerade and ungerade states of H_2 , there is a BCP in the the middle of the two nuclei, differing by the local symmetries of the electron density which determines the nature of the interaction, i.e. bonded or not bonded, respectively. In the former case, the LED isocontours show that the paired electrons are being shared between the two atoms. For the ungerade state the same isocontour values appear as a set of unconnected spherical shells centred around each nucleus, indicating that the electrons are confined within their respective atomic regions, with a marginal sharing of electronic presence at the BCP that is about twice the electron density at the same distance for the isolated atom.

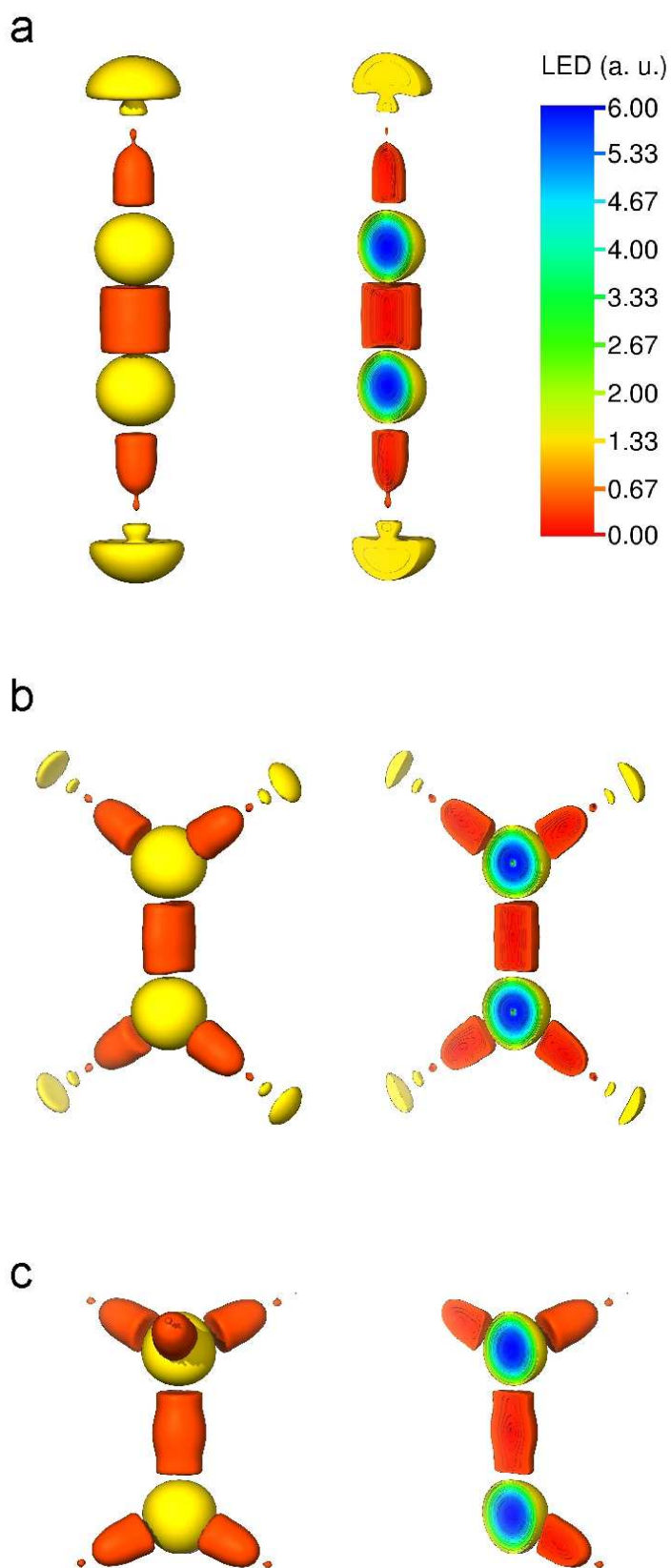
6.3.1 Single, Double and Triple Bonded Atoms

The 3D plots of \tilde{P} for the carbon series C_2H_n permit a comparative study of the single (n=6), double (n=4) and triple bonding interactions (n=2). Isocontours at $\tilde{P} = 1.30$ a.u. (yellow) and $\tilde{P} = 0.40$ a.u. (orange) are shown in Fig. 6.3 for the three molecules.

(supported by NIH P41 RR-01081).¹⁰⁶ QTAIM computations were done using AIMAll (Version 09.11.08), by Todd A. Keith, 2009 (aim.tkgristmill.com).

^xELF does not provide information about this simple case due to the fact that $\tau = \tau_w$ for the H_2 molecule, and hence $\eta = 1$ everywhere.

Figure 6.3: \tilde{P} isocontours for C_2H_2 (a), C_2H_4 (b), and C_2H_6 (c). Isocontours at $\tilde{P} = 1.30$ a.u. (yellow) and $\tilde{P} = 0.40$ a.u. (orange) are shown for the three molecules. Cross sections are coloured as indicated.



The bond critical points are located at $\tilde{P} = 0$, and therefore the covalent regions around these points depict the bonding interactions. The covalent regions located between the carbon nuclei (in orange) reveal that the single-bonded carbons in ethane and triple-bonded carbons in acetylene exhibit cylindrical symmetry, while the double bond of ethylene shows elliptic cylindrical symmetry, with the major axis being perpendicular to the molecular plane (Fig. 6.3 c). The volume enclosed by these covalent regions suggests a progression from single to triple bonded carbons.

A bell-shaped form (in orange) with the top of the bell towards the hydrogen nucleus characterizes the C-H covalent regions. These isocontours are very similar along the series, as an indication of the transferable character of the C-H bonding interactions.

The cyan, green and blue regions enclosed by the yellow spherical isocontours represent the core electrons of carbon atoms. These values have a maximum at $\tilde{P} = 6$ a.u., which corresponds to the highest value for the electron momentum in the neighbourhood of a carbon atom, as given by Kato's cusp condition (Eq. 6.9).

The yellow isocontours located outside the hydrogen bonding regions in Fig. 6.3 show the electron polarization in the valence region around the hydrogen atoms. Their size decreases from a big semisphere in acetylene to a tiny spheroid for the aliphatic two-carbon alkane, following an inverse order with respect to the C-C bond order of the three molecules.

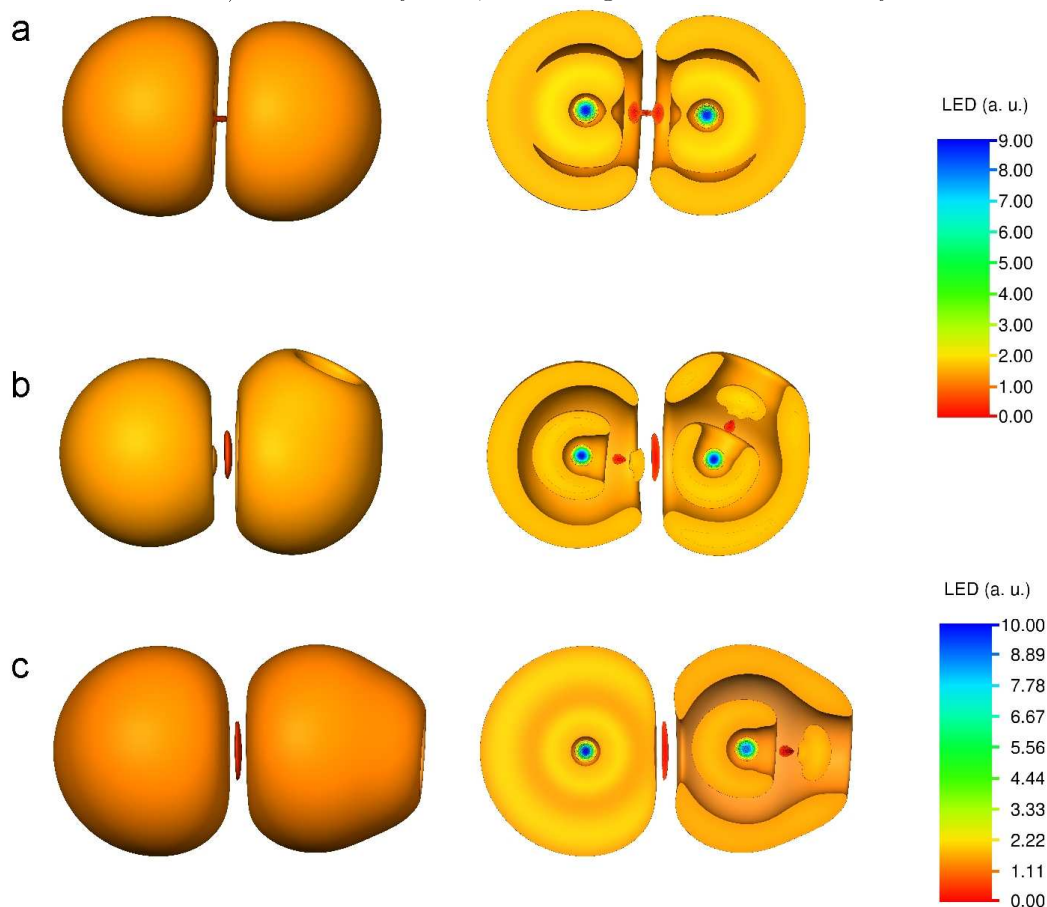
6.3.2 Hydrogen Bonded Systems

The increasing interest in hydrogen-bonded systems has created a need for theoretical tools that can visualize these important types of chemical interactions. ELF fails to provide such information, while the single-particle local momentum correctly detects the presence of hydrogen bonding and provides graphical insight into these systems, as illustrated here .

Extreme hydrogen bonding interactions can be adequately studied by the isoelectronic series (FHF)-, HF \cdots HF, and Ne \cdots HF, as suggested by Legon for IUPAC.¹⁰⁸ In (FHF)-, the binding energy borders a covalent bonding interaction with an energy of ~ 167 kJ/mol,¹⁰⁹ while HF \cdots HF is a typical hydrogen bond with a dissociation energy ~ 19 kJ/mol dominated by electrostatic forces,¹¹⁰ and in Ne \cdots HF there is a weak interaction with a very low binding energy ~ 3 kJ/mol,¹¹¹ dominated by dispersive and inductive forces. The LED plots for these three molecules correctly depict

the nature of their respective bonding interactions, as shown in Fig. 6.4.

Figure 6.4: LED isocontours for the 20-electron hydrogen-bonded isoelectronic series: (FHF)⁻ (a), FH···FH (b), and Ne···FH (c). The *covalent regions* are in red-orange ($\tilde{P} = 0.6$ a.u.) and the *valence regions* ($\tilde{P} = 1.6$ a.u.) are in dark yellow, according to the LED color key.



The valence regions, as given by the isocontours $\tilde{P} = 1.6$ a.u. (dark yellow), reveal significant similarities, in spite of their chemical differences. These contours are disjoint regions, each enclosing a ten-electron subsystem, with greater open sides toward the place where the bonding interactions take place. The symmetrical shapes of these valence regions suggest an agreement with the premise of VSEPR model, according to which the valence electron pairs surrounding an atom mutually repel each other, and therefore adopt an arrangement that minimizes this repulsion. In Fig. 6.4 this seems to be true for the inner valence regions located around the atoms in each HF subsystem, as well for the whole complexes. In other words, these results indicate that LED displays the expected behaviour of the valence electron pairs.

The covalent regions corresponding to the isocontour $\tilde{P} = 0.6$ a.u. are shown in Fig. 6.4 in red-orange. The covalent region of (FHF)- (Fig. 6.4 a)) reveals a different character to the other two molecules. In this molecule, each H-F covalent region resembles more a covalent bond, which can explain its high binding energy, and it is confirmed by the Laplacian value of $\nabla^2\rho_{bcp} = -1.7208$ ($\rho_{bcp} = 0.150$), characteristic of a shared interaction.

On the other hand, the hydrogen bonds in the other two molecules show closed-shell interaction symmetries, with $\nabla^2\rho_{bcp} = 0.1234$, $\rho_{bcp} = 0.022$ for HF \cdots HF, and $\nabla^2\rho_{bcp} = 0.0179$, $\rho_{bcp} = 0.002$ for Ne \cdots HF. This large decrease in the electron density at the BCP (one order of magnitude in each case) is consistent with the hydrogen bond energy strengths.

The cross sections show a conserved character of the HF bonding interaction in these molecules, while the valence shells of those atoms directly involved in the bonding interactions are visibly distorted. The neon atom seems to be only partially polarized in the same direction of the dipole moment of the hydrogen fluoride (Fig. 6.4 c), an indication of the electrostatic nature of this hydrogen bonding interaction.^{xi}

6.4 Conclusions

The variable \tilde{P} depicts the molecular regions of electronic overlap, as in the case of bonding regions and atomic shells. Unlike \mathcal{D} in ELF, the values of the vector field $\tilde{\mathbf{P}}$ are bounded, have direct (not relative) physical interpretation, and are easily obtainable without knowing the orbital expansion of the electron density. It has been found that \tilde{P} depicts the same kind of symmetries around the BCPs that are given by the Laplacian of the electron density. It means that \tilde{P} is able to show three-dimensionally the different kinds of bonding interactions identified by the topological analyses: the closed shell and the shared interactions. By studying the single-particle local momentum \tilde{P} , it has been shown that this variable provides a direct analysis of intra and extra molecular bonding interactions. This analysis agrees with the intuitive notion of the location of bonding interactions in molecules.

^{xi}In fact this effect is also very small as the dipole moment of Ne \cdots HF is $\mu = 1.963$ D, while the value for HF alone is $\mu = 1.946$ D.

Chapter 7

The Localized Electrons Detector as an *Ab Initio* Representation of Molecular Structures

Reprinted with kind permission from Wiley-VCH Verlag GmbH & Co.

H. J. Bohórquez, C. F. Matta and R. J. Boyd, *Int. J. Quantum Chem.*,
110 2418-2425 (2010)

7.1 Introduction

Many chemical observations are summarized succinctly in a molecular structure whereby a molecule is composed of atoms linked by a “network of bonds”.⁹ Bader introduced the quantum mechanical molecular structure in terms of the topology of the electron density in real space within the framework of the quantum theory of atoms in molecules (QTAIM). It is defined as the set of gradient paths connecting critical points.^{9,11} This molecular structure gives a graphical representation of the interaction lines inside a molecule or complex, and therefore facilitates the exploration of bonding interactions in all kinds of molecular systems.¹⁸

Recently, a measure of the bosonic character of the electron density that identifies the presence of paired electrons in atoms,⁹⁰ and molecules was proposed.¹¹² This property is given in terms of the electron density and its gradient only, by $-\hbar\nabla\rho/2\rho$, which is referred to as the *localized-electrons detector* (LED). A molecular representation emerges from this electron density property, where core shells, bonding and valence regions are clearly identifiable by separate isocontours. In this chapter it is shown that LED isocontours adequately represent the local symmetry of the electron density around the critical points of the electron density. This proposal is consistent with the tenets of the quantum theory of atoms in molecules.

In this chapter an additional classification of the critical points of the electron density based on the magnitude of the Hessian eigenvalues is introduced. LED isocontours

enclosing the critical points obey that classification, and therefore it is proposed that LED isocontours provide a topologically-correct spatial *ab initio* representation of the molecular structure. LED is orbital free and dependent on the electron density first derivative only, which opens its application to bonding studies for example in experimentally-determined electron densities.

This chapter reviews the main concepts leading to the LED definition, and provides several illustrative examples. This chapter is organized as follows: in the following section the main theoretical aspects of the proposal are discussed and then some textbook examples of bonding interactions are described. The chapter concludes with an outline of the future scope and applications of the present work.

7.2 Theory

The idea of a localized electron pair has been central in bonding theories since Lewis first postulated that the chemical bond is a consequence of electron pair sharing.⁹⁷ It is particularly remarkable that Lewis reached this important conclusion nine years before Uhlenbek and Goudsmit advanced their idea of electron spin¹¹³ and Pauli formulated his antisymmetry principle to the extent that he was forced to question the validity of Coulomb's law at small distances.¹¹⁴

According to Pauli's principle, the electronic wave function of a system of fermions must be antisymmetric with respect to the interchange of the space and spin coordinates of any two particles in the system. A region with a high probability density of an electron of a given spin excludes a same spin electron from this region but not electrons of the opposite spin due to Pauli's principle. But then the increased probability of the opposite spin electron also excludes other electrons with the same spin. The result is a region of space where there is an increased probability of finding a single pair of electrons. Recently, Samuelsson and Büttiker proved that a measure of correlation of pairs of electrons obeys Bose-Einstein statistics through the Hanbury-Brown and Twiss effect,¹⁰⁴ which supports experimentally that collective pairs of electrons behave as bosons, even though single electrons are fermions.

Electron pairs can be studied via the electron pair density distribution or some of its directly related functionals, such as the Fermi hole distribution. Bader and Heard found that local charge concentrations given by the Laplacian of the electron density $L(r)$ mirrors the Laplacian of the conditional pair density for same-spin electrons, $L(e^*, r)$, when computed at positions of maximum localization of the Fermi

hole.¹¹⁵ This mapping suggests that local concentrations of charge given by $L(r)$ correspond to regions in space where a partial condensation of the pair density towards individually localized electron pairs occurs. They also compared $L(e^*, r)$ with the electron localization function (ELF),⁴ finding no direct relationship between them, whereas an homeomorphic relationshipⁱ between ELF and the Laplacian was further suggested.²⁴ Bader and coworkers also studied the outer shell of charge concentration, called the *valence shell charge concentration* (VSCC) concluding that the Laplacian of the charge density also provides the physical basis for the VSEPR model.⁹

While the electron pairs are given explicitly by the pair density, its six-dimensional character limits to certain reference points its study in the coordinates space, as in the case of the the Laplacian of the conditional pair density for same-spin electrons, $L(e^*, r)$. Other approaches involve approximate expressions for the pair density, as in the case of ELF. In this paper, localized electrons are investigated in the three-dimensional space via the single particle kinetic energy, which depends on the electron density and its first spatial derivative only.

Within the single-particle approximation, the Weizsäcker kinetic energy density τ_w is the variance of the momentum fluctuation $\tilde{\mathbf{P}}$, at a given state ψ ,⁹⁰

$$\langle \tau_w(\mathbf{r}) \rangle_\psi = \frac{1}{2m_e} \text{Var}_\psi(\mathbf{P} - \bar{\mathbf{P}}) \quad (7.1)$$

where \mathbf{P} is the electron momentum operator and $\bar{\mathbf{P}}$ is its average real value. These two quantities are related by^{90,28,32}

$$\mathbf{P}(\mathbf{r}) = \bar{\mathbf{P}}(\mathbf{r}) + i\tilde{\mathbf{P}}(\mathbf{r}) \quad (7.2)$$

Equation 7.2 is the local representation of the single particle momentum operator, where $\tilde{\mathbf{P}}$ is the single-particle component of the momentum operator associated with the electron density ρ ,

$$\tilde{\mathbf{P}}(\mathbf{r}) = -\frac{\hbar}{2} \frac{\nabla \rho(\mathbf{r})}{\rho(\mathbf{r})} \quad (7.3)$$

A kinetic energy *local value per electron*, K can be defined by

$$K(\mathbf{r}) = \frac{1}{2m_e} \left| \tilde{\mathbf{P}}(\mathbf{r}) \right|^2 \quad (7.4)$$

ⁱTwo functions are said to be homeomorphic if one can be continuously transformed into the other.

which has the same form as the kinetic energy of a classical particle. Hence, the expectation value of the Weizsäcker kinetic energy density τ_w in terms of $\tilde{\mathbf{P}}(\mathbf{r})$ from Eq. 7.3 is

$$\langle \tau_w(\mathbf{r}) \rangle = \frac{1}{2m_e} \int \tilde{\mathbf{P}}(\mathbf{r}) \cdot \tilde{\mathbf{P}}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \quad (7.5)$$

The Weizsäcker kinetic energy density is a lower boundary for the total kinetic energy density τ , i.e. $\tau_w \leq \tau$. By taking the volume derivative of both sides of Eq. 7.5, we get the relationship between the *local value* $K(\mathbf{r})$ and its corresponding *density*, $\tau_w(\mathbf{r})$

$$\frac{d\langle \tau_w(\mathbf{r}) \rangle}{dV} = \tau_w(\mathbf{r}) = K(\mathbf{r}) \rho(\mathbf{r}) \quad (7.6)$$

Similarly, if τ is the *exact* kinetic energy density of a multielectronic system, its local value is τ/ρ . These local kinetic energy variables are discussed in more detail elsewhere,¹¹² in connection with the electron localization function (ELF) and the Laplacian. The main points are summarized in the following subsection.

7.2.1 The Single Particle Kinetic Energy and the Electron Localization

The kinetic energyⁱⁱ can be written in terms of the local representation of the momentum (Eq. 7.2) by

$$\tau(\mathbf{r}) \equiv \frac{|\mathbf{P}|^2}{2m_e} \rho(\mathbf{r}) \quad (7.7)$$

which, after replacing Eq. 7.2, leads to

$$\tau(\mathbf{r}) = \frac{\rho(\mathbf{r})}{2m_e} \overline{\mathbf{P}}(\mathbf{r})^2 + \tau_w(\mathbf{r}) \quad (7.8)$$

Therefore, the difference between the total kinetic energy density τ and the single-particle contribution κ is the kinetic energy density associated with the average momentum component $\overline{\mathbf{P}}$

$$\mathcal{D}(\mathbf{r}) \doteq \frac{\rho(\mathbf{r})}{2m_e} \overline{\mathbf{P}}(\mathbf{r})^2 = \tau(\mathbf{r}) - \tau_w(\mathbf{r}) \quad (7.9)$$

ⁱⁱThe kinetic energy density has no unique representation in terms of the electron density, but it can be computed exactly by orbital dependent terms, such as τ .

\mathcal{D} is the Pauli energy functional, which is proportional to the Fermi hole mobility function of Luken and Culberson,¹⁰⁰ and is related to the curvature of the Fermi hole as shown by Dobson.¹⁰¹ Becke and Edgecombe derived equation 7.9 from the spherically-averaged conditional same-spin pair probability density, and proposed \mathcal{D} as the main variable for defining the electron localization function (ELF).⁴ They identify the localization of an electron with the probability density to find an electron in the vicinity of a second same-spin reference electron. The smaller the probability density \mathcal{D} , the higher the electron localization. Gatti pointed out that the only measure of the electron localization present in ELF is the expression \mathcal{D} ,¹⁰² which is a relative measure of the *bosonic character* of the electron density, because \mathcal{D} is the excess kinetic energy electrons have compared to a system of bosons of the same density,¹⁰³ as discussed in more detail elsewhere.¹¹²

7.2.2 A Localized Electrons Detector (LED)

As an alternative to conventional electron localization functions, it is instructive to consider what Gatti refers to as the “real electron localization function” (\mathcal{D} in Eq. 7.9). In this equation, the Weizsäcker kinetic energy density κ accounts for the *bosonic character* of the system, and therefore it has been proposed that the study of its associated local momentum $\tilde{\mathbf{P}}$ (Eq. 7.1), may be useful for the determination of localized electrons in atoms and molecules. It has been shown that $\tilde{\mathbf{P}}$ is an orbital-free direct three-dimensional representation of localized pairs in atoms^{90,99} and molecules.¹¹² $\tilde{\mathbf{P}}$ is a vector field that runs anti-parallel to the gradient of the electron density $\nabla\rho$, and therefore depicts the same electron density gradient paths studied in QTAIM, as discussed in detail elsewhere.¹¹² The magnitude of the vector $\tilde{\mathbf{P}}$ (\tilde{P}), is bounded by theoretical limiting values:

- At the critical points of the electron density where $\nabla\rho = 0$ and hence $\tilde{P} = 0$, therefore providing all the critical points of QTAIM.
- A maximum value of \tilde{P} occurs in the neighborhood of the heaviest nucleus within the molecule (\mathbf{R}_A), for which $\lim_{\mathbf{r} \rightarrow \mathbf{R}_A} \tilde{P}(\mathbf{r}) = p_0 Z_A$, according to Kato’s cusp conditionⁱⁱⁱ. This speed value is used in the estimation of relativistic corrections for atomic systems.⁵⁸

ⁱⁱⁱ $p_0 = \hbar/a_0$ is the atomic unit of momentum.

- At long distances valence electron kinetic energies are limited by the ionization energy I_A , and hence $\lim_{|\mathbf{r}|\rightarrow\infty} \tilde{P}(\mathbf{r}) = \sqrt{2m_e I_A}$ for atomic systems, which leads to a experimentally-based atomic radii scale.⁹⁹ The exponential decay of the electron density imposes a similar behaviour to molecular systems.

These three limiting values define three different spatial regions of interest in any molecular system: *covalent* ($\tilde{P} \simeq 0$), located around a BCP, *core* ($\tilde{P} \simeq Z_A p_o$), located around the nuclear attractors, and *valence* regions ($\sqrt{2m_e I} < \tilde{P} \ll Z p_o$).

There are several practical advantages for studying the localized-electrons detector (LED), \tilde{P} . It depends exclusively on the electron density and its gradient. Hence this variable consistently fits within the conventional *atoms in molecules* analysis, by providing complementary information about localized electrons. \tilde{P} provides the correct electron shells and its population in atomic systems.^{90,60,61,116} Several fundamental theoretical relationships between QTAIM and LED have been discussed elsewhere.¹¹² Here it should be emphasized that LED correctly portrays the symmetry of the electron density at the bond critical points (BCP), and it is shown that the LED isocontours provide an *ab initio* molecular structure representation consistent with the topology of the electron density.

7.2.3 Electron Density Symmetry Around the Critical Points

Table 7.1: Rank three critical points classification according to the sign of the Hessian eigenvalues, $\lambda_1 \leq \lambda_2 \leq \lambda_3$.

label	curvatures	type	character
(3, -3)	$\lambda_i < 0, i = 1, 2, 3$	local maximum	attractor
(3, -1)	$\lambda_i < 0, i = 1, 2, \lambda_3 > 0$	saddle	bond critical point
(3, +1)	$\lambda_i > 0, i = 1, 2, \lambda_3 < 0$	saddle	ring critical point
(3, +3)	$\lambda_i > 0, i = 1, 2, 3$	local minimum	cage critical point

At the critical points the gradient of the density vanishes, and hence the local curvatures are given by the Hessian matrix. The sign of the eigenvalues of the Hessian at the critical points provide the typical classification used in QTAIM,⁹ which is summarized in Table 7.1. The Laplacian is the trace of the Hessian matrix, which is one of the topological invariants of the electron density. Additional information about the local symmetry of the electron density at the critical points can be extracted from the magnitudes of the eigenvalues.

The Hessian matrix encodes important shape information, which is often used for three-dimensional data extraction from scalar fields.¹¹⁷ It can differentiate between line-like, plane-like and sphere-like structures, for example. In Table 7.5 a classification of the eigenvalues according to the respective local three dimensional shapes of the electron density at the critical points, as adapted from a shape classification by Westin et al. is introduced.¹¹⁷

A direct inspection of the eigenvalues of the Hessian can be made in order to identify the type of symmetry involved, but is particularly interesting from the molecular structure perspective to be focused on the bond critical point symmetries. The Laplacian of the electron density, $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$, identifies the nature of the bond critical point type as shared or closed-shell interaction. The local virial theorem connects the Laplacian with the positive-defined kinetic energy density $G = \hbar^2/2m_e \nabla\psi^* \cdot \nabla\psi^*$ and the potential energy density \mathcal{V} by

$$\frac{\hbar^2}{4m_e} \nabla^2\rho(\mathbf{r}) = 2G(\mathbf{r}) + \mathcal{V}(\mathbf{r}) \quad (7.10)$$

Since G is always positive and \mathcal{V} is always negative, the sign of the Laplacian determines the relative magnitudes of the potential and kinetic energy densities. In regions where $\nabla^2\rho < 0$, the potential energy density is dominant. This leads to a local concentration of negative charge, and it is termed a *shared interaction*. In contrast, in regions where $\nabla^2\rho > 0$ the kinetic energy dominates and causes a depletion of negative charge, mainly on the interatomic surface, as can be inferred from Table 7.5.

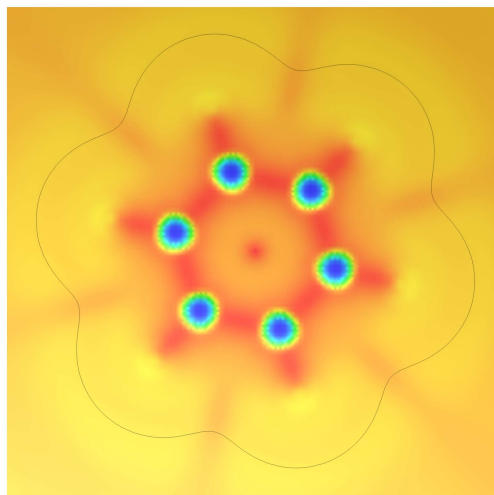
The local symmetries of the electron density, as provided by the curvatures are also reflected onto LED isocontours located around critical points, i.e. for values $\tilde{P} \simeq 0$. Therefore, depending on the type of critical point, there are different characteristic shapes to be found, as illustrated in the next section.

7.3 Results

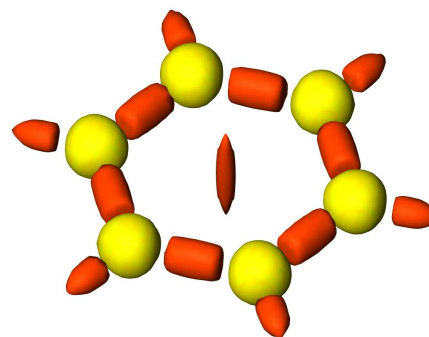
In order to illustrate the observations discussed in the previous section, Fig. 7.1 shows LED for benzene. The isocontour $\rho = 0.001$ au in Fig. 7.1a indicates the spatial extent of the electron density. Fig. 7.1a shows LED contour-plot values projected onto the molecular plane. The lowest values appear in red, ($\tilde{P} \simeq 0$ au), which include C-C and C-H bond critical points and the ring critical point in the middle of the benzene ring.

Figure 7.1: LED (Eq. 7.3) for benzene.

(a) LED values projected onto the molecular plane. The gray line corresponds to $\rho = 0.001$ au isocontour. The side of the square is 14 au.



(b) Molecular structure given by LED = 0.33 au (red-orange), and LED = 1.40 au. (yellow)



The covalent bonding regions follows the particular bonding symmetries around the BCPs, which is consistent with the local shape of the the electron density at the BCPs given in Table 7.5, according to which the shape of the electron density at the CH and CC bonds exhibit cylindrical symmetries, and at the RCP . The curvatures at the critical points for benzene appear in Table 7.3 as obtained at the MP2/6-311++G(2d,2p) level of theory (Gaussian 03¹¹⁸). QTAIM computations were done using AIMAll (Version 09.11.08), by Todd A. Keith, 2009 (aim.tkgristmill.com).

Electronic shells are given by the concentric isocontours around the LED highest value (in blue) of $\tilde{P} = 6$ au for present case, which agrees with the fact that the heaviest nucleus is carbon, which by Kato's cusp condition imposes $\tilde{P}_{max} = Zp_0$. Fig. 7.1b shows the almost spherical isocontour $\tilde{P} = 1.40$ au (in yellow) that encloses the inner shell electrons of carbon atoms, while the isocontour $\tilde{P} = 0.33$ au (in orange) depicts the bonding interactions. Notice the cylindrical shape of C-C bonds, and the peaked cylinder of H-C bonds, with the peak pointing to the H nucleus. At the RCP there is a thin cylinder perpendicular to the molecular plane with peaked ends.

Thus, the molecular structure representation of benzene in Fig. 7.1b agrees with the local curvatures of Table 7.5. There is a striking resemblance of these plots with

Table 7.3: Curvatures of benzene. All quantities in atomic units.

CP	Subsys.	λ_1	λ_2	λ_3	$\nabla^2\rho$	Shape
BCP-S	C-C	-0.7192	-0.5984	0.3118	-1.0058	cylindroid
BCP-S	C-H	-0.8179	-0.8047	0.4372	-1.1854	cylindroid
RCP	C ₆	-0.0144	0.0886	0.0886	0.1629	cylinder
NACP	C	-1.81×10^5	-1.81×10^5	-1.81×10^5	-5.44×10^5	sphere
NACP	H	-8.6177	-8.60757	-7.7727	-24.9979	spheroid

the curl of the first-order current density $\mathbf{J}^{(1)}(\mathbf{r})$ plots for the same molecule reported by Keith and Bader.¹¹⁹

Intermolecular interactions are also depicted by LED, as shown in Fig. 7.2 for the hydrogen bonded water dimer. The computations were done at the MP2/cc-pVDZ level of theory. The LED isocontour ($\tilde{P} = 0.62$ au) located around the BCP corresponding to the hydrogen bond appears as a planar region oriented perpendicularly to the gradient path in Fig. 7.2a. The planarity of the isocontours is directly related with the character of the interaction, and a planar symmetry indicates the presence of a closed shell type, according to the local shape classification (Table 7.5). This planar region encloses the O–H interatomic surface, which is shown in Fig. 7.2b. The local symmetries around each critical point for the water dimer are summarized in Table 7.4.

According to the present results and previous ones,¹¹² LED provides information about localized electrons in molecules that closely resembles the ones obtained by studying the Laplacian of the conditional pair density for same-spin electrons, $L(e^*, r)$ proposed by Bader et al.¹¹⁵ While these conditional pair density plots depend on a reference point to be computed, LED is a single-particle, orbital-free property that depends on the electron density and its first derivative only, which facilitates its computation from experimentally-derived electron densities, or from conventional computational tools, such as the Gaussian suite of programs^{iv}.

Table 7.4: Curvatures of water dimer.

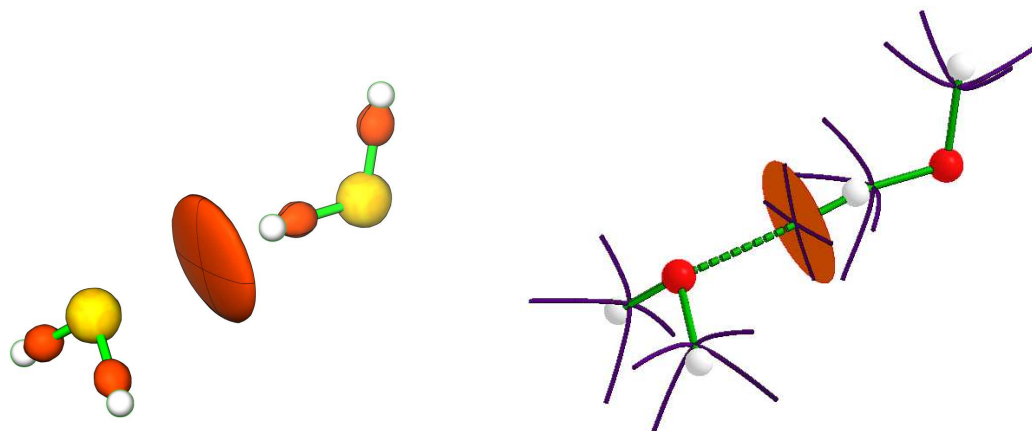
CP	Subsys.	λ_1	λ_2	λ_3	$\nabla^2\rho$	Shape
BCP-S	O-H	-1.8945	-1.8565	1.4126	-2.3385	spheroid
BCP-CS	O–H	-0.0287	-0.0281	0.1513	0.09457	planar
NACP	H	-4.2828	-4.2630	-2.7210	-11.2669	spheroid
NACP	O	-9.67×10^5	-9.67×10^5	-9.67×10^5	-2.90×10^6	sphere

^{iv}Via the CUBE option for example or using the routine CUBEGEN, which provides the gradient and the density that can be processed to get the values of Eq. 7.3.

Figure 7.2: Interatomic surfaces and LED isocontours for the water dimer. The molecular structure (CPK style) is included to facilitate the comparison.

(a) Molecular structure given by LED = 0.62 au (orange) and LED = 1.60 au (yellow). Perpendicular geodesics are plotted in black.

(b) IAS for the hydrogen bond (orange) and IAS paths (purple). A cut-off value of $\rho = 0.01$ au was used.



7.4 Conclusions

In this chapter it has been shown that the graphical representation of bonding interactions given by LED is consistent with the local curvatures of the electron density charge as given by the eigenvalues of the Hessian matrix.

The obtained plots for isocontours near the critical points, i.e. $\tilde{P} \simeq 0$, depict three dimensional shapes that agree with the eigenvalues of the Hessian symmetries at the critical points, which is summarized in Table 7.5. As a result, these plots are distinctive for BCPs with shared interactions from those with closed shell interactions, which might be of interest for the study of large molecular systems mainly, where keeping track of topological properties through tables becomes unpractical.

Table 7.5: Critical points classification according to the magnitude of eigenvalues, $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$.

curvature relations	shape	critical point type
$\lambda_1 \approx \lambda_2 \ll \lambda_3$	planar	BCP closed-shell interaction
$\lambda_1 \ll \lambda_2 \approx \lambda_3$	cylindrical	RCP
$\lambda_1 \approx \lambda_2 \approx \lambda_3$	spheroidal	BCP shared interaction; nuclear attractor

The molecular structure LED representation resembles the conditional pair density projections introduced by Bader and Heard,¹¹⁵ and the curl of the first-order current density $\mathbf{J}^{(1)}(\mathbf{r})$ plots by Keith and Bader.¹¹⁹ These two properties are related to the Laplacian of the electron density, and therefore all together provide information about localized electrons in molecules. The present proposal is mathematically simpler, and hence it has the potential to be a more accessible graphical tool than, for example, $\mathbf{J}^{(1)}(\mathbf{r})$ or $L(e^*, r)$.

Due to the theoretical connection of LED with QTAIM, including the location of bond critical points and the interatomic surface definition,¹¹² it is unnecessary to explore a partitioning scheme based on \tilde{P} . The main partition scheme of the molecule into atomic regions is guided by the virial theorem,¹²⁰ which yields the conventional QTAIM analysis of atomic properties.⁹ Nevertheless, \tilde{P} provides information about the electronic structure *inside* an atomic basin, so an intra-basin population analysis based on \tilde{P} isocontours may be of interest.

Chapter 8

Taxonomy of Chemical Bonds

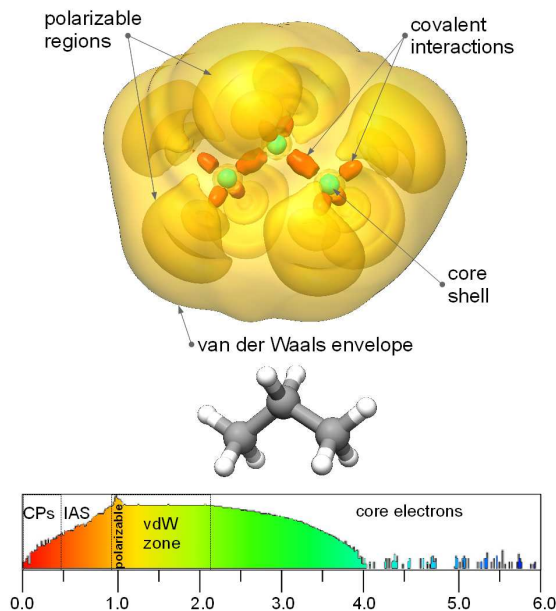
8.1 Introduction

The stability of molecular systems depends on a variety of interatomic interactions that are always mediated by the local behavior of the electron density. While covalent bonding is the strongest and most stable, several other interactions do not fit within the Langmuir-Lewis electron pair description. This is the case of van der Waals and noncovalent interactions, which are central to physicochemical processes such as the biomolecular recognition in proteins, nucleobase stacking in DNA, crystal packing, vapor–liquid condensation, polymer aggregation or fluid viscosity. Enzymatic and other important chemical processes also involve breaking and forming of various classes of hydrogen bonds and weaker interactions, and in certain cases their physical nature is not entirely understood. In fact, Pauling’s model can account for enzymatic reaction acceleration up to 11 orders of magnitude, but most enzymes exceed that proficiency, which leads to the hypothesis that nearly all enzymes speed reactions via covalent interactions.¹²¹ With these new findings a detailed quantum mechanical description of all sorts of stabilizing interatomic interactions is required.

Unfortunately, the conventional electron-density-based methods for the study of interatomic interactions have certain limitations. For instance, the electron localization function (ELF) does not describe hydrogen bonds or weaker interactions, mainly because it only captures the presence of spin-paired electrons,⁴ i.e. strong covalent bonds. On the other hand, the *quantum theory of atoms in molecules* (QTAIM) identifies every atomic interconnectivity as a *bond critical point* (BCP) disregarding the nature of the interaction involved.⁹ As Ciowlowski pointed out almost two decades ago: “the topological approach does not provide a fundamental distinction between the *normal* strong bonds and bonds resulting from weak interactions”.¹²² These observations initiated a controversy that has not been solved yet.¹²³

In the present chapter new advances in the topological analysis of electron density are introduced. From the single-particle local momentum an accurate representation of the local symmetry of the electron density, around the critical points and within the

Figure 8.1: Electronic regions for propane. Top: LED isocontours (0.40, 1.13 and 4.00). Bottom: LED values distribution within the range [0,6] with the respective electronic zones: critical points (CPs), interatomic surfaces (IAS), steric regions, van der Waals (vdW) zone and nuclear electronic shells.



polarization regions, is obtained. The proposed analysis provides a three-dimensional visual insight of the stabilizing interatomic interactions in molecular systems. The new topological criteria clearly differentiates covalent from noncovalent interatomic interactions, and provides a universal classification of the stabilizing interactions that may occur between atoms, ranging from van der Waals to metallic ones, in an unified way. The proposed theoretical tool complements conventional analyses of the electron density such as QTAIM, while it is independently derived.

8.2 Theory

The main variable of study is the local momentum $\tilde{\mathbf{p}} = -\hbar\nabla\rho/2\rho$ that captures the single-particle or *bosonic* character of the electron density, and it is termed *localized electrons detector* (LED) because it distinctively localizes the stable electronic regions in atoms,^{90,99} molecules,¹¹² and complexes.¹²⁴

The vector field $\tilde{\mathbf{p}}$ can be naturally incorporated into the conventional quantum chemical topology analysis, as shown elsewhere.⁹⁰ The classical kinetic energy expression associated to this local momentum, $\tilde{k} = \tilde{p}^2/2m_e$, is essentially the Weizsäcker

kinetic energy density, $\tau_w = \tilde{k}\rho$, which is a common gradient correction in DFT models. Recently, Vydrov has proven that if \tilde{k} is used as a *local polarizability* term in a DFT model, it accurately accounts for the dispersion contribution to the molecular energy.¹²⁵ In addition, Liu interprets τ_w as a quantitative measure of steric effects within conceptual DFT.⁵ These works indicate that the bosonic local momentum $\tilde{\mathbf{p}}$ may establish a theoretical link between QTAIM and DFT, as proposed at the 5th European Charge Density Meeting in 2008, where the first results in molecular systems using LED were made public.

LED has certain advantages for the graphical analysis of electronic regions, as has been shown in a series of papers.^{90,99,112,124} In general, LED has an absolute scale (i.e. transferable) whose limits are theoretically known for any given molecule, and it is given by $\tilde{p} \in [0, Z_{max}]$ (in atomic units), with Z_{max} the highest atomic number in the system. Thus LED is a totally bounded variable, i.e. it has finite boundaries, unlike the family of gradient functions $\nabla\rho/\rho^m$ that diverges for $m > 1$. This is an important practical advantage of LED over other graphical analyses whose variables are unbounded.

LED is obtainable from standard numerical wavefunctions or from experimental electron densities; for the computations reported in this chapter a modified version of the FORTRAN code *CheckDen* developed by Pacios was used.¹²⁶ The cube files were visualized with the program *Chimera*,¹⁰⁶ which is capable of coloring LED isocontours according to the range values, i.e. $[0, Z_{max}]$. The color code used here is the same of previous works: red \rightarrow yellow \rightarrow green \rightarrow blue, as shown at the bottom of 8.1. The color code spans LED values in a consistent and transferable way, as can be verified by comparing the plot in 8.1 with previously reported results for molecules and complexes.^{112,124} In all the results discussed here a fixed range $[0, 6]$ was used (unless otherwise stated).ⁱ

The 3D inspection of LED reveals three main electronic regions with theoretically known limiting values: core ($\tilde{p} \sim Z_A$), saddle critical points (including *bond critical points* and *ring critical points*) ($\tilde{p} \sim 0$) and polarizable regions ($\tilde{p} \sim \sqrt{2I}$), with I the first ion, as shown in 8.1 for propane. The first two regions depict an *ab initio* molecular structure that is very similar to the ball-and-stick ones: covalent bonds are represented by isocontours (orange) with axial symmetry connecting the bonded nuclei, and core regions appear as spherical isocontours enclosing closed-shell electrons

ⁱFor molecules including nuclei with $Z > 6$ the only information that is missing from the graphics are the inner core shells, which are not relevant for the present discussion.

(green). The transparent-yellow isocontour in 8.1 is a van der Waals (vdW) envelopeⁱⁱ, inside of which several internal isosurfaces appear, such as the semi-spheres around hydrogens, which indicate the polarization of the electron density in that region. More details on these observations can be found elsewhere.^{90,112,124}

For noncovalent interactions, the isocontours $\tilde{p} \sim 0$ usually run parallel to the interatomic surface (IAS) separating both atoms, and therefore these isocontours appear as *local patches* of the respective IAS. This observation has been recently reported for the water dimer and other weakly bonded systems,¹²⁴ and here its general validity is demonstrated.

The main focus of the present paper is the derivation of a new topological classification of the critical points induced by LED that leads to quantitative and graphical representations of the different interatomic interactions, ranging from covalent to noncovalent ones. The optimized geometries were obtained at the MP2/aug-cc-pVDZ level of theory with Gaussian09.¹⁰⁵ QTAIM properties were computed with AIMALL.¹²⁷

8.2.1 Symmetry of the Electron Density at the Stationary Points

Around the stationary points of the electron density, (i.e. the critical points $\nabla\rho = 0$), LED isocontours are simple-connected shapes that follow the local curvature of the electron density. Within the neighborhood of any critical point, \tilde{p} goes to zero and the topology of the isocontours $\tilde{p} \sim 0$ is dominated by $|\nabla\rho| \sim 0$.ⁱⁱⁱ From the analytical expansion of the electron density, the approximate expression of the gradient around a critical point $\mathbf{r}_0 = (x_0, y_0, z_0)$ is given in terms of the second derivatives by

$$\nabla\rho \approx (x - x_0) \lambda_1 \mathbf{u}_1 + (y - y_0) \lambda_2 \mathbf{u}_2 + (z - z_0) \lambda_3 \mathbf{u}_3 \quad (8.1)$$

where $\{\lambda_1 \leq \lambda_2 \leq \lambda_3\}$ are the eigenvalues (or *principal curvatures*) of the Hessian, and $\{\mathbf{u}_i\}$ their respective eigenvectors (or *principal directions*); where $\mathbf{r} = (x, y, z)$ is the position vector near the critical point \mathbf{r}_0 and $|\mathbf{r} - \mathbf{r}_0| \rightarrow 0$. It is assumed that both position vectors are given in terms of the orthonormal basis set $\{\mathbf{u}_i\}$. After taking the square of the gradient vector (Eq. 8.1) and dividing by $|\nabla\rho|^2$, the equation of an

ⁱⁱThis isocontour is nearly identical in shape and volume to the the isocontour $n = 0.0002$ au.

ⁱⁱⁱThis is true in general for every function of the form $\nabla n/n^m$ as can be obtained by applying L'Hopital's theorem to the ratio, and using the fact that the electron density is a continuous function with continuous derivatives.

ellipsoid centered at \mathbf{r}_0 is obtained

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} + \frac{(z - z_0)^2}{c^2} \approx 1 \quad (8.2)$$

whose principal axis are given by $a = |\nabla\rho|/|\lambda_1|$, $b = |\nabla\rho|/|\lambda_2|$ and $c = |\nabla\rho|/|\lambda_3|$.^{iv}

Equation 8.2 demonstrates that the isocontours $\tilde{p} \sim 0$ are essentially ellipsoids centered at the critical points that are aligned with the local principal directions. In other words, at the critical points, the eigenvalues of the Hessian can be translated into a graphical representation of the local symmetry of the electron density via the isocontours $\tilde{p} \sim 0$. The main result emanated from this assertion is that the topological properties of the electron density can be studied by the *shapes* of the scalar field \tilde{p} , thanks to the correspondence between the isocontours and the curvatures established by Eq. 8.2. Therefore, the spherical symmetry of core electrons around the nuclei plotted by LED is explained by Eq. 8.2 because the three curvatures are identical in most of the cases ($a \approx b \approx c$), i.e. for atoms with at least the first shell filled (or $Z \geq 2$), as shown for carbon atoms by the green isocontours in 8.1. It is important to note that the scalar field \tilde{p} has removable discontinuities exactly at the nuclei, i.e. at every nuclear position \mathbf{R}_A , $\tilde{p}(\mathbf{R}_A) = 0$ and $\tilde{p}(\mathbf{R}_A + \delta\mathbf{r}) = Z_A$ with $\delta\mathbf{r}$ a vector with infinitesimally small magnitude $|\delta\mathbf{r}| > 0$. It means that all the points in the neighborhood of the nucleus behave according to Kato's cusp condition.²⁹

The ellipsoid given by Eq. 8.2 also explains the spike-shaped ellipsoids found at the *ring critical point* (RCP) of benzene and similar structures, for which the curvatures on the plane are smaller than the perpendicular one, i.e. $a \approx b \ll c$.

The size and shape of LED isocontours at the BCP indicates its character, i.e. the relative values of the axes a , b and c . All the observations on the symmetry of the electron density at the BCPs as induced by LED are summarized in 8.2, where it is shown how the relative values of the ellipsoid axes yields to the characteristic shapes depicted by LED for the different bonding interactions. In these diagrams the shaded region represents the local patch of the IAS separating the atomic basins of the hypothetical interacting atoms, A and B. The eigenvectors of the Hessian are also indicated (in blue), in the usual order imposed by the eigenvalues. Hence, three main shapes are found, each corresponding to a different type of bonding interaction. This succinct classification complements previous topological classifications achieved

^{iv}Assuming that all the critical points are rank-three, i.e. $\lambda_i \neq 0$ for $i = 1, 2, 3$.

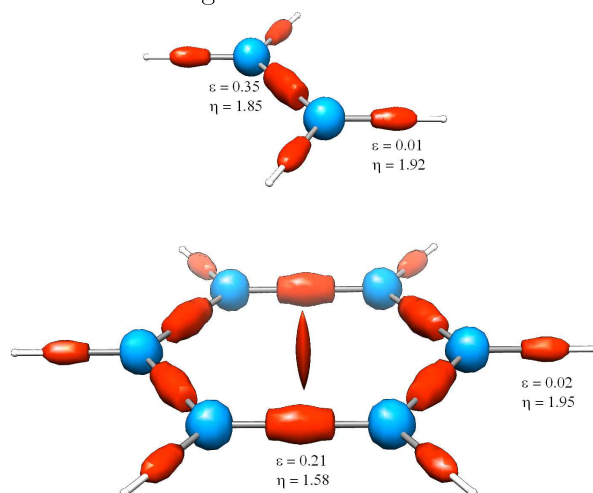
Figure 8.2: BCPs classification according to the local symmetry. Ellipsoids main axes are indicated in red and the respective eigenvectors in blue. The hatched plane represents the IAS *local patch*.

Type of bond	Ellipsoid relations	Graphical representation
Covalent	$c > a$	
Closed shell	$c < a$	
Noncovalent	$c \ll a \approx b$	

by QTAIM that are discussed in detail elsewhere.¹²⁸

In the following paragraphs a set of selected molecular systems exemplify certain types of bonding interactions. Their respective electronic regions are discussed in detail from the perspective of the analysis here introduced.

Figure 8.3: LED for ethene (top) and benzene (bottom). The isocontours correspond to $\tilde{\rho} = 0.25$ au (red) and $\tilde{\rho} = 5$ au (blue). The color code used is the same as in Fig. 8.1.



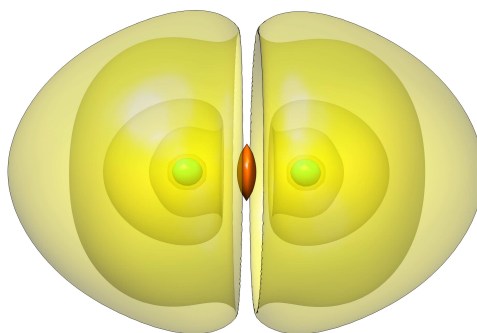
8.3 Covalent Bonds

Two interacting nuclei are connected by a BCP disregarding the type of bonding interaction occurring between them.⁹ The location of the BCP is always closer to the less electronegative atom, as a consequence of the electronic pressure of the more electronegative one.¹⁰⁷ Hence, the type of chemical interaction between two atoms interconnected by a BCP reflects the covalent regions involved. If there is an electron-pair formation, as in a typical covalent bond, the measure of local concentrations of charge provided by the Laplacian ($\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$) suffices as an indicator of covalency. Consistently, $\nabla^2\rho < 0$ for most covalent interactions, which in terms of the curvatures is $|\lambda_3| < |\lambda_1| + |\lambda_2|$. For single-bonds $\lambda_1 \approx \lambda_2$ and therefore a simple relationship between the ellipsoid axes emerges, $c > 2a$, meaning that for a single covalent bond the axis of the ellipsoid along the bond path is at least twice the perpendicular one. In consequence, single bonds are represented by elongated ellipsoids linking both atoms, as shown in Figs. 8.1 and 8.3 for CH and CC bonds (red). LED represents the BCPs symmetries by ellipsoid isocontours with a flat side facing the closest valence shell, as the CH bonds in hydrocarbons (Fig. 8.3) or the CO bond in formaldehyde (Fig. 8.5). In the case of CH, the hydrogen atom lacks inner closed shells, and therefore the flat side faces the carbon atom only. In the CO bond of formaldehyde the flat side faces the valence shell of the carbon, which is closer than the more populated oxygen. In the case of the CC bonds in propane (Fig. 8.1), ethene and benzene (Fig. 8.3) both carbons are equidistant from the BCP, and hence there are two flat side ends for the respective ellipsoids.

Topologically, BCPs are characterized by the transverse and longitudinal ellipticity measures $\epsilon = (\lambda_1/\lambda_2 - 1)$ and $\eta = |\lambda_1|/|\lambda_3|$, respectively. These parameters can be expressed in terms of the ellipsoid axes at the BCPs by $\epsilon = (b/a - 1)$ and $\eta = c/a$, that are represented directly by LED isocontours. An ellipticity $\epsilon \neq 0$ (or $a > b$) indicates the π character of bonding interactions, while the longitudinal ellipticity provides an estimate of the *covalency degree* (of covalent bonds) at the BCP: $\eta > 1$ ($c > a$) indicates concentration of density toward the bond path through the BCP, and is characteristic of shared or covalent interactions, while $\eta < 1$ ($c < a$) indicates a polarization of the electron density along the bond path away from the BCP and denotes the existence of charge transfer between the bounded atoms, like in closed-shell or electrostatic interactions. In Fig. 8.3 it is shown how LED graphically represents the ellipticity ϵ in the carbon-carbon bonds of ethene and benzene. While η is very

similar for the C-C bonds in both molecules (1.86 and 1.85), the ellipticity of ethene is ~ 1.7 times of that in benzene (0.35 and 0.21, respectively), which is consistent with the π character of these bonds. A value $\epsilon \approx 0$ is characteristic of single and triple bonds which have cylindrical symmetry around the bond path, as the CH single bonds shown in Fig. 8.3.

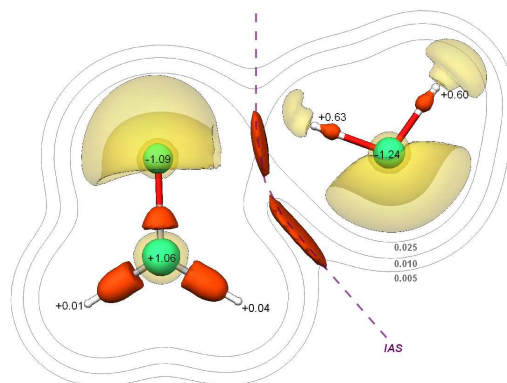
Figure 8.4: LED for fluorine molecule. LED isocontours plotted are 0.70, 1.50 and 4.00 au.



A negative Laplacian is not an universal indicator of covalency because certain type of covalent bonds called *closed shell* (or charge shift, according to Shaik et al.) (CS) bonds have a positive Laplacian and high electron density at the BCP.¹²⁹ While the topological characterization of the CS bond is rather new, its occurrence is quite common. These are *closed-shell* interactions (in QTAIM terminology), for which $c < a$, and they are characterized by the acceptor atom having orbital vacancies equal to or greater than the number of valence electrons on the donor, like in the ionic bonds of LiF, NaCl, LiH, LiO and BeO. Other examples of CS bonds are the single bond in F₂ and the CO double bond in formaldehyde (Fig. 8.5). These bonds are composed by at least one element with high electronegativity that causes a significant *valence shell repulsion* that affects the covalent bond. Hence, while for F₂ there is a significant charge concentration within the two nuclei forming the bond, the outermost filled shells are not involved in the bond, but instead they contribute to its weakening. In Fig. 8.4 LED isocontours show exactly the situation for the fluorine molecule: the red isocontour marks a small bonding region, while the yellow (transparent) ones depict the outermost shells as disjoint regions centered around each nucleus. Similar analysis can be made for the formaldehyde, for which the CO bond is being affected by the oxygen excess of one electron that causes a great pressure on the bonding region. But, on the other hand, the carbon atom is positively charged and therefore attracts local

concentrations of charge from the coordinated atoms, resulting at the end in a very stable covalent bond.

Figure 8.5: LED for the formaldehyde-water complex. LED isocontours are 0.45, 1.30 and 4.00 au. Electron density isocontours and atomic charges are given in au.



It is important to recall that in an equilibrium geometry there are no net forces on the nuclei, because all the possible intramolecular interactions are balanced. The valence repulsion does not mean the contrary, but as illustrated in the figures, the otherwise spherical (i.e. isolated atom) valence shells are polarized. In the mentioned figures, the yellow regions represent the deformation of the valence shells that indicates their polarization. It is inferred that the more stable the valence shells are, the weaker the stabilizing interactions may occur. In this way, it can be deduced that the valence repulsion may be stronger in bonding interactions involving two closed-shell systems, like noble gases, for instance.

8.4 Weak Interactions

The mechanism of complexation via noncovalent interactions is different from the covalent bonding formation in that there is no electron pair being formed. According to Feynman,⁷ the proximity of the two species produces a polarization of the electrons around each atom, which implies a displacement of the respective negative center of charge, and, as a result, each nucleus is attracted by the distorted charge distribution of its *own electrons*, giving rise to the attractive force between the monomers. This mechanism explains the stabilization of molecular complexes via noncovalent interactions, and somehow it contradicts the traditional intuition that states that once two closed shells get in contact then an inevitable repulsion must occur. That may be true

only if the involved atoms have a very low polarizability, i.e. if their electron densities behave as point charges. Hence, closed-shells certainly exert valence repulsion (via electron-pair pressure, mainly expressed by the kinetic energy density), but the polarizability makes possible that closed shells get deformed and in closer contact than expected, as indeed is found in many chemical systems. The BCP found in the neon dimer, for instance, reveals the presence of a noncovalent interaction and the stability of the complex is accounted entirely by the energy of the system, including the dispersion term that gives rise to an attractive mutually induced dipolar interaction.

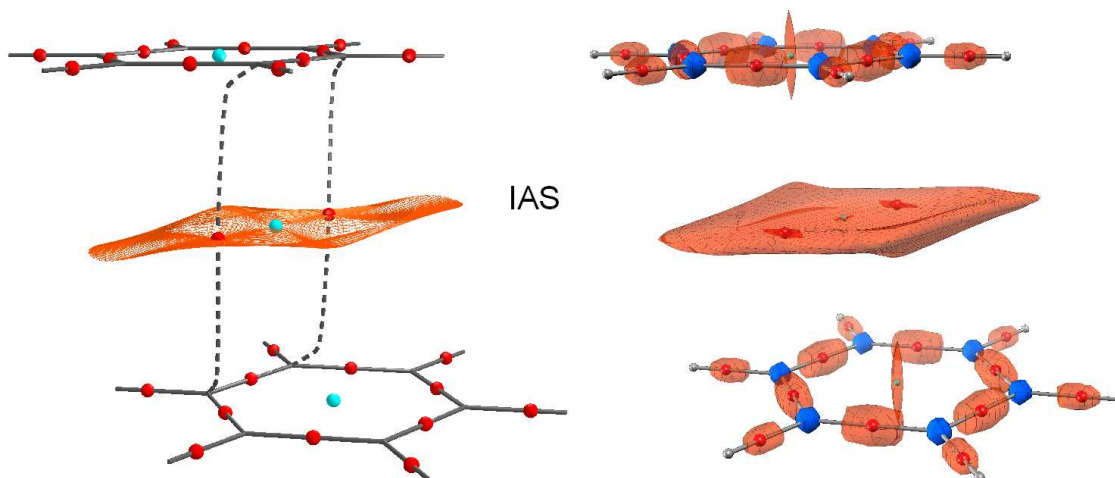
The Laplacian at the BCP is always positive for noncovalent interactions, $|\lambda_3| > |\lambda_1| + |\lambda_2|$, making the axis of the ellipsoid along the bond path (c) smaller than the two transverse axes, i.e. $c \ll a \approx b$. As a result, LED isocontours around noncovalent BCPs are large disc-shaped regions. By considering the zero-flux condition that locally defines the IAS, $\nabla\rho \cdot d\mathbf{a} = 0$, it is easy to show that the isocontours $\tilde{p} \sim 0$ for noncovalent interactions are discs parallel to the IAS originated at the BCP.^v It must be recalled that the IAS is a bundle of gradient lines, and hence $\nabla\rho$ is always parallel to the IAS at the neighboring points. This observation is key to understand the critical points located between noncovalently-bonded molecules. The isocontours $\tilde{p} \sim 0$ in these regions follow the interatomic surfaces which pass through the adjacent saddle critical points, including ring critical points. The benzene dimer shown in Fig. 8.6 illustrates the situation described and reveals the similarity between LED and the intermolecular surface. The respective QTAIM molecular structure is also included with the BCPs represented as red spheres and the RCPs as cyan spheres. The same representation of these critical points is included with the molecular structure underneath the LED isocontours. It is clear the similarity between the intermolecular isocontour $\tilde{p} \sim 0$ and the IAS dividing the two monomers. The noncovalent BCPs of the benzene dimer have a very low electron density ($\rho_{BCP} = 0.003 e$) that is almost the same for the central RCP ($\rho_{RCP} = 0.002 e$), besides they are separated by a distance of $2.08 a_0$ that yields an estimate of $\rho' \sim 0.0005$ and hence $\tilde{p} \sim \rho'/2\rho \sim 0.08 au$, which is almost a fourth of the plotted isocontour ($0.35 au$).

^vBecause of the expansion of the gradient around the BCP in terms of the principal directions (8.1), the local flux condition implies

$$\lim_{\mathbf{r} \rightarrow \mathbf{r}_0} \nabla n(\mathbf{r}) \cdot d\mathbf{a} \rightarrow \lim_{z \rightarrow z_0} (z - z_0) = 0$$

for any $x = x_0 + \delta x$ and $y = y_0 + \delta y$. This is, for a noncovalent bond, the isocontours around the BCP are elliptic regions adhered to both sides of the IAS, because the ellipsoid equation 8.2 has symmetrical solutions for $\pm z$.

Figure 8.6: QTAIM and LED for benzene dimer. LED isocontours are 0.35 (red) and 5 (blue) au.



Hence, the intermolecular regions are characterized by a very low and almost constant electron density, implying that $\tilde{\rho} \sim 0$ are flat isocontours parallel to the IAS containing adjacent saddle points (RCPs and BCPs). These interface regions of almost constant density separates the two monomers, and hence they may have physical implications for their dissociation. Further research is required to fully understand how the molecular counterparts use these constant-density interfaces as a mechanism for complexation and decomplexation.

In some cases, the presence of intermolecular regions of locally-constant density indicates an incipient formation of an additional BCP, or the proximity to a of *catastrophe* structural conformation. The formation and breaking of BCPs have been studied extensively by Bader with the use of *dynamical systems theory*, where these catastrophe conformations are defined.⁹ Some configurations, not necessary at the geometrical equilibrium, give rise to new BCPs and other critical points. That is the case of the formaldehyde-water complex, which only has one critical point between the oxygen of the formaldehyde and one hydrogen of water, but another configuration with two hydrogen bonds can be found. At the equilibrium geometry, LED shows a portion of the interatomic surface between the oxygen of water and one of the hydrogens of formaldehyde (Fig. 8.5). Nevertheless, from the polarizable regions (transparent yellow) it is easy to tell which of the intermolecular regions involve the formation of a hydrogen bond: the broken valence regions (in yellow) of the formaldehyde oxygen and the adjacent hydrogen of water clearly indicates that their interaction includes charge transfer, while the oxygen of the water molecule has an almost intact (symmetrical)

valence region (yellow). Therefore, these interface regions (in red) are topologically related to the stabilization of the molecule. In Fig. 8.5 the electron density isocontours complements these observations: the isocontour lines that pass through the intermolecular BCP cross the LED point, while the nearest isocontour $n \sim 0.01$ au is almost parallel to the LED isocontour. The outermost isodensity $\rho \sim 0.005$ is half of the previous one, within a distance of $\sim 1 a_0$ meaning that LED is ~ 0.25 au, which is close but smaller than the plotted isocontour value $\tilde{p} = 0.45$ au.

The hydrogen-hydrogen BCP found at the *bay region* of phenanthrene is an example of an unusual type of interaction.¹²⁸ Some authors argue the stabilizing role of this bonding interaction, because it contradicts *chemical intuition*. While the existence of a bond path linking the two hydrogens is not questioned, the corresponding stabilizing energy is. In brief, the argument is that the system cannot be stabilized due to “nonbonded steric repulsion between these hydrogen atoms”.¹³⁰ Nevertheless, energy terms accounting for the “steric repulsion” are not uniquely defined, which leads to an endless discussion. The classical explanation for the enhanced stability of kinked polycyclic benzenoids over their linear isomers is given by Clar’s π -sextets rule, but from the topological standpoint an additional argument can be drawn.

An inspection to the LED isocontours (Fig. 8.7) reveals that, once the two bay hydrogens are forced to be close, they get polarized in response. The dark-yellow isocontours in Fig. 8.7 show the polarization regions of the hydrogens, with a noticeable shrinking of the interconnected ones. Such polarization decreases an otherwise repulsive interaction, as it is reflected in certain physical parameters. The bay hydrogens have smaller basin volumes than the other hydrogens of the molecule, which increases its average electron density^{vi}. By using the density at the H–H BCP as a reference ($\rho_b = 0.01305$ au.), a relative scale for the average densities within the hydrogen basins can be plotted, as shown in Fig. 8.8a. It is clear that the average density within the bay hydrogen is significantly greater than at the BCP ($\sim 125\%$), and higher than at any other hydrogen. This result means that the probability of finding an electron at the basin of any bay hydrogen is about 25% greater than finding it at the BCP between the two hydrogens. The rise of the average density within the basin of H_{bay} is the result of the compression of the electronic population, that in turn shields the nuclear charge and therefore reduces the nuclear-nuclear repulsive effect between hydrogens.

^{vi}The average density is the ratio between the atomic population and the respective integration volume, $\bar{\rho} = N_{\Omega}/V_{\Omega}$. The IAS were limited by the isocontour $\rho = 0.0004$ au.

Figure 8.7: LED for phenanthrene. The isocontours are 0.35 au (red), 0.70 au (orange) and 1.15 au (dark yellow). The gray isocontours correspond to the electron density at the H–H BCP, and the outer limit of the atomic basins, with their respective values as indicated.

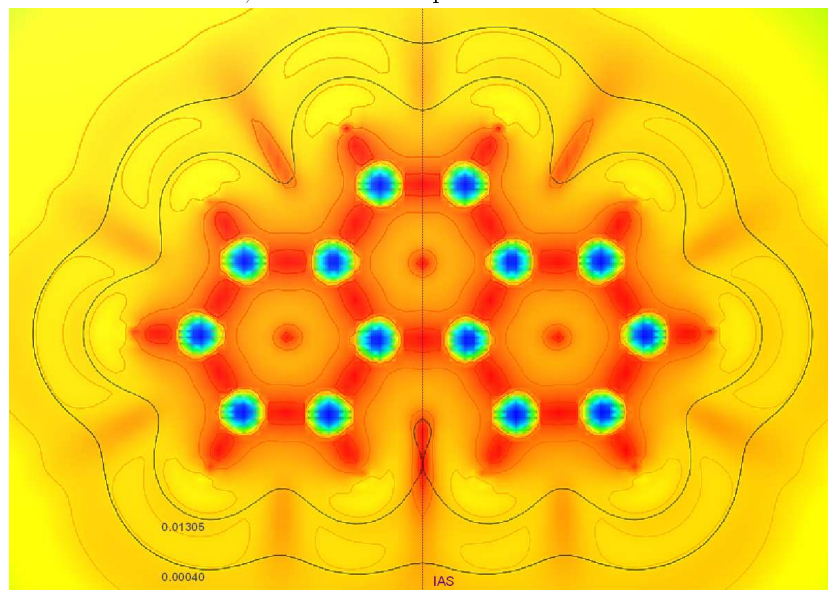
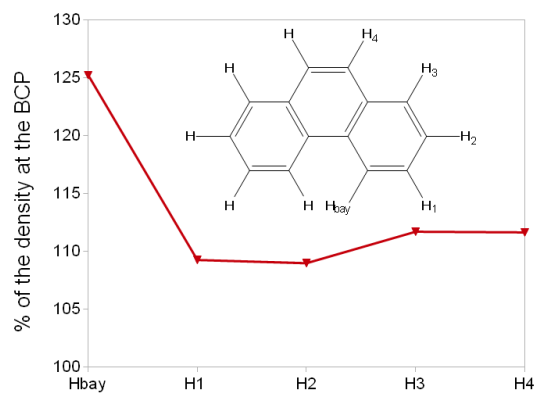
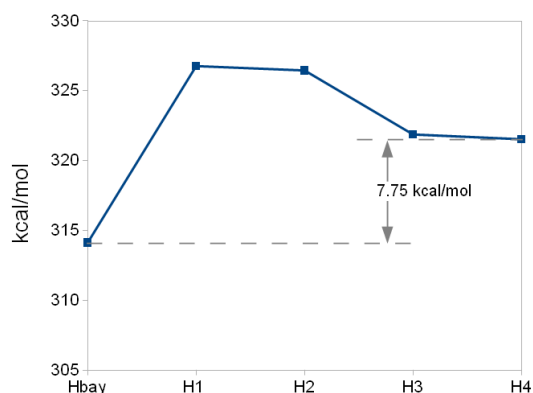


Figure 8.8: Relative density and steric pressure at hydrogen atoms in phenanthrene.

(a) Average density at hydrogens.



(b) Steric pressure at hydrogens.



The steric pressure can be quantitatively estimated by the Weizsäcker kinetic energy density, according to Liu.⁵ The expectation value of this term, $K_w = \int_{\Omega} \tau_w dv$, decreases significantly for the bay hydrogen when compared with the others, as shown in Fig. 8.8b. The HH interaction makes K_w for H_{bay} to be -7.75 kcal/mol lower than the lowest K_w value of the remaining hydrogens (H₄). The hydrogen atoms H₃ and H₄

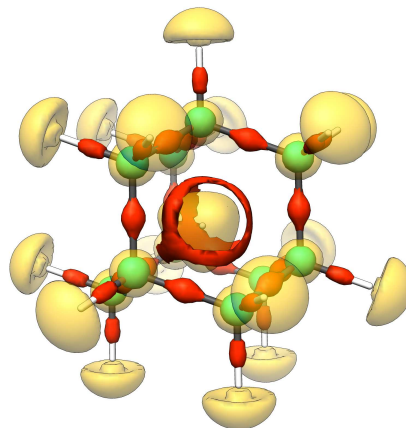
in Fig. 8.8a are the second pair at close proximity, and a near-constant-density region appears between them, as indicated by the red zone in Fig. 8.7. The bonds CH_3 and CH_4 are mutually parallel (1.37°), which forces the hydrogens to have a decrease of the steric pressure of -4.9 kcal/mol with respect to the values for H_1 or H_2 , the most benzene-like of the phenanthrene hydrogens (Fig. 8.8b).

Evidently, not all the stabilization of the molecule can be attributed to the K_w term, but it directly accounts for the polarization of the electron density in the atomic basins. The total energy difference between H_{bay} and H_4 is -5.49 kcal/mol. If the stability of phenanthrene over its linear isomer (anthracene) is accounted only by their hydrogen atoms, an estimate based only on the atomic energies of the mentioned atoms favor the phenanthrene with a lower energy of ~ 11 kcal/mol^{vii}, which is close to previous estimates of ~ 12 kcal/mol, which were obtained at the RHF/6-31G(d,p) level with a different approach.¹²⁸ The decreasing of the steric pressure at the bay hydrogens and the slight increment of the electron density at the BCP of the H_{bay}C bond indicates an electron density polarization away from the BCP between the bay hydrogens, in response to their close contact. This mechanism is consistent with the cited mechanism for weak interactions introduced by Feynman.

A final example of noncovalent interactions that has been the source of certain debate about the BCP concept in QTAIM is the He trapped inside adamantane (He@Ada), an example of what Wang, Qiu and Schwarz called a “confinement bonding”.¹³¹ The helium in adamantane is unable to interact covalently with the electrons of adjacent atoms, and hence it is trapped inside a noncovalent cage, whilst its electrons get squeezed-in as a reaction to the pressure of the surrounding electrons, as it is illustrated by LED in Fig. 8.9.

^{vii}An estimate of the energy can be done by assuming that there are three types of hydrogen atoms in phenanthrene, namely two H_{bay} , four H_{\parallel} and four H_{β} while anthracene has six H_{\parallel} and four H_{β} . Hence, the difference of hydrogen energies is given by $\Delta E_H = 2E_{H_{bay}} + 4E_{H_{\parallel}} + 4E_{H_{\beta}} - (6E_{H_{\parallel}} + 4E_{H_{\beta}}) = 2(E_{H_{bay}} - E_{H_{\parallel}}) = -10.98$ kcal/mol.

Figure 8.9: LED for He@adamantane. The plotted isocontours are 0.28 au (red), 1.21 au (transparent yellow) and 3.00 au (green).



The BCPs between the helium and the surrounding carbon atoms make a tetrahedral cage that also encloses the RCPs (in red). This region is characterized by a near constant electron density, with $\rho_b = 0.079$ au for the BCP and $\rho_r = 0.060$ au for the RCP. The red cage around the helium visualizes the *confinement bonding* and exemplifies a case for which the critical points are less informative than the continuous representation provided by LED.

The trapped helium gets strongly polarized, as it is shown by the tetrahedral shape of the surrounding yellow region, with a marginal increase in its electronic population ($0.075e$). As in the case of the HH interaction in phenanthrene, the trapped helium shrinks in response to the steric pressure made by the adjacent carbon atoms. The free volume of helium is estimated as $V_{free} = 14.89 \text{ \AA}^3$, which is about seven times bigger than the volume at the adamantane $V_{@Ada} = 2.26 \text{ \AA}^3$. The average density is altered in a similar proportion, which for the free atom is $\bar{\rho}_{free} = 0.0198$ au and for the trapped one $\bar{\rho}_{@Ada} = 0.1372$ au. Consequently, the electrostatic interaction made by the helium nucleus is being strongly shielded by the polarization. The steric pressure given by the expectation value of the Weizsäcker energy is lowered from the free atom to the trapped one by $\Delta K_w = -333.71$ kcal/mol. It is remarkable that the change in steric pressure is very close in magnitude to the reported ‘‘Pauli’’ contribution to the dissociation energy of 319.31 kcal/mol.¹³¹ The dissociation energy for the system is -140.68 kcal/mol, which means that the steric pressure is $\sim 237\%$ the magnitude of the dissociation energy.

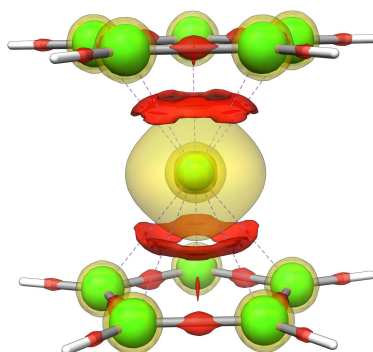
The counterintuitive behavior of the steric pressure term K_w can be understood by

considering the shell contraction in atoms: as the outermost shells get more populated, the inner shells get contracted in response, shielding the nuclear charge. The bay hydrogens in phenanthrene and the confined helium within the adamantane react similarly to the confinement conditions imposed by their respective noncovalent interactions. Simultaneously, by reducing the volume, the effective surface is reduced, reducing therefore the contact area, and hence the steric effect. Notice that here Liu’s interpretation of the kinetic energy density term associated with the local momentum \tilde{p} is used, and no reference to any sort of “force” is necessary as the systems studied are in equilibrium and no net forces on the nuclei might appear.

8.5 Metallic Bonding

The last kind of stable interatomic interaction to be discussed here is the metallic bond, that is abundant in solid state matter and in several organic systems like metalloenzymes.

Figure 8.10: LED for ferrocene. The plotted isocontours are 0.10 au and 0.20 au (red), 1.38 au (transparent yellow) and 3.00 au (green).



Organometallic bonding is a class of stable interaction that has significant differences from covalent interactions, while keeping the same essential topological features.¹³² The multiple-coordinated bond (i.e. involving several BCPs) in metal-ring systems, for instance, form a kind of *collective* bonding interaction with the nuclei involved in the π delocalization of the ring. Finding those topological features might be challenging for certain systems, according to recent experimental and theoretical works on transition metals and carbocyclic rings (MC_{ring}) by Farrugia et al.: “The

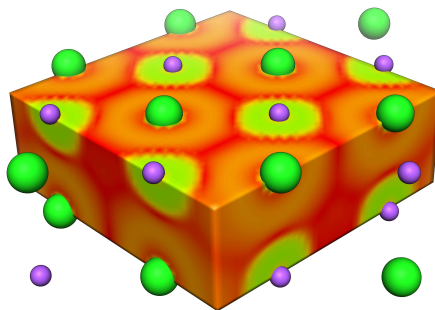
interaction of a transition metal with a closed or open π -hydrocarbyl ligand generates a zone of very flat ρ and very low $\nabla\rho$ in which any MC bond critical points must be located. This almost inevitably leads to a MC_{ring} topology that is close to a catastrophe situation, and as a result, the number of bond paths cannot be predicted with any certainty. In such circumstances, a localized description of the chemical bonding in terms of the individual MC_{ring} bond paths is not completely satisfactory. Continuous QTAIM indicators, which do not rely on the presence of a BCP, such as the delocalization indexes and the source function, provide a picture of the chemical interactions that is more in line with chemical experience.”¹³³

LED can be considered as a continuous QTAIM indicator, because it locates the critical points and their respective regions by the local symmetry of the electron density, as it is illustrated in Fig. 8.10 for ferrocene. An inspection to the BCPs involving the iron atom reveals that the BCPs and RCPs are located almost at the same distance from the iron, and they have very low but similar electron densities of 0.1129 au and 0.1079 au, respectively. The fact that the electron density between the metal and the carbon atoms is almost constant implies that $\tilde{p} \sim 0$ creates isosurfaces around the iron that enclose all the adjacent BCPs and RCPs, creating a sort of electronic shell that represents the collective metal-ring bond. The polarization region (transparent-yellow) also shows the strong influence the π rings have on the iron atom. The iron core shell (green) is visibly smaller compared to those of the carbon atom: a contraction of the inner shell that is expected for a heavy nucleus.

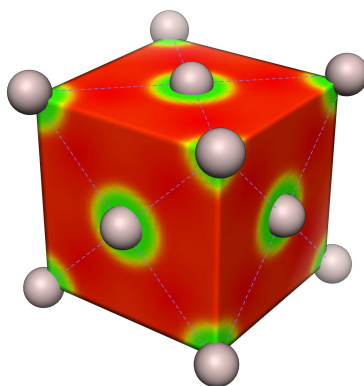
With the advent of nanotechnology a renewed interest in cluster chemistry has emerged in recent years. One interesting question is the appearance of bulk versus surface properties such as the *metallicity*. Topologically, the metallic bonding is not different than other covalent interactions, except for the ability of metals to form multiple-coordinated stable bonds. In Fig. 8.11 plots for $\text{Na}_{16}\text{Cl}_{16}$ and of Al_{14} clusters are shown, both computed at the LSDA/cc-pVDZ level of theory. In the ionic crystal, the sodium atoms are positively charged (0.88 e), and hence they attract the surrounding electron density, forming characteristic square red zones around the big green-yellow zones in Fig. 8.11a. The negatively charged chlorine have a octagonally-shaped red region that connects different critical points, since the electron density in those regions is almost constant: 0.0056 au for ClCl BCPs and 0.0052 au for the RCPs. These regions are joined to the interatomic surfaces between Na and Cl, by BCPs with an electron density just twice bigger than the other BCPs of 0.0116 au. The red regions are therefore zones where electrons are likely to be found, regions

Figure 8.11: LED for clusters involving metals. The color code is the same as in 8.1.

(a) NaCl cluster.



(b) Aluminum cluster.



where delocalization takes place, but they are mainly restricted to the boundaries of the interatomic surfaces, as shown in Fig. 8.11a.

The aluminum cluster shown in Fig. 8.11b, on the other hand, shows a practically total delocalization of the valence electrons. In fact, the electron density at BCPs is 0.029 au while for RCPs is 0.028 au and for CCPs is 0.026 au. In other words, for most of the interatomic regions the electron density is very low and practically constant, a particular feature of conduction electrons which are located within an "electron sea", clearly visible by LED red regions. The nuclei are attracted to this delocalized electronic presence by electrostatic interactions. Hence, LED illustrates consistently the electronic regions of metallic clusters, and further research is required to address the differences between bulk and crystal electronic regions with more detail. Here it

is shown that LED is able to detect these differences, but quantitative estimates and the comparison with conventional solid state procedures is matter of future work.

8.6 Discussion and Conclusions

In the present chapter a variety of bonding interactions ranging from noncovalent to metallic has been discussed. It has been illustrated how LED provides a consistent and intuitively evident description of the different electronic regions that can be found in molecules, complexes and crystals. The study of the critical points as given by LED provides new criteria for their classification that complements the conventional QTAIM approach. The new criteria makes emphasis on the symmetry of the electron density at the critical points, as accounted by the local curvatures. Such analysis is generally valid for continuous fields of the form $\nabla\rho/\rho^m$ for whose numerator goes to zero faster than the denominator at the critical points. The studied variable, i.e. the local momentum $\tilde{\mathbf{p}} = -\hbar\nabla\rho/2\rho$, has convenient physical boundary limits, which simplifies its graphical and conceptual analysis.

LED can be conceptually defined independently from DFT or QTAIM, as has been shown before,⁹⁰ and, interestingly, the local momentum $\tilde{\mathbf{p}}$ (i.e. LED) is directly connected with the Weizsäcker kinetic energy, which links both theories in a simple way. With the present work, a theoretical bridge between DFT and QTAIM can be formally established. Indeed, the use of expressions like LED or its explicit kinetic energy form have been increasingly included in different DFT models. A remarkable effort to rationalize the role of this variable, and the Laplacian in DFT, has been addressed by Liu in his interpretation of the *local quantum potential* as a measure of the molecular stericity.⁵ The results reported here support such a view by explaining the counter-intuitive nature of HH interactions in phenanthrene and the confinement bonding of He@Ada. In addition, Vydrof uses the local kinetic energy term associated with $\tilde{\mathbf{p}}$ as an explicit *local polarizability* that is responsible for the long-range dispersion interactions in molecules, which he successfully tested recently on a DFT model.¹²⁵ LED assumes identical interpretation and this, in fact, is the central argument for the definition of the experimental atomic radii table that have been recently proposed.⁹⁹ In the present work the polarizable regions found in different bonding contexts complements QTAIM critical point analysis by providing information about the deformation of the atomic electron shells after the bonding interaction is being formed.

This renewed interest in kinetic energy terms that depend on the ratio of the gradient over the density ($\tilde{\mathbf{p}} = -\nabla\rho/2\rho$) indicates that this apparently simple variable unveils a wealth of chemical information that is useful to model all sorts of interatomic interactions, as it is shown here. Recently Yang et al. reported an analysis of the the *unit-less gradient*, s ,¹³⁴ that is defined as the ratio between the local momentum \tilde{p} and the Fermi momentum $p_F = \hbar(3\pi^2)^{1/3}\rho^{1/3}$, or $s = \tilde{p}/p_F$. Relative variables have certain relevance in DFT, but for graphical analysis they represent both conceptual and, sometimes, practical challenges. For instance, the core variable in ELF is a relative quantity: it is the difference between the total kinetic energy density and the Weizsäcker kinetic energy density, divided by the Thomas-Fermi kinetic energy density. In order to be plotted, it is mapped onto the $[0, 1]$ range via a sigmoid function.⁴ In the study proposed by Yang et al., the unit-less gradient is used because it is one of the pervasive variables in DFT formulae. While the divergent behavior of this variable for the tail region is not discussed (i.e. $s \rightarrow \infty$ for $r \gg$), in their implementation of the numerical approach they have to trim the values of the function in order to get useful visualizations. They focus on the representation of weak interactions for $s \rightarrow 0$, but the obvious connection of these regions with the critical points analysis in QTAIM is missing. As a result, they color the interatomic regions according to the *sign* of the second eigenvalue of the Hessian, which plays by itself a minor role in distinguishing noncovalent interactions. They claim that $sign(\lambda_2)$ differentiates attractive from nonbonded interactions. In fact, the only difference that can be seen by this choice is between BCPs ($\lambda_2 < 0$) and RCPs ($\lambda_2 > 0$). As was observed in recent publications (prior to Yang’s paper),^{112,124} and as it is discussed here extensively, these two types of saddle points are simply connected by $\nabla\rho \sim 0$ isocontours, which are parallel to the interatomic surface that separates the complexed monomers (e.g. as shown in Fig. 8.6 for the benzene dimer). In summary, the claims by Yang et al. are better understood in terms of the properties of the critical points here exposed, which are induced by the study of LED carried out by the authors during the past five years.

The present chapter highlights the relevance that the atomic polarization has on the equilibrium geometries and bonding interactions. The polarizability regions are distinctively plotted by LED, a feature not available with other continuous analyses. For certain systems, such as phenanthrene, the polarization on hydrogens is the key to understand the HH counter-intuitive stabilizing role. Consistently, a quantitative estimate of the atomic “steric effect” is given by the average value of the single-particle

kinetic energy density, $\int_{\Omega} \tau_w dv$. The results described here explain how LED constitutes an insightful and easy-to-interpret tool for the exploration of new stabilizing interactions for all kind of molecular systems. Because LED is given in terms of the electron density alone it is independent of the level of theory used and can be applied to experimentally-obtained electron densities as well.

Chapter 9

Quantum Molecular Similarity in Terms of Physical Observables

9.1 Introduction

Recent advances in quantum similarity measures have paved the way for the *quantum quantitative structure-activity relationship* (QQ-SAR). Molecular similarity is a theoretical tool for the virtual screening of molecules by the approximate estimation of their relative properties such as absorption, distribution, metabolism, excretion, and toxicity, among others. It is also useful for the prediction of several physicochemical properties: solubility, water-octanol partitioning coefficient, etc.¹³⁵ Given the current status of computational quantum mechanics, drug designers may be increasingly interested in these methods.

According to Carbo, the *maximum* overlap of the electron densities of two molecules gives their quantum similarity (QS) measure.¹³⁶ The central computational task involved in this quantum similarity measure, i.e., the mutual alignment of the two compared molecules is CPU expensive. This maximization procedure consumes most of the overall CPU cost involved in a QSAR study based on this definition of quantum similarity.

The similarity measure can be transformed into an equivalent equation that does not require the molecular alignment step. This is a promising procedure as the obtained expressions are directly connected to the quantum theory of atoms in molecules (QTAIM) and, therefore, the molecular similarity theory can benefit from its formalism. The formal aspects of the quantum similarity definition are the subject of this chapter, and how the derived expressions might reduce the actual CPU cost involved in QSAR studies. The molecular similarity can be reduced to the comparison of observable variables computed at the critical points of the electron density and to their averaged values. Hence this work connects the electron density topology with the molecular similarity concepts.

9.2 Quantum Molecular Similarity

The *molecular similarity principle* states that structurally similar molecules exhibit similar physicochemical and biological properties. The similarity I_{AB} between two molecular systems characterized by their respective electron densities ρ_A and ρ_B was first proposed by Carbó as the quantity¹³⁶

$$I_{AB} = \max \left\{ \frac{S_{A,B}}{\sqrt{S_{A,A}S_{B,B}}} \right\} \quad (9.1)$$

where the overlapping integrals S_{AB} and S_{AA} are defined as

$$S_{AB} = \int \int \rho_A(\mathbf{r})G(\mathbf{r}, \mathbf{r}')\rho_B(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad (9.2)$$

$$S_{AA} = \int \int \rho_A(\mathbf{r})G(\mathbf{r}, \mathbf{r}')\rho_A(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad (9.3)$$

Here, $G(\mathbf{r}, \mathbf{r}')$ is a weight function such as the Coulomb operator, or other density function. By using the Kronecker delta

$$G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \quad (9.4)$$

the original equations proposed by Carbó-Dorca are obtained, i.e.

$$S_{AB} = \int \rho_A(\mathbf{r})\rho_B(\mathbf{r})d\mathbf{r} \quad (9.5)$$

$$S_{AA} = \int \rho_A^2(\mathbf{r})d\mathbf{r} \quad (9.6)$$

Clearly, $I_{AB} \in [0, 1]$. The molecular similarity measure (Eq. 9.1) requires an optimal alignment between molecules A and B . But such requirement is hard to guarantee. From the classical mechanics of rigid bodies it is clear that the general solution for the alignment requires exactly two group transformations,¹³⁷ a rotation $\hat{R}_{\mathbf{r},\mathbf{r}'}$ and a translation $\hat{T}_{\mathbf{r},\mathbf{r}'}$ that transforms the vectors defined in the system of coordinates o onto the coordinates o' . This computational task has no obvious theoretical procedures and numerical solutions are usually required. The alignment can be included

within the operator $G(\mathbf{r}, \mathbf{r}')$

$$G(\mathbf{r}, \mathbf{r}') = \hat{R}_{\mathbf{r}, \mathbf{r}'} \hat{T}_{\mathbf{r}, \mathbf{r}'} F(\mathbf{r}, \mathbf{r}')$$

where the general weight function $F(\mathbf{r}, \mathbf{r}')$ is transformed by the rotation and translation operators that aligns the two molecules. An exhaustive search of an optimal alignment determines the success on evaluating Eq. 9.1. Due to the lack of analytical procedures for solving the transformation $\hat{R}_{\mathbf{r}, \mathbf{r}'} \hat{T}_{\mathbf{r}, \mathbf{r}'}$, the implemented numerical methods entail two main approximations: a) a finite-grid exhaustive overlap maximization search, which requires that b) the electron densities are approximated by promolecular ones. Hence, in practice, the method dictates the use of approximate electron densities in order to be efficient. The validity of promolecular approximations is beyond the present work's boundaries, but it is sufficient to mention that a promolecular density mimics the *ab initio* one but will lose its topological features which are critical in QTAIM.

9.3 On the CPU Cost of Molecular Similarity

The overlap of the two densities as a measure of quantum similarity is a definition that entails several practical shortcomings. An estimate of the computational demands of this similarity measure when implemented in QQ-SARs studies is provided.

In order to compare N molecules, a total of $N - 1$ alignments are required to maximize their mutual overlaps (Eq. 9.1). It can be assumed that the molecules are all equally centred and it only remains to perform the rotation of the two densities to complete the alignment. In addition, it can be assumed that all the computations have the same grid resolution, given by m^3 , with m being the number of points considered in each direction of the cube file containing the density. An unsupervised numerical alignment between two densities implies a number of steps that is at least of the order of the combinatorial number of possible orientations the two grids can have, which is practically infinite, as the required trial rotations are not restricted *a priori*.

The minimal number of unsupervised rotations required can be obtained by assuming the simplest case of all the different 90 degree rotations. Hence a minimum of $6(N - 1)$ overlap integrals are required to get an approximate estimate of the correct alignments. The similarity matrix for this set requires a total of $N^2/2$ similarities (the diagonal elements do not require an alignment), hence a total of $3N^2(N - 1)$

integrals must be done. Numerical integrals are reduced to sums when cube grids are used, and therefore the number of simple arithmetic operations involved on building the matrix is finally $3N^2(N-1)m^3$. Thus, the CPU cost grows cubically with the size of the grid and the number of molecules simultaneously. Small grid sizes typically involve $m \sim 50$, and a small set is $N \sim 20$, which translates into a total of $\sim 2.85 \cdot 10^9$ arithmetic operations.

Therefore around 3 billion arithmetic operations are required in order to build the similarity matrix for a set of 20 molecules. Any practical computation requires at least two orders of magnitude more computations to find the correct alignment between the molecules. Accordingly, it is safe to say that on average the quantum similarity requires around $\sim 10^{12}$ of these basic arithmetic operations. Obviously this is merely an estimate in order to illustrate the CPU cost involved, and not the usual practice which involves more efficient methods.

If, on the other hand, the similarity matrix is built by using the averaged properties computed *separately* for each molecule, this task requires approximately $Nm^3 + gN(N+1)/2$ operations, where g is the number of scalar properties used for characterizing the moleculesⁱ. An estimate of 2,501,900 arithmetic operations is expected if ten properties are computed, (i.e. assuming $g = 10$) and using the same parameters than previously. Therefore, a total of ~ 2.5 million of simple arithmetic operations are required in order to build the similarity matrix. This is a difference of at least six orders of magnitude in CPU cost between the similarity measure and a direct comparison of the averaged properties independently computed for each property.

What motivates the use of such a CPU-demanding QQ-SAR? Equation 9.1 has a quantum form that seems to be very compelling over other formulations. As it is shown in the following section, Eq. 9.1 can be mathematically transformed into an equivalent expression that does not require the alignment and hence its CPU cost is reduced consequently.

9.4 From Quantum Similarity to Observable Comparison

The quantum similarity index I_{AB} (Eq. 9.1) depends on the term $\rho_A(\mathbf{r})\rho_B(\mathbf{r}')$ which can be equivalently written as

$$\rho_A(\mathbf{r})\rho_B(\mathbf{r}') = \frac{\rho_A^2(\mathbf{r})}{2} + \frac{\rho_B^2(\mathbf{r}')}{2} - \frac{(\rho_A(\mathbf{r}) - \rho_B(\mathbf{r}'))^2}{2} \quad (9.7)$$

ⁱHere it is assumed that the averaged properties are obtained from grid integrals.

After replacing previous equation into the overlapping expression, Eq. 9.2, it becomes

$$S_{AB} = \int \int R(\mathbf{r}, \mathbf{r}') T(\mathbf{r}, \mathbf{r}') \left\{ F(\mathbf{r}, \mathbf{r}') \delta(\mathbf{r}, \mathbf{r}') \left(\frac{\rho_A^2(\mathbf{r})}{2} + \frac{\rho_B^2(\mathbf{r}')}{2} - \frac{(\rho_A(\mathbf{r}) - \rho_B(\mathbf{r}'))^2}{2} \right) d\mathbf{r} d\mathbf{r}' \right\}$$

which leads to

$$S_{AB} = \frac{S_{AA}}{2} + \frac{S_{BB}}{2} - \int \int \hat{R}_{\mathbf{r}, \mathbf{r}'} \hat{T}_{\mathbf{r}, \mathbf{r}'} \left\{ F(\mathbf{r}, \mathbf{r}') \frac{(\rho_A(\mathbf{r}) - \rho_B(\mathbf{r}'))^2}{2} d\mathbf{r} d\mathbf{r}' \right\} \quad (9.8)$$

if $\Delta\rho(\mathbf{r}; \mathbf{r}') = \rho_A(\mathbf{r}) - \rho_B(\mathbf{r}')$ then the third term is the self-overlapping of the difference density $\Delta\rho(\mathbf{r}; \mathbf{r}')$, $S_{\Delta\Delta}$, and hence

$$I_{AB} = \frac{1}{2} \frac{1}{\sqrt{S_{AA} S_{BB}}} \{S_{AA} + S_{BB} - S_{\Delta\Delta}\}$$

Clearly, the similarity I_{AB} is a relative measure of the self-overlapping of the densities difference with respect to the sum of the two separate self-overlapping S_A and S_B . If $\Delta\rho(\mathbf{r}; \mathbf{r}') = 0$ and $\mathbf{r} = \mathbf{r}'$, then $S_{\Delta\Delta} = 0$, $S_{AA} = S_{BB}$ and $I_{AB} = 1$. If $\Delta\rho(\mathbf{r}; \mathbf{r}') \neq 0$ and $S_{\Delta\Delta} \gtrsim 0$, then $I_{AB} \approx \frac{1}{2} \frac{1}{\sqrt{S_{AA} S_{BB}}} \{S_{AA} + S_{BB}\}$, i.e. I_{AB} is some kind of geometric mean of the two self-overlapping values. It appears that $S_{\Delta\Delta}$ has a more direct interpretation than I_{AB} , and therefore it is more convenient to work with $S_{\Delta\Delta}$ as a measure of similarity.

The computation of the integral $S_{\Delta\Delta}$ also requires the alignment between $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r}')$, but can be transformed into a linear operation. By defining a general weight function as $M(\mathbf{r}, \mathbf{r}') = (\rho_A(\mathbf{r}) - \rho_B(\mathbf{r}')) F(\mathbf{r}, \mathbf{r}')$, the third term of Eq. 9.8 becomes

$$S_{\Delta\Delta} = \int \int \hat{R}_{\mathbf{r}, \mathbf{r}'} \hat{T}_{\mathbf{r}, \mathbf{r}'} \{M(\mathbf{r}, \mathbf{r}') (\rho_A(\mathbf{r}) - \rho_B(\mathbf{r}')) d\mathbf{r} d\mathbf{r}'\} \quad (9.9)$$

The integral can be now split into two separate integrals whose values should be aligned if the properties are tensors,

$$S_{\Delta\Delta} = \hat{R}_{\mathbf{r}, \mathbf{r}'} \hat{T}_{\mathbf{r}, \mathbf{r}'} \left\{ \int M(\mathbf{r}) \rho_A(\mathbf{r}) d\mathbf{r} - \int M(\mathbf{r}') \rho_B(\mathbf{r}') d\mathbf{r}' \right\} \quad (9.10)$$

It is more practical to work with the two separated integrals of Eq. 9.10, each in its own coordinate system, than with the original overlap proposed by Carbo. In fact,

$S_{\Delta\Delta} = 0$ if both molecules exhibit identical values for the integrand. If the result of the integral is a tensor variable, the averaged values may be aligned *a posteriori*, which simplifies substantially the similarity measure. It is important to notice that the weight function $M(\mathbf{r})$ is a property density and hence the integral value is just its expectation value with respect to the density coordinate system

$$\langle M_A \rangle_{\Omega} = \int_{\Omega} M(\mathbf{r}) \rho_A(\mathbf{r}) d\mathbf{r} \quad (9.11)$$

where Ω represents the integration region, in general it is the whole molecular volume. Equation 9.11 also suggests that $M(\mathbf{r})$ may be an *observable* property density, with its obvious conveniences. This observation leads to the important result that the similarity measure can be reduced formally to a difference between the expectation values of a given molecular property M computed independently for each molecule

$$S_{\Delta\Delta} = \hat{R}_{\mathbf{r},\mathbf{r}'} \hat{T}_{\mathbf{r},\mathbf{r}'} \{ \langle M_A \rangle_{\Omega} - \langle M_B \rangle_{\Omega'} \} \quad (9.12)$$

In summary, the quantum similarity index can be reduced to a direct comparison of average physical properties computed independently for each of the two molecules. If they are scalar properties, such as the ionization energy or the net charge, there is no need for a mutual alignment at all. In general, physical properties are either scalars or vectors, which are always much easier to align than solving the transformation $\hat{R}_{\mathbf{r},\mathbf{r}'} \hat{T}_{\mathbf{r},\mathbf{r}'}$ for the two electron densities.

Bader has shown that the properties defined by the zero-flux condition optimizes atomic (and fragment) chemical transferability.¹³⁸ This observation guarantees that when a property is assigned to an atom or fragment, these values are found to be very similar whenever the atom or molecule is placed in a similar molecular bonding context. It makes possible the creation of theoretical databases for atomic and fragment properties, which is of prime importance in molecular screening in drug design.

9.5 A Maximum Overlapping Implies Molecular Structure Similarity

The quantum similarity measure depends on the mutual alignment as stated in Eq. 9.1, which guarantees the maximal overlap; in this section, some topological consequences derived from the maximal overlap are outlined.

The overlap maximization of Eq. 9.1 can be transformed into an equivalent local

condition according to which the product $\rho_A(\mathbf{r})\rho_B(\mathbf{r})$ is locally maximal, a condition that translates into $\nabla(\rho_A(\mathbf{r})\rho_B(\mathbf{r})) = 0$ and $\nabla^2(\rho_A(\mathbf{r})\rho_B(\mathbf{r})) < 0$ which leads to

$$\frac{\nabla\rho_A(\mathbf{r})}{\rho_A(\mathbf{r})} + \frac{\nabla\rho_B(\mathbf{r})}{\rho_B(\mathbf{r})} = 0$$

and

$$\frac{\nabla^2\rho_A(\mathbf{r})}{\rho_A(\mathbf{r})} + \frac{\nabla^2\rho_B(\mathbf{r})}{\rho_B(\mathbf{r})} + 2\frac{\nabla\rho_A(\mathbf{r})}{\rho_A(\mathbf{r})} \cdot \frac{\nabla\rho_B(\mathbf{r})}{\rho_B(\mathbf{r})} < 0$$

respectively. These equations admit a set of solutions at the critical points given by

$$\nabla\rho_A(\mathbf{r}) = \nabla\rho_B(\mathbf{r}) = 0$$

that exhibits local accumulation of charge, i.e.,

$$\nabla^2\rho_A(\mathbf{r}) + \nabla^2\rho_B(\mathbf{r}) < 0$$

Hence, the critical points of both molecules must be located at *exactly the same* local maxima positions. In QTAIM language, this means that both molecules must exhibit the same *molecular structure*. This result is directly derived from the similarity measure, which, interestingly, was introduced as an *ad hoc* definition and independently from the quantum topology theory. When this local condition is applied to the case $\rho_B(\mathbf{r}) = \rho_A(\mathbf{r})$, the second condition implies $\nabla^2\rho_A(\mathbf{r}) < 0$, which holds for the regions in space where the negative charges are locally concentrated, e.g. around the nuclei and bond critical points.

Therefore, if the density overlap is taken as a local condition, it implies the coincidence in the molecular structures. In QTAIM the molecular structure is quantized, i.e., for every molecular system in an equilibrium geometry there is a finite number of topological elements linked by gradient paths, as accounted for by the Hopfman-Kopp theorem.¹¹ It is expected that in a comparative study involving a set of molecules, those having more critical points in common are more similar. In other words, similar structures are closer to each other according to Carbo's similarity. But structural changes may occur between a molecule and one of its conformers after a simple geometry optimization, for instance. In order to avoid a spatial dependence in the comparison, the molecular structure can be accounted for the critical points of the electron density. Popelier partially exploits this idea in his QSAR proposal,^{139,140}

which is focused on the bond critical points only, i.e. a subset of the critical points. The results shown here suggest that nuclear attractors and ring critical points also must be included.

In summary, the quantum similarity implies that two molecules may be compared by the properties of their critical points, which is equivalent to comparing their molecular structures.

9.6 Discussion and Conclusions

Chemical design rests on the idea that atoms and groups of atoms exhibit similar local properties within similar bonding environments. It means that functional groups exhibit *transferable* properties. Clearly, molecular similarity and group transferability are interconnected concepts.

Present analysis of the quantum-based similarity definition leads to a topological *ab initio* molecular similarity that is expressed by two complementary analysis: a) a direct comparison of the molecular critical points properties (the molecular structures) and b) the similarity of averaged physical properties, either for the whole molecule or one of its fragments.

The extensive search of molecular alignment required by the quantum similarity measure is unnecessary. Consequently, the computational cost involved in Carbo's similarity measure can be reduced by at least three orders of magnitude. In the present approach the main computational cost is focused on the determination of the molecular wave function, its physical variables and the respective topological properties. In this way, the molecular properties are obtained *ab initio*, not at the promolecular level as in several quantum-similarity-based QSARs.

The equations introduced here provide theoretical support to several QTAIM molecular similarity studies. For example, QSARs using topological properties at the bond critical points have been shown to be useful.^{141, 139, 140}

The author previously explored the quantum similarity of molecular fragments by using physical properties, as derived in previous sections. A direct comparison of the side chains of the amino acids by their electronic polarization and energies as given by QTAIM yields a theoretical classification identical to the experimentally-determined biochemical classification.¹⁴² The results discussed in this chapter indicate that a complementary analysis based on the critical points can be of practical interest as well.

In summary, in the present chapter a crucial step involved in the quantum molecular similarity measure is questioned. As an alternative, a direct comparison of averaged molecular properties is shown to be more cost effective because it does not require the iterative search for an optimal alignment. A practical advantage is the fact that all standard quantum mechanical software packages provide a set of average physical properties, and hence its comparison is just part of a post-processing analysis. In the next two chapters are discussed several methods for post-processing ab initio computations in amino amino acids and small peptides.

Chapter 10

Methods in Biocomputational Chemistry: a Lesson From the Amino Acids

Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

H. J. Bohórquez, C. Cárdenas, C. F. Matta, R. J. Boyd and M. E. Patarroyo, in *Quantum Biochemistry*, 403-421, Ch. **13**, ed. C. F. Matta (2010)

Computer-aided drug design (CADD) requires accurate and fast methods to identify and characterize molecules with potential therapeutic use. While quantum mechanics (QM) provides the best available theoretical framework to predict molecular properties, it is computationally expensive for biologically-relevant molecules, molecules that are usually composed of hundreds of atoms such as proteins and nucleic acids. This practical limitation dictates the use of approximate methods that are fast enough to screen large sets of biochemical compounds. Ideally, these methods are designed to identify molecules with a specific biological activity *in silico*. Predictive methods employed in drug design such as statistical analysis (SA) and molecular mechanics (MM), address different levels of detail of the molecular problem. Statistical methods, based mainly on database records, are designed to provide averaged molecular properties such as secondary-structure propensities or the hydrophobic character of a polypeptide chain. Molecular mechanics (MM) methods provide information about specific functional groups and their interactions in terms of Newtonian (classical) mechanics through force fields parametrized for a given class of biomolecules. This parametrization is based on the results of quantum mechanical computations. Hence, QM plays an indirect but crucial role in approximate MM biocomputational methods. In more recent years, and when a specific reactive centre is known, one can combine QM and MM in a single calculation in what has become known as QM/MM (and its variants).

Here are explained three strategies developed for studying peptides that include statistical analysis over quantum mechanical data to characterize amino acids functional similarities and activity in proteins. These strategies were developed in stages,

each one focusing on a different aspect of the problem. The first stage addresses the question: which theoretical variables describe the conformational trends of the amino acids best? This question underlies the structure-activity relationship (SAR) paradigm, according to which similar structures yield similar bioactivity.¹⁴³ Hence, from this standpoint, it would be advantageous to select the parameters that optimally describe similarity in the construction of quantitative structure-activity relationship (QSAR) models. A principal result from this research is that electrostatic variables sufficiently discriminate structural and chemical trends.¹⁴⁴ This work is summarized in the first section of the present chapter. The second stage of the study examines a smaller set of capped amino acids (AA) model [(HC=O)-AA-NH₂], in two conformations, i.e., 40 molecules in total. The quantum theory of atoms in molecules (QTAIM) was used to analyze the resulting electron densities. The variables studied are the electronic energy and the multipole moments (polarizations) of the amino acid side chains. This set of variables defines a 10-dimensional space. The similarities in this 40-molecule/10-variable system are determined in two ways. The first method is graphical, based on a multidimensional projection known as the Andrews plot.¹⁴⁵ The second is an unbiased pairing method (neighbour joining) that determines similarities on the basis of the distance between the vectors representing each side chain. Remarkably, this procedure is able to replicate the standard biochemical classification of the genetically-encoded amino acids, providing an *ab initio* classification of amino acids, perhaps the first reported so far.¹⁴² The second section provides details about this work and its future extensions.

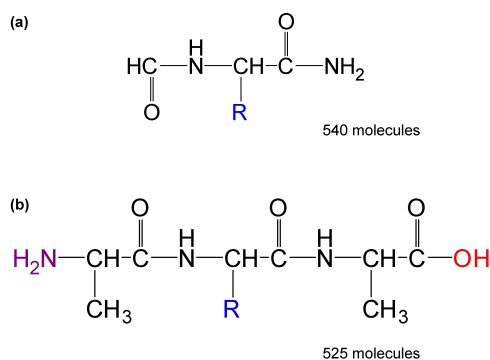
The third stage of this research illustrates the practical application of the previously mentioned findings through a method that incorporates the electrostatic variables for the study of peptide-host interactions.¹⁴⁶ In the last part of the present chapter the advantages of using a Mulliken multipole-based approach to the study of MHC-antigenic peptide complexes are illustrated. Comments about the strengths and future directions of this approach conclude this chapter.

10.1 Conformers, Rotamers and Physicochemical Variables

The number of possible molecules formed from a given set of atoms is determined by the combinatorial number of allowed stable bonding interactions between these atoms. If only amino-acid based penta-peptides are counted, for example, the number of possible molecular structures is about 20^5 , i.e., 3.2×10^6 molecules. This number is based

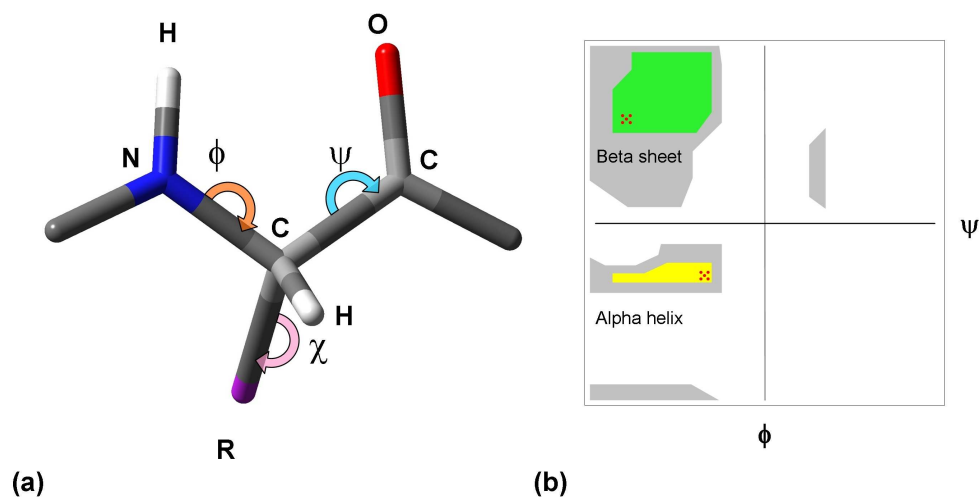
on the 20 genetically-encoded amino acids only, a subset of the approx. 300 amino acids found in living systems, excluding unnatural amino acids. Not surprisingly, the idea of drug design appears to be a hopeless quest. How can this number be effectively reduced to a manageable set that, eventually, will display the desired drug properties? A first step consists of selecting a set of variables that can be obtained consistently for every molecule. Each molecule is then represented by a vector whose components are the selected variables. Each variable should be well-defined and, at least in principle, be also a measurable property. Each molecule is represented by its respective set of properties in the multidimensional space, i.e., by a vector $V \in R^N$. The representation of every molecule in this multidimensional vector space, R^N , enables one to define an Euclidean distance, d_{AB} , between two molecules A and B. Ideally, two molecules separated by the shortest distance in this vector space are also the most chemically-similar among the set. This hypothesis is based on the realization that similar molecules must exhibit similar molecular physicochemical properties. In this approach, molecular design implies the identification of similarities in this vector space.

Figure 10.1: Capped amino acid residues used in this study. Each molecule is represented by the variables listed in Table 10.1. The total number of molecules studied using each model is indicated.



The biochemical behaviour of a protein is encoded in its primary structure, i.e., the amino acid sequence which determines its functionality via the secondary and tertiary structures. In the study of the genetically-encoded amino acids it is important to determine which variables account for their idiosyncratic biochemical features. Within the context of protein-based drug design, the following question is addressed: What theoretical variables better represent the highly-specific yet overlapping biochemical

Figure 10.2: Amino acids structure. (a) A stick model of an amino acid residue with the standard dihedral angles defining the main chain (ϕ, ψ) and the side chain (χ) conformations. (b) Ramachandran plot with the studied α -helical and β -sheet conformation regions highlighted (red dots).



functions displayed by each of the genetically-encoded amino acids? In order to answer this question, two models were built that mimic the electronic environment of an amino acid residue inside a peptide chain. The two models differ in the capping groups for the N- and the C-terminuses as shown in Fig.10.1. The non-zwitterionic amino acid models studied are: $\text{H}(\text{C}=\text{O})|\text{AA}|\text{NH}_2$ and $\text{Ala}|\text{AA}|\text{Ala}$. With the second model reveals the effect of neighbouring amino acids on the properties of the central amino acid. Each side chain has preferred side chain torsion angles.^{147,148} Three of these side chain conformers or rotamers were selected: *gauche*(+) = -66.7° , *gauche*(-) = 64.1° and *trans* = 183.68° . The main chain conformers were set at five α -helical and five β -sheet conformations (defined by ϕ and ψ angles as shown in Fig.10.2). The five backbone torsion angles corresponding to an α -helical conformation are $-65^\circ \pm b4^\circ$ and $-39^\circ \pm b4^\circ$, and those corresponding to the β -strand conformations are $-130^\circ \pm b5^\circ$ and $120^\circ \pm b5^\circ$, with $b \in (0, 1)$. These conformations are indicated by the red dots in the Fig.10.2(b).

A total of 40 theoretical variables were considered: (a) 19 graph descriptors, i.e., connectivity descriptors of the molecular structure;^{149,150} and (b) 21 physicochemical variables obtained from quantum mechanical calculations. The variables representing the amino acids are listed in Table 10.1. A total of 1065 molecules representing the twenty amino acids in different conformations and capping models was studied.

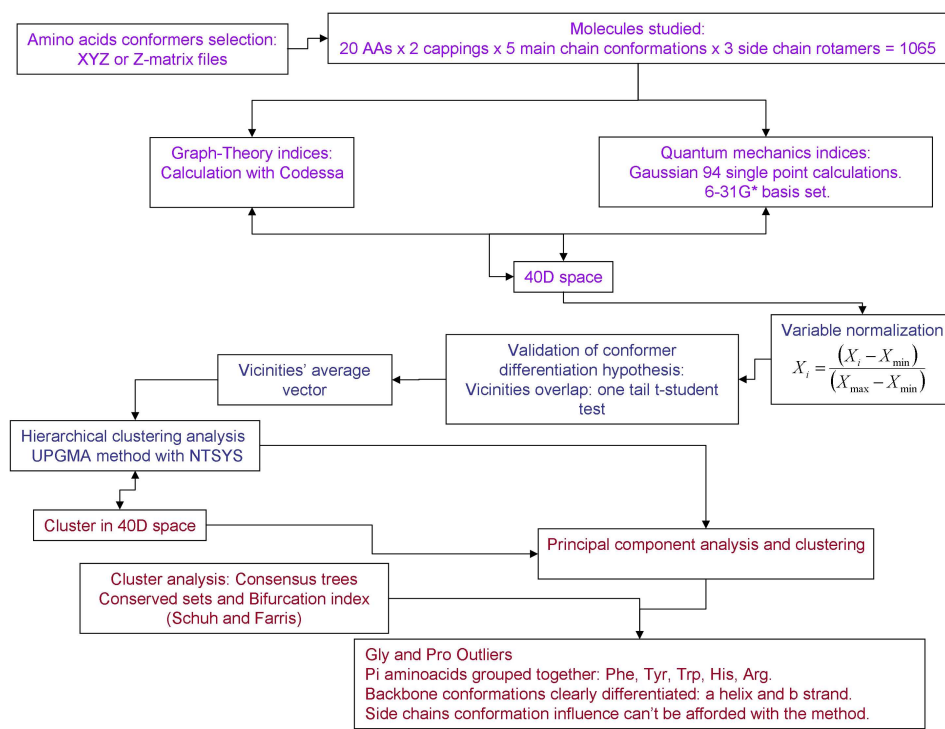
Table 10.1: Properties selected for the representation of the amino acids.

Graph theory indices	Quantum variables
Wiener index	Moment of inertia
Randic indices of order 0-3	Molecular weight of the amino acid residue
Kier and Hall connectivity indices of order 0-3	Electronic spatial extent
Kier and Hall shape indices of order 1 through 3	Nuclear repulsion energy
Shadow indices	Total energy
	Highest occupied molecular orbital energy, HOMO
	Lowest unoccupied molecular orbital energy, LUMO
	Mulliken partial charges
	Sum of the Mulliken partial charges for the side chain atoms
	Total dipole moment
	Electric potential
	Sum of the electric potential for the side chain atoms
	Quadrupole norm

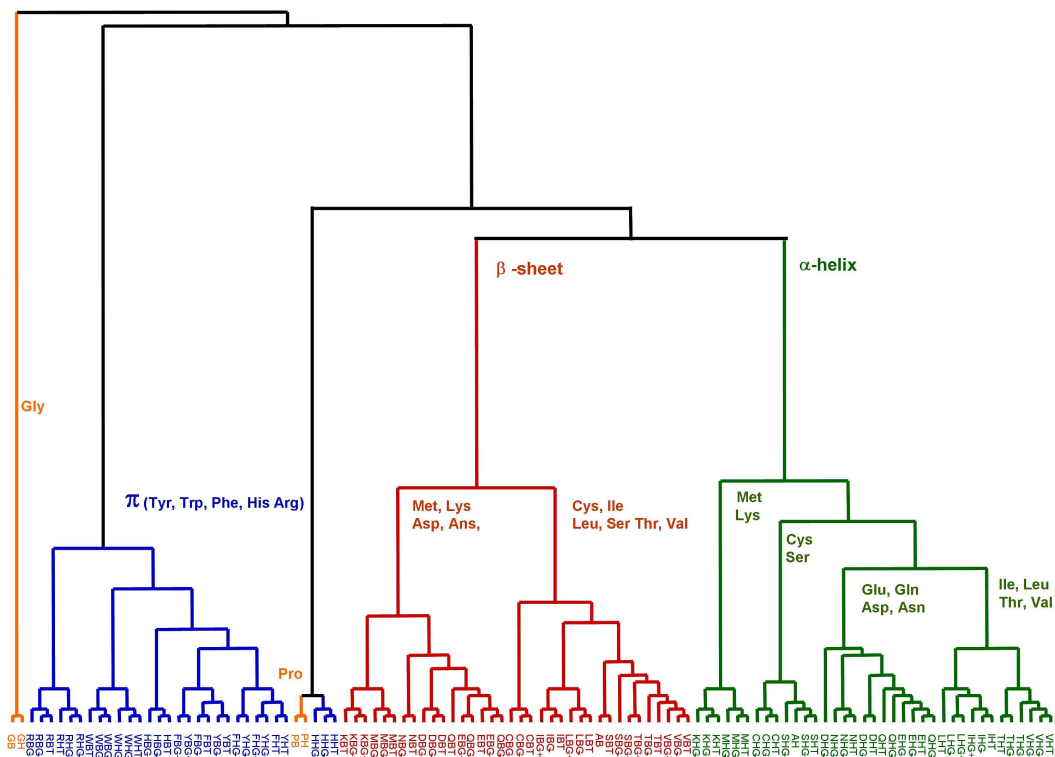
The graph-theory indices were calculated with Codessa,¹⁵¹ and the QM computations were done at HF/6-31G(d) level, with polarization functions on heavy atoms capable of forming hydrogen bonds (N, O and S). A hierarchical cluster analysis was done with NTSYS program,¹⁵² with the unweighted pair group method with arithmetic mean (UPGMA) method. A schematic representation of the steps followed in this strategy is depicted in Fig.10.3. Principal components analysis (PCA) was used to determine the principal variables that specify the similarity between the amino acids. An important result is that the amino acids are separated into statistically-disjoint groups, and those groups are segregated mainly by their electrostatic properties alone. Fig.10.4 shows the classification obtained from the PCA. The group of amino acids containing π electrons is clearly identifiable, which includes the aromatic amino acids, such as Phe, Tyr and Trp as well as His and Arg. Two amino acids, Gly and Pro, are clearly outliers in this classification, which reflects their particular biochemical behaviour: the first have the smallest side chain (a hydrogen atom) and the second is an imino acid, i.e., its side chain is cycled over the backbone. The analysis was able to automatically distinguish between the two large groups differing only in the conformation of the backbone, namely, the α and β -conformers of all of the amino acids separate into two large statistically-distinguishable groups, as can be seen in Fig.10.4. In general, side chain rotamers are grouped closely to each other according to their respective amino acid. Rotamers of isoelectronic side chains, such as the Asn-Asp or Gln-Glu isoelectronic pairs are also located closely in the 40-dimensional

vector space, which indicates that the method generates valid results according to intrinsic (total) molecular properties, but misses details on the specific smaller functional groups. The groups obtained for each amino acid conformer are conserved

Figure 10.3: Schematic diagram of the steps followed to determine the theoretical variables responsible for the amino acids' main structural propensities. The key factor in this approach is the representation of every molecule by a set of properties in a multidimensional (40D) vector space for performing a PCA and a clustering analysis.



across the two capping groups, which indicates that the selected variables capture the intrinsic nature of amino acid properties. These groups can be reproduced by eight variables only (five quantum, three from graph theory), as indicated by the PCA analysis. In conclusion, the classification of the set of amino acids studied is driven principally by electrostatic properties. It is clear from the clustering shown in Fig.10.4 that the main groups are those side chains containing electrons with symmetry, and the two backbone conformations, α and β , whose multipole moments are oriented in different directions, and therefore they have different electrostatic interactions. The results reviewed here show that the structural features of amino acids are sufficiently accounted for by electrostatic variables alone. Some quantum QSARs come to a similar conclusion. For example, Brink et al. studied about 100 QM-based variables to

Figure 10.4: Amino acids classification based on 8 principal components.

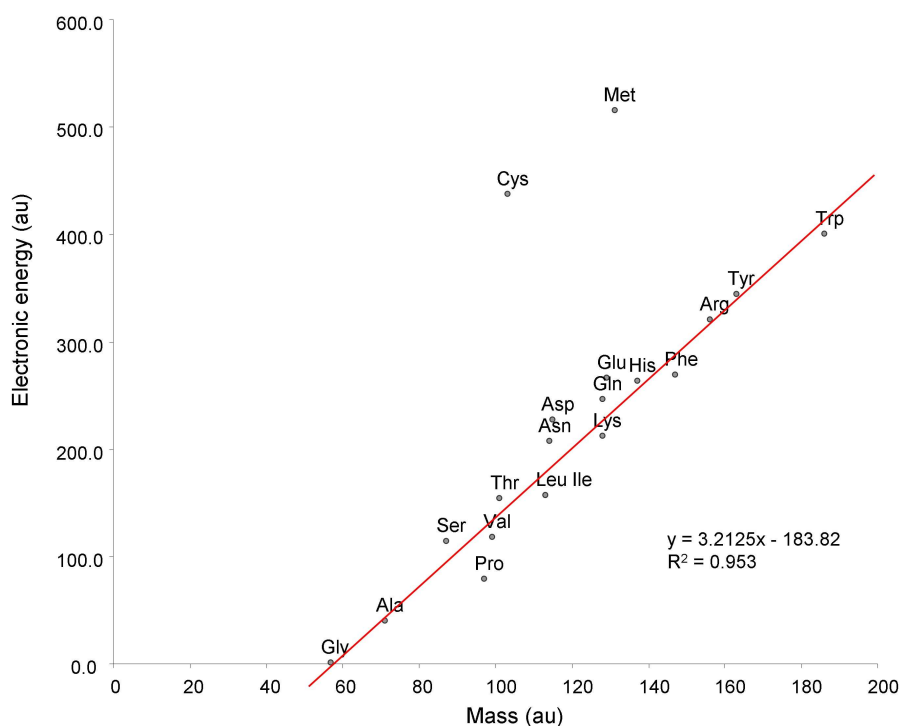
predict the water-octanol partition coefficient (Po/w) from the molecular wave functions. These authors concluded that three electrostatic variables, surface area, the surface electrostatic potential and the spatial minima of the electrostatic potential, respectively, can give good correlations with $\log Po/w$ for a number of molecules with biological and pharmacological interest.¹⁵³ This is a remarkable result in the sense that $\log Po/w$ is an experimentally-determined biochemical property usually measured at standard conditions yet the thermodynamic factors are not included in the quantum computation. These results as well as those reviewed in this chapter suggest that the electrostatic variables are good descriptors regardless of the thermodynamic conditions, providing support to the validity of the isolated (“gas phase”) QM model.

10.2 Side Chain Polarizations and the Theoretical Classification of Amino Acids

From the results reviewed earlier, it is clear that the amino acids can be adequately described in terms of the electrostatic variables. In this section the use of QTAIM for

characterizing the genetically-encoded amino acids is described. QTAIM partitions the molecular properties into additive atomic contributions, and, in doing so, characterizes molecular transferable fragments such as the amino acids' side chains. Details about the theory are omitted here as they can be found elsewhere.^{9,18} The focus of this section will be placed on the similarity of the amino acids within the context of this theory. The earlier work of Bader and coworkers on peptides and amino acids has

Figure 10.5: Energy magnitude versus mass, as provided by QTAIM, for the genetically-encoded amino acid side chains. The linear fit excludes the sulfur-containing side chains, Cys and Met.



been extended in greater detail by Matta and Bader more recently.^{154,155,156} The local structural properties of each amino acid determine the overall tertiary structure of a protein, and the side chains are responsible for its specific bioactivity. Therefore it is vital to characterize the physicochemical properties of the amino acid side chains in order to understand and predict the bioactivity of peptides and proteins. The amino acid model studied is shown in Fig.10.1(a), which was initially studied by Bader.^{157,158} Two main backbone angles and a single rotamer per amino acid were studied, giving a total of 40 molecules, including a total of 888 atoms at the HF/6-31G(d) level with polarization functions on N, O and S. The atomic properties were computed with the AIMPAC suite of routines from Bader's group.¹⁵⁹ In order to compare tensor and

vector properties (which are origin- and orientation-dependent), all the amino acids were properly aligned by the common atoms of the backbone and the first atom in the side chain. The origin of the coordinates was placed at the α carbon atom.

Figure 10.6: Andrews plots for the 10 QTAIM variables on 40 amino acid side chains. Molecular similarities appear as similar colors and shapes.

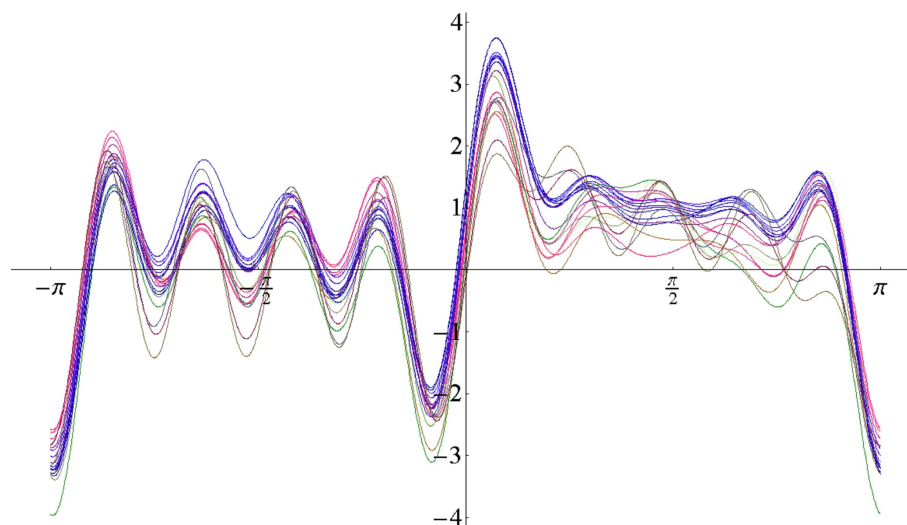
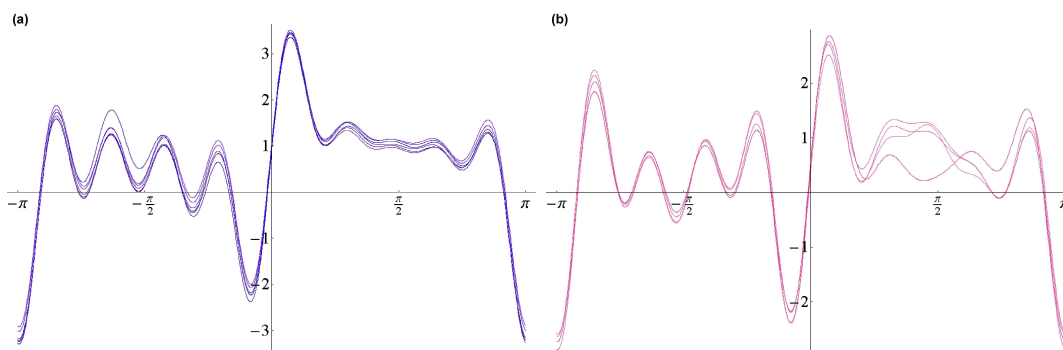


Figure 10.7: Similar side chains as revealed by their corresponding Andrews plots (color and shape).

(a) Gly, Ala, Val, Ile and Leu.

(b) Ans, Gln, Asp and Glu; the later one exhibits a different pattern than the others.



Each amino acid was represented by the three first terms in the multipole moment

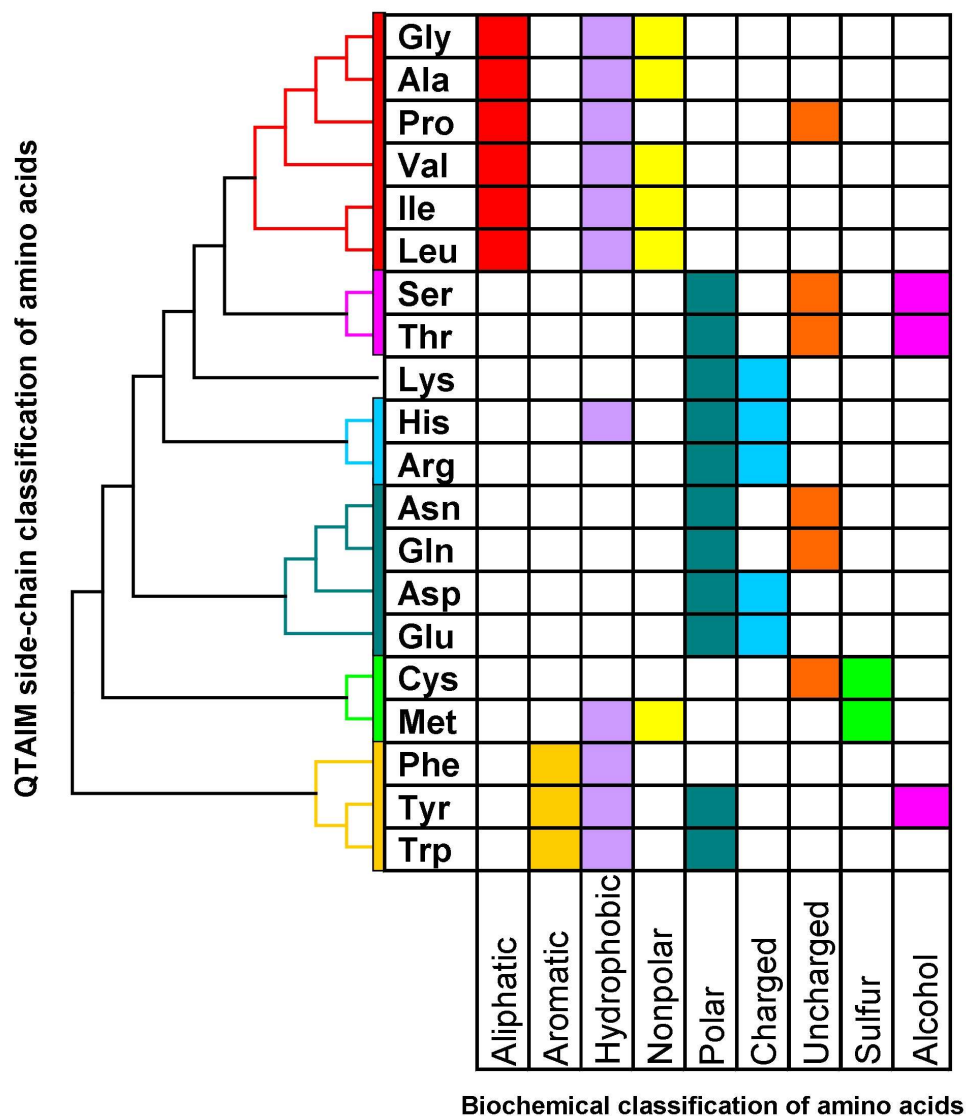
expansion of the side chain charge density (side chain charges (monopoles), and side chain dipolar and quadrupolar polarizations) and the side chain total electronic energy. These electronic multipole moments (polarizations) should not be confused with the total multipole moments of the amino acids. Multipole moments provide a basis for a general procedure to systematically extract the symmetries of a continuous distribution, such as the charge density, and hence they characterize its shape. They depend on the origins and the relative orientations of the coordinate system and therefore the molecules were pre-aligned as described above.

The energy of the side chains measures their size, as illustrated in Fig.10.5. This figure shows that the side chain energy magnitude can be linearly fitted to the side chain mass with a correlation $R^2 = 0.95$, for the non-sulfur side chains. The electronic energy of the side chains involving only elements located in the first two rows of the periodic table exhibit a linear correlation with mass, the same correlation does not apply for the side chains involving a third row atom, such as sulfur. It is desirable to visualize similarities between the molecules under study before performing any further statistical survey, but any multi-dimensional molecular representation always entails a graphical challenge. The Andrews plots (AP) are a useful tool for addressing this task. As illustrated in Fig.10.6, each molecule can be represented by a single strand that is easily obtainable from the following formula:

$$\begin{aligned}
 A = & E \sin(t) + M_x \cos(t) + M_y \sin(2t) + & (10.1) \\
 & M_z \cos(2t) + Q_{xx} \sin(3t) + Q_{xy} \cos(3t) + \\
 & Q_{xz} \sin(4t) + Q_{yy} \cos(4t) + \dots
 \end{aligned}$$

where E is the side chain energy, M_i and Q_{ij} are the dipole and quadrupole polarization components, respectively, and $t \in [-\pi, \pi]$. The values used were standardized as explained in detail in reference.¹⁴² Each strand represents a side chain as a smooth function, with coefficients equal to the corresponding physical properties. A color code was also added to each strand by assigning each component of the color code to the (standardized) magnitudes of the energy, dipole and quadrupole moment. The final color is a combination of three basic tones: red, green and blue. Each tone is defined as a number within the interval. For the present case, $RGB = [1 - M, Q, E]$, where

Figure 10.8: Quantum theoretical classification of the genetically-encoded amino acids. This classification was obtained after applying a clustering procedure to the side chain properties: energy, dipolar polarization and quadrupolar polarization as provided by QTAIM computations at HF 6-31G(d) level of theory. The table highlights the typical physicochemical properties of the side chains. The main clusters were colored according to these properties.



M , Q , E are the normalized magnitudes of the dipolar polarization, quadrupolar polarization and energy, respectively (i.e. each of these variables lies within the interval $[0, 1]$). Therefore, similar shapes and colors that correspond to similar molecules can be identified. Fig.10.7 shows the APs of the 40 side chains studied. The distinctive shape in blue groups the aliphatic side chains {Gly, Ala, Pro, Val, Ile, Leu}, while the group {Asn, Gln, Asp, Glu} exhibit a similar red color. This simple analysis reveals the existence of underlying similarities within the set of amino acids.

The graphical analysis shows the existence of similarities between the side chains, but in order to quantitatively determine these similarities a systematic classification procedure is required. Consequently, a multivariate classification of the side chains in the $10D$ vector space that is based on the distance between elements in this vector space was used. The neighbor joining method applied over a two-fold distance measure provides the amino acids classification shown in Fig. 10.8. It is clear that the main biochemical features coincide with several of the groups obtained. This theoretical classification of the amino acids (the first quantum theoretical classification that the author is aware of), provides a rich variety of clearly identifiable biochemical groups on the sole basis of transferable properties provided by QTAIM. On the other hand, experimentally-based classifications tend to emphasize certain molecular features and downplay others, which explains why the classification resulting from their associated matrices coincides with the biochemical classification only for major groups such as “aliphatic AAs” or “charged AA”, while several amino acids appear as outliers,^{160,161,162} as recently reported by Esteve and Falceto.¹⁶³ The successful classification of the amino acids *in silico* can be attributed to the quality of the atomic and group properties provided by the quantum theory of atoms in molecules. It has been shown how one can use QTAIM group properties in conjunction with clustering analysis to recover a well-known biochemical classification of a set of functionally-related molecules (the amino acids). Amino acid classification based on the electrostatic moments is superior to those obtained by scoring matrices widely used in protein biostatistics. One key advantage of the theoretically-based classification over experimentally-based ones is the homogeneity of the quality of the input dataset. Experimentally-based amino acid properties that serve as a basis for the replacement matrices for example, involve a variety of data sources with different precisions, compromising the outcome of the analysis. As an extension of this work, a theoretical amino acid replacement matrix for bioinformatics can be readily developed, that can potentially overcome several of the drawbacks faced by the empirical ones. The

methodology outlined here can be replicated for any other set of molecules, and it emerges as an alternative to QSAR methods in the sense that it provides unbiased quantitative similarities among the studied set.

10.3 Quantum Mechanical Studies of Peptide-Host Interactions

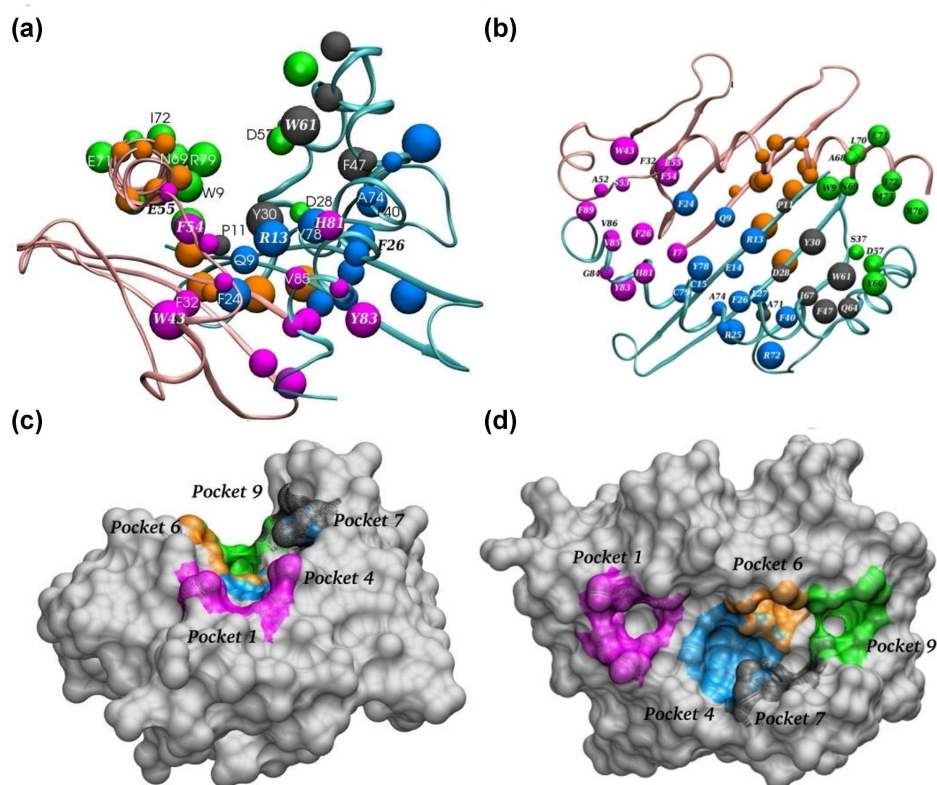
In the previous section it was shown that calculated QM electrostatic properties provide a biochemical classification of the amino acids consistent with their known chemical and physical properties. The main hypothesis for the study of peptide-host interactions is that these interactions are, to a large extent, dominated by the electrostatic properties of the AA residues constituting the interacting peptides. This research program is motivated by the development of a synthetic anti-malarial vaccine at the Fundacion Instituto de Immunologia de Colombia (FIDIC).^{164,165} A key step for developing a specific immune response against a pathogen is the formation of a stable complex between the major histocompatibility complex (MHC) molecule and antigenic peptides, capable of bringing information to the T-cell receptor (TCR) molecules necessary to trigger an immune response against the pathogen. The details about this process are omitted, in order to focus on the QM/MM hybrid approach used for the peptide-host interaction studies of the anti-malarial design of a synthetic peptide-based vaccine. The MHC-peptides (MHC-P) interaction is a prototypical ligand-receptor interaction, and hence the approach outlined here can be used to study other similar biochemical complexes. The extended peptide (~ 9 amino acids) forms a non-covalent complex with the host MHC protein at the peptide binding region (PBR) through certain spots that act as anchoring sites, known as pockets. In Fig.10.9 a MHC Class II PBR is shown, with the pockets in colour, as obtained from the Protein Data Bank (PDB).

According to the present hypothesis, the MHC-P interaction can be described by the quantum-based electrostatic potential,^{166,167} which in terms of the multipole expansion has the form

$$V = \frac{1}{4\pi\epsilon_0} \left[\sum \frac{q_k}{r} + \sum p_k \delta_k \frac{1}{r^2} + \frac{1}{2} \sum \frac{1}{3} Q_{ij} \delta_i \delta_j \frac{1}{r^3} \dots \right] \quad (10.2)$$

where the index k runs over all the host atoms involved in the interaction. Therefore a partitioning scheme is necessary to provide atomic contributions for each multipole moment that appears in this expansion. Unfortunately, the number of atoms involved

Figure 10.9: Peptide binding region (PBR) of the (MHCII-P). LA-DR β 1*1501 molecule with the α -chain as a pink ribbon and the β -chain as a light blue ribbon. (A) Frontal view and (B) top view. Pocket amino acids are represented as spheres with different sizes and colours: Pocket 1 (magenta), Pocket 4 (dark blue), Pocket 7 (grey) and Pocket 9 (green). Molecular surface showing (C) a frontal view of the PBR and (D) the top view showing the relative depth of the different pockets. P1 and P9 are deeper whereas Pocket 4, 6 and 7, are more superficial, lying towards the walls of the groove. (Graphic reprinted from the Ref.[30] under the Creative Commons Attribution License (CCAL)).



in the screening of MHC-P interactions exceeds the practical application of QTAIM, which would be an ideal partition schemeⁱ. Instead of QTAIM, point-charge multipoles derived from the Mulliken population analysis were obtained from standard quantum mechanical calculations (as provided by programs such as Gaussian). Accordingly, the dipolar and quadrupolar moments and their respective norms are

$$\mathbf{p} = \sum_{k=1}^N q_k \mathbf{r}_k \quad d = \sqrt{\mathbf{p} \cdot \mathbf{p}} \quad (10.3)$$

$$Q_{ij} = \sum_{k=1}^N q_k (3x_i x_j - r_k^2 \delta_{ij}) \quad C = \sqrt{\sum_{i,j}^3 Q_{ij}^2} \quad (10.4)$$

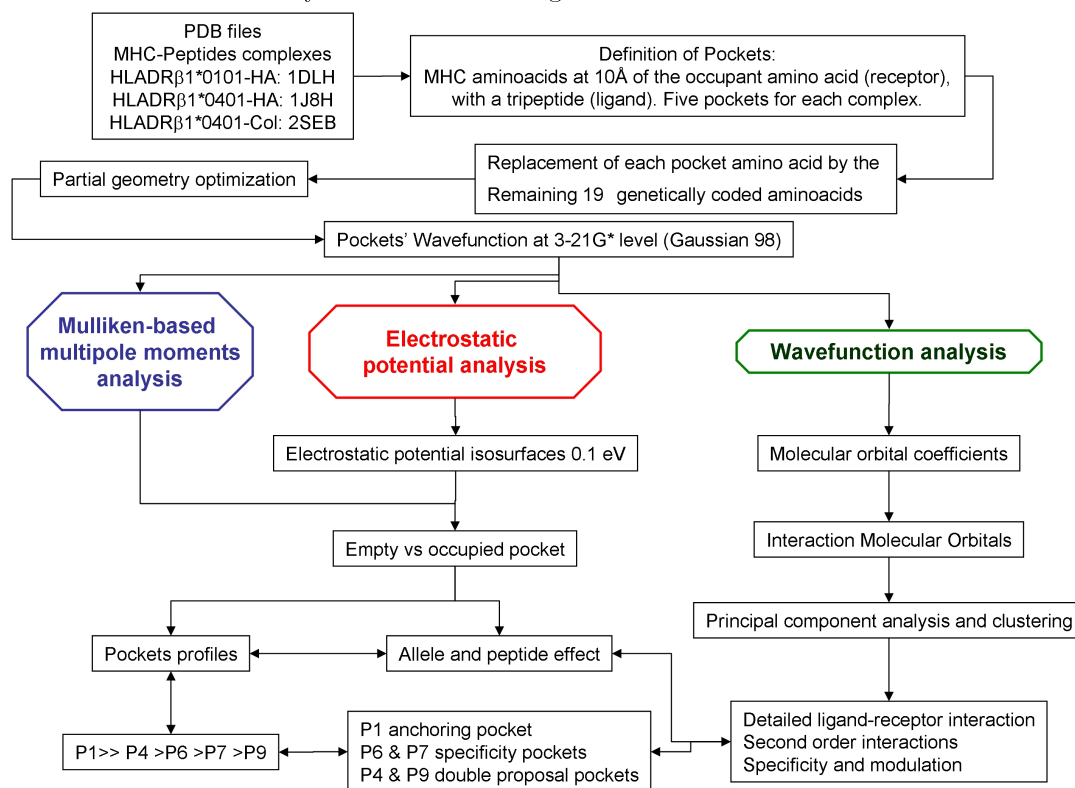
with q_k and \mathbf{r}_k the charge and position of the k^{th} atom, respectively. Each amino acid residue involved in the complex formation was systematically replaced by each of the remaining 19 genetically-encoded amino acids in order to quantitatively determine its relevance in the MHC-P complex stability. A detailed account of the steps followed in this approach is shown in Fig.10.10. The changes observed after each replacement were estimated by examining three aspects:

- Multipole moments (Fig.10.11); in order to estimate the effect that each specific amino acid exerts in the pockets, the Mulliken-derived electrostatic multipoles are evaluated over the pockets. The isolated complexing peptide is used as the reference system for evaluating the changes in the multipoles.
- Electrostatic potential as projected over a molecular surface (Fig.10.12); a traditional study of the QM potential projected over an electron density surface guides the analysis for the atoms directly involved in the complex.
- Identification of those orbitals contributing directly to the complex formation; the orbital expansion coefficients are classified according to pocket and peptide contributions, by a statistical analysis, as schematically explained in Fig.10.13.

While a graphic study of the electrostatic field reveals some of the details of the

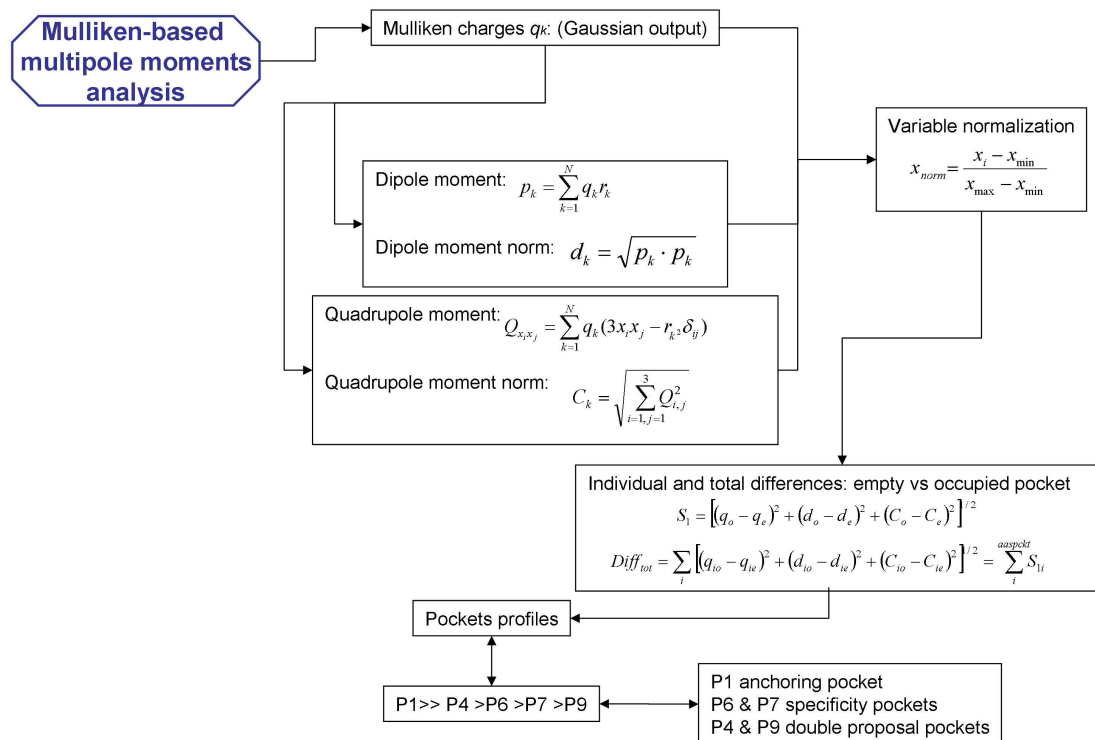
ⁱQTAIM requires optimized geometries for each system under study at the same level of theory of the energy computations. A failure to this request may lead to errors in the atomic energies and forces, among other important properties. Hence, QTAIM is a method that is not suitable for being used with MM geometries.

Figure 10.10: Diagram of the quantum study of MHC-P complexes. Each specific analysis is detailed in figs. 11-13.



peptide-host complex formation, only the multipole and wavefunction analyses provide a hierarchy of relevance among the pocket sites, which is highly correlated with the experimentally-observed one.^{168, 169, 170} For example, in a MHCII-peptide complex study, the prevalence for aromatic amino acids in pocket #1 was unambiguously determined (see Table 2 in Ref.¹⁴⁶). Such specific prevalence for aromatic side chains plays a significant role in the complex stability as this pocket works as an anchoring site for the guest peptide. In this way, the traditional direct visual study of the electrostatic potential of the MHC-P complexes provides merely a complementary analysis that verifies the quantitative classification given by the other two methods. The success obtained so far in the description of the essential amino acids that are responsible for the complex stability and their respective synonymous replacements validates the overall proposal reviewed in this section.

Figure 10.11: Diagram of the quantum study of MHC-P complexes using the Mulliken-based multipole method.

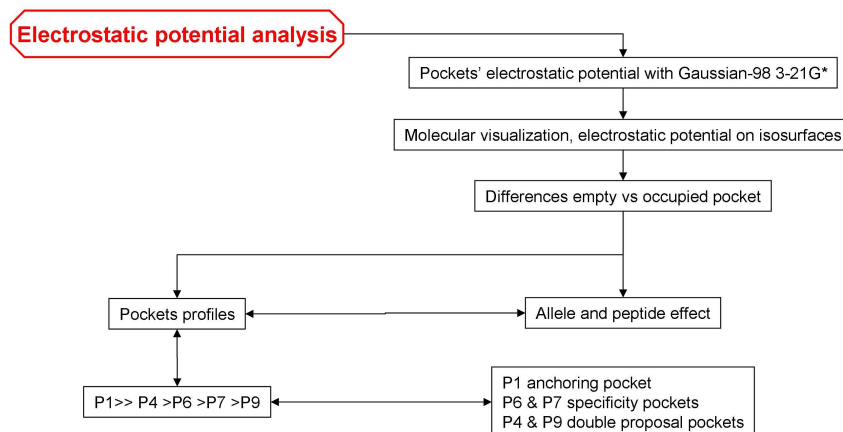


10.4 Conclusions

The work reviewed in this chapter supports the idea that much of the biochemical information carried by the amino acids is encoded in an electrostatic language. Initially, a principal components analysis (PCA) over a set of amino acid conformers identifies those *ab initio* variables that best describe amino acids' features. As revealed by the success in describing amino acid similarities^{144,142} and peptide-host complexes,^{146,168,170,166,167} the multipole moments sufficiently account for the characteristic features of the side chains and their interactions.

Here are explained several of the ideas proposed as an answer to the question about the theoretical study of bioactivity in molecules. The methods and strategies discussed are used to study small peptides but can also be applied to other sets of molecules. The present discussion is an attempt to clarify the overall strategies and

Figure 10.12: Diagram of the quantum study of MHC-P complexes using the electrostatic potential.

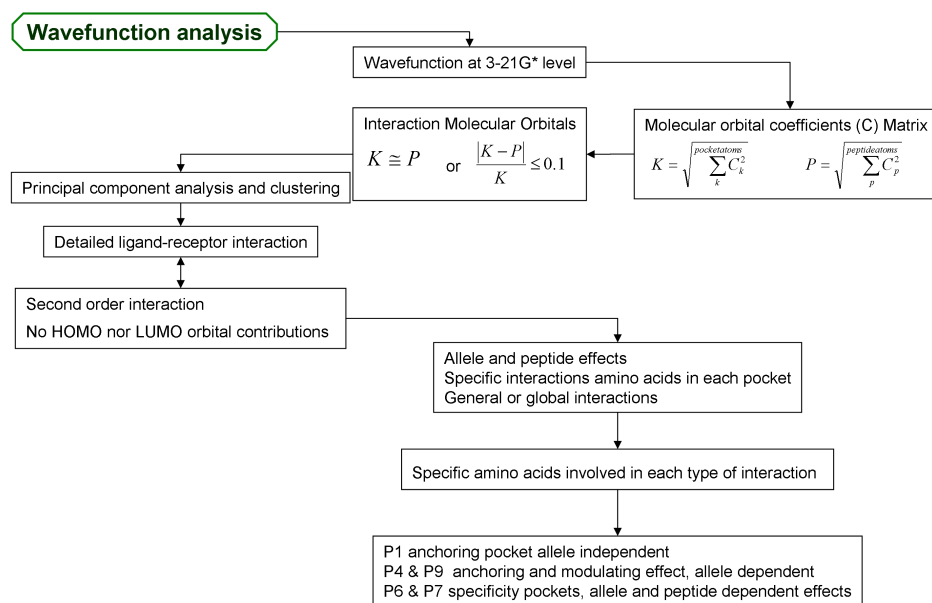


methods involved in the FIDIC's QM approach to molecular design of a synthetic anti-malarial vaccine, and hence is a complementary work to the original papers published over the past ten years.

Here are illustrated the strengths of original analysis for identifying biochemical propensities in peptides and proteins. These studies reveal that, in order to get insightful information about the relative physicochemical properties in biomolecules, there is no need to compute unphysical descriptors, as used in some QSAR studies. Instead, physical variables sufficiently account for similarities and functionality of the amino acids, as shown by a PCA survey of over 1065 amino acids models.

With the increasing computational power available to laboratories around the world, QM methods are getting involved in more and more studies of biomolecular systems. The main message outlined in this chapter is not focused on how to get the wave function of a biomolecule, but rather how to extract relevant biochemical information out of it.

Figure 10.13: Diagram of the quantum study of MHC-P complexes using a wavefunction analysis.



Chapter 11

Ab Initio Study of the Structure, Energy and Polarizability of Amino Acids

11.1 Introduction

The α -helix is a very common secondary structure pattern found in proteins. The formation of α -helices (AHs) involves cooperative self-assembling between adjacent amino acids in the sequence, and hence they have an important role in protein folding. In order to understand and predict protein structure formation, α -helices have been the objective of numerous experimental and theoretical studies. In addition, these naturally-structured patterns are of great interest not only for their biological relevance, but also from the molecular design perspective. The mechanical and optimal-packaging properties of AHs have made this a privileged structural pattern for amino acids sequences.

One of the stabilizing factors in α -helices is the cooperative effect of the peptide bonds thanks to the mutual alignment of the corresponding dipole moments that are responsible for the total and relatively large dipole moment of the helices (about 20 times that of water for a pentapeptide). While other kinds of helices are also found naturally, like the polyproline that produces a tighter structure, α -helices are abundant patterns responsible for tertiary and quaternary structures of proteins.

It is ideal to study the secondary structure in proteins with time-dependent quantum mechanics (TD-QM), but the size of a quantum system grows with the number of electrons, which limits its use to short peptides of about a hundred atoms with conventional tools. Static computations are possible for equilibrium geometries of small peptides, and several of its observable properties can be calculated. While most of the quantum-mechanical studies of peptides have been done *in vacuum*, the solvent plays a significant role in the formation and stability of secondary structures. For instance, in CD and NMR studies the presence of trifluoroethanol (TFE) as a cosolvent increases the population of α -helix and β -sheet content in secondary-structure-forming peptides.¹⁷¹ Hence, TFE-NMR studies are used to assess the potential *structurability* of a sequence rather than to determine its native conformation under biological conditions.

In a similar way, QM simulations of small peptides can provide the structurability of a given sequence and, if the solvent is adequately modelled, QM even can provide reliable information about *denaturation*ⁱ processes due to pH or temperature changes.

In the present work the structural stability of α -helices upon mutation of the central amino acid is investigated at a state-of-the-art QM level, including solvent effects and dispersion interactions. A comparative analysis of the relative stability of a set of α -helices is reported. In addition, a model for the approximate computation of amino acids polarizability is introduced.

A realistic account of the solvent effect on peptides requires the explicit inclusion of solvent molecules in the simulation, but the number of explicit water molecules required exceeds by a few orders of magnitude the number of electrons that can be computed with conventional QM tools. A single hydration shell can demand more computations than the peptide itself. Hybrid models such as ONIOM still carry some limitations. For instance, modelling the solvent with MM and the solute with QM does not provide a good representation of the hydrogen bonds between the solvent and the peptide, limiting therefore the simulation capabilities.

One way to efficiently model solvent effects in QM is by the use of *effective mean field* models, where the solvent is replaced by a perturbation term in the Hamiltonian that responds self-consistently to the electronic distribution of the molecule. These methods are generically called *cavity models* or solvent reaction fields, and with them the solvent is regarded as a dielectric continuum. Hence the surrounding is characterized mainly by its dielectric constant and densityⁱⁱ. The calculation is performed by placing the solute in a cavity within the solvent reaction field; the dielectric medium is polarized by the solute, and this polarization creates a reaction field that causes a perturbation of the solute itself. A very CPU-cost-effective model for the solvent is the polarizable continuum model (PCM) that creates the solute cavity via a set of overlapping spheres. It was initially proposed by Tomasi and coworkers^{172,173} and Pascual-Ahuir and coworkers.¹⁷⁴ One of the advantages in using PCM over the explicit solvent is the very low cost added to standard computations, yet it is very responsive for electrostatic interactions like dipole-dipole interactions, as those responsible for

ⁱDenaturation is a process in which proteins or nucleic acids lose their tertiary structure and secondary structure by application of some external stress or compound, such as a strong acid or base, a concentrated inorganic salt, an organic solvent (e.g., alcohol or chloroform), or heat.

ⁱⁱIt is important to recall that, unlike classical mechanics, in QM the dielectric constant does not provide enough information for a solvent reaction field (SRF) model, and additional parameters are required.

the stability of α -helices.

Another important factor determining structural stability in proteins is the non-covalent interactions: hydrogen bonds and van der Waals interactions. The later have been recently included in several QM methods, mainly as empirically-calibrated terms added to conventional DFT models. These are weak but long-range interactions that become relevant when cooperative effects emerge in the system, such as side-chain to main-chain interactions, for instance. Hence these are determining terms to be included in a QM study of an alpha-helix.

The recent addition of very efficient PCM solvent models plus the dispersion corrected DFT methods motivated the theoretical level of the present work. This approach improves substantially the level of theory of previous reports on alpha helices by Dannenberg's group¹⁷⁵ and recent QTAIM studies on α -helices by the author.^{142, 176}

11.2 Methods

A preliminary QTAIM analysis performed on an α -helix modelⁱⁱⁱ revealed details of the intramolecular hydrogen bond network upon mutation of the central residue.¹⁷⁶ In that work, a set of 19 peptide sequences were fully-optimized. The structures correspond to the α -helix conformation of the polypeptide $\text{COH}(\text{Ala})_6\text{X}(\text{Ala})_6\text{NH}_2$ where X is any of the amino acids (excluding proline). Polyalanine peptides form stable α -helices and hence are important models for the study of secondary structure stability. The effect of amino acid substitution at the central position of the peptide on the hydrogen bond network of the α -helix was assessed, and the strength of the hydrogen bond network was determined with QTAIM, as found elsewhere.¹⁷⁶

These preliminary results lead to some important lessons about QM modelling of α -helices. First, all the 13-amino acid peptides were affected by the selected capping groups, which have partial charges that induce a bend in the helix axis. In addition, the inability of the COH capping group to induce α -helical-like hydrogen bonds with the first alanine originated the funnel-like structures that were obtained after the geometry optimization. These optimized structures are somewhat distant from the ideal helix initially intended to be studied. The use of these capping groups was motivated by previous QTAIM results on shorter peptides by Bader and Popelier,¹⁵⁸

ⁱⁱⁱThe corresponding paper was a collaborative effort of different members of the Boyd research group, including the author. The particular results on the electron density analysis were performed by S. Lapointe and are not part of the present thesis. The original idea of performing the α -helices to study the mutations was proposed by the author.

and Bohórquez et al.¹⁴² In order to prevent the bending of the helix axis and speed up the computations, for the present study the helices include only seven amino acids in total, and the initial capping group is the acetyl group (C=O)CH₃, that mimics an alanine residue and is able to make hydrogen bonds in a similar way to the rest of amino acids in the chain. The helix backbone makes two complete turns. In addition, the solvent effects were included via the PCM model with a dielectric constant of $\epsilon = 78.3553$. In order to include van der Waals effects, the fully optimized geometries were computed using the B97D density functional developed by Grimme et al. that includes dispersion interactions.¹⁷⁷ A D95(d,p) Dunning/Huzinaga full double zeta basis set was used for all the computations.

11.3 Optimized Structures

The optimized geometries and wave functions of nineteen α -helices of the form acetyl(Ala)₃X(Ala)₃NH₂ were obtained, where X represents the genetically-encoded amino acids except proline. Proline was excluded from the study because it is an iminoacid that disrupts the secondary structure by forming a covalent bond between the side chain and the backbone. X=Ala defines the reference structure because polyalanine is known to form stable α -helices. The capping groups studied here have been used before by Dannenberg's group.¹⁷⁵

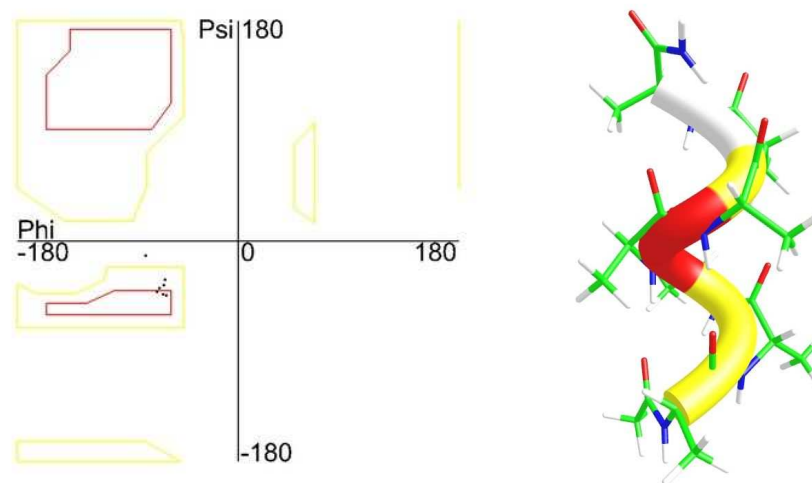
The optimized geometry of the polyalanine helix was used as the initial geometry for the remaining helices. For all the cases an energy minimum was reached, with zero imaginary vibrational frequencies. All the molecules keep identical helical structures, showing minor deviations from the ideal α -helix, mainly at the terminal residue of the sequence, as it is shown by the Ramachandran plot for the polyalanine acetyl(Ala)₇NH₂ in Fig. 11.1a. One hydrogen of the NH₂ capping group is not aligned with the precedent NH in the chain, which puts the last dihedral angle outside the helical region, while all the remaining dihedral angles are close enough to the ideal α -helix torsion angles values.

11.4 Energetic Results

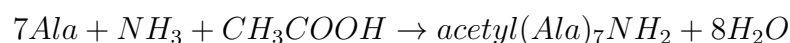
To study the effect of a single substitution on the stability of a secondary structure (like the α -helix) a reference structure must be carefully defined. In the present study the reference structure is the polyalanine peptide acetyl(Ala)₇NH₂. The energy

Figure 11.1: Optimized geometry of acetyl(Ala)₇NH₂ at the PCM-B97D/D95(d,p) level of theory. The red zone within the negative quadrant indicates the region of observed α -helices.

- (a) Ramachandran plot for the optimized polyaniline α -helix. (b) Molecular and secondary structures.



of mutation can be estimated from the component amino acids, which can be well defined via an imaginary condensation polymerization involving the required amino acids that yield the polypeptide plus one water molecule per peptide bond formed.¹⁷⁵ The capping groups are also incorporated into the energetic calculation. Hence the reaction for the formation of the acetyl(Ala)₇NH₂ would be



The relative energy of the peptide with respect to the components would be

$$E_{rel} = E_{peptide} + 8E_{water} - 7E_{Ala} - E_{ammonia} - E_{acetic\ acid} \quad (11.1)$$

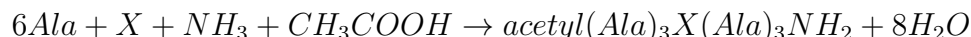
The energy differences upon the mutation of the central Ala by another amino acid in the reference peptide can be evaluated in many different ways. This substitution can be estimated by evaluating the hypothetical reaction



where X is another amino acid. This equation provides a measure of the energetic effect of the amino acid mutation for each structure. The relative energy of the amino acid mutation would be

$$E_{rel} = E_{polyAla} + E_x - E_{peptide} - E_{Ala} \quad (11.2)$$

Another method consists of the hypothetical condensation reactions for the formation of the helices



The corresponding relative energy is

$$E_{rel} = E_{peptide} + 8E_{water} - 6E_{Ala} - E_X - E_{ammonia} - E_{acetic\ acid} \quad (11.3)$$

Both methods yield estimates of the formation energy and give an estimate of helix destabilization by computing the differences in energy for helix formation by condensation and substitution. The interest here is to use the polyalanine as a reference system, and hence the reported data in Table 11.1 corresponds to the values obtained with Eq. 11.2. The energies used here are the values with zero-point vibrational corrections (ZPVCs).

The data clearly reveals the effect the substitution has on the stability of the reference polyalanine helix. Eight substitutions have higher energies than the polyalanine, and ten have lower energies. Substitution by aspartic acid and glutamic acid destabilizes the helix the most, in almost equal amounts (2.98 kcal/mol). On the other hand, arginine (-7.51 kcal/mol) stabilizes the system by almost twice as much as asparagine (-3.87 kcal/mol), which has the second largest effect. The great flexibility of arginine may explain its relative stability. These results are consistent with similar computations reported by Wiczorek and Dannenberg for Gly, Leu, Val, Phe, and Ser at ONIOM B3LYP/D95(d,p) /AM1 level.¹⁷⁵

An almost perfect linear correlation between the residue energies and peptide energies is found, $E_{peptide} = -1615.638059 + E_{residue}$, with correlation of $R^2 = 1.0 - 1.11 \times 10^{-9}$. This correlation is intriguing due to the fact that the energies correspond to independently optimized geometries, whose final geometries do not necessarily match. This result might be used as a reference relationship for developing approximate peptide models. At this point is unclear if it holds for larger peptide chains. But it clearly

Table 11.1: Energy and dipole moment for the residues and peptides. $\Delta\Delta E$ is the relative energy with respect to the polyaniline alpha helix. (Eq. 11.2).

Amino acid	E_X^a	$E_{peptide}^a$	$\Delta\Delta E^a$	$\Delta\Delta E^b$	μ_X^c	$\mu_{peptide}^c$
Ala	-323.551093	-1939.187802	0.0000	0.00	3.06	37.74
Arg	-606.122633	-2221.747371	-0.0120	-7.51	9.37	41.12
Asn	-492.177634	-2107.808168	-0.0062	-3.87	7.75	39.52
Asp	-512.044739	-2127.686191	0.0047	2.98	6.77	38.67
Cys	-721.739240	-2337.372429	-0.0035	-2.21	8.73	39.52
Gln	-531.438793	-2147.078457	0.0030	1.85	4.21	33.80
Glu	-551.313268	-2166.954719	0.0047	2.98	5.31	38.05
Gly	-284.284046	-1899.915116	-0.0056	-3.54	1.64	37.47
His	-548.418423	-2164.052573	-0.0026	-1.61	8.58	37.50
Ile	-441.350339	-2056.988030	0.0010	0.62	3.01	38.97
Leu	-441.351431	-2056.987940	-0.0002	-0.13	2.82	38.51
Lys	-496.660155	-2112.299642	0.0028	1.74	3.86	36.61
Met	-800.268715	-2415.900935	-0.0045	-2.82	2.93	40.76
Phe	-554.384064	-2170.021836	0.0011	0.67	2.29	37.60
Ser	-398.742996	-2014.376595	-0.0031	-1.95	8.28	38.47
Thr	-438.012535	-2053.644413	-0.0048	-3.03	5.34	40.36
Trp	-685.859771	-2301.491789	-0.0047	-2.94	7.96	39.47
Tyr	-629.586367	-2245.225782	0.0027	1.70	2.10	38.02
Val	-402.084433	-2017.723980	0.0028	1.78	2.98	37.86

^aau

^bkcal/mol

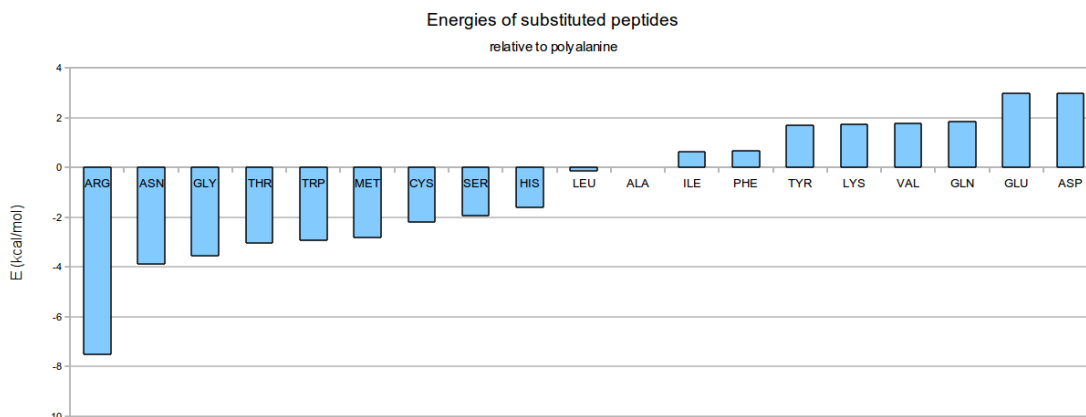
^cD

indicates that a partition of the energy into fragment contributions can be as accurate as other methods.

One remarkable result from the optimized geometries is the resilient prevalence of the dipole moment of the α -helix upon the central amino acid substitution. An average value of 38.42 D ($\sigma = 1.42$ D) for the nineteen helices indicates the great stability of aligned peptide bonds, with a minor perturbation done caused most of side chains.

The overall effect of the central amino acid substitution seems more like a small perturbation to the helix, rather than a real structural change. The specific flexibility and polarizability of the side chain seem to be the main factors determining the relative stability of the final geometry. Because all the peptides can reach an equilibrium

Figure 11.2: Comparison of the energetic effects of the mutation Ala→X for the 4th residue in acetyl(Ala)₇NH₂ (11.2).



energy in α -helical conformation, they are possible conformations that might be altered due to intermolecular side chain interactions rather than due to intramolecular ones.

11.5 Polarizability of the Amino Acids

In previous works by the author, it has been shown how the polarizability of the amino acid side chains is responsible of several of their distinctive biochemical features.^{142,144,179} The polarizability is an extensive property of the system, and several works explore its explicit dependence with atomic volumes.^{77,87,180,181} The polarizability depends on the electron density, particularly on the valence electrons. Certain simple relationships for atoms, in terms of the atomic number, have been known for a long time, as recently reviewed by Schwerdtfeger.¹⁸²

In principle, the precise computation of polarizability requires a higher level of theory than that employed here, (eg, a quantum dynamical approach). Nevertheless, the static polarization can be written as

$$\alpha = \frac{\hbar^2 e^2}{m_e} \sum_{n \neq 0} \frac{f_{n0}}{\Delta E_{n0}^2} \quad (11.4)$$

where the f_{n0} are the oscillator strengths which can be determined from the intensities of the electronic transitions of the molecule and the energies ΔE_{n0} from the frequencies where these transitions occur. This expression can be approximated by replacing the numerator by the number of electrons N_e , according to the Kuhn-Thomas sum rule.

Table 11.2: Polarizability of the amino acids in au.

Amino acid	α^a	α^b
Ala	58.98	52.78
Arg	119.02	117.11
Asn	68.32	75.41
Asp	71.96	70.05
Cys	67.46	76.88
Gln	87.06	88.94
Glu	84.83	85.08
Gly	44.41	39.93
His	98.98	101.61
Ile	99.10	92.08
Leu	98.57	93.06
Lys	101.06	102.62
Met	95.63	105.06
Phe	108.12	123.90
Ser	58.30	56.62
Thr	75.01	70.41
Trp	153.39	156.94
Tyr	125.60	130.60
Val	86.11	77.85

^aFrom Eq. 11.5

^bFrom Ref.¹⁷⁸

By a convenient ansatz, the denominator can be approximated by $\Delta E_{n0}^2 = \Delta E_{gap} \Delta \bar{E}$, where $\Delta E_{gap} = E_{LUMO} - E_{HOMO}$ and $\Delta \bar{E}$ is an averaged value representing the occupied eigenvalues ϵ_i , as obtained from the numerical wavefunction solution^{iv}.

Hence the following polarizability expression can be easily evaluated

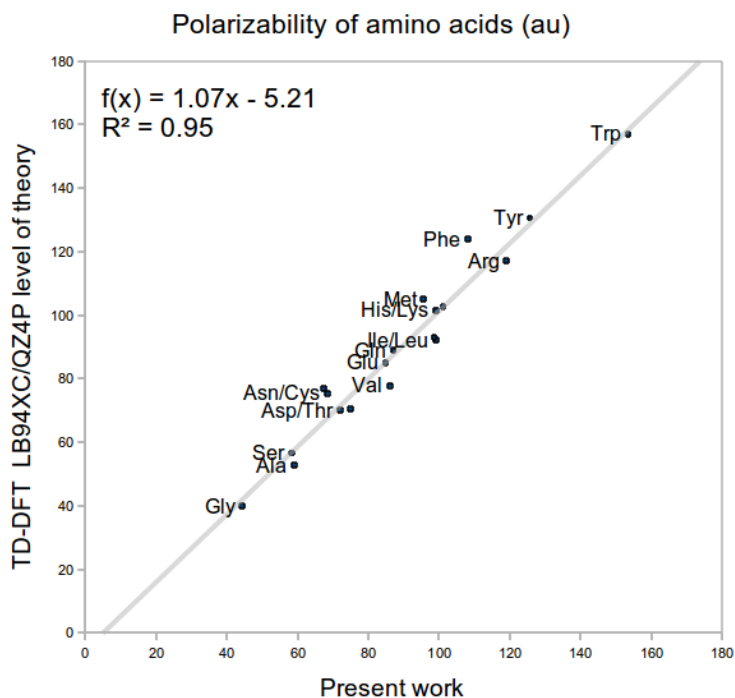
$$\alpha \cong \frac{\hbar^2 e^2}{m_e} \frac{N_e}{\Delta E_{gap} \Delta \bar{E}} \quad (11.5)$$

The computed values for the amino acids are listed in Table 11.2 together with the recent computations by Swart et al., which were obtained at the time-dependent DFT (TD-DFT) level with the LB94 XC-potential in the QZ4P basis, and hence are among the most precise available.¹⁷⁸ The respective values appear in Table 11.2 and

^{iv}

$$\bar{E}_{occupied} = 2 \sum_{i=0}^{N-1} \frac{\epsilon_i}{N! - i!}$$

where N is the number of occupied states.

Figure 11.3: Theoretical polarizability of the amino acids.

are also plotted in Fig 11.3.

It must be noticed that Swart et al. used NH_2CHRCOH as a model for the amino acids, while the model used for the computations of the previous section is $\text{NH}_2\text{CHRCOOH}$. This discrepancy seems to have little effect on the polarizability, because the linear relationship between the two set of values has a surprisingly high correlation ($R^2 = 0.95$), considering the simplicity of the method. The average difference between the two sets of values is 6% ($\sigma = 4\%$). This result confirms the validity of the approximation here introduced with Eq. 11.5, which competes with direct reaction field (DRF) approaches, i.e. approximate values of the polarizability from atomic polarizabilities.

The key element in the equation for the polarization is the involvement of the energy ground levels through the term $\Delta\bar{E}$. Since many response functions in QM can be written in the form $f_{n0}/\Delta E_{n0}^2$, it will be interesting to test the same approximation for other observables. Additional improvements may arise from more accurate values of the term ΔE_{n0}^2 and a benchmarking of the method might validate its wide applicability.

11.6 Conclusions

A variety of properties of the alpha helices can be computed *ab initio*, as has been shown here. The structural results suggest that the level of theory selected for the present work is adequate for the study of oligopeptides of about 100 atoms. The optimized geometries were obtained in relatively short times (~ 50 h CPU time).^v The effect of the polarizability in the stability of the α -helices has been addressed by using the PCM method for the solvent and the dispersion terms in the Hamiltonian. The great stability of the helix upon mutation of the central amino acid explains somehow why this pattern is commonly found in proteins. A relative scale of stability centred at the polyalanine structure was determined. These values can be used as input parameters for statistically-assessed helical propensities of individual amino acids. Due to the cost-effective approach, it can be used for the study and classification of *de novo* mutations in protein-related molecules.

In addition, a novel idea for the *ab initio* computation of the molecular polarizability of amino acids was tested. The agreement with higher levels of theory is encouraging, especially considering that no fitting parameters were required. From the equation 11.5 and the Koopmans' theorem an inverse relationship between the ionization energy and the polarizability is found

$$\alpha \cong \frac{\hbar^2 e^2}{m_e} \frac{N_e}{I \Delta \bar{E}} \quad (11.6)$$

This result questions the validity of approximate expressions according to which $\alpha \propto 1/I^2$.⁷⁷

The small discrepancies between the computed results with Eq. 11.5 and the ones reported in the literature may be attributed to structural differences. More research is required to validate a wide applicability of Eq. 11.5. For the purposes of the amino acids study, this method provide a unique and cost-effective approach to one of the fundamental properties of molecules.

^vBut still far from the CPU times required for being practical in routinely drug design.

Chapter 12

Final Remarks and Conclusions

In the present thesis a theoretical tool for the study of the electronic regions in molecules has been explored. It is based on the single particle momentum which is the momentum associated with the Weizsäcker kinetic energy density. A quantum statistical theorem linking both quantities has been demonstrated (Chapter 3). A more direct relationship can be established by assuming that the classical kinetic energy expression is also valid for the local quantum momentum.

This variable condenses two of the important theorems about the electron density: Kato's cusp condition and the tail exponential behaviour. Both theorems impose physical limits to the speed an electron can achieve inside an atom. Kato's theorem gives an upper limit for the electron speed, which is equal to Z in atomic units. This limit is indeed used to estimate relativistic corrections in chemistry. Here such an idea is further explored via the local momentum to determine the extent of the relativistic effects beyond the core electrons by a lower estimate of relativistic electron-mass corrections (Chapter 4). With this method, a non heuristic assessment of effective core potentials can be addressed.

The second theorem indicates that the kinetic energy of an electron in the outermost shell of an atom is the first ionization energy. This fact suggests an alternative way to approximate the ionization energies, as is shown in Chapter 3. The local momentum also describes the electronic shell structure of atoms, whose boundaries are given by the topological criterion here introduced. The electronic population of each shell agrees with the Aufbau principle and the results compete with other electron localization methods. The behaviour of the local momentum in the tail region suggests a confinement region for the atom, which can be uniquely defined in terms of a physical observable, the ionization energy (Chapter 5). A rather simple relationship between the atomic radius and the ionization energy emerges that fulfils all the expected periodic trends, including the lanthanoid-actinoid contraction -a pure relativistic effect. The new atomic radius is well defined for every atom in the periodic table, unlike most of the conventional atomic radii definitions. In addition, the new

atomic radius holds linear correlations ($R > 0.9$) with five different theoretical atomic radii simultaneously, suggesting its universal validity. The experimentally-determined atomic radius is perhaps a first step for the standardization of a variable that is non-observable in the first place, but from whose definition depend important observable quantities, like the atomic polarization.

For molecular systems, the single-particle momentum reveals its practical utility: it depicts an *ab initio* 3D representation of core electrons and bonding interactions (Chapters 6-7). The only input required is the electron density of the molecule, which makes the method representation independent, in principle. The local momentum is directly related to the topology of the electron density, as it is shown here in different instances. Through the current density operator, a general formulation of the zero flux condition is obtained which is central for the quantum theory of atoms in molecules. In addition, a new set of topological properties of the electron density at the stationary points is found. When the single-particle momentum is used as a tool for the identification of electronic regions in atoms and molecules it is called LED (localized-electrons detector). LED has been explored for numerous molecules, revealing a complementary view that guides the chemical analysis in an objective way. For instance, while aromaticity is not defined, several of the topological properties of the aromatic systems are readily plotted by LED: aromatic bonds have large ellipticities and usually form rings perpendicular to the ellipse main axis, as was discussed for benzene and ethene in Chapters 6 to 8. Hence LED identifies aromatic groups with non obvious orientation, as in the case of organometallic compounds, for example.

In the present work conclusive evidence of the stabilizing role of hydrogen-hydrogen interactions at the bay region of phenanthrene has been presented. The arguments given in Chapter 8 add to the discussion a mechanism under which these two atoms decrease their exposed areas, therefore lowering the steric effect, ultimately making this molecule more stable than its extended isomer. The atomic contributions to the Weizsäcker energy give an estimate of the stability due to the hydrogen atoms. The value found coincides with other computations on the same system by different procedures. The mechanism that makes the bay hydrogens of phenanthrene to be close together is described in terms of their electronic polarizations. The central part of such a mechanism involves the polarizability as accounted for by the kinetic energy associated with the single-particle momentum.

LED also works even where QTAIM is expected to experience difficulties. In particular, metallic solids have almost constant electron density in the conduction

region, and hence the concept of a critical points analysis becomes somehow less relevant for those regions, while a continuous variable is more suitable for their study. LED shows in a colourful way the presence of conduction regions in metallic and semiconductor systems, while also revealing the atomic shells of the involved atoms. Hence LED can be used to explore surface effects of bulk electrons, materials defects, substitutions, etc, by providing a 3D picture of the systems that agrees with the chemistry and the physics of the system.

The results discussed here indicate that the relevance of the Weizsäcker term in multielectronic systems is well beyond a mere “gradient correction”, as has been historically considered in DFT. With the present results, the most significant contribution achieved here is that of revealing the polarization regions in molecules, given its relevance for noncovalent and dispersion interactions. The electronic polarizability is responsible for the similarity of amino acids, as reviewed in Chapter 9. Hence the present work might be relevant for the study of macromolecules of biological interest, like proteins. The study of similarity measures (Chapter 9) and the simulation of alpha helices (Chapter 11) are steps in that direction. Along with the reexamination of the quantum similarity measure emerges a notion that chemically-relevant information can be derived from physical observables, without the necessity of ill defined variables. The simulation of the alpha helices provides information about the stability of these polypeptides via a substitution of the central amino acid. A state-of-the-art quantum mechanical simulation allows the optimization of these geometries. The solvent is simulated with a continuous-polarizable medium and the Hamiltonian includes dispersion interactions. One result that indicates the successful simulation is the resilient persistence of the dipole moment held by all the helices, despite the size or composition of the central residue. The results also indicate that the longer side chains are able to further stabilize the helical structure, a fact attributable mainly to electrostatic interactions between the side chain and the dipole of the main chain. The obtained relative stability agrees with the experimentally observed trends. In addition, a new method for the ab initio computation of the polarizability in molecules is introduced (Chapter 11). The results are comparable to those obtained with higher levels of theory.

From the results discussed here, a general conclusion may be drawn: polarizability is perhaps, after electron-pairing, the second most important mechanism involved in bonding formation, and the local momentum is a useful theoretical tool for studying both bonding mechanisms.

12.1 Outlook

Some of the potential applications of the analysis introduced here have been already explored but others still can be further developed. For instance, a combination of approximate theoretical models can use LED for the automatic assessment of bonding interactions in *de novo* obtained experimentally-derived electron densities. Most of the identification of molecular groups in protein crystals is done by hand, and LED offers a new possibility to assess that information automatically, because, thanks to Kato's cusp condition, it locates in the 3D space the nuclear coordinates automatically. In addition, strong covalent bonds can be detected with LED, which can reduce substantially the effort to decipher X-ray electron densities. Another possibility is the further exploration of the polarization regions in molecules, as a way to understand its role in bonding formation. Polarization interactions are increasingly involved in different areas of theoretical chemistry like DFT, conceptual chemistry and newer classical mechanics methods. LED offers a unique opportunity to explore and validate these different approaches. All of these discussions can be centrally addressed with the local momentum formulation described in this thesis, mainly because, instead of competing with these other developments, the present work supports and complements them. Perhaps where the applications of the present developments are less understood is in the area of time-dependent quantum mechanics. The idea of involving a momentum quantity that is directly related to the electron regions may suggest that the simulations of structural changes induced by photons may be formulated in a simpler way than the conventional radiation-matter formalism. The implementation of LED with time dependent variables may be useful for the study of chromophores and other light harvesting systems. For now, the essential developments discussed in the present thesis guarantee that any further exploration involving the local momentum variable may start with solid theoretical grounds, which were scarce and dispersed before the present research was initiated.

Bibliography

- [1] Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864–871.
- [2] Bader, R. F. W. *Can. J. Chem.* **1962**, *40*, 1164–1175.
- [3] Kohn, W. *Rev. Mod. Phys.* **1999**, *71*, 1253–1266.
- [4] Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397–5403.
- [5] Liu, S. *J. Chem. Phys.* **2007**, *126*, 244103–244103.
- [6] Hoffmann, R. *J. Mol. Struct.* **1998**, *424*, 1 – 6; A Faithful Couple: Qualitative and Quantitative Understanding of Chemistry.
- [7] Feynman, R. P. *Phys. Rev.* **1939**, *56*, 340–343.
- [8] Szabo, A.; Ostlund, N. *Modern Quantum Chemistry*; Dover, 1996.
- [9] Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Oxford University Press, 1990.
- [10] Bader, R. F. W. *Phys. Rev. B* **1994**, *49*, 13348–13356.
- [11] Nguyen-Dang, T. T.; Bader, R. F. W. *Physica A* **1982**, *114*, 68–73.
- [12] Srebrenik, S.; Bader, R. F. W.; Nguyen-Dang, T. T. *J. Chem. Phys.* **1978**, *68*, 3667–3679.
- [13] Walker, P.; Mezey, P. *J. Am. Chem. Soc.* **1994**, *116*, 12022–12032.
- [14] Pichon-Pesme, V.; Lecomte, C.; Wiest, R.; Bonard, M. *J. Am. Chem. Soc.* **1992**, *114*, 2713–2715.
- [15] Savin, A.; Nesper, R.; Wengert, S.; Fässler, T. *Angew. Chem. Int. Ed.* **1997**, *36*, 1808–1832.
- [16] Jelsch, C.; Pichon-Pesme, V.; Lecomte, C.; Aubry, A. *Acta Cryst.* **1998**, *D54*, 1306–1318.
- [17] Breneman, C. M.; Thompson, T. R.; Rhem, M.; Dung, M. *Comp. Chem.* **1995**, *19*, 161–173; Third Conference on Computers in Chemistry.
- [18] Matta, C. F.; Boyd, R. J. *The Quantum Theory of Atoms in Molecules. From Solid State to DNA and Drug Design*; Wiley-VCH, 2007.
- [19] Murphy, J. *Scientific Computing World* **2003**, *May-Jun.*

- [20] *The Nobel Prize in Physics 1965*. http://nobelprize.org/nobel_prizes/physics/laureates/1965/.
- [21] Wiberg, K.; Rablen, P. *J. Comp. Chem.* **1993**, *14*, 1504–1518.
- [22] Angyan, J.; Jansen, G.; Loos, M.; Hattig, C.; Hess, B. *Chem. Phys. Lett.* **1994**, *219*, 267.
- [23] Lobayan, R. M.; Boichichio, R. C.; Lain, L.; Torre, A. *J. Chem. Phys.* **2005**, *123*, 144116.
- [24] Bader, R. F. W.; Johnson, S.; Tang, T. H.; Popelier, P. L. A. *J. Phys. Chem.* **1996**, *100*, 15398–15415.
- [25] Bamzai, A. S.; Deb, B. M. *Rev. Mod. Phys.* **1981**, *53*, 95.
- [26] Spackman, M.; Howard, J.; Destro, R. *IUCr Newsletter* **2000**, *8*, 2.
- [27] Bader, R. F. W.; et. al. *Faraday Discuss.* **2007**, *135*, 125.
- [28] Luo, S. L. *Int. J. Theor. Phys.* **2002**, *41*, 1713–1731.
- [29] Kato, W. A. *Commun. Pure. Appl. Math* **1957**, *10*, 151–177.
- [30] Hoffmann-Ostenhof, M.; Hoffmann-Ostenhof, T. *Phys. Rev. A* **1977**, *16*, 1782–1785.
- [31] March, N. H. *Electron density of atoms and molecules*; Academic Press, 1992.
- [32] Mosna, R. A.; Hamilton, I. P.; Site, L. D. *J. Phys. A: Math. Gen.* **2005**, *38*, 3869–3878.
- [33] Hamilton, I. P.; Mosna, R. A.; Site, L. D. *Theor. Chem. Acc.* **2007**, *118*, 407–415.
- [34] Nelson, E. *Quantum Fluctuations*; Princeton Univ. Press, 1985.
- [35] Ghosh, S. K.; Deb, B. M. *Phys. Rep.* **1982**, *92*, 1–44.
- [36] Ghosh, S. K. *Curr. Sci.* **1983**, *52*, 769–774.
- [37] Kan, K.-K.; Griffin, J. J. *Phys. Rev. C* **1977**, *15*, 1126–1151.
- [38] Salesi, G.; Recami, E. *Found. Phys.* **1998**, *28*, 763–776.
- [39] Holland, P. R. *The Quantum Theory of Motion*; Cambridge Univ. Press, 1993.
- [40] Wan, K. K.; Sumner, P. J. *Phys. Lett. A* **1988**, *128*, 458–462.
- [41] Cohen, L. *Phys. Lett. A* **1996**, *212*, 315–319.

- [42] Fisher, R. A. *Proceed. Camb. Phil. Soc.* **1925**, *22*, 700–725.
- [43] Bohm, D. *Phys. Rev.* **1952**, *85*, 166–179.
- [44] Bohm, D. *Phys. Rev.* **1952**, *85*, 180–193.
- [45] Hirschfelder, J. O.; Christophe, A.; Palke, W. E. *J. Chem. Phys.* **1974**, *61*, 5435–5455.
- [46] Site, L. D. *Int. J. Mod. Phys. B* **2000**, *14*, 771–781.
- [47] Site, L. D. *Int. J. Mod. Phys. B* **2000**, *14*, 1891–1901.
- [48] Sears, S. B.; Parr, R.; Dinur, U. *Isr. J. Chem.* **1980**, *19*, 165–173.
- [49] Jaramillo, J.; Scuseria, G.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 1068–1073.
- [50] Kaupp, M.; Bahmann, H.; Arbuznikov, A. V. *J. Chem. Phys.* **2007**, *127*, 194102.
- [51] Garcia-Aldea, D.; Alvarillos, J. E. *J. Chem. Phys.* **2007**, *127*, 144109.
- [52] Garcia-Aldea, D.; Alvarillos, J. E. *Phys. Rev. A* **2008**, *77*, 022502.
- [53] Hunter, G. *Int. J. Quant. Chem.* **1986**, *29*, 197–204.
- [54] Hunter, G. *Can. J. Chem.* **1996**, *74*, 1008–1013.
- [55] Becke, A. D. *Int. J. Quantum Chem.* **1989**, *23*, 599–609.
- [56] Becke, A. D.; Dickson, R. M. *J. Chem. Phys.* **1990**, *92*, 3610–3612.
- [57] March, N. H. *Phys. Rev. A* **1986**, *33*, 88–89.
- [58] Pyykko, P. *Chem. Rev.* **1988**, *88*, 563–594.
- [59] Bohórquez, H. J.; Matta, C. F.; Boyd, R. J. **2011**; To be submitted.
- [60] Kohout, M.; Savin, A.; Preuss, H. *J. Chem. Phys.* **1991**, *95*, 1928–1942.
- [61] Nagy, A.; March, N. H. *Mol. Phys.* **1997**, *90*, 271–276.
- [62] Martin, W. C.; Musgrove, A.; Kotochigova, S.; Sansonetti, J. E.; *Table of Ground Levels and Ionization Energies for the Neutral Atoms*; NIST Standard Reference Database 111; 2003.
- [63] Cabrera-Trujillo, R.; Apell, P.; Oddershede, J.; Sabin, J. R. *AIP Conference Proceedings* **2003**, *680*, 86–89.
- [64] Kohout, M.; Savin, A. *Int. J. Quant. Chem.* **1996**, *60*, 875–882.

- [65] Navarrete-Lopez, A. M.; Garza, J.; Vargas, R. *J. Chem. Phys.* **2008**, *128*, 104110–104110.
- [66] Schmider, H. L.; Becke, A. D. *J. Mol. Struct. (Theochem)* **2000**, *527*, 51–61.
- [67] Harcourt, R. D. *Found. Chem.* **1999**, *1*, 295–296.
- [68] Boyd, R. J. *J. Phys. B: Atom. Molec. Phys.* **1976**, *9*, L69–L72.
- [69] Shi, Z.; Boyd, R. J. *J. Chem. Phys.* **1988**, *88*, 4375–4377.
- [70] Bader, R. F. W.; Essen, H. *J. Chem. Phys.* **1984**, *80*.
- [71] Eickerling, G.; Reiher, M. *J. Chem. Theory. Comput.* **2008**, *4*, 286–296.
- [72] Wang, W. P.; Parr, R. G. *Phys. Rev. A* **1977**, *16*, 891–902.
- [73] Haiduke, R. L. A.; de Paiva Martins Filho, H.; da Silva, A. B. F. *Chem. Phys.* **2008**, *348*, 89 – 96.
- [74] Dolg, M.; Stoll, H. *Handbook on the Physics and Chemistry of Rare Earths*; K.A. Gschneidner, J.; Eyring, L., Eds.; Elsevier: Amsterdam, 1996; Chapter Electronic structure calculations for molecules containing lanthanide atoms, p 607.
- [75] Nagle, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 4741–4747.
- [76] Ghosh, D. C.; Gupta, K. *J. Theor. Comp. Chem.* **2006**, *5*, 895–911.
- [77] Politzer, P.; Jin, P.; Murray, J. S. *J. Chem. Phys.* **2002**, *117*, 8197–8202.
- [78] Putz, M.; Russo, N.; Sicilia, E. *J. Phys. Chem. A* **2003**, *107*, 5461–5465.
- [79] Ellenbogen, J. C. *Phys. Rev. A* **2006**, *74*, 034501.
- [80] Slater, J. C. *J. Chem. Phys.* **1964**, *41*, 3199–3204.
- [81] Froese-Fischer, C. *Atomic Data* **1972**, *4*, 87–99.
- [82] Waber, J. T.; Cromer, D. T. *J. Chem. Phys.* **1965**, *42*, 4116–4123.
- [83] Clementi, E.; Raimondi, D. L.; Reinhard, W. *J. Chem. Phys.* **1967**, *47*, 1300–1307.
- [84] Boyd, R. J. *J. Phys. B* **1977**, *10*, 2283–2291.
- [85] Zhang, M. B.; Zhao, D. X.; Yang, Z. Z. *J. Theor. Comput. Chem.* **2005**, *4*, 281–288.

- [86] Bader, R. F. W.; Henneker, W. H.; Cade, P. E. *J. Chem. Phys.* **1967**, *46*, 3341–3363.
- [87] Ghosh, D. C.; Biswas, R. *Int. J. Mol. Sci.* **2002**, *3*, 87–113.
- [88] Boyd, R. J. *Nature* **1974**, *250*, 566–567.
- [89] Luo, S. *J. Phys. A* **2002**, *35*, 5181–5187.
- [90] Bohórquez, H. J.; Boyd, R. J. *J. Chem. Phys.* **2008**, *129*, 024110.
- [91] Martin, W.; Musgrove, A.; Kotochigova, S.; Sansonetti, J. *Ground levels and ionization energies for the neutral atoms*; NIST, Ed.; NIST: <http://physics.nist.gov>, 2003; Vol. Physical Reference Data.
- [92] Eve, A. S. *Nature* **1921**, *107*, 552–553.
- [93] Saha, M. *Nature* **1921**, *107*, 682–683.
- [94] Ahrens, L. H. *Geochim. Cosmochim. Acta* **1952**, *2*, 155–169.
- [95] Roy, R. S. *Spectrochim. Acta B* **1972**, *B 27*, 323–326.
- [96] Tamura, S. *J. Mat. Sci. Lett.* **1996**, *15*, 1678–1679.
- [97] Lewis, G. N. *J. Am. Chem. Soc.* **1916**, *38*, 762–785.
- [98] Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683–686.
- [99] Bohórquez, H. J.; Boyd, R. J. *Chem. Phys. Lett.* **2009**, *480*, 127–131.
- [100] Luken, W. L.; Culberson, J. C. *Int. J. Quantum Chem.* **1982**, *16*, 265–276.
- [101] Dobson, J. F. *J. Chem. Phys.* **1991**, *94*, 4328–4333.
- [102] Gatti, C. *Z. Kristallogr.* **2005**, *220*, 399–457.
- [103] Silvi, B.; Savin, A. *Min. Mag.* **1994**, *58A*, 842–843.
- [104] Samuelsson, P.; Büttiker, M. *Phys. Rev. Lett.* **2002**, *89*, 046601.
- [105] Frisch, M. J.; et al.; *Gaussian 09 Revision A.1*; Gaussian Inc. Wallingford CT 2009.
- [106] Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *13*, 1605–1612.
- [107] Bader, R. F. W. *J. Phys. Chem. A* **2009**, *113*, 10391–10396.
- [108] Arunan, E.; Scheiner, S. *Chem. Int.* **2007**, *29*, 1.

- [109] Elgobashi, N.; González, L. *J. Chem. Phys.* **2006**, *124*, 174308.
- [110] Klopper, W.; Quack, M.; Suhm, M. *J. Chem. Phys.* **1998**, *108*, 10096.
- [111] Meuwly, M.; Hutson, J. M. *J. Chem. Phys.* **1999**, *110*, 8338.
- [112] Bohórquez, H. J.; Boyd, R. J. *Theor. Chem. Acc.* **2010**, *127*, 393–400; 10.1007/s00214-010-0727-5.
- [113] Uhlenbeck, G.; Goudsmit, S. *Naturwissenschaften* **1925**, *47*, 953.
- [114] Pauli, W. *Zeitschrift für Physik* **1925**, *31*, 765–783.
- [115] Bader, R. F. W.; Heard, G. L. *J. Chem. Phys.* **1999**, *111*, 8789–8798.
- [116] Sperber, G. *Int. J. Quantum Chem.* **1971**, *5*.
- [117] Westin, C.-F.; Bhalerao, A.; Kikinis, R.; Knutsson, H.; IEEE Computer Society: Los Alamitos, CA, USA; p 794.
- [118] Frisch, M. J.; et al.; *Gaussian 03, Revision C.02*; Gaussian, Inc., Wallingford, CT, 2004.
- [119] Keith, T. A.; Bader, R. F. W. *J. Chem. Phys.* **1993**, *99*, 3669.
- [120] Bader, R. F. W. *J. Chem. Phys.* **1980**, *73*, 2871–2883.
- [121] Zhang, X.; Houk, K. N. *Acc. Chem. Res.* **2005**, *38*, 379–385; PMID: 15895975.
- [122] Cioslowski, J.; Mixon, S. T.; Edwards, W. D. *J. Am. Chem. Soc.* **1991**, *113*, 1083–1085.
- [123] Popelier, P. In *Structure and Bonding*; Wales, D., Ed.; Springer-verlag, 2005; Vol. 115, Chapter I, Quantum chemical topology: on bonds and potentials, pp 1–56.
- [124] Bohórquez, H. J.; Matta, C. F.; Boyd, R. J. *Int. J. Quantum Chem.* **2010**, *110*, 2418–2425.
- [125] Vydrov, O. A.; Van Voorhis, T. *Phys. Rev. A* **2010**, *81*, 062708.
- [126] Pacios, L. F.; Fernandez, A. *J. Mol. Graphics Modell.* **2009**, *28*, 102 – 112.
- [127] Keith, T. A.; *AIMAll (Version 10.06.21)*; 2010. aim.tkgristmill.com.
- [128] Matta, C. *Hydrogen Bonding-New Insights*; Grabowski, S., Ed.; Springer, 2006; Chapter 9, pp 337–375.
- [129] Shaik, S.; Danovich, D.; Wu, W.; Hiberty, P. C. *Nature Chem.* **2009**, *1*, 443–449.

- [130] Poater, J.; Visser, R.; Sola, M.; Bickelhaupt, F. M. *J. Org. Chem.* **2007**, *72*, 1134–1142.
- [131] Wang, S.-G.; Qiu, Y.-X.; Schwarz, W. H. E. *Chem-A Eur. J.* **2009**, *15*, 6032–6040.
- [132] Bader, R.; Matta, C. *Inorg. Chem.* **2001**, *40*, 5603–5611.
- [133] Farrugia, L. J.; Evans, C.; Lentz, D.; Roemer, M. *J. Am. Chem. Soc.* **2009**, *131*, 1251–1268.
- [134] Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.
- [135] Bender, A.; Glen, R. C. *Org. Biomol. Chem.* **2004**, *2*, 3204 – 3218.
- [136] Carbó, R.; Layda, L.; Arnau, M. *Int. J. Quantum Chem.* **1980**, *17*, 1185–1189.
- [137] Goldstein, H. *Classical Mechanics*; Addison-Wesley, 1980.
- [138] Bader, R. F. W.; Becker, P. *Chem. Phys. Lett.* **1988**, *148*, 452–458.
- [139] Popelier, P. L. A. *J. Phys. Chem. A* **1999**, *103*, 2883–2890.
- [140] O'Brien, S.; Popelier, P. *Can. J. Chem.* **1999**, *77*, 28–36.
- [141] Popelier, P. L. A.; Aicken, F. M.; O'brien, S. E. *Chemical modelling: application and theory*; Cambridge: Royal Society of Chemistry, 2000; Vol. 1, Chapter 3, pp 143–198.
- [142] Bohórquez, H. J.; Obregón, M.; Cárdenas, C.; Llanos, E.; Suárez, C.; Villaveces, J. L.; Patarroyo, M. E. *J. Phys. Chem. A* **2003**, *107*, 10090–10097.
- [143] Kumar, D. A. *Mini Rev. Med. Chem.* **2001**, *1*, 187–195.
- [144] Cárdenas, C.; Obregón, M.; Llanos, E.; Machado, E.; Bohórquez, H. J.; Villaveces, J. L.; Patarroyo, M. E. *J. Comp. Chem.* **2002**, *26*, 631–646.
- [145] Andrews, D. *Biometrics* **1972**, *28*, 125–136.
- [146] Cárdenas, C.; Villaveces, J. L.; Bohórquez, H.; Llanos, E.; Suárez, C.; Obregón, M.; Patarroyo, M. E. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 1265–1277.
- [147] Lee, K.; Xie, D.; Freire, E.; Amzel, L. *PROTEINS Struct. Funct. Gen.* **1994**, *20*, 68–84.
- [148] Ho, B. K.; Agard, D. A. *BMC Struc. Biol.* **2008**, *8*.
- [149] Trinastic, N. *Chemical Graph Theory*; CRC Press Inc., 1992.

- [150] Garcia-Domenech, R.; Galvez, J.; de Julian-Ortiz, J. V.; Pogliani, L. *Chem. Rev.* **2008**, *108*, 1127–1169.
- [151] SemiChem; the University of Florida; *CODESSA: Comprehensive Descriptors for Structural and Statistical Analysis*; 1995.
- [152] Rohlf, F. J. *Numerical Taxonomy and Multivariate Analysis System (NTSYS)*, Ver. 1.8; Exeter Publishing, Ltd.: Setauket, NY., 1992.
- [153] Haeberlein, M.; Brinck, T. *J. Chem. Soc-P. Trans. 2* **1997**, 289–294.
- [154] Matta, C. F.; Bader, R. F. W. *Proteins Struct. Funct. Genet.* **2000**, *40*, 310–329.
- [155] Matta, C. F.; Bader, R. F. W. *Proteins Struct. Funct. Genet.* **2002**, *48*, 519–538.
- [156] Matta, C. F.; Bader, R. F. W. *Proteins Struct. Funct. Genet.* **2003**, *52*, 360–399.
- [157] Chang, C.; Bader, R. *J. Phys. Chem.* **1992**, *96*, 1654–1662.
- [158] Popelier, P.; Bader, R. *J. Phys. Chem.* **1994**, *98*, 4473–4481.
- [159] Biegler-Konig, F.; Bader, R.; Tang, T. *J. Comp. Chem.* **1982**, *3*, 317.
- [160] Kawashima, S.; Pokarowski, P.; Pokarowska, M.; Kolinski, A.; Katayama, T.; Kanehisa, M. *Nucleic Acids Res.* **2008**, *36*, D202–5.
- [161] Huang, J.; Kawashima, S.; Kanehisa, M. *Genome Inform.* **2007**, *18*, 152–161.
- [162] Kidera, A.; Konishi, Y.; Oka, M.; Ooi, T.; Scheraga, H. *J. Protein Chem.* **1985**, *4*, 265–297.
- [163] Esteve, J. G.; Falceto, F. *Biophys. Chem.* **2005**, *115*, 177–180.
- [164] Patarroyo, M.; Romero, P.; Torres, M.; Clavijo, P.; Moreno, A.; Martinez, A.; Rodriguez, R.; Guzman, F.; Cabezas, E. *Nature* **1987**, *328*, 629–632.
- [165] Patarroyo, M.; Amador, R.; Clavijo, P.; Moreno, A.; Guzman, F.; Romero, P.; Tascon, R.; Franco, A.; Murillo, L.; Ponton, G.; Trujillo, G. *Nature* **1988**, *332*, 158–161.
- [166] Agudelo, W. A.; Galindo, J. F.; Ortiz, M.; Villaveces, J. L.; Daza, E. E.; Patarroyo, M. E. *PLoS ONE* **2009**, *4*.
- [167] Agudelo, W. A.; Patarroyo, M. E. *Mini-Rev. Med. Chem.* **2010**, *10*, 746–758.
- [168] Cárdenas, M. B. A. V. J., C; Ortiz; Patarroyo, M. E. *Biochem. Biophys. Res. Comm.* **2005**, *330*, 1162–1167.

- [169] Cárdenas, J. S. C. M. J. L., C; Villaveces; E.; , P. M. *J. Struct. Biol.* **2005**, *149*, 38–52.
- [170] Balbin, C. V. J., A; Cardenas; Patarroyo, M. E. *Biochimie* **2006**, *88*, 1307–1311.
- [171] BUCK, M. *Quarterly Reviews of Biophysics* **1998**, *31*, 297–355.
- [172] Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129.
- [173] Barone, V.; Cossi, M.; Tomasi, J. *The Journal of Chemical Physics* **1997**, *107*, 3210–3221.
- [174] Pascual-Ahuir, J. L.; Silla, E.; Tunon, I. *J. Comput. Chem.* **1994**, *15*, 1127–1138.
- [175] Wieczorek, R.; Dannenberg, J. *J. Am. Chem. Soc.* **2005**, *127*, 17216–17223.
- [176] LaPointe, S. M.; Farrag, S.; Bohórquez, H. J.; Boyd, R. J. *J. Phys. Chem. B* **2009**, *113*, 10957–10964.
- [177] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- [178] Swarta, M.; Snijdersb, J. G.; van Duijnenb, P. T. *J. Comp. Meth. Sci. Eng.* **2004**, *4*, 419–425.
- [179] Bohórquez, H. J.; Cárdenas, C.; Matta, C. F.; Boyd, R. J.; Patarroyo, M. E. *Quantum Biochemistry*; Matta, C. F., Ed.; Wiley-VCH, 2010; Chapter 13, pp 403–421.
- [180] Ghanty, T. K.; Ghosh, S. K. *J. Phys. Chem.* **1996**, *100*, 17429–17433.
- [181] Ghosh, D. C. *JTCC* **2005**, *4*, 21 – 33.
- [182] Schwerdtfeger, P. *Atoms, molecules and clusters in electric fields: theoretical approaches to the calculation of electric polarizability*; Maroulis, G., Ed.; Imperial College Press, 2006; Chapter I, pp 1–32.

Appendix A

Copyright Permission Letters

The following pages include the copyright permission letters from the publishers of the papers included in this Thesis.

The annexed letters correspond to the following chapters and their respective publications:

1. **Chapter 3:** Bohórquez, H. J. and Boyd, R. J., On the local representation of the electronic momentum operator in atomic systems, *J. Chem. Phys.* (2008), pp. 024110.
2. **Chapter 5:** Bohórquez, H. J. and Boyd, R. J., Is the size of an atom determined by its ionization energy?, *Chem. Phys. Lett.* (2009), pp. 127-131.
3. **Chapter 6:** Bohórquez, H. J. and Boyd, R. J., A localized electrons detector for atomic and molecular systems, *Theor. Chim. Acta.* (2010), pp. 393-400.
4. **Chapter 7:** Bohórquez, H. J.; Matta, C. F. and Boyd, R. J., The localized electrons detector as an ab initio representation of molecular structures, *Int. J. Quantum Chem.* (2010), pp. 2418-2425.
5. **Chapter 10:** Bohórquez, H. J.; Cárdenas, C.; Matta, C. F.; Boyd, R. J. and M. E. Patarroyo, Methods in biocomputational chemistry: a lesson from the amino acids, in *Quantum Biochemistry*, Ch. 13, Wiley-VCH (2010), pp. 403-421.

**AMERICAN INSTITUTE OF PHYSICS LICENSE
TERMS AND CONDITIONS**

Jan 17, 2011

This is a License Agreement between Hugo J Bohorquez ("You") and American Institute of Physics ("AIP") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Institute of Physics , and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2591400933212
License date	Jan 17, 2011
Licensed content publisher	American Institute of Physics
Licensed content publication	Journal of Chemical Physics
Licensed content title	On the local representation of the electronic momentum operator in atomic systems
Licensed content author	Hugo Bohórquez, Russell Boyd
Licensed content date	Jul 10, 2008
Volume number	129
Issue number	2
Type of Use	Thesis/Dissertation
Requestor type	Author (original article)
Format	Electronic
Portion	Excerpt (> 800 words)
Will you be translating?	No
Title of your thesis / dissertation	Local Quantum Chemistry
Expected completion date	Jan 2011
Estimated size (number of pages)	160
Total	0.00 USD

Terms and Conditions

American Institute of Physics -- Terms and Conditions: Permissions Uses

American Institute of Physics ("AIP") hereby grants to you the non-exclusive right and license to use and/or distribute the Material according to the use specified in your order, on a one-time basis, for the specified term, with a maximum distribution equal to the number that you have ordered. Any links or other content accompanying the Material are not the subject of this license.

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Apr 07, 2011

This is a License Agreement between Hugo J Bohorquez ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Hugo J Bohorquez
Customer address	1339 EDWARD ST HAIFAX, NS B3H3H5
License number	2643781481791
License date	Apr 07, 2011
Licensed content publisher	Elsevier
Licensed content publication	Chemical Physics Letters
Licensed content title	Is the size of an atom determined by its ionization energy?
Licensed content author	Hugo J. Bohórquez, Russell J. Boyd
Licensed content date	28 September 2009
Licensed content volume number	480
Licensed content issue number	1-3
Number of pages	5
Start Page	127
End Page	131
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Local Quantum Chemistry
Expected completion date	Apr 2011

**SPRINGER LICENSE
TERMS AND CONDITIONS**

Jan 17, 2011

This is a License Agreement between Hugo J Bohorquez ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2591420116669
License date	Jan 17, 2011
Licensed content publisher	Springer
Licensed content publication	Theoretical Chemistry Accounts
Licensed content title	A localized electrons detector for atomic and molecular systems
Licensed content author	Hugo J. Bohórquez
Licensed content date	Jan 1, 2010
Volume number	127
Issue number	4
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	5
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Title of your thesis / dissertation	Local Quantum Chemistry
Expected completion date	Jan 2011
Estimated size(pages)	160
Total	0.00 USD
Terms and Conditions	

**JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS**

Apr 07, 2011

This is a License Agreement between Hugo J Bohorquez ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	2643791385510
License date	Apr 07, 2011
Licensed content publisher	John Wiley and Sons
Licensed content publication	International Journal of Quantum Chemistry
Licensed content title	The localized electrons detector as an ab initio representation of molecular structures
Licensed content author	Hugo J. Bohórquez,Chérif F. Matta,Russell J. Boyd
Licensed content date	Nov 5, 2010
Start page	2418
End page	2425
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Order reference number	
Total	0.00 USD

Terms and Conditions

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at <http://myaccount.copyright.com>)

Date: Tue, 10 May 2011 16:41:16 +0100 [05/10/11 12:41:16 ADT]
 From: Rights DE <RIGHTS-and-LICENCES@wiley-vch.de>
 To: 'Hugo J Bohorquez' <HUGO@dal.ca>
 Subject: AW: AW: Form: Permission request

Dear Hugo J Bohorquez,

Thank you for your email.

We hereby grant permission for the requested use expected that due credit is given to the original source.

If material appears within our work with credit to another source, authorisation from that source must be obtained.

Credit must include the following components:

- Books: Author(s)/ Editor(s) Name(s): Title of the Book. Page(s). Publication year. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

- Journals: Author(s) Name(s): Title of the Article. Name of the Journal. Publication year. Volume. Page(s). Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

With kind regards

Bettina Loycke

Bettina Loycke

Senior Rights Manager

Wiley-VCH Verlag GmbH & Co. KGaA

Boschstr. 12

69469 Weinheim

Germany

Phone: +49 (0) 62 01- 606 - 280

Fax: +49 (0) 62 01 - 606 - 332

Email: rights@wiley-vch.de

 Wiley-VCH Verlag GmbH & Co. KGaA

Location of the Company: Weinheim

Chairman of the Supervisory Board: Stephen Michael Smith

Trade Register: Mannheim, HRB 432833

General Partner: John Wiley & Sons GmbH, Location: Weinheim

Trade Register Mannheim, HRB 432296

Managing Directors : Christopher J. Dicks, Bijan Ghawami, William Pesce