

Lecture Notes in Chemistry 112

Chérif F. Matta
Paul W. Ayers
Ronald Cook

Electron Localization- Delocalization Matrices

 Springer

Lecture Notes in Chemistry

Volume 112

Series Editors

Barry Carpenter, Cardiff, UK

Paola Ceroni, Bologna, Italy

Katharina Landfester, Mainz, Germany

Jerzy Leszczynski, Jackson, USA

Tien-Yau Luh, Taipei, Taiwan

Eva Perlt, Bonn, Germany

Nicolas C. Polfer, Gainesville, USA

Reiner Salzer, Dresden, Germany

Kazuya Saito, Department of Chemistry, University of Tsukuba, Tsukuba, Japan

The series Lecture Notes in Chemistry (LNC), reports new developments in chemistry and molecular science - quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge for teaching and training purposes. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research. They will serve the following purposes:

- provide an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas,
- provide a source of advanced teaching material for specialized seminars, courses and schools, and
- be readily accessible in print and online.

The series covers all established fields of chemistry such as analytical chemistry, organic chemistry, inorganic chemistry, physical chemistry including electrochemistry, theoretical and computational chemistry, industrial chemistry, and catalysis. It is also a particularly suitable forum for volumes addressing the interfaces of chemistry with other disciplines, such as biology, medicine, physics, engineering, materials science including polymer and nanoscience, or earth and environmental science.

Both authored and edited volumes will be considered for publication. Edited volumes should however consist of a very limited number of contributions only. Proceedings will not be considered for LNC.

The year 2010 marked the relaunch of LNC.

Chérif F. Matta · Paul W. Ayers · Ronald Cook

Electron Localization-Delocalization Matrices

 Springer

Chérif F. Matta
Department of Chemistry and Physics
Mount Saint Vincent University
Halifax, NS, Canada

Department of Chemistry
Saint Mary's University
Halifax, NS, Canada

Département de chimie
Université Laval
Québec, QC, Canada

Department of Chemistry
Dalhousie University
Halifax, NS, Canada

Ronald Cook
Molecular Design Innovations
Golden, CO, USA

Paul W. Ayers
Department of Chemistry & Chemical
Biology
McMaster University
Hamilton, ON, Canada

ISSN 0342-4901

Lecture Notes in Chemistry

ISBN 978-3-031-51432-6

<https://doi.org/10.1007/978-3-031-51434-0>

ISSN 2192-6603 (electronic)

ISBN 978-3-031-51434-0 (eBook)

© Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

This book is dedicated to four scientists we admire and whose oeuvre has laid the foundations for the work presented in this monograph:

*To the memories of
Richard F. W. Bader (1931–2012) and Robert G. Parr (1921–2017).*



and to Todd A. Keith and Ismat Sumar.



Foreword

Quantum chemistry has come a very long way from being restricted to small molecules and specialist applications to the major research tool it is today. In fact tens of thousands of quantum chemists use quantum chemistry every day in academia as well as industry in order to solve real-life chemistry problems, to obtain inspiration for new experiments, or to interpret their experimental data. While a big part of this success is the increasingly accurate numerical predictions offered by contemporary quantum chemistry, one should never forget that obtaining accurate numbers is not synonymous with understanding the chemical problem. Hence, it is a vital and sometimes somewhat neglected part of our profession to turn the results of quantum chemical calculations into chemical context and chemical language. This is the purpose of this monograph. Quantum chemistry and Chemical Graph Theory are two well-developed but disparate branches of theoretical chemistry. This monograph is an enlightening first step for their union taking Richard F. W. Bader's Quantum Theory of Atoms in Molecules (QTAIM) as the bridge between them. The authors provide a concise summary of their contributions to this field over the past decade and have included—especially in the later chapters—several new methods and unpublished results of considerable interest. The examples used in the exposition cover a large range of timely problems with practical solutions where the results of electronic structure calculations are summarized into the electron Localization and Delocalization Matrices (LDMs) which are then analyzed and compared by a whole battery of methods. The predictive power of such LDM analysis is promising. I commend the

authors, Chérif F. Matta, Paul W. Ayers, and Ronald Cook, all three of whom are prominent scientists in their respective lines of research, for pooling their efforts to bring to the light this important monograph.

Prof. Dr. Frank Neese

Director

Department for Molecular Theory and Spectroscopy

Max-Planck-Institut für Kohlenforschung

(Max Planck Institute for Coal Research)

Mülheim an der Ruhr, Germany

Preface

“The molecular descriptor is the final result of a logic and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment [Todeschini and Consonni, 2000].

Attention is paid to the term “useful” with its double meaning: it means that the number can give more insight into the interpretation of the molecular properties and/or is able to take part in a model for the prediction of some interesting property of other molecules.

Why must we also accept “or”?

It should not be thought that molecular descriptors are good only if they show an evident link to some information about molecular structure, that is, they are easily interpretable from a structural/ chemical point of view.

It often happens that interpretation of molecular descriptors could be weak, provisional, or completely lacking, but their predictive ability or usefulness in application to actual problems should be a strong motive for their use. On the other hand, descriptors with poor predictive ability may be usefully retained in models when they are theoretically well founded and interpretable due to their ability to encode structural chemical information.”

Roberto Todeschini and Viviana Consonni, *Molecular Descriptors for Chemoinformatics*, Wiley-VCH, Weinheim (2009)

This monograph introduces a new method that converts computational chemistry results into a set of tools to predict the activities and/or properties of a series of compounds. It uses Richard F. W. Bader’s Quantum Theory of Atoms in Molecules (QTAIM) starting from a topographical and topological analysis of the molecular electron density to provide a bookkeeping of electrons localized in, and delocalized between, atoms in a molecule, a complex, or a crystal. The method outlined in this monograph extracts molecular descriptors from the QTAIM numerical integration results which are then used to solve concrete experimental chemistry problems.

A principal theme in this book is the use of “molecular descriptors” in “Quantitative Structure-Activity/Property Relationship (QSAR/QSPR)” modeling. The method introduced here uses computational quantum chemistry to obtain molecular electronic wavefunctions and electron densities which are then analyzed to obtain the QTAIM localization and delocalization indices (LIs and DIs) which, in turn,

are organized to construct the corresponding electron Localization-Delocalization Matrices (LDMs). LDMs are then either compared directly or manipulated to extract matrix invariants as is done in Chemical Graph Theory (CGT) or using statistical methods such as Principal Component Analysis (PCA) to predict molecular activities/properties. The approach outlined in this book can be considered as a bridge between two branches of theoretical chemistry, that is, Chemical Graph Theory (CGT) on one hand and quantum chemistry, in particular QTAIM, on the other. We will term the overall approach presented in this work as *LDM Analysis*.

The *applications* of LDM Analysis are not bound to any specific field. As will be apparent, these applications span physical organic chemistry, corrosion inhibitors research, insect control, enzymology, physicochemical properties, and quantitative molecular comparisons, to name just a few examples. It is hoped that this approach will be of value to those interested in research in such domains as materials science, physical organic chemistry, and drug design, where *in silico* predictions can guide and/or shed light on experiments.

The level of exposition should be accessible to upper-level undergraduate students, graduate students, and researchers in physical chemistry, biophysics, molecular modeling, bioinformatics, and related fields of specialization.

In closing, this book is a *monograph* in the sense that it is a topical review that focuses primarily on the contributions and interests of the authors. It is not at all meant to be a comprehensive all-encompassing review of the literature on molecular descriptors, QSAR/QSPR, or computational quantum chemistry. The exposition is far from complete or comprehensive and can be taken as a “teaser” since the field is young and wide open.

Halifax, NS, Canada
Hamilton, ON, Canada
Denver, CO, USA

Chérif F. Matta
Paul W. Ayers
Ronald Cook

Acknowledgments

R. C. thanks his spouse Diane Cook for her unwavering support. P. W. A. is grateful for his parents and sister. C. F. M. is much indebted to his parents (Farid A. Matta and Nabila N. Abdel-Nour) and to Maged F. Matta, Hebatallah S. Habib, Sara and Nadine Matta, and Sami and Boa Habib.

The authors are indebted to several mentors, colleagues, and students among whom we particularly thank Richard F. W. Bader, Robert G. Parr, Lou Massa, John C. Polanyi, Frank Neese, Claude Lecomte, Russell J. Boyd, Ismat Sumar, Todd A. Keith, André D. Bandrauk, Sason Shaik, Sonja Nikolić, Weitao Yang, Tina A. Harriott, Daniel Majaess, Rafik A. Matta, Ashraf W. El-Miniawy, Federico Rosei, Nagwa El-Badry, Mahmoud Abdel-Aty, Hugo Bohórquez, Eva Knoll, Ángel Martín-Pendás, Labiba K. El-Khordagui, Abdel-Aziz Saleh, Ignacy Cukrowski, Jesús Hernández-Trujillo, Fernando Cortés-Guzmán, Aurora Costales Castro, Olimpia Lombardi, Eric Scerri, James S. M. Anderson, Lázaro A. M. Castanedo, Peyman Fahimi, Ihab Gad, Essam F. Khamis, Islam K. Matar, Halis Seuret, Samir E.-G. Gabrial, Ashraf S. Abousaif, Sherif F. Boutrous (Borgzanno), and Hani B. Maher.

The authors acknowledge with gratitude the Springer's team, its editorial and production offices, and the handling editors for supporting this book.

The completion of this monograph would not have been possible if it wasn't for the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Foundation for Innovation (CFI), Compute Canada, Mount Saint Vincent University, and Saint Mary's University.

Contents

1	An Introduction to Electron Localization-Delocalization	
	Matrices	1
	References	4
2	The Physics of Electron Localization and Delocalization	7
2.1	What do we mean by Electron (De)Localization?	7
2.2	Pauli's Exclusion Principle and the Formation of Localized Electron Pairs	13
2.3	Conclusions	20
	References	20
3	The Quantum Theory of Atoms in Molecules and Electron Localization and Delocalization	21
3.1	The Electron Density: Introductory Concepts	21
3.2	From the Topography of the Electron Density to the Topology of a Bonded Structure	25
3.3	The Zero-Flux Surface as a Condition for a Well-Defined Quantum Subsystem	29
3.4	Examples of Properties of Atoms in Molecules	32
	3.4.1 Examples of Bond Properties	33
	3.4.2 Examples of Atomic Properties	37
3.5	Localization of Electrons within and their Delocalization between Atoms in a Molecule	39
3.6	The Delocalization Indices and Chemical Bonding by Example	44
	3.6.1 Bonding to Titanium	44
	3.6.2 Delocalization Indices and NMR Spin-Spin Coupling Constants	46
3.7	The Electron Localization-Delocalization Matrix (LDM)	51
3.8	The Interacting Quantum Atoms (IQA) Matrix (IQAM)	52
3.9	Closing Remarks	53
	References	54

4	Localization-Delocalization Matrix Representation of Molecules	57
4.1	Networks, Graphs, and Matrices: General Considerations	57
4.2	The Molecule as an Incomplete Network of Bond Paths	60
4.3	The Molecule as Complete Network of Exchange Channels	61
4.4	Dissimilarity of Two Molecules as the Distance between their LDMs	63
4.5	Problems and Solutions of the LDM Representation of a Molecule	68
4.5.1	The Non-Uniqueness of an LDM in Specifying a Chemical Structure	68
4.5.2	Unequally-Sized Molecules or Molecules not Sharing a “Recognizable Common” Skeleton	70
4.5.3	Insensitivity to Optical Isomerism	79
4.5.4	Conformational Flexibility	79
4.6	Closing Remarks	80
	References	80
5	Molecular Fingerprinting using Localization-Delocalization Matrices: Computational Aspects	83
5.1	Computational Implementation of LDM Analysis	83
5.2	The Triangle Inequality	85
5.2.1	Test Case 1: The First Four Saturated Aliphatic Hydrocarbons	87
5.2.2	Test Case 2: The Iso-Electronic Series with N = 50 Electrons	88
5.2.3	Test Case 3: Halogenated Benzoic Acids	90
5.3	Basis Set Effects	91
5.4	Closing Remarks	94
	References	94
6	Principal Component Analysis of Localization-Delocalization Matrices	97
6.1	The Localization-Delocalization Matrix as Tabulated Data	97
6.2	The Statistical Concepts of Covariance and Correlation	98
6.3	Some Basics of Principal Component Analysis (PCA)	103
6.3.1	A Numerical Example: H ₂ O	108
6.3.2	A More Elaborate Numerical Example: Imidazole	109
6.4	Factor Analysis (FA) Contrasted with Principal Component Analysis (PCA)	112
6.5	General Protocol for Applying PCA and FA to LDMs	114
6.6	Closing Remarks	115
	Appendix	116
	References	118

7	Localization-Delocalization Matrices Analysis for Corrosion Inhibition	121
7.1	Corrosion Inhibition Efficiency	121
7.2	The Exorbitant Economic Costs of Corrosion	122
7.3	First Case Study: Quantum Chemical Prediction of Corrosion Inhibitors' Efficiency	123
7.3.1	LDM Prediction of Corrosion Inhibitors' Efficiency Outperforms a Traditional Approach	123
7.3.2	Methodology of the LDMs Modeling of Corrosion Inhibitors' Efficiency	128
7.3.3	LDMs' Prediction of Corrosion Inhibitors' Efficiency	128
7.4	Second Case Study: Predicting Stress Corrosion Cracking Inhibitors' Efficiency	132
7.4.1	Stress Corrosion Cracking (SCC) Costs Lives and Treasury	132
7.4.2	Structure–Activity Relationship of Stress Corrosion Cracking Inhibitors	133
7.4.3	Methodology of the LDMs Modeling of Stress Corrosion Cracking Inhibitors' Efficiency	133
7.4.4	QSAR and Mechanistic Insight into the Active Molecular Species of SCC Inhibitors	135
7.5	Third Case Study: PCs of the LDMs for the Modeling of General Corrosion Inhibition	138
7.5.1	Introductory Remarks	138
7.5.2	Heterocyclic Diazoles Corrosion Inhibitors for Iron in a Strong Acidic Medium	139
7.5.3	Heterocyclic Diazoles Corrosion Inhibitors for Aluminum Alloys in a Strong Acidic Medium	140
7.5.4	Classifying Corrosion Inhibitors by a Latent Class Analysis of the PCs of the LDMs	142
7.6	Closing Remarks	143
	References	147
8	Localization-Delocalization Matrices Analysis in Predicting Mosquito Repellency	151
8.1	Insect Olfaction and the Mitigation of Mosquito-Transmitted Diseases	151
8.2	Steps for an LDMs Predictive Modeling of Mosquito Repellency	153
8.3	Modeling of Semiochemical Toxicants for Mosquitoes	154
8.4	Modeling of Semiochemical Toxicity Against <i>Aedes Aegypti</i>	156
8.5	Modeling of Odorant Binding by Odorant Binding Proteins (OBPs)	160

8.6	Modeling Mosquito Repellent Response of the Odorant Receptor of <i>Culex quinquefasciatus</i>	166
8.7	Closing Remarks	170
	Appendix	170
	References	177
9	Modeling Enzyme-Substrate Interaction with Localization-Delocalization Matrices	179
9.1	Enzyme Catalysis: Some Generalities	179
9.2	Method	181
9.3	Modeling Carbamates Inhibitors of Acetylcholinesterase as Alzheimer Disease Slowing Agents	182
9.4	Organophosphorus Hydrolysis by Phosphotriesterase	184
9.5	Closing Remarks	187
	References	187
10	Localization-Delocalization Matrices of Large Systems	191
10.1	The Scaling Problem of Quantum Chemical Calculations	191
10.2	The Kernel Energy Method (KEM)	193
10.3	Fragmentation Without Capping	197
10.4	Closing Remarks	204
	References	208
11	Closing Remarks	213
	Index	217

About the Authors



Chérif F. Matta, B.Pharm.Sci. (Alexandria University), Dipl. (Sadat Academy), PhD (McMaster University), HDR (Université de Lorraine), FRSA, FRSB, FInstP, FRSC, FAAS, FAAAS is Professor and Chair/Head of the Department of Chemistry & Physics at Mount Saint Vincent University and an Adjunct (Honorary) Professor at Dalhousie, Laval, and Saint Mary's Universities (all in Canada), and holds frequent Visiting Professorships at Zewail City for Science and Technology (Egypt). After receiving his doctorate from McMaster University under Professor Richard F. W. Bader, Matta has been a Postdoctoral Fellow at the University of Toronto (with Nobel Laureate Professor John C. Polanyi) and later an Izaak W. Killam Postdoctoral Fellow at Dalhousie University (with Professor Russell J. Boyd). Chérif Matta's research ranges from theoretical and computational chemistry to mitochondrial biophysics. He has introduced and developed with the co-authors of this book the concept of the electron Localization-Delocalization Matrices (LDMs) as predictive descriptors in materials and bio-pharmacological research. He discovered an unusual role of the enzyme ATP synthase over and above catalysis as a modulator of the reaction energy itself and not only of the barrier to the reaction. He is also the co-discoverer of the Hydrogen-Hydrogen bonding interaction. Matta has given more than 250 lectures in over 30 countries and has *ca.* 200 publications to his credit including four books. Matta has been named "Distinguished Visiting Professor" by the Polish Ministry of Science and Higher Education 2022–2023, was the

2021–2022 “Science Atlantic Speaker of the Year”, and has been awarded Mount Saint Vincent University’s Award for Research Excellence (2017), the Lady Davis Fellowship (2016–2017), the NSERC-Acfas “Preuve par l’image” Prize for 2016, the Molecular Graphics and Molecular Simulation Society Silver Jubilee Prize for 2009, the John C. Polanyi Prize in Chemistry for 2004, and the Izaak Walton Killam Fellowship in 2002–2004. Prof. Matta is an elected Fellow of the American Association for the Advancement of Science, of the African Academy of Sciences, of the Royal Society of Chemistry, of the Institute of Physics, of the Royal Society of Biology, and of the Royal Society of the Arts. He has served (2021–2022) as Chair/Head of the Interdisciplinary Adjudication Committee of the Canada Research Chairs Programme of the Tri-Council of Canada and is currently serving as Director of Accreditation & Member of the Board of Directors of the Canadian Society for Chemistry. In 2022–2023 he has served as a Member of the Multidisciplinary Selection Board of Canada Excellence Research Chairs of the Tri-Council and is currently serving a second term as an elected Member of the Commission for Quantum Crystallography of the International Union of Crystallography (IUCr) and as a Member of the Canadian National Committee for Crystallography (CNCC).



Paul W. Ayers is the Canada Research Chair in Theoretical Chemistry at McMaster University. Paul’s PhD was supervised by the late Robert G. Parr at the University of North Carolina. After postdoctoral training with Weitao Yang at Duke University, Paul became a (Tier 2) Canada Research Chair at McMaster University. Paul’s research focusses on developing new qualitative and quantitative methods for describing molecular electronic structure and reactivity, focussing on the theory and applications of density (matrix) functional theory. Paul has co-authored more than 350 peer-reviewed publications, with an h-index of 75 and more than 20,000 cumulative citations, has won various awards (including the Dirac Medal, the Rutherford Medal, the Steacie Prize, and the IAQMS medal) and is a Fellow of the Royal Society of Chemistry (U.K.) and a Member of the College of Young Scholars of the Royal Society of Canada. Paul co-founded the QC-Devs international software consortium

for free and open-source software in the computational sciences.



Ronald Cook, PhD was a Principal Scientist at TDA Research, Inc., from 1982 to 2021 and carried and continues to carry out research in computational materials science and materials discovery. He has led numerous modeling efforts, building on previous successful efforts to identify improved corrosion-inhibiting formulations through modeling. This work has included the development of quantum mechanical modeling using the quantum theory of atoms in molecules (QTAIM) computed parameters. Dr. Cook has, in cooperation with Prof. Chérif F. Matta, developed new types of universal descriptors that are generated by computing the Principal Components of, or the GINI indices of the Localization-Delocalization Matrices that are accessible from QTAIM calculations. Dr. Cook has a high level of expertise in converting the descriptors to descriptive/predictive equations using Symbolic Regression (ChaosHunter™), Latent Class Regression, Correlated Component Regression, Self-Organizing Maps, Neural Networks, and Genetic Algorithms. The universal descriptors that are generated by computing the Principal Components of Localization-Delocalization Matrices have been applied to successfully predict corrosion inhibition rates for small molecules that were not predictable through standard Density Functional Theory descriptors. Dr. Cook has extended the use of the universal descriptors described above to successfully predict the inhibition of acetylcholinesterase, the binding of vapor molecules to odorant binding proteins, the prediction of the effectiveness of mosquito repellents and the binding of antifilarial antimycin analogues. Dr. Cook has retired from TDA Research, Inc., and is currently pursuing new developments in QTAIM modeling. Dr. Cook has authored more than 44 peer-reviewed papers and 15 patents.

Chapter 1

An Introduction to Electron Localization-Delocalization Matrices



Give us insight not numbers [1] (see also [2]).

Charles A. Coulson (1910–1974)—(1959)

Give us insight and numbers [1].

Frank Neese, Mihail Atanasov, Giovanni Bistoni, Dimitrios Maganas,
and Shengfa Ye—(2019)

This book has been written with the practicing chemist in sight. It is *not* a book on theoretical chemistry but rather one that is primarily focused on making practical predictions starting from theoretical chemistry calculations. It is true that predicting and explaining is not “the same thing”, as argued along the length of a book that René Thom¹ titles: “*Prédire n’est pas expliquer*” (Predicting is Not Explaining) [3]. However, as it is hoped it will become apparent in the following pages that, in making numerical predictions, one can, in fact, gain insight. In other words, by exploiting the panoply of existing tools from quantum chemistry, to graph theory, and ending up with modern statistics one can, in many cases, go beyond Coulson’s wish and indeed get “*insight and numbers*” out of electronic structure calculations as the modern revision of this wish formulated six decades later by Neese, Atanasov, Bistoni, Maganas, and Ye in 2019 [1].

What will be explored in detail in this book is a new method to extract, summarize, and use the information obtained from quantum chemical (electronic structure) calculations for a series of molecules to uncover trends in sought-for properties for the purpose of making useful predictions. Sometimes such predictions can also give us mechanistic insights into the chemistry of the problem. The core of the method

¹ René Thom (1923–2002) was an eminent French mathematician and 1958 Fields Medalist credited for several noted contributions to topology including being the principal developer of Catastrophe Theory (along with British Mathematician Prof. Sir E. Christopher Zeeman (1925–2016)) who widely popularized the theory).

that will be expanded on the pages of this book is the abstract linear algebraic representation of a molecular electronic structure in terms of an electron Localization-Delocalization Matrix, or LDM—which is the subject of analysis as a powerful molecular descriptor.

In performing such an LDM analysis, one starts from electronic structure calculations, typically of the Density Functional Theory type [4–6] (with a moderate-to-relatively-large basis set as much as the compromise of speed and accuracy allows). The resulting electronic structure is then summarized at an atomic/atomic pair level of “coarseness” [7] using Bader’s Quantum Theory of Atoms in Molecules (QTAIM) [8, 9] by extracting the so-called localization and delocalization indices (LIs and DIs, respectively) [10]. These indices describe the electronic structure in terms of the number of electrons that are, on average, localized within atomic basins and those that are shared (delocalized) between every atomic pair in the molecule whether chemically bonded or not. The sharing is naturally more significant when there exists a bonding interaction linking the two atoms but it is not zero otherwise. One can describe the bond path (lines of maximal electron density that link bonded atoms in real space) as “privileged exchange channels” [11] since electron sharing is mediated by the exchange symmetry of the wavefunction.

These LIs and DIs can clearly be arranged in matrix format not dissimilar to the *charge and bond order matrix*, an archaic name to what is now more commonly known as the *density matrix*, listing the bond orders as its off-diagonal elements leaving the diagonal ones for atomic populations [12–15]. The arrangement of the LIs and DIs of a given molecule in matrix format yields what we term the electron Localization-Delocalization Matrix (LDM). LDMs are rich in coded chemical information since they capture and summarize elegantly the electron distribution reflecting, at once, aspects of both the molecular electron density (e.g., atomic electron populations, atomic charges, atomic contributions to higher molecular multipoles) and of the electron *pair* density distributions (e.g., the delocalization/sharing of electrons between atoms through the mechanism of exchange) at an atomic-diatomic resolution.

Chemical Graph Theory has a long tradition of summarizing molecular topology into topological (connectivity) matrices [16–29]. Meanwhile, any matrix representing a set of connected objects—is numbering/labeling dependent since there exist $n!$ way to arrange n distinguishable objects (except when symmetry reduces this number). Chemical graph theorists extract matrix invariants from these matrices as a means to uniquely describe molecules [16–29]. Instead of re-inventing the wheel, we follow the chemical graph theorists’ footsteps to extract invariants from the LDMs, albeit by introducing new ways to do so such as the Frobenius distance between matrices or using statistical methods such as Principal Component Analysis.

Balasubramanian, in a far-sighted article [23], proposed three decades ago the “[i]ntegration of graph theory and quantum chemistry” and underscored the correspondence of graph-theoretical graphs whether within Chemical Graph Theory or as the complete set bond paths defined within QTAIM [8, 30]. This book is an attempt to fulfill Balasubramanian’s vision by extending and expanding the now decade-old

proposal to achieve such a merger [31]. The antecedents of this book have appeared in the following key references: [31–39].

The book starts with a review of the physical description of electron localization and delocalization in molecular quantum mechanics (Chapter 2). The exposition then moves to review, in Chapter 3, the basic tenets of Bader’s QTAIM, especially with regard to how it defines LIs and DIs in terms of the six-dimensional integrals of the Fermi hole density. The chapter emphasizes the strong connection between these 2-electron mathematical indices and several experimentally measured quantities. This connection reinforces the idea that these indices condense much chemical information. Chapter 4 defines the LDMs and related matrices and explores methods to define molecular similarity or proximity in this context. The chapter then discusses briefly some of the challenges to this kind of representation of molecular structure and proposes some possible solutions. The chapter shows how solutions to these challenges can lead to the modeling of molecular behavior in terms of the LDMs by predicting properties such as boiling points, pK_a ’s of weak acids, and λ_{\max} of substituted benzoic acids. Chapter 5, discusses the approach to molecular fingerprinting using LDMs and several technical and programming details to automate the implementation of this type of analysis by introducing the programme AIMDELOC with examples of application. In anticipation of the following application-centered chapters, Chapter 6 reviews the basics of Principal Component Analysis (PCA) since this analysis is applied in subsequent chapters to extract invariant descriptors from the LDMs. A general protocol for how to apply PCA to LDMs in property modeling is introduced in this chapter. The following chapters (Chapters 7–9) are all about the use of PCA to extract invariant molecular descriptors from the LDMs. Chapter 7 outlines a PCA of the LDMs of corrosion inhibitors. The analysis results in robust predictive modeling of corrosion inhibitors’ efficiency and, further, provides valuable insight into the *mechanism* of some of the studied corrosion inhibitors pinpointing what may have been overlooked in the literature as the *active species* of (at least) three stress corrosion cracking inhibitors containing sulfhydryl groups (R–SH). Chapter 8 applies this analysis to mosquito repellency showing how it can be used to *classify* odorants and in shedding light into the modes of action of the repellents. Chapter 9 demonstrates the usefulness of the LDM-PC analysis in predicting enzyme–substrate interaction, for example, the action of acetylcholine esterase inhibitors as Alzheimer’s Disease retarding drugs. LDM-PC analysis is also shown of value in the predictive modeling of the bioremediation approach to clean organophosphorous environmental toxicants and chemical warfare agents. Chapter 10 explores the extension of LDM Analysis to very large molecules *via* the two fragmentation approaches aiming at increasing the speed of the calculation by recombining the LDMs of small fragments to approximate the LDM of the large molecule. Chapter 11 concludes this book by summarizing the main points to be retained and possible future developments.

In closing, it is stressed that the topic of this book is *a work in progress*. The field is still young but certainly appears promising in providing “insight *and* numbers”. The LDMs, in themselves, are elegant mathematical summaries of complex molecular electronic structures.

References

1. Neese F, Atanasov M, Bistoni G, Maganas D, Ye S (2019) Chemistry and quantum mechanics in 2019: Give us insight and numbers. *J Am Chem Soc* 141:2814–2824
2. Coulson CA (1960) Present state of molecular structure calculations. *Rev Mod Phys* 32:170–177
3. Thom R (1993) *Prédire n'est pas expliquer: Entretiens avec Emile Noël*. Champs Science, Flammarion, Paris, France
4. Parr RG, Yang W (1989) *Density-functional theory of atoms and molecules*. Oxford University Press, Oxford
5. Koch W, Holthausen MC (2001) *A chemist's guide to density functional theory*, 2nd edn. Wiley-VCH, New York
6. Scholl DS, Steckel JA (2009) *Density functional theory: a practical introduction*. Wiley, Hoboken, NJ
7. Matta CF, Lombardi O, Jaimes Arriaga J (2020) Two-step emergence: the quantum theory of atoms in molecules as a bridge between quantum mechanics and molecular chemistry. *Found Chem* 22:107–129
8. Bader RFW (1990) *Atoms in molecules: a quantum theory*. Oxford University Press, Oxford, UK
9. Matta CF, Boyd RJ (Eds.) (2007) *The quantum theory of atoms in molecules: from solid state to DNA and drug design*. Wiley-VCH, Weinheim
10. Fradera X, Austen MA, Bader RFW (1999) The Lewis model and beyond. *J Phys Chem A* 103:304–331
11. Martín Pendás Á, Francisco E, Blanco MA, Gatti C (2007) Bond paths as privileged exchange channels. *Chem Eur J* 13:9362–9371
12. Coulson CA, Longuet-Higgins HC (1947) The electronic structure of conjugated systems. I. General theory. *Proc Roy Soc (London) A* 191:39–60
13. Lennard-Jones JE (1952) The spatial correlation of electrons in molecules. *J Chem Phys* 20:1024–1029
14. Löwdin P-O (1955) Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Phys Rev* 97:1474–1489
15. Chirgwin BH (1957) Summation convention and the density matrix in quantum theory. *Phys Rev* 107:1013–1025
16. Balaban AT (1976) *Chemical applications of graph theory*. Academic Press, New York
17. Hall LH, Kier LB (1976) *Molecular connectivity in chemistry and drug research*. Academic Press, Boston
18. King RB, Rouvray DH (1977) *Chemical applications of group-theory and topology*. 7. Graph-theoretical interpretation of bonding topology in polyhedral boranes, carboranes, and metal clusters. *J Am Chem Soc* 99:7834–7840
19. Bonchev D, Trinajstić N (1977) Information theory, distance matrix, and molecular branching. *J Chem Phys* 67:4517–4533
20. Dmitriev IS (1981) *Molecules without chemical bonds (English Translation)*. Mir Publishers, Moscow
21. Randić M (1990) *Design of molecules with desired properties: a molecular similarity approach to property optimization. Concepts and applications of molecular similarity*, Wiley, New York
22. Bonchev D, Rouvray DH (1991) *Chemical graph theory: Introduction and fundamentals*. OPA, Amsterdam
23. Balasubramanian K (1994) Integration of graph theory and quantum chemistry for structure-activity relationship. *SAR & QSAR Environ Res* 2:59–77
24. Diudea MD, Gutman I, Lorentz J (1999) *Molecular topology*. Nova Science Publishers Inc., Hauppauge, NY
25. Bonchev D, Rouvray DH (2003) *Complexity in chemistry: introductions and fundamentals*. Taylor and Francis, London

26. Janežić D, Milicević A, Nikolić S, Trinajstić N (2007) Graph theoretical matrices in chemistry (Mathematical Chemistry Monographs, vol 3). MCM and the University of Kragujevac, Kragujevac
27. Mandado M, González-Moa MJ, Mosquera RA (2007) Chemical graph theory and n -center electron delocalization indices: a study on polycyclic aromatic hydrocarbons. *J Comput Chem* 28:1625–1633
28. Todeschini R, Consonni V (2009) Molecular descriptors for chemoinformatics, vol I and II. 2nd edn. Wiley-VCH Weinheim, Weinheim
29. García-Revilla M, Francisco E, Popelier PLA, Martín Pendás Á (2013) Domain-averaged exchange-correlation energies as a physical underpinning for chemical graphs. *Chem Phys Chem* 14:1211–1218
30. Runtz GR, Bader RFW, Messer RR (1977) Definition of bond paths and bond directions in terms of the molecular charge distribution. *Can J Chem* 55:3040–3045
31. Matta CF (2014) Modeling biophysical and biological properties from the characteristics of the molecular electron density, electron localization and delocalization matrices, and the electrostatic potential. *J Comput Chem* 35:1165–1198
32. Sumar I, Ayers PW, Matta CF (2014) Electron localization and delocalization matrices in the prediction of pK_a 's and UV-wavelengths of maximum absorbance of p -benzoic acids and the definition of super-atoms in molecules. *Chem Phys Lett* 612:190–197
33. Sumar I, Cook R, Ayers PW, Matta CF (2015) AIMLDM: a program to generate and analyze electron localization-delocalization matrices (LDMs). *Comput Theor Chem* 1070:55–67
34. Matta CF, Sumar I, Cook R, Ayers PW (2015) Localization-delocalization and electron density-weighted connectivity matrices: a bridge between the quantum theory of atoms in molecules and chemical graph theory. In: Chauvin R, Silvi B, Alikhani E, Lepetit C (eds) Applications of topological methods in molecular chemistry. Springer, pp 53–88
35. Sumar I (2016) Chemical applications of electron localization-delocalization matrices (LDMs) with an emphasis on predicting molecular properties. M.Sc. Thesis, Saint Mary's University, Halifax
36. Sumar I, Cook R, Ayers PW, Matta CF (2016) Aromaticity of rings-in-molecules (RIMs) from electron localization-delocalization matrices (LDMs). *Phys Script* 91:013001
37. Cook R (2017) Principal components of localization-delocalization matrices: new descriptors for modeling biological activities of organic compounds. Part I: Mosquito insecticides and repellents. *Struct Chem* 28:1527–1538
38. Matta CF (2018) Molecules as networks: a localization-delocalization matrices approach. *Theor Comput Chem* 1124:1–14
39. Matta CF (2023) Electron localization-delocalization matrices (LDMs). In: Boyd RJ, Yáñez (eds) Comprehensive computational chemistry Elsevier, The Netherlands (Volume 2, pp. 82–94)