

WAVEFUNCTIONS TO CHEMICAL REACTIONS

**FROM WAVEFUNCTIONS TO CHEMICAL REACTIONS:
NEW MATHEMATICAL TOOLS FOR PREDICTING THE REACTIVITY
OF ATOMIC SITES FROM QUANTUM MECHANICS**

By

**JAMES S. M. ANDERSON,
BSCH (Biomolecular Chemistry), BSC (Mathematics)**

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

© Copyright by James Stewart Murray Anderson, November 2010

DOCTOR OF PHILOSOPHY (2010)
(Chemistry)

McMaster University
Hamilton, Ontario, Canada

TITLE: From wavefunctions to chemical reactions: new
mathematical tools for predicting the reactivity of atomic
sites from quantum mechanics.

AUTHOR: James S. M. Anderson BSC –Mathematics
BSCH –Biomolecular Chemistry
(Queen’s University)

SUPERVISOR: Professor Paul W. Ayers

NUMBER OF PAGES: xviii, 271

ABSTRACT

Solving the electronic Schrödinger equation for the molecular wavefunction is the central problem in theoretical chemistry. From these wavefunctions (possibly with relativistic corrections), one may completely characterise the chemical reactivity and physical properties of atoms, molecules, and materials. Unfortunately, there are very few systematic approaches for obtaining highly-accurate molecular wavefunctions. The approaches that do exist suffer from the so-called curse of dimensionality: their computational cost grows exponentially as the number of particles increases. Furthermore, even after obtaining an accurate wavefunction, partitioning the molecule into atoms is not straightforward. This is because the kinetic energy operator is a differential operator in spatial coordinates. This is a source of ambiguity in the definition of an atom-in-a-molecule and the associated atomic properties. Even after selecting an appropriate definition of an atom and obtaining the atoms from the wavefunction, the atom's intrinsic reactivity cannot be completely characterised without considering every possible reaction partner. This is because each set of two molecules produces a new wavefunction that is more complicated than the products of the wavefunctions of the separate molecules.

This thesis presents methods for addressing the three challenges raised in the previous paragraph: computing atomic properties (e.g. chemical reactivity), partitioning molecules into atoms, and computing accurate molecular wavefunctions. The first challenge is addressed by developing a general-purpose reactivity indicator to quantify the reactivity of an atom within a molecule. This indicator quantifies the reactivity of any point of the molecule using only the electrostatic potential and Fukui potential at that point. The key idea is to include only a vague description of an incoming molecule and compute an approximate interaction with the incoming object; this ensures that the general-purpose reactivity indicator is simple enough to be useful. Practically, this indicator is most useful when it is used to compute the reactivity of the atomic sites in the molecule of interest.

Partitioning a molecule into atoms is not straightforward because of the inherent nonlocality of quantum mechanics. In the context of molecular electronic structure, this nonlocality arises from the nature of the kinetic energy operator. The quantum theory of atoms in molecules (QTAIM) is a popular method that partitions molecules into atoms. QTAIM resolves the problem of ambiguity for all permissible forms of the kinetic energy operator. In this thesis the characterisation of an atom provided by QTAIM is extended to include relativistic contributions in the zero-order regular approximation (ZORA). The intrinsic ambiguity arising from the kinetic energy operator is also examined in detail.

Computing atomic or molecular properties (including computing the general-purpose reactivity indicator) almost always requires a wavefunction. For this reason, obtaining accurate wavefunctions is the central hurdle of quantum chemistry. This thesis proposes algorithms for finding high-accuracy molecular wavefunctions without exponentially exploding computational cost. To do this, tools for exploiting the smoothness of electronic wavefunctions are crafted. Computational methods that use these tools can break the curse of exponential scaling without sacrificing accuracy. Specifically, the computation cost of these new methods grows only as some polynomial of the electron number. The wavefunctions obtained from these methods are much simpler than those from conventional approaches of similar accuracy, and are therefore ideal for computing the electron density and atomic properties.

To my parents, Stewart and Margaret Anderson,
my brothers Teddy, Thomas, Christopher, and Geoffrey Anderson,
and my Grandmother Honore Edmunds

For all of their love, support, and wisdom
May God Bless Them

ACKNOWLEDGEMENTS

From my time as a graduate student in the chemistry department at McMaster University there are many people to whom I owe a great deal of gratitude. Here, I humbly thank each of them and recognise how each of them have moulded and influenced my life and kept me going.

First, I would like to thank McMaster University and its chemistry department for giving me the opportunity to pursue a degree. Particularly I would like to thank Doug Welch, Carol Dada and Lynda Fraser. Doug Welch showed me how important it is to find alternatives and to give the benefit of the doubt. You have motivated me more than you will know. Carol Dada and Lynda Fraser both have always made themselves available whenever I needed anything, and always came through. Especially Carol, whom went out of her way to grant a very special request that I made.

I would like to thank Donald Sprung and Randy Dumont whom served on my Ph.D. committee. Particularly I would like to thank Randy Dumont for all of our discussions. I have learned much from you.

I would like to thank Richard Bader and Vedene Smith. Richard, your passion and enthusiasm for research and life are admirable. I hope that one day I can embody the same passion and excitement. Vedene, for your kindness and enthusiasm that you showed me. I miss our discussions and may you rest in peace.

I would like to thank the Ayers' group. I would like to thank the postdocs Bijoy Dey, David Thompson, Utpal Sarkar, Steven Burger, Carlos Cardenas, Lourdes Romero, and Alfredo Guevara. I would like to thank the graduate students Juan Rodriguez, Annie Liu, Ivan Vinogradov, Debajit Chakraborty, Rogelio Cuevas, Pavel Kulikov, Ahmed Kamel, Sandra Rabi, Paul Johnson, and Farnaz Heidarzadeh. Particularly, I would like to thank David Thompson, Juan Rodriguez, Annie Liu, and Steven Burger.

David, for always being there to help with proofreading, always being positive, and having a way of just "cutting through the crap" and saying not to worry that I'll "be fine." You helped me work through some of those early obstacles directly while you were here, and when you left you kept in touch and always helped.

Juan, thank you for showing me the true meaning of "tranquilo" a word that I did not even know existed before I met you. You showed me that it is possible for one can relax no matter how high pressure the situation is. You have helped me in some ways that you will never even realise.

Annie, you are my theoretical fraternal twin sister. We experienced all the same ups and downs, particularly during our first year of graduate school. You have a way of dealing with adverse situations head on that I admire. Whenever there was something you wanted done it seemed to me you didn't complain and just went on and did it. I wish you the best of luck finishing your studies.

Steven Burger, you have really showed me what it means to be a Protestant Christian. Though I do not always agree with this perspective I admire your dedication to it. I appreciate all of your help during the time you have been here.

To all my friends, thank you for all of your kindness and support. Too numerous to mention all of you, but each of you mean the world to me. To my best friend John Santiago, thank you for being there at every moment I needed you. We have been through much together, all the ups and downs of life, and more to come. I look forward to the next adventures.

I would like to thank my family whom I love dearly. Even when I seem distant, you are always in my heart.

Daisy I miss your unconditional love and loyalty. May you rest in peace.

I would like to thank my Grandma. I wish you were alive to see this achievement. Thank you for your encouragement and may you rest in peace.

I would like to thank my parents Margaret and Stewart for their unending unconditional love and support. I know I can always count on you to help me or give me my space when I need it. I appreciate every moment of it and am lucky to have parents like you.

I would like to thank my brothers. We are very similar and very different at the same time. One thing I know for sure is that all of you will be successes and I am proud to have you in my family.

Teddy, I want to thank you for how you kept the family together during one of our weaker times. I know your charm and decisiveness to succeed will serve you well and I do have no doubt that you will be successful. I admire your unwavering disposition that things will happen the way you want them to.

Thomas, I admire the great bravery you showed when you were sick. You are back to full health and have taken control of your life to a level that I admire. You are definitely the most organised of us all.

Christopher, I admire how you do not worry too much and forgive easily. Your ease at being the “popular guy” and your ability take control of the crowd when you choose is something I admire.

Geoffrey, I admire your raw intelligence. You impress me with your ability to instantly multiply large integers in your head or solve linear equations, despite not even being interested in formal math. I admire your calmness and your refusal to rush, something I can learn from.

Grandmother, on your centennial, I thank you for the wisdom you have provided me and the care you have given. I thank you for the summers that I lived with you. I have enjoyed learning about your life and look forward to the years to come.

Lastly, I would like to thank my supervisor Paul Ayers. From the beginning you have always been there to advocate for me. From research questions, advice on applications, paper writing, and even influence on this thesis, I am greatly in your debt. You are both my friend and mentor. You have shown me the importance of giving. I appreciate everything you have done for me. You have done so much for me, too numerous to mention here. From our nightly walks to the “casino”, solving and sketching problems on the blackboard, our late night car rides, dealing with your car getting smashed, or my cars getting stolen, we have been through it all. You have encouraged my broad interests and I have always been in awe of your personal knowledge. Most of all, I know that I can always count on you when I need you. I have tried to do my best to do the same for you and have tried to not let you down. If I become a professor I hope that I am able to emulate these qualities to my future students. Thank you for this experience.

TABLE OF CONTENTS

ABSTRACT	iii
THESIS DEDICATION	v
ACKNOWLEDGEMENTS	vi
LIST OF FIGURES	xii
LIST OF TABLES	xiv
PREFACE	xvi

Chapter I Background

I.I.	Introduction	2
I.II.	Postulates of Quantum Mechanics	3
I.III.	The Schrödinger Equation	4
I.IV.	Approximate Wavefunctions and the Variational Principle	6
I.V.	Hartree-Fock Method	8
I.VI.	Configuration Interaction Method	10
I.VII.	Computational Complexity of the Schrödinger Equation	11
I.VIII.	Sparse Grids for the Electronic Schrödinger Equation	13
I.IX.	Boys' Collocation Method for the Scaled Schrödinger	18
I.X.	Density-Functional Theory	19
I.XI.	The Kohn-Sham Equations	21
I.XII.	Conceptual Density-Functional Theory	24
I.XIII.	Atomic Partitioning	26
I.XIV.	Summary of Ph.D. Work	28
I.XV.	References	30

Chapter II Perturbative Perspectives on the Chemical Reaction Prediction Problem

II.I.	Statement of the Problem	36
II.II.	Introduction	36
II.III.	Motivation	37
II.IV.	Theoretical Development	38
II.V.	Regioselectivity	43
II.VI.	Electrophilicity and the Quality of Leaving Groups	53
II.VII.	Conclusion	61
II.VIII.	Appendix	61
II.IX.	References	65

Chapter III Conceptual Density-Functional Theory for General Chemical Reactions, Including Those That Are Neither Charge Nor Frontier-Orbital Controlled **I. Theory and Derivation of a General-Purpose Reactivity Indicator**

III.I.	Statement of the Problem	70
--------	--------------------------	----

III.II.	Introduction	70
III.III.	Theoretical Background	71
III.IV.	A General-Purpose Model for Chemical Reactivity	79
III.V.	Recapitulation	95
III.VI.	Appendix A. Derivation of the Electrostatic Potential Contribution	98
III.VII.	Appendix B. Error Analysis for the Reactive Site Interaction Model	100
III.VIII.	References	104
Chapter IV	Conceptual Density-Functional Theory for General Chemical Reactions, Including Those That Are Neither Charge Nor Frontier-Orbital Controlled	
	II. Application to Molecules Where Frontier Molecular Orbital Theory Fails	
IV.I.	Statement of the Problem.	110
IV.II.	Introduction	110
IV.III.	Computational Methods	113
IV.IV.	Results and Discussion	114
IV.V.	Summary	140
IV.VI.	References	141
Chapter V	Predicting the Reactivity of Ambidentate Nucleophiles and Electrophiles Using a Single, General-Purpose, Reactivity Indicator	
V.I.	Statement of the Problem.	146
V.II.	Background	146
V.III.	Computational Methodology	149
V.IV.	Results	151
V.V.	Discussion	167
V.VI.	References	168
Chapter VI	Relativistic Quantum Theory of Atoms in Molecules: Results for the ZORA Hamiltonian	
VI.I.	Statement of the Problem.	174
VI.II.	Introduction	174
VI.III.	Background	176
VI.IV.	The ZORA Hamiltonian	178
VI.V.	The SR-ZORA Extension of QTAIM.	179
VI.VI.	Conclusion	181
VI.VII.	Appendices	182
VI.VIII.	References	185

Chapter VII How Ambiguous is the Local Kinetic Energy?

VII.I.	Statement of the Problem.	190
VII.II.	Introduction	190
VII.III.	Local Kinetic Energy from the Local Electronic Stress Tensor	195
VII.IV.	Local Kinetic Energy from the Quasiprobability Distributions	199
VII.V.	Analysis from Virial Theorems for Density Functionals	207
VII.VI.	Discussion	210
VII.VII.	References	213

Chapter VIII Breaking the Curse of Dimension for the Electronic Schrödinger Equation with Functional Analysis

VIII.I.	Statement of the Problem.	222
VIII.II.	Limited Configuration Interaction Calculations	222
VIII.III.	Motivation: Balancing Errors in Configuration Interaction	223
VIII.IV.	Background: Complexity Theory	229
VIII.V.	Background: Convergence Rate of Basis Set Expansions	231
VIII.VI.	Algorithm Details: Griebel-Knapek CI by Degrees	234
VIII.VII.	Discussion	238
VIII.VIII.	Summary	244
VIII.IX.	Appendix	245
VIII.X.	References	246

Chapter IX Approaching the Theoretical Limits of Computational Efficiency with the Local Schrödinger Equation

IX.I.	Statement of the Problem.	252
IX.II.	Motivation	252
IX.III.	Boys Collocation for the Local Schrödinger Equation	254
IX.IV.	Algorithm: Sparse Numerical Integration Grids	257
IX.V.	Algorithm: Free Iterative Complement Basis Functions	262
IX.VI.	Summary	264
IX.VII.	References	265

Chapter X Conclusions and Prospects for Future Work

X.I.	Conclusions	268
X.II.	Prospects for Future Work on Reactivity	270
X.III.	Prospects for Future Work on Kinetic Energy for Partitioning Atoms	270
X.IV.	Prospects for Future Work on Grid Techniques	271

LIST OF FIGURES

Chapter II Perturbative Perspectives on the Chemical Reaction Prediction Problem

Figure 1	The electrostatic potential of molecule a in Table 1, plotted on the van der Waals surface of the molecule	46
Figure 2	The Fukui function from above of molecule b in Table 1, plotted on the van der Waals surface of the molecule	48
Figure 3	(a) The electrostatic potential and (b) the Fukui function from below of molecule c in Table 1	50
Figure 4	(a) The electrostatic potential and (b) the Fukui function from below of molecule d in Table 1	52
Figure 5	The quadratic model for the dependence of energy on the molecular charge	56
Figure 6	Comparison between the nucleofugality of the halides and rate of solvolysis of halogenated 1-phenylethyl esters in 80% aqueous ethanol at 75°C	60

Chapter IV Conceptual Density-Functional Theory for General Chemical Reactions, Including Those That Are Neither Charge Nor Frontier-Orbital Controlled II. Application to Molecules Where Frontier Molecular Orbital Theory Fails

Figure 1	(a) The square magnitude of the HOMO orbital, (b) the Fukui function from below are plotted on the isodensity surface isoquinoline	116
Figure 2	The atomic charges (condensed electrostatic potential) of isoquinoline	119
Figure 3	The atomic charges (condensed electrostatic potential) of <i>protonated</i> isoquinoline	122
Figure 4	The condensed Fukui function of protonated isoquinoline	123
Figure 5	The (a) atomic charges and (b) condensed Fukui functions of isoquinoline with a sodium “spectator cation.”	125
Figure 6	The (a) atomic charges and (b) condensed Fukui functions of 10-hydroxy-10,9-bozaroarphenanthrene	128
Figure 7	The (a) “hydrogen summed” atomic charges and (b) “hydrogen summed” condensed Fukui functions of 10-hydroxy-10,9-bozaroarphenanthrene	131

Chapter VIII Breaking the Curse of Dimension for the Electronic Schrödinger Equation with Functional Analysis

Figure 1	Representation of Slater determinants used in a configuration interaction wavefunction. (a) The reference Slater determinant in the truncated basis set. (b) A highly excited Slater determinant that is included in a full configuration interaction (FCI) calculation in the truncated basis set. (c) A Slater determinant included in FCI on an extended basis set that makes a larger contribution to the wavefunction than the Slater determinant in (b)	226
----------	---	-----

Figure 2 The number of Slater determinants in various types of Griebel-Knapek CI by degrees, for singlet states with various numbers of electrons, N , and sizes of the spatial-orbital basis, K . Different values of $T_{\alpha\alpha} = T_{\beta\beta} = T$ are used; $T_{\alpha\beta} \rightarrow \infty$. The values of o^2v^4 and o^3v^5 are also given; these numbers characterize the scaling of coupled-cluster singles and doubles (CCSD) and coupled-cluster singles, doubles, and triples (CCSDT), respectively. ($o = N_o$ is the number of occupied states and $v = (K - N_o)$ is the number of virtual states.) (a) 30 spatial orbitals; (b) 50 spatial orbitals; (c) 100 spatial orbitals; (d) 200 spatial orbitals 242

LIST OF TABLES

Chapter II Perturbative Perspectives on the Chemical Reaction Prediction Problem

Table 1	Structures and orientations of molecules in figures 1-4	45
Table 2	The electrofugality of alkyl-mercury compounds	58

Chapter IV Conceptual Density-Functional Theory for General Chemical Reactions, Including Those That Are Neither Charge Nor Frontier-Orbital Controlled II. Application to Molecules Where Frontier Molecular Orbital Theory Fails

Table 1	Reactivity Transition Tables for 10-hydroxy-10,9-bozaroarphenanthrene using natural population analysis for the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice (b) Second-Choice	134
Table 2	Reactivity Transition Tables for 10-hydroxy-10,9-bozaroarphenanthrene using CHelpG population analysis with the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice (b) Second-Choice	135
Table 3	Reactivity Transition Tables for 10-hydroxy-10,9-bozaroarphenanthrene using Merz-Singh-Kollman population analysis with the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice (b) Second-Choice	136
Table 4	Reactivity Transition Tables for 10-methyl-10,9-bozaroarphenanthrene using natural population analysis with the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice (b) Second-Choice	139

Chapter V Predicting the Reactivity of Ambidentate Nucleophiles and Electrophiles Using a Single, General-Purpose, Reactivity Indicator

Table 1	Condensed general-purpose reactivity indicator for electrophilic attack on SCN^- using cHelpG charges	152
Table 2	Condensed general-purpose reactivity indicator for electrophilic attack on SeCN^- using cHelpG charges	154
Table 3	Condensed general-purpose reactivity indicator for electrophilic attack on NO_2^- using cHelpG charges	156
Table 4	Condensed general-purpose reactivity indicator for electrophilic attack on SO_3^{2-} using cHelpG charges	158
Table 5	Condensed general-purpose reactivity indicator for nucleophilic attack on dimethyl carbonate using cHelpG charges	160
Table 6	Condensed general-purpose reactivity indicator for nucleophilic attack on MNTS using cHelpG charges	162

Table 7	Condensed general-purpose reactivity indicator for nucleophilic attack on CNB using cHepG charges.	164
Table 8	Condensed general-purpose reactivity indicator for nucleophilic attack on CNB using cHepG charges	166

Chapter VIII Breaking the Curse of Dimension for the Electronic Schrödinger Equation with Functional Analysis

Table 1	The electron configurations included in a doubly-occupied configuration interaction calculation on Beryllium and Neon, in a double- ζ basis set $(3s, 2p, 1d)$, with $T_{\alpha\alpha} = T_{\beta\beta} = 0$; $T_{\alpha\beta} \rightarrow \infty$	239
Table 2	The electron configurations included in a doubly-occupied configuration interaction calculation on Beryllium and Neon, in a triple- ζ basis set $(4s, 3p, 2d, 1f)$, with $T_{\alpha\alpha} = T_{\beta\beta} = 0$; $T_{\alpha\beta} \rightarrow \infty$	240

PREFACE

This thesis contains some of the papers that emerged from the topics I investigated during my time at McMaster. Given the nature of modern science, few projects are completed individually. In this section my specific contributions to each chapter of this thesis is explained. The reader will notice that though my interests are broad, (and there are a half-dozen papers that did not seem to fit in this thesis) the underlying theme of my research is “using mathematics to explain chemical phenomena.” Indeed, when I began graduate school I was of the very rigid opinion that mathematics was the only way to explain chemistry. I have softened this opinion considerably and have come to see the value of physical and chemical intuition. Still, I think that mathematics is interesting, enjoyable and can provide much insight for explaining and predicting chemistry. Furthermore, even at this point in the history of chemistry, I feel that the usefulness of mathematics within chemistry has not been completely exploited, and I look forward to seeing and contributing to the further development of mathematical chemistry.

My thesis consists of an introduction, eight journal articles, and conclusions. The introduction, Chapter I, provides background details so that the thesis is self-contained. Chapters II-IX of this thesis include reprints of articles that are published or in preparation for submission for publication. The first section in each of these chapters is an additional statement of the problem, placing the journal article in the broader context of this thesis. Chapter X provides conclusions and suggests some future directions for research.

Chapter II is a reprint of the article “Perturbative perspectives on the chemical reaction prediction problem”, published in *International Journal of Quantum Chemistry*. (P. W. Ayers, J. S. M. Anderson, L. J. Bartolotti *Int. J. Quantum Chem.* **2005**, *101*, 520-534.) I am the second author of this article and provided computational assistance. The section on leaving groups is primarily my work. My coauthors are Paul W. Ayers and Libero J. Bartolotti. Paul Ayers wrote the first draft of this review article, I revised this to form the second draft, then he and I exchanged drafts until we were both satisfied.

Chapter III is a reprint of the article “Conceptual density-functional theory for general chemical reactions, including those that are neither charge- nor frontier-orbital-controlled. 1. Theory and derivation of a general-purpose reactivity indicator”, published in *Journal of Chemical Theory and Computation*. (J. S. M. Anderson, J. Melin, P. W. Ayers *J. Chem. Th. Comp.* **2007**, *3*, 358-374.) I identified the research problem that is addressed in this paper, had the key insight that lead to the general-purpose reactivity indicator, and wrote the first draft. My preceptor, Paul Ayers, provided mathematical details and historical

perspective. I am the first author of this article. My coauthors are Junia Melin and Paul W. Ayers.

Chapter IV is a reprint of the article “Conceptual density-functional theory for general chemical reactions, including those that are neither charge- nor frontier-orbital-controlled. 2. Application to molecules where frontier molecular orbital theory fails”, published in *Journal of Chemical Theory and Computation*. (J. S. M. Anderson, J. Melin, P. W. Ayers *J. Chem. Th. Comp.* **2007**, *3*, 375-389.) I performed all of the computations in this paper, did most of the data analysis, and wrote the initial draft. I am the first author of this paper. My coauthors are Junia Melin and Paul W. Ayers.

Chapter V is a reprint of the article “Predicting the reactivity of ambidentate nucleophiles and electrophiles using a single, general-purpose, reactivity indicator”, published in *Physical Chemistry Chemical Physics*. (J. S. M. Anderson, P. W. Ayers *Phys. Chem. Chem. Phys.* **2007**, *9*, 2371-2378.) I selected the molecules to be studied, performed the calculations, analyzed the data, and wrote the first draft. I am the first author of this paper. My coauthor is Paul W. Ayers.

Chapter VI is a reprint of the article “Quantum Theory of Atoms in Molecules, SR-ZORA treatment”, in preparation to be submitted for publication in a special issue of *Journal of Chemical Physics*. I started this project, entirely independently, as part of a graduate course I was taking. I performed the mathematical derivations in this paper and wrote the first draft. I am the first author of this article. My coauthor is Paul W. Ayers.

Chapter VII is a reprint of the article “How ambiguous is the local kinetic energy?”, published in *Journal of Physical Chemistry A*. (J. S. M. Anderson, P. W. Ayers, J. I. Rodriguez Hernandez *J. Phys. Chem. A* **2010**, *114*, 8884-8895.) This project grew out of my comprehensive examination, so all of the key results were derived by me, entirely independently. I wrote the initial draft. My coauthors are Paul W. Ayers, and Juan I. Rodriguez-Hernandez.

Chapter VIII is a reprint of the article “Breaking the curse of dimension for the electronic Schrödinger equation with functional analysis”, in preparation to be submitted for publication in *The Journal of Chemical Physics*. This article emerged from a set of research notes that I prepared for Hiroshi Nakatsuji on sparse grids and information-based complexity. I am the first author of this article. My coauthor is Paul W. Ayers.

Chapter IX is a reprint of the article “Approaching the theoretical limits of computational efficiency with the local Schrödinger equation”, in preparation to

be submitted for publication in *The Journal of Chemical Physics*. This article emerged from the aforementioned set of research notes written for Hiroshi Nakatsuji. I prepared the numerical integration grids and directed this project. I am the first author of this article. My coauthors include Paul W. Ayers, Hiroyuki Nakashima, and Hiroshi Nakatsuji.

The author of this thesis did the majority of the work in these articles notwithstanding the inclusion of coauthors. The author performed all computations, provided key insight, and performed most of the derivations. With the exception of chapter II, I wrote the first draft of all of the included papers. Chapters II-IX were principally guided by Paul W. Ayers.

Chapter I

Background

I.I. Introduction

When one thinks of chemistry, the first thing that springs to mind is mixing chemicals in beakers. One rarely thinks of mathematics. In fact, one might think that mathematics in chemistry is anathema. This belief stubbornly persists, even among some chemists. However, with the advent of quantitative thermodynamics in the 1870's and dawn of quantum mechanics in the 1920's, the mathematical equations for describing chemistry emerged. By solving these equations one can predict chemical phenomena to, and even beyond, experimental accuracy!

Sound too good to be true? In many ways it is. First, solving the equations provides numbers, not chemical intuition. Merely manipulating the equations is applied mathematics; it need not lead to chemical insight. In addition, the equations can only be solved exactly for the simplest systems. As Dirac put it:¹

“The underlying physical 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 the exact application of these laws leads to equations much too complicated to be soluble.”

The very first part of this statement encourages mathematical approaches to chemistry: all the mathematical tools needed for chemistry are available. Specifically, the framework for computing molecular properties is provided by quantum mechanics. However, Dirac concludes by stating that the framework is practically useless. This seems discouraging, but Dirac follows this statement with:

“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.”

This is the goal of this thesis.

Although people usually quote only Dirac's first sentence, it is the redemptive second sentence that has guided the research of a large part of the theoretical physics and the whole of the theoretical chemistry. Theoretical chemists use quantum mechanics by solving the electronic Schrödinger equation

in an approximate way, obtaining an approximate wavefunction. Using this wavefunction, molecular properties are described, and the mathematical underpinnings of qualitative chemical concepts are revealed.

I.II. Postulates of Quantum Mechanics

The following postulates provide the mathematical framework of quantum mechanics.^{2,3} For the purposes of chemistry, they provide the prescription for computing the properties of molecules, molecular aggregates, and materials.

Postulate 1. The state of a quantum mechanical system is completely specified by its wavefunction, $\Psi(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t)$. The wavefunction depends on the coordinates of the particle(s), their spin, and the time. $\Psi^*(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) \Psi(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) d\vec{\tau} dt$ is the probability that, in the time interval dt , the particle(s) are in the volume element $d\vec{\tau}$ located at $\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N$ at time t . Since probability distribution functions are normalised, the wavefunction is normalised according to

$$\int \Psi^*(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) \Psi(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) d\vec{\tau} = 1. \quad (1)$$

Postulate 2. To every observable from classical mechanics there corresponds a linear, Hermitian, operator in quantum mechanics. Because the operator is Hermitian, computing the observable always gives a real number.

Postulate 3. When measuring the observable associated with the Hermitian operator \hat{A} , the only values that will ever be observed in a precise observation are the eigenvalues a of the operator. The eigenvalues satisfy the equation

$$\hat{A}\Psi = a\Psi \quad (2)$$

The third postulate of quantum mechanics reveals the possible results of experimental measurements. Because most Hermitian operators have an infinite number of distinct eigenvalues, most experiments have an infinite number of possible outcomes. Finding all these outcomes is impossible. Fortunately one usually only wants to know the average value of the property.

Postulate 4. If a system is in a state described by a normalised wavefunction Ψ , then the expected value of the observable corresponding to the operator \hat{Q} is given by

$$Q = \sum_{s_i} \iint \cdots \int \Psi^* (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N) \hat{Q} \Psi (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N) d\vec{r}_1 \cdots d\vec{r}_N \quad (3)$$

The fourth postulate of quantum mechanics tells how to use Hermitian operators to compute the average (mean) values of molecular properties.

Postulate 5. The wavefunction of a system evolves in time according to the time-dependent Schrödinger equation⁴

$$\hat{H}\Psi (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) = i\hbar \frac{\partial}{\partial t} \Psi (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) \quad (4)$$

The Schrödinger equation is the equation that must be solved to determine the wavefunction. Once the wavefunction is known, the previous postulates provide a way to model chemical phenomena.

Postulate 6. The total wavefunction must be antisymmetric with respect to the interchange of all coordinates (space and spin) of one fermion with those of another. Electronic spin must be included in this set of coordinates.

Quantum chemistry is mainly concerned with the interactions between electrons with each other and with atomic nuclei. One of the simplest ways to ensure that postulate 6 is satisfied is to use a Slater determinant (described later) to approximate the wavefunction.⁵

I.III. The Schrödinger Equation

The Schrödinger equation is the fundamental equation of quantum mechanics. It is usually expressed as

$$\hat{H}\Psi (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t) = i\hbar \frac{\partial}{\partial t} \Psi (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t). \quad (5)$$

where \hat{H} is the quantum mechanical Hamiltonian operator, $i = \sqrt{-1}$, \hbar is Planck's constant divided by 2π , and t is the time. $\Psi (\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N, t)$ is

the N -particle wavefunction, the solution to the Schrödinger equation. The position and spin of the j^{th} particle are denoted \vec{r}_j and s_j , respectively.

If the Hamiltonian does not depend on time then Schrödinger equation simplifies to

$$\hat{H}\Psi(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N) = E\Psi(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N) \quad (6)$$

where E is the energy. The energy is an eigenvalue of the Hamiltonian. The first postulate of quantum mechanics indicates that the solutions to this equation provide all of the information required to completely describe the microscopic system specified by the choice of \hat{H} . Equation (6) is called the time-independent Schrödinger equation.

In chemistry, one is primarily interested in obtaining information about atoms and molecules. The general molecular Hamiltonian for N electrons and P nuclei, in atomic units, is

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \frac{1}{2} \sum_{\alpha=1}^P \frac{\nabla_{\alpha}^2}{M_{\alpha}} + \sum_{i=1}^N \sum_{\alpha=1}^P \frac{-Z_{\alpha}}{|\vec{r}_i - \vec{R}_{\alpha}|} + \sum_{i=1}^{N-1} \sum_{j=1+i}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{\alpha=1}^{P-1} \sum_{\beta=\alpha+1}^P \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|} \quad (7)$$

where $\{\nabla_i^2\}_{i=1}^N$ is the Laplacian acting on the electronic coordinates, $\{\nabla_{\alpha}^2\}_{\alpha=1}^P$ is the Laplacian acting on the nuclear coordinates, M_{α} and Z_{α} are the mass and atomic number (charge) of the α^{th} nucleus, $\{\vec{R}_{\alpha}\}_{\alpha=1}^P$ are the nuclear coordinates, specifying the positions of the atomic nuclei, and $\{\vec{r}_i\}_{i=1}^N$ are the electronic coordinates, specifying the positions of the electrons. The first term in Eq. (7) is the kinetic energy for the N electrons in the molecule; the second term is the kinetic energy for the P nuclei in the molecule. The third term is the potential energy operator for the nuclear-electron attraction; the fourth term is the potential energy operator for the electron-electron repulsion; the last term is the potential energy operator for the nuclear-nuclear repulsion.

The full molecular Hamiltonian in Eq. (7) is seldom used. Atomic nuclei are much more massive than electrons. (The lightest nucleus, the nucleus of the hydrogen atom, is already 1832 times more massive than an electron.) Because the nuclei are much more massive than the electrons, they move much more slowly. Hence, neglecting the motion of the nuclei should not affect our ability to model the motion of the electrons. To this end, Born and Oppenheimer proposed neglecting the coupling between electronic and nuclear motion. Within the Born-Oppenheimer approximation, the molecular Hamiltonian becomes

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N \sum_{\alpha=1}^P \frac{-Z_{\alpha}}{|\vec{r}_i - \vec{R}_{\alpha}|} + \sum_{i=1}^{N-1} \sum_{j=1+i}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{\alpha=1}^{P-1} \sum_{\beta=\alpha+1}^P \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|}. \quad (8)$$

This is known as the electronic Hamiltonian. The corresponding Schrödinger equation is referred to as the electronic Schrödinger equation. The nuclear-nuclear repulsion term in the electronic Hamiltonian is constant for a given molecular geometry and is often omitted.

The solutions to the electronic Schrödinger equation are the electronic wavefunctions. They provide key information about thermochemistry, molecular properties, spectroscopy, as well as every other chemically relevant observable. Because this thesis is primarily concerned with the electronic Hamiltonian and the electronic Schrödinger equation in this thesis, the word “electronic” will be omitted whenever the context is clear.

I.IV. Approximate Wavefunctions and the Variational Principle

As mentioned previously, Dirac indicated that making quantum mechanics useful for chemistry requires approximations. Approximating the molecular Hamiltonian as the electronic Hamiltonian results in a second-order partial differential equation that still depends on $3N$ spatial and N spin variables. Making a practical theory requires either making further simplifications to the Hamiltonian (giving rise to semiempirical quantum mechanics methods)⁶ or approximating the wavefunction.

From mathematics it is known that any well-behaved function can be expanded in a complete basis,^{7,8} that is

$$\Psi = \sum_{i=1}^{\infty} c_i \varphi_i \quad (9)$$

where φ_i is a basis function and c_i is the projection of the wavefunction, Ψ , onto this basis function. The basis functions are assumed to be orthogonal. There are many reasons to represent a wavefunction as a basis set expansion. For example, one representation is sometimes more convenient for illustrating formal properties. In addition, while the analytic form of Ψ is typically unknown, one can write an explicit formula for φ_i . Finally, one can use an incomplete set of φ_i to *approximate* the wavefunction. In other words, one can approximate the wavefunction as

$$\Psi \approx \sum_{i=1}^K c_i \varphi_i . \quad (10)$$

If the φ_i 's form a complete basis ($K \rightarrow \infty$), then the wavefunction is recovered exactly. Otherwise the wavefunction is recovered only approximately (excluding some trivial exceptions).

The most common method for approximating the wavefunction with a basis set is the variational principle. Recall that the ground-state energy is the lowest eigenvalue of the Schrödinger equation,

$$\hat{H}\Psi_{gs} = E_{gs} \Psi_{gs} . \quad (11)$$

Ψ_{gs} is the ground state wavefunction of the Hamiltonian \hat{H} . The variational principle indicates that approximate wavefunctions always have higher energy than the exact ground state wavefunction,

$$\frac{\left\langle \sum_{i=1}^K c_i \varphi_i \left| \hat{H} \right| \sum_{i=1}^K c_i \varphi_i \right\rangle}{\left\langle \sum_{i=1}^K c_i \varphi_i \left| \sum_{i=1}^K c_i \varphi_i \right\rangle} \geq E_{gs} = \frac{\langle \Psi_{gs} | \hat{H} | \Psi_{gs} \rangle}{\langle \Psi_{gs} | \Psi_{gs} \rangle} \quad (12)$$

The equality holds if Ψ_{gs} lies within the space spanned by the basis set, $\{\varphi_i\}_{i=1}^K$; otherwise the inequality holds. Generally, the equality only holds only if the basis is complete ($K \rightarrow \infty$). The variational principle indicates that given two approximate wavefunctions, the one with the lower energy is more accurate. This motivates the Rayleigh-Ritz method: the best ground state energy for a given set of φ_i 's is obtained by minimizing the left-hand-side of (12) with respect to the coefficients. At the minimum, one has an approximation to the ground-state wavefunction as a linear combination of the φ_i 's.

Postulate 6 indicates that the wavefunction must be antisymmetric with respect to exchange between pairs of electronic coordinates. This suggests that one should choose the φ_i 's as Slater determinants of one-electron wavefunctions, $\phi_k(\vec{r})$,⁵

$$\varphi_i \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\vec{r}_1)\alpha(1) & \phi_1(\vec{r}_1)\beta(1) & \dots & \phi_{N/2}(\vec{r}_1)\beta(1) \\ \phi_1(\vec{r}_2)\alpha(2) & \phi_1(\vec{r}_2)\beta(2) & \dots & \phi_{N/2}(\vec{r}_2)\beta(2) \\ \vdots & \vdots & & \vdots \\ \phi_1(\vec{r}_N)\alpha(N) & \phi_1(\vec{r}_N)\beta(N) & \dots & \phi_{N/2}(\vec{r}_N)\beta(N) \end{vmatrix} \quad (13)$$

The one-electron wavefunctions, $\phi_k(\vec{r})$, are called orbitals; they are normalised and orthogonal to each other. $\alpha(j)$ and $\beta(j)$ denote the two choices (up-spin and down-spin) for the spin of electron j . The prefactor ensures the wavefunction is normalised. Because a determinant of a matrix changes sign when two rows are interchanged and is zero when two rows are the same, the Slater determinant wavefunction is antisymmetric with respect to interchange of electrons and is zero when two electrons with the same spin are at the same location. This implies the Pauli exclusion principle: the probability of observing two electrons with the same spin and the same location is zero. Applying the Rayleigh-Ritz method to a single Slater determinant ($K = 1$ in Eq. (12)), then minimizing the energy with respect to the form of the orbitals, defines the Hartree-Fock method.

I.V. Hartree-Fock Method

Hartree-Fock theory is one of the most fundamental methods in quantum chemistry. It is an approximate wavefunction method based on utilising the variational principle to find the lowest-energy Slater determinant. The Hartree-Fock method is a mean-field method because each electron feels an average potential from the remaining electrons. Because the electrons only feel the average potential of the other electrons, the electrons move quasi-independently, each with its own individual wavefunction. The concept of orbitals emerges from this model because in this model electrons dwell within independent wavefunctions.

The Hartree-Fock equations are derived by applying the Rayleigh-Ritz method to a single Slater determinant. This gives a set of $N/2$ coupled integro-differential equations,

$$\left(-\frac{1}{2}\nabla^2 - \sum_{A=1}^M \frac{Z_A}{|\vec{r} - \vec{R}_A|} + \hat{J}(\vec{r}) + \hat{K}(\vec{r}) \right) \phi_j(\vec{r}) = \varepsilon_j \phi_j(\vec{r}); \quad j = 1, \dots, N/2 \quad (14)$$

where $\hat{J}(\vec{r})$ is the Coulomb operator defined as

$$\hat{J}(\vec{r})\phi_j(\vec{r}) = 2 \left(\sum_{i \neq j}^{N/2} \int \frac{\phi_i^*(\vec{r}')\phi_i(\vec{r}')d\vec{r}'}{|\vec{r} - \vec{r}'|} \right) \phi_j(\vec{r}) \quad (15)$$

and $\hat{K}(\vec{r})$ is the exchange operator defined as

$$\hat{K}(\vec{r})\phi_j(\vec{r}) \equiv \left(\sum_{i \neq j}^{N/2} \int \frac{\phi_i^*(\vec{r}')\phi_j(\vec{r}')d\vec{r}'}{|\vec{r} - \vec{r}'|} \right) \phi_i(\vec{r}). \quad (16)$$

Solving the Hartree-Fock equations exactly would give the exact Hartree-Fock orbitals, $\phi_j(\vec{r})$, and their corresponding orbital energies ε_j . If the Hartree-Fock wavefunction were the exact solution to the Schrödinger equation, then the orbital energies would be exactly equal to the molecule's electron-removal energies (ionization potentials) and electron-attachment energies (electron affinities). Even though the Hartree-Fock wavefunction is only approximate, the first $N/2$ orbital energies are useful approximations to the ionization potentials.

The Hartree-Fock equations are coupled by $\hat{J}(\vec{r})$ and $\hat{K}(\vec{r})$, which model the average electron repulsion that an electron in orbital $\phi_j(\vec{r})$ feels from the other electrons. Because determining $\hat{J}(\vec{r})$ and $\hat{K}(\vec{r})$ requires knowledge of the orbitals, but Eq. (14) can only be solved for the orbitals when $\hat{J}(\vec{r})$ and $\hat{K}(\vec{r})$ are known, these equations must be solved self-consistently. That is, starting from an initial guess for the Hartree-Fock orbitals, $\phi_j(\vec{r})$, the potentials $\hat{J}(\vec{r})$ and $\hat{K}(\vec{r})$ are computed. Then Eqs. (14) are solved, generating a new set of orbitals. If these new orbitals are close enough to the input orbitals, then the equations are considered solved. Otherwise, new orbitals are guessed. Typically, these orbitals are guessed from some combination of previously guessed orbitals and the output orbitals. (The simplest guess is to take the previous output orbitals, but this may not always converge.) The new guess orbitals are used to compute $\hat{J}(\vec{r})$ and $\hat{K}(\vec{r})$ and Eqs. (14) are solved again. This process repeats until the orbitals that were used to construct $\hat{J}(\vec{r})$ and $\hat{K}(\vec{r})$ are close enough to the orbitals obtained from solving Eqs. (14). At that point the orbitals in the Hartree-Fock equations are "self-consistent" and the equations are considered solved.

Hartree-Fock is generally easily solved within a basis set ($O(N^4)$ computational scaling),⁹ and typically provides reliable results for equilibrium positions of atomic nuclei and, with appropriate scaling corrections, vibrational spectra.¹⁰⁻¹² Unfortunately, Hartree-Fock is not generally reliable for describing chemical processes in which chemical bonds are broken or formed. However, because of error cancellation, Hartree-Fock is often accurate for isodesmic reactions, where the number and type of chemical bonds present before and after a chemical reaction are the same.

The difference between the true ground state energy E_{gs} and the Hartree-Fock energy, E_{HF} , defines the correlation energy,

$$E_{corr} \equiv E_{gs} - E_{HF} \quad (17)$$

Despite this correlation error, the Hartree-Fock method is still very instructive. Much of the current understanding of chemistry is based on the “molecular orbital theory” built from the Hartree-Fock orbitals.¹³

Correcting the electron correlation error remedies the inherent errors in the Hartree-Fock approach and allows chemical phenomena to be accurately predicted. The most straightforward approach is to apply the variational principle to a linear combination of the Hartree-Fock wavefunctions. (I.e., the φ_i in Eq. (10) come from the ground state Hartree-Fock Slater determinant and its excited states). This approach is known as configuration interaction. Other approaches include using perturbation theory, cumulants, and a combination of these two starting from a Hartree-Fock ansatz. These methods have computational scalings ranging from $O(N^5)$ to $O(N!)$; they are referred to as post-Hartree-Fock methods.^{11,14,15} The fastest post-Hartree Fock method is only applicable to molecules with up to about 100 atoms.

I.VI. Configuration Interaction Method

Consider a wavefunction that is a linear combination of the Hartree-Fock ground state wavefunction and various excited-state Hartree-Fock wavefunctions. Minimizing the expansion coefficients of this wavefunction, as prescribed by the Rayleigh-Ritz principle, defines the configuration interaction method. If the basis set includes the ground state Slater determinant and all of its singly-excited electron configurations, the method is referred to as configuration interaction singles (CIS). If the basis set includes the ground state and all of its doubly-excited configurations, the method is referred to as configuration interaction doubles (CID). Obviously one can combine these two approaches, including both singly- and doubly-excited electron configurations to obtain configuration interaction singles and doubles (CISD). Most generally, the configuration interaction method has the expansion

$$\Psi \approx \underbrace{c_0 \Phi_0}_{\substack{\text{No Excitations} \\ \text{Hartree-Fock} \\ \text{Ground State}}} + \underbrace{\sum_{i=1}^N \sum_{a=N+1}^K c_i^a \Phi_i^a}_{\text{Single Excitations (S)}} + \underbrace{\sum_{i>j=1}^N \sum_{a>b=N+1}^K c_{ij}^{ab} \Phi_{ij}^{ab}}_{\text{Double Excitations (D)}} \quad (18)$$

$$+ \underbrace{\sum_{i>j>k=1}^N \sum_{a>b>c=N+1}^K c_{ijk}^{abc} \Phi_{ijk}^{abc}}_{\text{Triple Excitations (T)}} + \underbrace{\sum_{i>j>k>l=1}^N \sum_{a>b>c>d=N+1}^K c_{ijkl}^{abcd} \Phi_{ijkl}^{abcd}}_{\text{Quadruple Excitations (Q)}} + \dots$$

where the Φ 's are the solutions to the Hartree-Fock equations. Φ_0 is the ground state Hartree-Fock wavefunction. Φ_i^a is the singly excited Hartree-Fock wavefunction obtained by exciting an electron from the i^{th} occupied spin-orbital to

the a^{th} unoccupied spin-orbital of the ground-state Hartree-Fock wavefunction, Φ_0 . Φ_{ij}^{ab} is the doubly-excited Hartree-Fock wavefunction obtained by exciting electrons from the i^{th} and j^{th} occupied spin-orbitals to the a^{th} and b^{th} unoccupied spin-orbitals. Configuration interaction singles and doubles (CISD) stops at this point, and does not include terms from the second line of Eq. (18). Including further terms gives better results, but the computational scaling for including up to m^{th} excitations is $O(N^{2(m+1)})$, so highly-excited CI calculations are rare.

Because the energy is invariant under unitary transformations of the Hartree-Fock orbitals, CIS provides no correction to the Hartree-Fock ground state energy in a complete orbital basis set. Even in practical incomplete basis set calculations, CIS provides only a very small correction. By contrast, the doubly excited determinants almost always account for more than half of the correlation energy, and provide the largest correction of all the terms in Eq. (18).

If Eq. (18) is not truncated, and all possible excited-state Slater determinants are included, the exact solution, for a given orbital basis set, is obtained. This is known as full configuration interaction (FCI). This is both the most accurate and most computationally costly method for a given orbital basis. The computational scaling of FCI is $O(N!)$, so it is rarely used except for benchmarking other methods. FCI is only reliable for larger basis sets and it is limited to molecules with 10 or fewer electrons. Often CISD with a large orbital basis set can produce more accurate energies, more quickly, than FCI with a small orbital basis set.

I.VII. Computational Complexity of the Schrödinger Equation

The computational cost of solving the Schrödinger equation can be studied using the language and mathematical formalism of information-based complexity theory.¹⁶⁻¹⁸ In complexity theory, the computational cost is equal to the number of floating-point operations (additions and multiplications) that are required to obtain a given accuracy, ϵ . The prohibitive $O(N!)$ scaling of FCI is expressed as

$$\text{cost}_{\text{FCI}} \sim \epsilon^{-N} \quad (19)$$

This means that if one wishes to double the accuracy of our calculation ($\epsilon \rightarrow \frac{1}{2}\epsilon$ in Eq. (19)), then the cost of the calculation increases exponentially with increasing numbers of electrons, ($\text{cost} \rightarrow \text{cost} \cdot 2^N$). Computational methods whose cost rises exponentially with increasing problem size are said to suffer from the curse of dimension. Such methods are called intractable. A computational method is tractable if its cost does not grow exponentially with

electron number. (I.e., $\text{cost}_{\text{tractable}} \sim f(N)\varepsilon^{-x}$, where x does not depend on the number of electrons.) A method is strongly tractable if its cost is independent of the number of electrons.

The notation $\text{cost}(\varepsilon) \sim f(\varepsilon)$ indicates that the computational cost is proportional to $f(\varepsilon)$ in the high-accuracy ($\varepsilon \rightarrow 0$) limit. There are usually higher-order terms in Eq. (19) and computational prefactors; these subtleties can have practical significance for small calculations at the relatively low level of accuracy that is often required in chemical applications. One could include these factors, and replace Eq. (19) with

$$\text{cost}_{\text{FCI}} = \sum_{j=0}^N a_j \varepsilon^{-j}, \quad (20)$$

but the resulting expression is too complicated to be very useful.

Information-based complexity theory provides a useful mathematical framework for comparing the cost of existing numerical algorithms and developing new ones. It is used to develop new computational methods for solving the Schrödinger equation in chapters VIII and IX of this thesis.

Chapters VIII and IX of this thesis address the fundamental question: Is there a way to recover the accuracy of FCI with a computational method that is more efficient? In particular, can we formulate a strongly tractable method for solving the Schrödinger equation?

For general Hamiltonians, one can show that FCI is the most efficient possible method. However, the molecular electronic structure problem is special, because the potential-energy operator is smooth except for simple poles when two particles (either an electron and an atomic nucleus or two electrons) are at the same position. The mathematically “nice” nature of the Hamiltonian reveals itself in the mathematical “niceness” of the electronic wavefunctions for molecules.¹⁹⁻²¹ In particular, molecular wavefunctions have bounded mixed derivatives.²² That is, there exists a number, $t > 0$, for which

$$\left| \frac{\partial^{t_x^{(1)}} \partial^{t_y^{(1)}} \partial^{t_z^{(1)}} \partial^{t_x^{(2)}} \partial^{t_y^{(2)}} \partial^{t_z^{(2)}} \dots \partial^{t_x^{(N)}} \partial^{t_y^{(N)}} \partial^{t_z^{(N)}} \Psi(x_1, y_1, z_1, s_1; x_2, y_2, z_2, s_2; \dots; x_N, y_N, z_N, s_N)}{\partial x_1^{t_x^{(1)}} \partial y_1^{t_y^{(1)}} \partial z_1^{t_z^{(1)}} \partial x_2^{t_x^{(2)}} \partial y_2^{t_y^{(2)}} \partial z_2^{t_z^{(2)}} \dots \partial x_N^{t_x^{(N)}} \partial y_N^{t_y^{(N)}} \partial z_N^{t_z^{(N)}}} \right| < \infty \quad (21)$$

as long as all of the indices, $\{t_x^{(j)}, t_y^{(j)}, t_z^{(j)}\}_{j=1}^N$, are less than or equal to t . This is a much stronger result than merely requiring that the t^{th} derivatives of Ψ exist because it requires that certain very-high-order derivatives exist. For example, the t^{3N} -th derivative

$$\left| \frac{\partial' \partial' \partial' \dots \partial' \partial' \partial' \Psi(x_1, y_1, z_1, s_1; x_2, y_2, z_2, s_1; \dots; x_N, y_N, z_N, s_N)}{\partial x_1' \partial y_1' \partial z_1' \dots \partial x' \partial y' \partial z'} \right| < \infty \quad (22)$$

exists. In fact, Eq. (21) requires that $(t+1)^{3N}$ derivatives exist. Because the electronic wavefunctions are “exponentially differentiable,” information-based complexity tells us that there exist methods for solving the molecular Schrödinger equation that do not have exponential cost. In particular, for a wavefunction with t bounded mixed derivatives,

$$\text{cost} \sim \varepsilon^{-1/t}. \quad (23)$$

Strongly tractable methods for solving the electronic Schrödinger equation exist. It is important to emphasize that this result is specific to molecular Hamiltonians: for general Hamiltonians FCI is, in some sense, optimal.

Equation (23) says that efficient computational methods for solving the molecular Schrödinger equation exist. It does not provide insight into how to construct such methods. Chapters VIII and IX propose two approaches.

The first approach, in Chapter VIII, is based on truncation of the configuration-interaction expansion in Eq. (18). In particular, the method proposed there breaks through the $O(N!)$ computational scaling of the FCI method. The new method has a complicated computational scaling, but one can show that the scaling is bound by a polynomial, $O(N^k)$.^{23,24} Despite its much more favourable computational scaling, the proposed method is, theoretically, just as accurate as FCI. Even more surprisingly, it seems that in many cases the proposed method may be even more computationally efficient than popular approximations to FCI methods like coupled-cluster singles and doubles (CCSD) and coupled-cluster singles, doubles, and triples (CCSDT).

I.VIII. Sparse Grids for the Electronic Schrödinger Equation

To this point, every approach we have mentioned for approximating the wavefunction has been based on expanding the wavefunction in one (Hartree-Fock) or more (configuration interaction) Slater determinants. It is difficult to break the curse of dimension and achieve the optimal computational cost in Eq. (23) with such methods. In Chapter IX, we explore whether computational methods that are not based on Slater determinants might be preferable.

Much of the mathematical development of information-based complexity theory has been focussed on the problem of many-dimensional numerical

integration. By contrast, the problem that is directly relevant for solving the Schrödinger equation—solving many-dimensional partial differential equations—has been much less studied.²⁵⁻³⁰ This motivated us to explore whether we could develop an algorithm for solving the many-electron Schrödinger equation that requires only numerical integration.

The problem of developing efficient numerical integration techniques for electronic structure problems was something I worked on a great deal during the early part of my graduate studies.^{31,32} The resulting papers are not included in this thesis, so I will summarise the main results here.

Our approach for many-electron integrals is based on the Smolyak approach to multi-dimensional quadrature.³³ The basic idea is to construct multidimensional quadratures from Cartesian products of one-dimensional quadrature grids.

Suppose that one is given a sequence of nested one-dimensional quadrature grids for the unit interval,

$$\int_0^1 f(x) dx \approx \sum_{i=1}^{n(q)} w_i^{(q)} f(x_i^{(q)}) \quad q = 1, 2, \dots \quad (24)$$

The number q is called the effort of the grid; the number of points, $n(q)$, in the quadrature formula is a nondecreasing function of the effort. The points (or nodes) in the quadrature formula are denoted $\{x_i^{(q)}\}_{i=1}^{n(q)}$. The integration weights are denoted $\{w_i^{(q)}\}_{i=1}^{n(q)}$. The basic form in Eq. (24) can be applied to any linear operator on $f(x)$ (not just integration) by using the appropriate weights.

The most computationally efficient method for numerical integration in one dimension is Gaussian quadrature, which has computational cost

$$\text{cost} \sim \varepsilon^{-1/t} \quad (25)$$

for functions with bounded t^{th} derivatives. In higher dimensions, Gaussian quadrature formulas need not exist and, in general, the computational scaling of methods for finding optimal quadrature formulae grows as the factorial of the number of dimensions.³⁴ For functions with t bounded mixed derivatives, however, there exist multidimensional integration methods with the same asymptotic computational cost as one-dimensional integration, Eq. (25).³⁵⁻³⁹

Although the optimal d -dimensional numerical integration formulae are not known, and arguably not knowable, there are practical methods for numerical integration that achieve the computational cost scaling in Eq. (25).^{23,26,33,40,41} These methods are less efficient than multidimensional Gaussian quadrature, but

the difference is only in the computational prefactor. In general, these numerical integration methods contain several times more points than the optimal quadrature formula. Even so, these practical methods are efficient enough to break the curse of dimension.

To construct many-electron quadrature methods, we first construct the one-dimensional difference grids,

$$\Delta^{(q)} f(x) = \sum_{i=1}^{n(q)} w_i^{(q)} f(x_i^{(q)}) - \sum_{i=1}^{n(q-1)} w_i^{(q-1)} f(x_i^{(q-1)}) = \sum_{i=1}^{n(q)} v_i f(x_i^{(q)}). \quad (26)$$

The $q = 0$ grid has zero points. We then construct multidimension grids as a direct (i.e., Cartesian) product of lower-dimensional grids. For example, for two dimensions,

$$\Delta_x^{(q_x)} \otimes \Delta_y^{(q_y)} f(x, y) = \sum_{i=1}^{n(q_x)} \sum_{j=1}^{n(q_y)} v_i^{(q_x)} v_j^{(q_y)} f(x_i^{(q_x)}, y_j^{(q_y)}). \quad (27)$$

In the general d -dimensional case

$$\begin{aligned} \bigotimes_{i=1}^d \Delta_{x_i}^{(q_i)} f(x_1, x_2, \dots, x_d) &= \Delta_{x_1}^{(q_1)} \otimes \Delta_{x_2}^{(q_2)} \otimes \dots \otimes \Delta_{x_d}^{(q_d)} f(x_1, x_2, \dots, x_d) \\ &= \sum_{i_1=1}^{n(q_1)} \sum_{i_2=1}^{n(q_2)} \dots \sum_{i_d=1}^{n(q_d)} v_{i_1}^{(q_1)} v_{i_2}^{(q_2)} \dots v_{i_d}^{(q_d)} f(x_{i_1}^{(q_1)}, x_{i_2}^{(q_2)}, \dots, x_{i_d}^{(q_d)}) \quad (28) \\ &= \sum_{i_1=1}^{n(q_1)} \sum_{i_2=1}^{n(q_2)} \dots \sum_{i_d=1}^{n(q_d)} \left(\prod_{k=1}^d v_{i_k}^{(q_k)} \right) f(x_{i_1}^{(q_1)}, x_{i_2}^{(q_2)}, \dots, x_{i_d}^{(q_d)}) \end{aligned}$$

The efficient multidimensional grids that fulfill the optimal asymptotic cost condition, Eq. (25), all have the general form

$$\sum_{\{\bar{q} | \mathcal{I}(\bar{q}, d) \leq Q\}} \bigotimes_{i=1}^d \Delta_{x_i}^{(q_i)} f(x_1, x_2, \dots, x_d) \quad (29)$$

The sum includes all $\bar{q} = [q_1, q_2, \dots, q_d]^T$ that satisfy a specified condition on the efforts, $\mathcal{I}(\bar{q}, d) \leq Q$. Q is the effort of the multidimensional integration. The simplest restriction on the one-dimensional efforts is simply

$$\max_{i=1,2,\dots,d} (q_i) \leq Q. \quad (30)$$

This is the full direct-product grid; it corresponds to grid-based FCI and it does not achieve the favourable computational cost in Eq. (25). Grids based on Eq. (29) that include fewer points than this direct-product grid are said to be sparse.

The first, and simplest, efficient multidimensional integration grid was proposed by Smolyak,³³

$$\mathcal{I}(\bar{q}, d) = 1 - d + \sum_{i=1}^d \bar{q}_i. \quad (31)$$

Except for a logarithmic factor, Smolyak integration achieves the optimal computation cost from Eq. (25).^{35,38} The logarithmic deficiency of the Smolyak method is eliminated if one uses Petras’s delayed one-dimensional quadrature formulae.⁴⁰

It turns out that molecular wavefunctions not only have bounded mixed derivatives; they have a property called dominating mixed smoothness.^{22,23,42,43} Using this property, one can choose an index set that is even more restrictive than the one prescribed by Smolyak, namely,

$$\mathcal{I}(\bar{q}, d) = \frac{1 - d \sum_{i=1}^d q_i - T \max_{i=1,2,\dots,d} (q_i)}{1 - T} \quad -\infty < T < 1 \quad (32)$$

The appropriate value of T depends on the differentiability of the function being integrated. $T = 0$ recovers the Smolyak choice (Eq. (31)). $T \rightarrow -\infty$ gives the full direct-product grid (Eq. (30)). For molecular wavefunctions, it is appropriate to choose $0 < T < -1$. The index set in Eq. (32) was proposed by Griebel and Knapek.²³ For $T > 0$, the Griebel-Knapek construction removes the logarithmic deficiency of the Smolyak grids and fulfills the optimal cost condition in Eq. (25). Many other definitions for sparse integration grids can be proposed. For example, Bungartz and Griebel derived a choice in which the error in the integration formula was measured not with the normal \mathbb{L}^2 norm, but with the Sobolev (energy) norm.^{26,44} Because this formula is designed to minimize the error in the energy, it might be especially useful for quantum mechanical applications. The final formula cannot be written in the same simple form as the previous ones: instead of $\mathcal{I}(\bar{q}, d) \leq Q$ one has the intimidating expression

$$\sum_{i=1}^d q_i - \frac{1}{5} \log_2 \left(\sum_{i=1}^d 4^{q_i} \right) \leq Q - \frac{1}{5} \log_2 (4^{Q-d+1} + 4d - 4). \quad (33)$$

The Bungartz-Griebel sparse grid also satisfies the optimal cost condition in Eq. (25).

Our approach for constructing numerical integration grids for N -electron integrals is based on Eq. (29). We have tried many different formulae for the one-dimensional integration formulae; Gauss-Chebyshev and Clenshaw-Curtis formulas work particularly well. Using these one dimensional formulae and Eq. (29) with $d = 3$, we obtain a three-dimensional grid on the unit cube, $[0,1]^3$. The resulting integration formula is

$$\int_0^1 \int_0^1 \int_0^1 f(\Theta_1, \Theta_2, \Theta_3) d\Theta_1 d\Theta_2 d\Theta_3 = \sum_{i=1}^{n(Q,3)} w_i^{(Q)} f(\Theta_{1,i}, \Theta_{2,i}, \Theta_{3,i}). \quad (34)$$

This formula is inappropriate for one-electron molecular integrals. This is because integrals lie in real space. To adapt Eq. (34) to molecular integrals, we use a transformation of coordinates,

$$\begin{aligned} \Theta_1(x) &= \frac{\int_{-\infty}^x \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(X, Y, Z) dXdYdZ}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(X, Y, Z) dXdYdZ} \\ \Theta_2(x, y) &= \frac{\int_{-\infty}^y \int_{-\infty}^{\infty} p(x, Y, Z) dYdZ}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, Y, Z) dYdZ} \\ \Theta_3(x, y, z) &= \frac{\int_{-\infty}^z p(x, y, Z) dZ}{\int_{-\infty}^{\infty} p(x, y, Z) dZ} \end{aligned} \quad (35)$$

The Jacobian determinant of this transformation is

$$\left| \frac{\partial(\Theta_1, \Theta_2, \Theta_3)}{\partial(x, y, z)} \right| = p(x, y, z), \quad (36)$$

so, Eq. (34) can be rewritten as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) p(x, y, z) dx dy dz = \sum_{i=1}^{n(Q,3)} w_i^{(Q)} f(x_i, y_i, z_i) \quad (37)$$

Equation (37) will be especially useful for one-electron integrals that can be written as $p(x,y,z)$ times a slowly-varying function. This suggests that we set $p(x,y,z)$ equal to the probability of observing an electron at a point in space, $\rho(x, y, z)$. Since the probability of observing an electron at a given point in space is known only after we have solved the Schrödinger equation, we approximate it with the sum of the atomic probabilities,⁴⁵

$$p(x, y, z) = \sum_{\alpha=1}^P \rho_{\alpha}^0(x - X_{\alpha}, y - Y_{\alpha}, z - Z_{\alpha}) \quad (38)$$

This is called the promolecular density. The quantities $\rho_{\alpha}^0(\bar{r} - \bar{R}_{\alpha})$ are just the ground-state electron densities of the isolated atoms, centred at the points $\{(X_{\alpha}, Y_{\alpha}, Z_{\alpha})\}_{\alpha=1}^P$.

Evaluating the transformation of coordinates requires performing the indefinite integrals in Eq. (35). This is computationally feasible only if these

integrals can be performed analytically. We therefore approximate the atomic densities as

$$\tilde{\rho}_\alpha^0(\vec{r}) = \sum_{i=1}^k c_i e^{-a_i r^2} + \sum_{i=k+1}^m \frac{c_i}{(r^2 + a_i)^2} \approx \rho_\alpha^0(\vec{r}). \quad (39)$$

The last term in Eq. (39)—the squared Lorentzian function—is required to enforce the boundary conditions on the Schrödinger equation. The integrands of interest in quantum chemistry fall off exponentially faster than cr^{-4} (but exponentially slower than e^{-ar^2}). Including at least one squared Lorentzian term (i.e., $m > k$ in Eq. (39)) is essential, because otherwise the function $f(\vec{r})$ in Eq. (37) has essential singularities as $r \rightarrow \infty$.

After performing the transformation of coordinates in Eq. (35), we have a set of nested grids for one-electron molecular integrals, cf. Eq. (37). Applying the sparse-grid generating expression again (cf. Eq. (29)) gives a formula for N -electron integrals,

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, y_1, z_1, \dots, x_N, y_N, z_N) \left(\prod_{k=1}^N p(x_k, y_k, z_k) \right) dx_1 dy_1 dz_1 \dots dx_N dy_N dz_N \\ &= \sum_{i=1}^{n(Q,3N)} w_i^{(Q)} f(x_{1,i}, y_{1,i}, z_{1,i}, \dots, x_{N,i}, y_{N,i}, z_{N,i}) \end{aligned} \quad (40)$$

The method to be discussed in chapter IX is based on our ability to efficiently perform many-electron integrals.

I.IX. Boys' Collocation Method for the Scaled Schrödinger Equation

Consider expanding the wavefunction in an arbitrary basis set, as in Eq. (10). Inserting this expression into the Schrödinger equation gives the equation,

$$\left(\hat{H} - E^{(n)} \right) \Psi(\vec{r}) = 0 \approx \left(\hat{H} - E^{(n)} \right) \sum_{i=1}^K c_i^{(n)} \phi_i(\vec{r}) \quad (41)$$

Here $E^{(n)}$ denotes the energy of the n^{th} excited state. There is generally no way to choose the expansion coefficients, $\{c_i^{(n)}\}_{i=1}^K$, to solve this equation exactly unless the basis set is complete. In the local Schrödinger equation approach, instead of choosing the $c_i^{(n)}$'s to minimize the energy, we choose them to minimize the error in Eq. (41).⁴⁶ To achieve this, multiply both sides the equation by a set of test

functions, $\{\psi_j^*(\bar{\tau})\}_{j=1}^K$, and integrate. This gives a secular equation to solve for the $c_i^{(n)}$'s:

$$0 = \sum_{i=1}^K c_i^{(n)} \left(\iint \cdots \int \psi_j^*(\bar{\tau}) (\hat{H} - E^{(n)}) \varphi_i(\bar{\tau}) d\bar{\tau} \right) \quad (42)$$

For very good basis functions like the free iterative complement functions of Nakatsuji *et al.*,⁴⁶ the integrals in Eq. (42) cannot be performed analytically. Using a numerical integration method gives,

$$0 = \sum_{i=1}^K \sum_{k=1}^{n(Q,3N)} c_i^{(n)} w_k^{(Q)} \psi_j^*(\bar{\tau}_k^{(Q)}) (\hat{H} - E^{(n)}) \varphi_i(\bar{\tau}_k^{(Q)}) \quad (43)$$

Equation (43) is the Boys collocation method.⁴⁷ It casts the Schrödinger equation as a numerical integration problem. As such, it provides a way to exploit the near-optimal many-electron integration grids we have developed.

Unfortunately, many-electron integration grids obtained from the sparse-direct-product formula, Eq. (29), contain points where two electrons are at the same position, $\bar{r}_i = \bar{r}_j$. The electron-electron repulsion potential diverges at these points, so there can be grid points where the integrand diverges. To remove this problem, multiply both sides of Eq. (41) by

$$U(\bar{\tau}) = \frac{1}{V_{ne}(\bar{\tau}) + V_{ee}(\bar{\tau})}. \quad (44)$$

Here $V_{ne}(\bar{\tau})$ and $V_{ee}(\bar{\tau})$ are the nuclear-electron and electron-electron potentials, respectively. Applying the same derivation as in Eqs. (41)-(43) to this scaled Schrödinger equation,

$$U(\hat{H} - E^{(n)})\Psi^{(n)}(\bar{\tau}) = 0 \quad (45)$$

gives

$$0 = \sum_{i=1}^K \sum_{k=1}^{n(Q,3N)} c_i^{(n)} w_k^{(Q)} \psi_j^*(\bar{\tau}_k^{(Q)}) \left[U(\hat{H} - E^{(n)}) \right] \varphi_i(\bar{\tau}_k^{(Q)}). \quad (46)$$

This equation is also due to Boys.⁴⁷ Chapter IX is based on solving Eq. (46) by combining the numerical integration grids developed in my earlier research^{31,32} with the free iterative complement basis functions^{46,48} I studied during a summer fellowship sponsored by Japan's Society for the Promotion of Science.

I.X. Density-Functional Theory

Instead of the electronic wavefunction, the electron density can be used as the fundamental descriptor of the system. The electron density, $\rho(\bar{r})$, is the

probability of finding an electron at point \vec{r} . The electron density can be directly computed from the wavefunction through:

$$\rho(\vec{r}, s) = N \sum_{s_i} \iint \cdots \int \Psi^*(\vec{r}, s, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N) \Psi(\vec{r}, s, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N) d\vec{r}_2 \cdots d\vec{r}_N. \quad (47)$$

Formulating quantum mechanics in terms of the electron density has the advantage that the electron density is measurable and only depends on three space coordinates. In contrast, the wavefunction is unphysical and depends on every electronic coordinate. The electron density is simpler, both conceptually and mathematically, than the wavefunction.⁴⁹

The mathematical foundations of density-functional theory were propounded in two theorems by Hohenberg and Kohn.⁵⁰

1st Theorem. The ground state's electron density determines all the properties of an electronic system. In particular, the ground state energy is a functional of the electron density, $E = E[\rho]$.

This establishes that the electron density encapsulates all the information from the ground-state electronic wavefunction. (Later authors proved this explicitly, by reconstructing the ground state wavefunction from the electron density.⁵¹) Practical applications of the 1st Hohenberg Kohn theorem, however, require methods for determining the ground state's electron density directly, without passing through the wavefunction. (One does not wish to use Eq. (47).) The 2nd Hohenberg-Kohn theorem provides a variational principle for finding the ground state density.⁵⁰

2nd Theorem. The N -electron ground state energy, E_{gs} , and ground state density, $\rho_{gs}(\vec{r})$, are obtained by minimizing the energy with respect to all N -electron densities,

$$E_{gs} = E[\rho_{gs}] = \underbrace{\min}_{\rho} E[\rho] \quad (48)$$

This is the analogue of the Rayleigh-Ritz variational principle for the wavefunction. Just as the Rayleigh-Ritz principle leads to the most popular computational approaches for the ground state wavefunction, the 2nd Hohenberg-Kohn theorem leads to the most popular approaches for the ground-state density. The ground-state energy and other chemical properties then follow. Because the electron density only depends on three coordinates, the domain of the functions

used to solve the electronic structure problem has been reduced from a space of $3N$ dimensions (plus spin) to a space of only 3 dimensions.

Sound too good to be true? The first Hohenberg-Kohn theorem is merely an existence theorem: it states that the electron density can be used to determine molecular properties; it doesn't say how to do this. In fact, it is usually much easier to extract the values of molecular properties from the wavefunction than from the electron density.

Using the variational principle requires writing the energy as a functional of the electron density. Hohenberg and Kohn demonstrated that the energy functional can be written as a sum of two terms:

$$E[\rho] = F[\rho] + \int \rho(\vec{r}) v^{ext}(\vec{r}) d\vec{r}. \quad (49)$$

The first term is unknown, and is often referred to as the Hohenberg-Kohn density functional. The second term represents the interaction potential that binds the electrons to the system. This is typically the nuclear-electron attraction potential,

$$v^{ext}(\vec{r}) = - \sum_{A=1}^M \frac{Z_A}{|\vec{r} - \vec{R}_A|}. \quad (50)$$

The Hohenberg-Kohn density functional represents the electronic contributions (kinetic energy, electron-electron potential energy) to the energy. It is "universal;" it does not depend on the types or positions of the atomic nuclei, so the same functional can be used for every molecule. In density-functional theory, the problem of approximating $3N$ -dimensional wavefunctions is replaced by the (equally difficult) problem of approximating $F[\rho]$.

Although density-functional theory is exact in principle, there can be no equation for determining the exact electron density that is not at least as difficult as solving the Schrödinger equation directly.⁵² Fortunately there exist methods, most of which are based on the Kohn-Sham formulation of density-functional theory,⁵³ which lends themselves to approximation.⁵⁴

I.XI. The Kohn-Sham Equations

Most modern density-functional theory calculations use the decomposition of $F[\rho]$ that was proposed by Kohn and Sham in 1965.⁵³ They wrote the Hohenberg-Kohn functional as a sum of three terms,

$$F[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho]. \quad (51)$$

$T_s[\rho]$ is the kinetic energy of a system of independent, noninteracting, electrons with same electron density, $\rho(\vec{r})$, as the system of interest. $J[\rho]$ is the classical electrostatic self-repulsion of the electron density,

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\vec{r})\rho(\vec{r}')d\vec{r}d\vec{r}'}{|\vec{r} - \vec{r}'|}. \quad (52)$$

The exchange-correlation energy functional, $E_{xc}[\rho]$, contains everything else. In particular, $E_{xc}[\rho]$ includes the reduction of the electron-electron repulsion potential energy between electrons of the same spin because of the Pauli principle (exchange energy), the reduction in the electron-electron repulsion energy due to the correlated mutually-avoiding motions of the electrons (correlation energy), and the increase in kinetic energy from electron correlation (correlation-kinetic energy). The correlation-potential energy is negative and the correlation-kinetic energy is positive because when electrons swerve to avoid each other, their potential energy of interaction is lowered, but their kinetic energy increases. In the Kohn-Sham-density-functional theory approach, only the exchange-correlation energy functional, $E_{xc}[\rho]$, has to be approximated.

The Kohn-Sham approximation reformulates density-functional theory as an independent electron problem, similar to Hartree-Fock. The derivation of the Kohn-Sham equations starts from the expression for the density of a system with N independent fermions,

$$\rho(\vec{r}) = \sum_{j=1}^N |\phi_j(\vec{r})|^2. \quad (53)$$

The $\phi_j(\vec{r})$ are one-electron spin-orbitals. Inserting the Kohn-Sham expression for $F[\rho]$ into the Hohenberg-Kohn variational principle and taking the functional derivative leads to the Kohn-Sham equations,

$$\left(-\frac{1}{2}\nabla^2 + v^{ext}(\vec{r}) + J[\rho, \vec{r}] + V^{xc}[\rho; \vec{r}] \right) \phi_j(\vec{r}) = \epsilon_j \phi_j(\vec{r}) \quad (54)$$

The Coulomb potential,

$$\begin{aligned} J[\rho, \vec{r}] \phi_j(\vec{r}) &= \left(\int \frac{\rho(\vec{r}')d\vec{r}'}{|\vec{r} - \vec{r}'|} \right) \phi_j(\vec{r}) \\ &= \left(\int \frac{\sum_{i=1}^N |\phi_i(\vec{r}')|^2 d\vec{r}'}{|\vec{r} - \vec{r}'|} \right) \phi_j(\vec{r}), \end{aligned} \quad (55)$$

is almost the same as in the Hartree-Fock method (Eq. (15)) but there is an important difference. In the Hartree-Fock method no orbital feels repulsion from

itself, but this is not true in Eq. (55). The difficulty of correcting this self-interaction-error is a major source of error in density-functional theory.⁵⁵⁻⁵⁷ The exchange-correlation potential, which includes effects of both exchange and correlation, is the functional derivative of the exchange-correlation energy,

$$V^{xc}[\rho(\vec{r});\vec{r}] = \frac{\delta E_{xc}[\rho]}{\delta \rho(\vec{r})} \quad (56)$$

If the exact $E_{xc}[\rho]$ were used in the Kohn-Sham equations, then Kohn-Sham density-functional theory would give the exact ground state energy and density.

Kohn and Sham expressed the exchange-correlation energy in terms of exchange-correlation energy density, per electron, at the point \vec{r} , $\varepsilon_{xc}[\rho, \vec{r}]$:

$$E_{xc}[\rho] = \int \rho(\vec{r}) \varepsilon_{xc}[\rho, \vec{r}] d\vec{r} \quad (57)$$

They simplified this expression by assuming that $\varepsilon_{xc}[\rho, \vec{r}]$ is a *function* of $\rho(\vec{r})$, $\varepsilon_{xc}(\rho(\vec{r}))$ instead of a functional. This is referred to as the local density approximation.⁵³ Finally, they approximated $\varepsilon_{xc}(\rho(\vec{r}))$ with the exchange energy density, per electron, of the uniform electron gas with density $\rho(\vec{r})$. Local density approximations like this one are the simplest possible density functional, but they do not give good results. More modern exchange-correlation functions depend on the gradients of the electron density, or even on the Kohn-Sham orbitals from Eq. (54).⁵⁸⁻⁶⁶

Although Kohn-Sham is similar to Hartree-Fock, it is usually better. This is mostly because the Kohn-Sham approach includes electron correlation. Moreover, it is easier to solve the Kohn-Sham equations than the Hartree-Fock equations because the Coulomb and exchange-(correlation) terms in Kohn-Sham are multiplicative operators, while in Hartree-Fock they are non-local integral operators. Traditional Kohn-Sham density-functional theory methods scale as $O(N^3)$, while Hartree-Fock scales as $O(N^4)$. Because density-functional approximations are usually more accurate than Hartree-Fock, when density-functional theory fails for a particular application one usually attempts to use a post-Hartree Fock method. This is not always possible: density-functional calculations can be performed with systems with thousands of atoms,⁶⁷⁻⁶⁹ while post-Hartree Fock methods are limited to systems with about 100 atoms.

Occasionally, the Hartree-Fock method is more reliable than Kohn-Sham density-functional theory. Even when Hartree-Fock is inaccurate, at least it always provides an upper bound to the ground-state energy. Though the Kohn-Sham method is based on the Hohenberg-Kohn variational principle, it only provides an upper bound to the ground-state energy if the exact exchange-

correlation functional is used. With approximate functionals, the energy is often below the true ground state energy.

Kohn-Sham density functional approximations are the most popular methods in modern computational chemistry: they have relatively low computational cost and they achieve reasonable accuracy for many chemical applications.

I.XII. Conceptual Density-Functional Theory

While the detailed mathematical and computational procedures may be daunting, computing molecular properties is conceptually simple: derive an approximation to the Schrödinger equation, then solve it. Extracting chemical insight from quantum mechanics is arguably more challenging. In particular, the quest to find quantum mechanical approaches to classical chemical concepts like atoms within molecules, electrophilicity, and electronegativity has proved extremely challenging. Chapters II-VII of my thesis represent some (but not all⁷⁰⁻⁷²) of my important contributions to this quest. All of my work in this regard has been focussed on methods that are based on the electron density. Chapters II-V are related to conceptual density-functional theory, and are primarily focussed on chemical reactivity. Chapters VI and VII present my work on some mathematical issues related to the definition of an atom within a molecule.

Although the wavefunction contains all the information required to compute molecular properties, it is not physically interpretable and it is difficult to compute accurately. On the other hand, the electron density contains all of the same information as the wavefunction, but it is experimentally measurable, physically interpretable, much simpler mathematically, and readily approximated computationally. The electron density is also highly transferable. For example, the electron density near an atomic nucleus changes little regardless of the other nuclei that are present.

Conceptual density-functional theory provides qualitative tools for interpreting and predicting chemical phenomena based on the electron density.⁷³⁻⁷⁵ The guiding principles of conceptual density-functional theory are that the descriptors of chemical phenomena should be (1) experimentally measurable observables (ergo, real physical quantities), (2) universal quantities, not quantities that arise as artefacts of a particular computational approach, and (3) mathematically rigorous, with firm theoretical underpinnings.

Consider the effect an approaching reagent, B , has on another molecule, A . The reagent may donate electrons to, or accept electrons from, A . So the number of electrons in A might change (ΔN). In addition, the electrons in molecule A will feel attracted to the nuclei, and repelled by the electrons, of the approaching reagent. So the electrons in A feel an effective external potential from reagent B , $w_B^{eff}(\vec{r})$. The change in the energy of A due to reagent B can then be computed using the functional Taylor series expansion of the energy with respect to the number of electrons N and the external potential $v(\vec{r})$,⁷⁶⁻⁷⁸

$$\begin{aligned}
 E[v_A(\vec{r}) + w_B^{eff}(\vec{r}); N_A + \Delta N] &= E[v_A(\vec{r}); N_A] \\
 &+ \left(\Delta N \cdot \left(\frac{\partial E[v; N]}{\partial N} \right)_{\substack{v(\vec{r})=v_A(\vec{r}) \\ N=N_A}} + \frac{(\Delta N)^2}{2} \cdot \left(\frac{\partial^2 E[v; N]}{\partial N^2} \right)_{\substack{v(\vec{r})=v_A(\vec{r}) \\ N=N_A}} + \dots \right) \\
 &+ \left(\int \left(\frac{\delta E[v; N]}{\delta v(\vec{r})} \right)_{\substack{v(\vec{r})=v_A(\vec{r}) \\ N=N_A}} w_B^{eff}(\mathbf{r}) d\mathbf{r} \right. \\
 &\quad \left. + \frac{1}{2} \iint w_B^{eff}(\vec{r}') \left(\frac{\delta^2 E[v; N]}{\delta v(\vec{r}) \delta v(\vec{r}')} \right)_{\substack{v(\vec{r})=v_A(\vec{r}) \\ N=N_A}} w_B^{eff}(\vec{r}) d\vec{r} d\vec{r}' + \dots \right) \\
 &+ \left(\Delta N \int \left(\frac{\delta E[v; N]}{\delta v(\vec{r}) \delta N} \right)_{\substack{v(\vec{r})=v_A(\vec{r}) \\ N=N_A}} w_B^{eff}(\vec{r}) d\vec{r} + \dots \right)
 \end{aligned} \tag{58}$$

The energetic response of A to this perturbation determines whether the interaction with B is favourable (ergo, possibly reactive) or not.

The coefficients of ΔN and $w_B^{eff}(\vec{r})$ in Eq. (58) are identified as reactivity indicators. Probably the most popular reactivity indicator is the Fukui function, which is listed on the last line of Eq. (58).⁷⁹⁻⁸¹ The Fukui function can be equivalently written as a functional derivative of the electronegativity, χ ,⁸² or as a partial derivative of the electron density,

$$f(\vec{r}) = - \left[\frac{\delta \chi}{\delta v(\vec{r})} \right]_N = \left[\frac{\partial \rho(\vec{r})}{\partial N} \right]_{v(\vec{r})} \tag{59}$$

The Fukui function captures information about the propensity of a molecule to accept/donate electrons at the point \vec{r} . As such, the Fukui function indicates where a nucleophile or an electrophile is most likely to bind to a reagent. The Fukui function is particularly effective for reagents that are very good electron donors or acceptors: if the reagent behaves similarly to an ideal electron donor or acceptor, then the other terms in Eq. (58) are usually negligible.

While the Fukui function is arguably the most popular reactivity indicator of conceptual density-functional theory, there are many others. For example, conceptual density-functional theory provides qualitative indicators for the chemical hardness,^{83,84} electronegativity,⁸² polarisability,⁸⁵ electrophilicity,^{86,87} nucleofugality,^{70,72} etc.

For a more detailed introduction to conceptual density-functional theory, the reader is referred to Chapter II. Chapters III-V introduce and test a new reactivity indicator called the general-purpose indicator.⁸⁸⁻⁹⁰ The general-purpose reactivity indicator was designed to work in cases where the Fukui function gives incorrect predictions for molecules' reactivity. Unlike all previous reactivity indicators, the general-purpose reactivity indicator is derived by starting from the rigorous and exact Eq. (58), and then introducing pragmatic, but controlled, approximations.

I.XIII. Atomic Partitioning

This chapter began with a quote of Paul Dirac, who stated that the pathway to progress in theoretical chemistry is to find approximate methods that “can lead to an explanation of the main features of complex atomic systems without too much computation.” Or, in the words of Charles Coulson, we wish to find “primitive patterns of understanding:” simple models that capture the main features of molecular phenomena.⁹¹ There is no feature of molecular systems so fundamental, or so ubiquitous, as this: there exist within molecules atoms, and these atoms have distinct and assignable properties. This is why the periodic table is central to chemistry. Exploring the features of atoms within molecules is the topic of chapters VI and VII of the thesis.

Resolving molecules into atoms requires partitioning the molecular wavefunction; one can then compute the features of the atoms in a prescribed fashion. The prescription one uses depends on the still controversial definition of an atom in a molecule.^{92,93}

One reason why the periodic table is so important is that the properties of atoms, and certain characteristic “functional groups” of atoms, are highly transferable. A carboxyl group, -COOH, is acidic in any chemical setting. The essential character of the group does not change; the only change is one of degrees. Transferability is the *raison d'être* of atoms in molecules.

Among the many approaches to atoms in molecules, the one that makes the closest contact with the key concept of transferability is the quantum theory of

atoms in molecules (QTAIM).⁹⁴⁻⁹⁷ The AIM theory is based on partitioning the molecular electron density into atomic parts. Atoms are transferable because the atom-in-molecule electron densities are transferable.

In QTAIM, the molecules are divided into atomic regions along zero-flux surfaces of the electron density,

$$\nabla\rho(\vec{r})\cdot\vec{n}(\vec{r})=0 \quad (60)$$

where $\vec{n}(\vec{r})$ is the normal to the atomic surface at the point \vec{r} . The value of a property for an atom is then computed by integrating a local property-density over an atomic region. For example, the kinetic energy of an atom can be computed as the integral of the the kinetic-energy density over an atomic region,

$$T_\alpha = \int_{\Omega_\alpha} t(\vec{r}) d\vec{r}; \quad (61)$$

Eq. (60) is satisfied over the boundary of the atomic region, $\partial\Omega_\alpha$. The value of a physical property for an entire system is obtained by adding together the values of the atomic properties. For example, the total kinetic energy is

$$T = \sum_{\alpha \in \text{molecule}} T_\alpha \quad (62)$$

This feature is the key to transferability, because it allows one to capture empirical property-additivity schemes: similar atoms make similar contributions to molecular properties.

QTAIM is derived from nonrelativistic quantum mechanics by formulating a Lagrangian, then postulating that there exists a stationary action principle for open quantum subsystems. Atoms-within-molecules separated by zero-flux surfaces satisfy this postulate.^{95,98-101}

Chapter VI presents a generalization of QTAIM to relativistic quantum mechanics, in the context of the scalar-relativistic zero-order regular approximation (SR-ZORA).^{102,103} Somewhat surprisingly, the zero-flux condition (Eq. (60)) emerges again, even though the relativistic Lagrangian is different. QTAIM is clearly a very robust theory.

Computing properties whose classical counterparts depend on the momentum may be problematic, however. The Heisenberg uncertainty principle indicates that it is impossible to specify both the momentum and position of a particle at a particular point in space. It is not possible, then, to uniquely define a property density, $p(\vec{r})$, for a momentum-based property, because one cannot say precisely what the momentum is at the point \vec{r} . Chapter VII focuses on the

momentum-based property that is most important to QTAIM: the local kinetic energy density.

Because the quantum mechanical momentum operator is non-local, the kinetic energy can be written in many equivalent forms. Each form defines a different kinetic energy density (the function $t(\vec{r})$ in Eq. (61)). The total molecular kinetic energy is the same for all of these forms, but the atomic kinetic energies in Eq. (61) could give different answers for different definitions of $t(\vec{r})$. One motivation of QTAIM is that the two most common forms for $t(\vec{r})$ give the same atomic kinetic energies for the zero-flux partitioning of atoms (Eq. (60)); this is clearly *not* true for arbitrary methods of partitioning molecules into atoms. Chapter VII goes further, showing that all obvious forms for $t(\vec{r})$, and many not-at-all-obvious forms, also give the same atomic kinetic energies, but only if molecules are partitioned using zero-flux surfaces.

The entire family of kinetic energy densities that are consistent with quantum mechanics was derived by Cohen,^{104,105}

$$t_{f(\vec{\theta}, \vec{\tau})}(\vec{r}) = \sum_{p=1}^{\infty} n_p \left[\begin{array}{l} -\frac{\hbar^2}{8m} (\phi_p^*(\vec{r}) \nabla^2 \phi_p(\vec{r}) - \nabla \phi_p^*(\vec{r}) \cdot \nabla \phi_p(\vec{r})) + c.c. \\ -\left(\frac{1}{2\pi}\right)^3 \iint e^{-i\vec{\theta} \cdot (\vec{r}-\vec{u})} \left(\frac{1}{2m} |\phi_p(\vec{u})|^2\right) \left[\nabla_{\vec{\tau}}^2 f(\vec{\theta}, \vec{\tau}) \right]_{\vec{\tau}=0} d\vec{u} d\vec{\theta} \\ -\left(\frac{1}{2\pi}\right)^3 \iint e^{-i\vec{\theta} \cdot (\vec{r}-\vec{u})} \left[\nabla_{\vec{\tau}} f(\vec{\theta}, \vec{\tau}) \right]_{\vec{\tau}=0} \cdot i \vec{j}_p(\vec{u}) d\vec{u} d\vec{\theta} \end{array} \right] \quad (63)$$

Some of these very kinetic energy densities are very unusual, even if they are mathematically permissible by the postulates of quantum mechanics. Chapter VII explores the possible kinetic energy densities in detail, trying to understand whether there are physical requirements that one may impose to ensure that the atomic kinetic energy is the same for all “physically reasonable” kinetic energy densities.

I.XIV. Summary of Ph.D. Work

Chapter II contains a review of conceptual density-functional theory, emphasising the fundamental mathematical structure of the theory. It illustrates the utility of the perturbative perspective for describing chemical phenomena and discusses the importance of qualitative chemical concepts in theoretical chemistry.

Chapter III contains a detailed derivation of the general-purpose reactivity indicator. This indicator provides information about which location or atom site in a molecule is most reactive. The key idea is to use the vaguest possible description of an incoming reagent. This both simplifies the mathematics and expands the scope of the resulting equations.

Chapters IV and V present applications of the general-purpose reactivity indicator. Chapter IV shows how the general-purpose reactivity indicator can be used to predict the reactivity of a molecule in some cases where frontier molecular orbital theory (and the closely related Fukui function) fails. With most reactivity indicators, one locates the most reactive site in a molecule by finding where that reactivity indicator has its largest value. Such indicators can identify at most one reactive site; they fail for ambidentate molecules: molecules that have two different reactive sites, one which reacts with hard reagents and one which reacts with soft reagents. Chapter V shows that the general-purpose reactivity indicator works for ambidentate molecules.

During my Ph.D. I wrote three other papers on conceptual DFT that are not included in this thesis. One of them explored the mathematical relationships between third-order reactivity indicators; these are the higher-order terms that were not explicitly shown in Eq. (58).⁷¹ These reactivity indicators are even more complicated than the general-purpose reactivity indicator derived in Chapter III. The other two papers were concerned with developing reactivity indicators for predicting the quality of molecular leaving groups.^{70,72}

The quantum theory of atoms in molecules (QTAIM) partitions molecules into atoms using the molecular electron density. As such, QTAIM can be considered to be the natural companion theory to conceptual density-functional theory. Conceptual density-functional theory uses the electron density to elucidate molecules' reactivity. QTAIM uses the electron density to elucidate molecules' electronic structure.

Chapter VI presents a relativistic generalisation of QTAIM based on the scalar-relativistic zero-order regular approximation (SR-ZORA). The zero-flux surfaces dividing molecules into atoms are recovered. This suggests that atoms in molecules can be defined in exactly the same way in both relativistic and nonrelativistic quantum mechanics.

Chapter VII presents a detailed investigation of the full family of possible kinetic energy densities that are consistent with the postulates of quantum mechanics (cf. Eq. (63)). This study has implications for QTAIM and shows, among other things, that the kinetic energy of atoms in molecules is invariant to

the choice of kinetic energy density as long as one restricts oneself to the most reasonable, simple, and explicit formulae for the kinetic energy density.

All of the previously mentioned chapters use density-based or wavefunction-based approaches to quantum mechanics for developing qualitative tools for understanding chemistry. In the following two chapters of my thesis, I explore whether there are better ways to solve the Schrödinger equation. In particular, I explore whether there are ways to solve the Schrödinger equation that achieve the optimal asymptotic computational complexity. Chapter VIII presents one approach, based on truncating the configuration interaction expansion. Chapter IX builds upon my work on many-electron integration grids (not included in the thesis^{31,32}) to develop a new approach based on Boys' collocation method and the scaled Schrödinger equation.

I.XV. References

- 1 P. A. M. Dirac, Proc. R. Soc. London, Ser. A **123**, 714 (1929).
- 2 A. Messiah, *Quantum Mechanics*. (Dover, Toronto, 1999).
- 3 I. N. Levine, *Quantum Chemistry*. (Prentice Hall, Englewood Cliffs, 1999).
- 4 E. Schrödinger, Physical Review **28**, 1049 (1926).
- 5 J. C. Slater, Phys.Rev. **34**, 1293 (1929).
- 6 J. D. Roberts, (W. A. Benjamin, New York, 1962).
- 7 M. H. Stone, *Linear transformations in Hilbert space*. (American Mathematical Society, New York, 1932).
- 8 J. von Neumann, *Mathematical foundations of quantum mechanics*. (Princeton UP, Princeton, New Jersey, U.S.A., 1983).
- 9 C. C. J. Roothaan, Reviews of Modern Physics **23** (2), 69 (1951).
- 10 M. Levy and J. P. Perdew, J. Chem. Phys. **84** (8), 4519 (1986).
- 11 M. Head-Gordon, J. Phys. Chem. **100**, 13213 (1996).
- 12 T. Helgaker, P. Jørgensen, and J. Olsen, *Modern electronic structure theory*. (Wiley, Chichester, 2000).
- 13 T. A. Albright, J. K. Burdett, and M. H. Whangbo, *Orbital Interactions in Chemistry*. (Wiley-Interscience, New York, 1985).
- 14 K. Raghavachari and J. B. Anderson, J. Phys. Chem. **100**, 12960 (1996).
- 15 C. J. Cramer, *Essentials of computational chemistry*. (Wiley, Chichester, 2002).
- 16 J. F. Traub, G. W. Wasilkowski, and H. Wozniakowski, *Information-Based Complexity*. (Academic Press, New York, 1988).
- 17 J. F. Traub and H. Wozniakowski, Bull.AMS **26**, 29 (1992).
- 18 J. F. Traub and A. G. Werschulz, *Complexity and information*. (Cambridge UP, Cambridge, 1998).

- 19 T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- 20 R. T. Pack and W. B. Brown, *J. Chem. Phys.* **45**, 556 (1966).
- 21 S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and T. O. Sorensen, *Communications in Mathematical Physics* **289**, 291 (2009).
- 22 H. Yserentant, *Numerische Mathematik* **98** (4), 731 (2004).
- 23 M. Griebel and S. Knapek, *Constructive Approximation* **16**, 525 (2000).
- 24 M. Griebel and S. Knapek, *Mathematics of Computation* **78**, 2223 (2009).
- 25 M. Griebel, *Computing* **61**, 151 (1998).
- 26 H. J. Bungartz and M. Griebel, *Journal of Complexity* **15**, 167 (1999).
- 27 J. Garcke and M. Griebel, *J. Comput. Phys.* **165**, 694 (2000).
- 28 M. Griebel and J. Hamaekers, *Esaim-Mathematical Modelling and Numerical Analysis-Modelisation Mathematique Et Analyse Numerique* **41** (2), 215 (2007).
- 29 M. Griebel and J. Hamaekers, *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics* **224**, 527 (2010).
- 30 G. Avila and T. Carrington, *J. Chem. Phys.* **131**, 174103 (2009).
- 31 J. I. Rodriguez, D. C. Thompson, J. S. M. Anderson, J. W. Thomson, and P. W. Ayers, *Journal of Physics A* **41**, 365202 (2008).
- 32 J. S. M. Anderson, J. I. Rodriguez, D. C. Thompson, and P. W. Ayers, in *Quantum Chemistry Research Trends* (Nova, Hauppauge, NY, 2007).
- 33 S. A. Smolyak, *Dokl. Akad. Nauk* **4**, 240 (1963).
- 34 Y. Xu, *Common zeros of polynomials in several variables and higher dimensional quadrature*. (Wiley New York, 1994).
- 35 G. W. Wasilkowski and H. Wozniakowski, *Journal of Complexity* **11**, 1 (1995).
- 36 E. Novak, *Advances in Computational Mathematics* **12**, 1 (2000).
- 37 R. Cools, E. Novak, and K. Ritter, *Computing* **62**, 147 (1999).
- 38 E. Novak and K. Ritter, *Constructive Approximation* **15**, 499 (1999).
- 39 E. Novak, *Nonlinear Analysis: Theory Methods & Applications* **30**, 1439 (1997).
- 40 K. Petras, *Numerische Mathematik* **93**, 729 (2003).
- 41 K. Petras, *Advances in Computational Mathematics* **12**, 71 (2000).
- 42 H. Yserentant, *Numerische Mathematik* **105**, 659 (2007).
- 43 H. Yserentant, *Numerische Mathematik* **101** (2), 381 (2005).
- 44 H. Bungartz and M. Griebel, *Acta Numerica* **13**, 147 (2001).
- 45 F. L. Hirshfeld, *Theor. Chim. Act.* **44**, 129 (1977).
- 46 H. Nakatsuji, H. Nakashima, Y. Kurokawa, and A. Ishikawa, *Phys. Rev. Lett.* **99**, 240402 (2007).
- 47 S. F. Boys, *Proc. R. Soc. London, Ser. A* **309**, 195 (1969).
- 48 H. Nakatsuji and H. Nakashima, *Int. J. Quantum Chem.* **109**, 2248 (2009).
- 49 W. Kohn, *Reviews of Modern Physics* **71**, 1253 (1999).

- 50 P. Hohenberg and W. Kohn, *Phys.Rev.* **136**, B864 (1964).
51 M. Levy, *Proc. Natl. Acad. Sci.* **76**, 6062 (1979).
52 N. Schuch and F. Verstraete, *Nature Physics* **5**, 732 (2009).
53 W. Kohn and L. J. Sham, *Phys.Rev.* **140**, A1133 (1965).
54 W. Kohn, A. D. Becke, and R. G. Parr, *J.Phys.Chem.* **100**, 12974 (1996).
55 J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
56 P. Mori-Sanchez, A. J. Cohen, and W. T. Yang, *J. Chem. Phys.* **125**, 201102 (2006).
57 A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, *J. Chem. Phys.* **126**, 104102 (2007).
58 A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
59 A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
60 C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
61 J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
62 J. M. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
63 A. D. Becke, *J. Chem. Phys.* **122** (6), 064101 (2005).
64 A. D. Becke, *J. Chem. Phys.* **119** (6), 2972 (2003).
65 P. Mori-Sanchez, A. J. Cohen, and W. Yang, *J. Chem. Phys.* **124**, 091102 (2006).
66 J. P. Perdew, A. Ruzsinszky, J. M. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, *J. Chem. Phys.* **123**, 062201 (2005).
67 W. Yang, J. M. Perez-Jorda, and P. v. R. Schleyer, in *Encyclopedia of Computational Chemistry* (Wiley, New York, 1998).
68 F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, *Comput. Phys. Commun.* **140**, 303 (2001).
69 A. Nakano, R. K. Kalia, K. Nomura, A. Sharma, P. Vashishta, F. Shimojo, A. C. T. van Duin, W. A. Goddard, R. Biswas, and D. Srivastava, *Computational Materials Science* **38**, 642 (2007).
70 J. S. M. Anderson, Y. L. Liu, J. W. Thomson, and P. W. Ayers, *Journal of Molecular Structure-Theochem* **943**, 168 (2010).
71 C. Cardenas, E. Echeagaray, D. Chakraborty, J. S. M. Anderson, and P. W. Ayers, *J. Chem. Phys.* **130**, 244105 (2009).
72 P. W. Ayers, J. S. M. Anderson, J. I. Rodriguez, and Z. Jawed, *PCCP* **7**, 1918 (2005).
73 R. G. Parr and W. T. Yang, *Annu. Rev. Phys. Chem.* **46**, 701 (1995).
74 R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. (Oxford UP, New York, 1989).
75 P. Geerlings, F. De Proft, and W. Langenaeker, *Chem. Rev.* **103**, 1793 (2003).
76 P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.* **123**, 2007 (2001).

- 77 P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.* **122**, 2010 (2000).
78 P. W. Ayers, J. S. M. Anderson, and L. J. Bartolotti, *Int. J. Quantum Chem.* **101**, 520 (2005).
79 R. G. Parr and W. T. Yang, *J. Am. Chem. Soc.* **106**, 4049 (1984).
80 W. T. Yang, R. G. Parr, and R. Pucci, *J. Chem. Phys.* **81**, 2862 (1984).
81 P. W. Ayers and M. Levy, *Theor. Chem. Acc.* **103**, 353 (2000).
82 R. G. Parr, R. A. Donnelly, M. Levy, and W. E. Palke, *J. Chem. Phys.* **68**, 3801 (1978).
83 R. G. Parr and R. G. Pearson, *J. Am. Chem. Soc.* **105**, 7512 (1983).
84 P. W. Ayers, *Faraday Discuss.* **135**, 161 (2007).
85 Y. Simon-Manso and P. Fuentealba, *J. Phys. Chem. A* **102**, 2029 (1998).
86 R. G. Parr, L. von Szentpály, and S. B. Liu, *J. Am. Chem. Soc.* **121**, 1922 (1999).
87 P. K. Chattaraj, U. Sarkar, and D. R. Roy, *Chem. Rev.* **106**, 2065 (2006).
88 J. S. M. Anderson, J. Melin, and P. W. Ayers, *Journal of Chemical Theory and Computation* **3**, 358 (2007).
89 J. S. M. Anderson and P. W. Ayers, *PCCP* **9**, 2371 (2007).
90 J. S. M. Anderson, J. Melin, and P. W. Ayers, *Journal of Chemical Theory and Computation* **3**, 375 (2007).
91 C. A. Coulson, *Reviews of Modern Physics* **32**, 170 (1960).
92 R. G. Parr, P. W. Ayers, and R. F. Nalewajski, *J. Phys. Chem. A* **109**, 3957 (2005).
93 C. F. Matta and R. F. W. Bader, *J. Phys. Chem. A* **110**, 6365 (2006).
94 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*. (Clarendon, Oxford, 1990).
95 R. F. W. Bader and T. T. Nguyendang, *Adv. Quantum Chem.* **14**, 63 (1981).
96 R. F. W. Bader and P. M. Beddall, *J. Chem. Phys.* **56** (7), 3320 (1972).
97 R. F. W. Bader, Y. Tal, S. G. Anderson, and T. T. Nguyen-Dang, *Isr.J.Chem.* **19**, 8 (1980).
98 S. Srebrenik, R. F. W. Bader, and T. T. Nguyendang, *J. Chem. Phys.* **68** (8), 3667 (1978).
99 R. F. W. Bader, S. Srebrenik, and T. T. Nguyendang, *J. Chem. Phys.* **68** (8), 3680 (1978).
100 R. F. W. Bader, *Phys. Rev. B* **49**, 13348 (1994).
101 P. F. Zou and R. F. W. Bader, *Int. J. Quantum Chem.* **43**, 677 (1992).
102 C. Chang, M. Pelissier, and P. Durand, *Phys. Scr.* **34** (5), 394 (1986).
103 J. L. Heully, I. Lindgren, E. Lindroth, S. Lundqvist, and A. M. Martenssonpendrill, *J. Phys. B* **19**, 2799 (1986).
104 L. Cohen, *J. Chem. Phys.* **80**, 4277 (1984).
105 L. Cohen, *J. Chem. Phys.* **70**, 788 (1979).

Chapter II

Perturbative Perspectives on the Chemical Reaction Prediction Problem*

* The content of this chapter has been published: P. W. Ayers, **J. S. M. Anderson**, L. J. Bartolotti
“Perturbative Perspectives on the Chemical Reaction Prediction Problem”; *Int. J. Quantum
Chem.* **2005**, *101*, 520-534.

II.I. Statement of the Problem

This chapter is a self-contained review of conceptual density-functional theory (DFT). Unlike previous reviews, this work presents conceptual density-functional theory from a “perturbative perspective”. The perturbative perspective allows one to treat indices associated with conceptual DFT in a unified way. Largely because of this review, the perturbative perspective has become the most common approach to conceptual DFT. Emphasis is placed on the implications of the perturbative perspective for describing regioselectivity and certain global properties of molecules, specifically their electrophilicity, nucleofugality, and electrofugality. This chapter will provide the necessary background for chapters II, IV and V.

II.II. Introduction

Since the dawn of quantum mechanics, there has been an interest in how the mathematical formalism and physical concepts of quantum mechanics can be used to describe chemical phenomena. For example, one may cite the Lewis-dot structures (which follows from old quantum mechanics),¹ orbital hybridization and resonance,² and a plethora of results derived from molecular orbital theory.³ These tools and the associated concepts (among them the octet rule, promotion energy, ligand field splitting,⁴ electronegativity, the Woodward-Hoffmann rules,⁵⁻⁹ and HOMO/LUMO based descriptors for electrophilic/nucleophilic attack¹⁰⁻¹⁵) now pervade the language of chemistry. Unfortunately, the most useful models for qualitative studies are often very approximate in a quantitative sense. Thus, while valence-bond theory is in principle exact, the most useful qualitative descriptions use only a few resonance structures, while thousands of resonance structures are required for quantitative accuracy. Similarly, qualitative applications of molecular-orbital theory are usually based on single-determinant wave functions, so the simplest MO-based arguments are suspect when a system’s wave function has significant multi-determinantal character. One can achieve quantitative accuracy by including configuration interaction, but thousands of determinants are needed to achieve quantitative agreement with experimental results. One wishes to have a method that, while exact in principle, is also qualitatively useful—that is, *a method that combines quantitative accuracy and qualitative utility*. One possibility is the use of the Dyson orbitals, which are deduced from accurate wave functions and, due to their direct relationship to electron loss and gain, are conceptually relevant.¹⁶⁻¹⁹ The use of Dyson orbitals has yet to achieve popular acceptance, however, perhaps because the idea of non-

orthogonal orbitals is anathema to chemists familiar with molecular-orbital methods based on single determinants.

One may also use density-functional theory.^{20;21} Not only is density-functional theory exact in principle,²²⁻²⁵ density-functional theory lends itself to qualitative chemical descriptors.^{20;26-28} Among the successes are the chemical potential²⁹ (counterpart to the electronegativity) and the Fukui functions³⁰⁻³² (counterpart to the frontier orbitals), as well as other concepts that had not been successfully defined in a valence-bond or molecular-orbital theory context (e.g., the chemical hardness³³ and electrophilicity³⁴). This paper is concerned with the physical interpretation of these indices and, in particular, adopts what may be termed a perturbative perspective: each reactivity index in density-functional theory may be viewed as representing the response of a system to a certain “model” perturbation, where the “model” is chosen to encapsulate the essence of the attacking reagent. In section III the chemical motivation behind this perspective is sketched; section IV presents the requisite background information, and two theoretical applications (based on two specific choices for the model perturbation) are explored in sections V and VI.

II.III. Motivation

The perturbative perspective is motivated by the presence of systematic trends in chemistry. For example, one can organize acids and bases in order of their strength and infer, for example, that in almost any reaction environment, a Brønsted-Lowry base will accept a proton from a protonated weaker base, and can then be induced to forfeit the proton if a still-stronger base is added to the mixture. As a specific example, one finds that protonated pyrrole ($pK_a = 0.4$) will donate a proton to pyridine ($pK_a = 5.25$), which will donate a proton to ammonia ($pK_a = 9.26$). One can, from this sequence of events, infer that the pyrrolium ion is a stronger acid than the ammonium ion and so, if one adds ammonia to a solution with protonated pyrrole, proton transfer from pyrrole to ammonia will occur.

Many chemical syntheses can, in fact, be regarded as a sequence of Lewis acid/base reactions: over the course of a reaction or sequence of reactions, electrons move from good donors (strong Lewis bases) to good acceptors (strong Lewis acids). One is led to the conclusion that a reagent’s inherent acidity/basicity can be expressed in terms of the properties of the isolated molecule, for otherwise we would expect solvent effects and the identity of the chemical reaction partner to exert significant—rather than subsidiary—effects. Moreover, the fact that a molecule reacts in a qualitatively similar ways to any

stronger base suggests that there is something inherently similar about how a strong base interacts with—or perturbs—a weak base, and vice versa. That is, the fact that a specific molecule will interact in similar ways with a wide range of reagents that are, on the surface, dissimilar suggests that the essence of the interaction might be captured by a simple “model perturbation,” which captures the essential similarity of the range of reagents of interest.

Despite what at first glance appears to be intimidating complexity, underlying simplicity is also found in studies of regioselectivity. For example, nitrogen bases almost always interact with acids through the nitrogen atom, and boron acids almost always interact with bases through the boron atom. Similarly, there are a large number of ortho-meta-para rules for electrophilic (and, much less frequently, nucleophilic) attack on aromatic compounds. The underlying principle is that a reagent tends to be susceptible to nucleophilic attack at one or, at most, a few sites and similarly tends to be susceptible to electrophilic attack at one or possibly a few sites. These sites are largely independent of the nature of the electrophile/nucleophile, the presence of a solvent, or the particular reaction conditions. That is, regioselectivity preferences are properties of the isolated molecule and, presumably, there is some essential essence that is common to nucleophilic and electrophilic reagents and which, viewed as a perturbation, can describe the susceptibility of a reagent to nucleophilic and electrophilic attack.

The mathematical ideas behind the perturbative perspective are already contained in the early work of Nalewajski and Parr,³⁵ with subsequent significant elaboration by Liu and Parr,³⁶⁻³⁸ Senet,^{39;40} and others. At least with respect to density-functional theory, the specific perspective adopted in the previous paragraphs seems to have been first employed by Parr, von Szentpaly, and Liu,⁴¹ who employed the idea of a “perfect nucleophile” as a model system. Subsequent developments have stressed the use of the concept in regioselectivity.⁴²⁻⁴⁵

II.IV. **Theoretical Development**

As explained in the previous section, the essence of the perturbative perspective on chemical reactivity is to describe the effect of an attacking reagent, B , on a given molecule, A . There are two possible effects. First of all, the presence of B changes the external potential felt by the electrons in A —that is, electrons in A are now subject not only to the external potential due to the nuclei in A , $v_A(\mathbf{r})$, but also to an external potential due to the electrons and nuclei in B , $w_B^{eff}(\mathbf{r})$. Various approximations for $w_B^{eff}(\mathbf{r})$ have been considered.⁴⁶ The

simplest approximation for $w_B^{eff}(\mathbf{r})$ is just the additive inverse of the electrostatic potential from the nuclei and electrons of B ,⁴⁷

$$w_B^{eff}(\mathbf{r}) \approx v_B(\mathbf{r}) + \int \frac{\rho_B(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \quad (1)$$

where $\rho_B(\mathbf{r})$ denotes the electron density of B . Use of Eq. (1) neglects exchange and correlation effects that couple electrons in A and B , but this can be corrected for. A method for finding the exact effective external potential from B has been proposed.⁴⁸

Second, electrons may flow from A to B , or vice versa. Thus, the number of electrons in A may change from the number in the isolated molecule, N_A , to $N_A + \Delta N$. Regarding these changes as perturbations on the isolated molecule A , we employ a functional Taylor series to model the changes, thus

$$\begin{aligned} E[v_A(\mathbf{r}) + w_B^{eff}(\mathbf{r}); N_A + \Delta N] = \\ E[v_A(\mathbf{r}); N_A] + \Delta N \cdot \left(\frac{\partial E[v; N]}{\partial N} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} + \frac{(\Delta N)^2}{2} \cdot \left(\frac{\partial^2 E[v; N]}{\partial N^2} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} + \dots \\ + \int \left(\frac{\delta E[v; N]}{\delta v(\mathbf{r})} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} w_B^{eff}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint w_B^{eff}(\mathbf{r}') \left(\frac{\delta^2 E[v; N]}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} w_B^{eff}(\mathbf{r}) d\mathbf{r} d\mathbf{r}' + \dots \\ + \Delta N \int \left(\frac{\delta \partial E[v; N]}{\delta v(\mathbf{r}) \partial N} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} w_B^{eff}(\mathbf{r}) d\mathbf{r} + \dots \end{aligned} \quad (2)$$

The first term in the Taylor series is identified as the electronic chemical potential of A , μ_A , which is related to the electronegativity, χ_A , by²⁹

$$-\chi_A = \mu_A \equiv \left(\frac{\partial E[v; N]}{\partial N} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} \quad (3)$$

Equation (3) is the first, and quite possibly the most important, result in conceptual density-functional theory. Many important results, most notably with regard to electronegativity equalization,⁴⁹⁻⁵¹ immediately follow.

The second term in the Taylor expansion was identified with the chemical hardness,

$$\eta_A \equiv \left(\frac{\partial^2 E[v; N]}{\partial N^2} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} \quad (4)$$

by Parr and Pearson.³³ (The present definition differs from the initial definition by a factor of 2, but is usually considered preferable and has been adopted by many researchers, including Parr and Pearson, in subsequent work.) Important

related results include the maximum hardness principle^{45;52-56} and the theory of hard and soft acids and bases^{54;57-59} (HSAB theory).

The next two terms address the effect of a change in external potential in the absence of any electron transfer. This is often a good approximation when the product molecule, AB , is essentially ionic in character, with little charge transfer between the substituents. From the Hellmann-Feynman theorem,^{60;61} the first term is seen to be just the electron density,

$$\rho_A(\mathbf{r}) \equiv \left(\frac{\delta E[v; N]}{\delta v(\mathbf{r})} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} = \left\langle \Psi_0^A \left| \sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) \right| \Psi_0^A \right\rangle \quad (5)$$

while the second is the linear response of the electron density to a change in external potential and is often termed the polarizability kernel,

$$\begin{aligned} P_A(\mathbf{r}, \mathbf{r}') &\equiv \left(\frac{\delta^2 E[v; N]}{\delta v(\mathbf{r}') \delta v(\mathbf{r})} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} = \left(\frac{\delta \rho_A(\mathbf{r})}{\delta v(\mathbf{r}') } \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} \\ &= \sum_{i=1}^N \frac{\left\langle \Psi_0^A \left| \sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) \right| \Psi_i^A \right\rangle \left\langle \Psi_i^A \left| \sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}') \right| \Psi_0^A \right\rangle}{E_0^A - E_i^A} + c.c. \end{aligned} \quad (6)$$

where Ψ_i^A and E_i^A are the wave function and energy of the i^{th} excited state of A . The final term in the Taylor series, (2), represents the coupling between changes in the number of electrons and changes in external potential; it is termed the Fukui function,^{30;31;62}

$$f_A(\mathbf{r}) \equiv \left(\frac{\delta E[v; N]}{\delta v(\mathbf{r}) \delta N} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} = \left(\frac{\partial \rho[v; N, \mathbf{r}]}{\partial N} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} = \left(\frac{\delta \mu[v; N]}{\delta v(\mathbf{r})} \right)_{\substack{v(\mathbf{r})=v_A(\mathbf{r}) \\ N=N_A}} \quad (7)$$

Many other indices have been proposed in the context of conceptual DFT, but most important of them are simply related to the preceding quantities. (The exceptions are related to higher-order derivatives in Eq. (2).)

One commonly truncates the perturbation series at second order, yielding

$$\begin{aligned} &E[v_A(\mathbf{r}) + w_B^{\text{eff}}(\mathbf{r}); N_A + \Delta N] \\ &= E[v_A(\mathbf{r}); N_A] + \Delta N \cdot \mu_A + \frac{(\Delta N)^2}{2} \cdot \eta_A + \int \rho_A(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} \\ &+ \Delta N \int f_A(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint w_B^{\text{eff}}(\mathbf{r}') P_A(\mathbf{r}, \mathbf{r}') w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} d\mathbf{r}'. \end{aligned} \quad (8)$$

This is a good approximation when the perturbations are small and presupposes that the course of a chemical reaction is determined when the reacting molecules

are still far apart—that is, the transition state lies early along the reaction path. In practice, it seems that this approximation is much more robust. (This is related to the empirical success of hill-climbing methods for predicting the products of chemical reactions; one can often locate the transition state of a chemical reaction by proceeding, step by step, from the reactant configuration in the least-steep direction (lowest frequency normal mode).) It should be emphasized that the second-order truncation is not mandatory: for example, Taylor-series with remainder forms (see the appendix) can be used to obtain higher accuracy. Alternatively, one can include third and higher order terms, in which case issues related to the convergence of the Taylor expansion become increasingly important. The convergence of the Taylor series is considered in the appendix.

In Eq. (2), derivatives with respect to the number of electrons are used. This implies that one can define what it means for a system to have a non-integer number of electrons. Several approaches have been proposed, including using the zero-temperature grand canonical ensemble^{63;64} and Fock space constructions.⁶⁵ An alternative to these is to make P duplicates of one's system and arrange these on the surface of a sphere of radius R . Next, add $M < P$ electrons to the system, and then take the limit as R approaches infinity. Using the most symmetric wave function for the system, each subsystem has $N_A + \frac{M}{P}$ electrons. If, then, we use size consistency arguments, we deduce that the properties of a system with $N_A + \frac{M}{P}$ electrons are given by the appropriate linear combination between the N_A - and $N_A + 1$ -electron systems:⁶⁶

$$Q[v_A(\mathbf{r}); N_A + \frac{M}{P}] = (1 - \frac{M}{P})Q[v_A(\mathbf{r}); N_A] + \frac{M}{P}Q[v_A(\mathbf{r}); N_A + 1]. \quad (9)$$

The zero-temperature grand canonical ensemble and Fock-space formulations give the same results.

Note that $\frac{\partial Q}{\partial N}|_{N=N_A}$ does not exist. Rather, there are different derivatives for “number increasing” and “number decreasing” variations. What is more, the linearity of Eq. (9) means that all terms in the Taylor series with a quadratic or higher-order dependence on ΔN are identically zero. Thus, the truncated Taylor series, Eq. (8), can be rewritten as

$$\begin{aligned} & E[v_A(\mathbf{r}) + w_B^{\text{eff}}(\mathbf{r}); N_A + \Delta N] \\ &= E[v_A(\mathbf{r}); N_A] + \Delta N \cdot \mu_A^+ + \int \rho_A(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} \\ & \quad + \Delta N \int f_A^{\pm}(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint w_B^{\text{eff}}(\mathbf{r}') P_A(\mathbf{r}, \mathbf{r}') w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} d\mathbf{r}'. \end{aligned} \quad (10)$$

Here a superscript “+” indicates differentiation from above and a superscript “-” indicates differentiation from below. From Eq. (9), the Fukui functions are given by

$$f_A^\pm(\mathbf{r}) = \pm(\rho_{N_A \pm 1}(\mathbf{r}) - \rho_{N_A}(\mathbf{r})) \quad (11)$$

and the chemical potentials from above and below are related to the electron affinity, A , and the ionization potential, I , by

$$\begin{aligned} \mu_A^+ &\equiv -A_A \\ \mu_A^- &\equiv -I_A. \end{aligned} \quad (12)$$

In solution, a molecule occurs in many different solvation environments, and so the “isolated molecule” construction for fractional numbers of electrons is suspect. In practice, it is found that a quadratic interpolation, with

$$\begin{aligned} &\frac{\partial Q[v, N]}{\partial N} \Big|_{\substack{v=v_A \\ N=N_0}} \\ &\equiv \frac{N_0 - (N_A - 1)}{2} (Q[v_A, N_A + 1] - Q[v_A; N_A]) + \frac{N_A + 1 - N_0}{2} (Q[v_A, N_A] - Q[v_A; N_A - 1]) \end{aligned} \quad (13)$$

is often effective. This gives the well-known Mulliken electronegativity,⁶⁷

$$\chi_A^{(\text{Mulliken})} = -\mu_A^{(\text{quadratic})} = \frac{I_A + A_A}{2} \quad (14)$$

and the Parr-Pearson definition of the hardness,³³

$$\eta_A^{(PP)} \equiv I_A - A_A. \quad (15)$$

The success of the quadratic model is difficult to rationalize. The use of “smooth curves” rather than curves with discontinuous derivatives may be rationalized since molecules in solutions occur in many different solvation environments, and thus exist in many different states. The average state of a molecule in solution, then, is a weighting of many states—not just the anion, cation, and neutral states of the isolated molecule, but many “perturbed” variants thereof. To be sure, the zero-temperature ensemble average fails to capture the richness of this phenomenon, but why the primitive quadratic model (which contains even less “physics” than Eq. (13)) captures this richness is not at all understood. The quadratic model is practically useful and mathematically simple, however; an example of its utility is given in section V.

With this background we can restate the “perturbative perspective” in a mathematical form. Note that each of the “coefficients” in the Taylor series (Eq. (8) or Eq. (10)) is a property of the isolated molecule; the only effect of changing the attacking reagent, B , is to change the appropriate value of ΔN and $w_B^{\text{eff}}(\mathbf{r})$. One might suppose that chemically similar reagents contribute similar values for ΔN (for example, ΔN is always positive for nucleophilic attack and negative for electrophilic attack) and $w_B^{\text{eff}}(\mathbf{r})$ (for example, nucleophiles tend to be negatively charged and electrophiles tend to be positively charged). Thus, a molecule’s reactivity with respect to a certain class of reagents—throughout which the main features of ΔN and $w_B^{\text{eff}}(\mathbf{r})$ are conserved—can be expressed in terms of the properties of the isolated molecule and, in particular, the coefficients of terms in

the Taylor series expansion of the energy. The coefficients of terms in the Taylor series, then, are readily identified as reactivity indicators.

One can discern the particular relationship between the reactivity of a molecule, A , and its Taylor-series coefficients by introducing a model perturbation. The essential idea is that since it has been empirically observed that a molecule reacts with chemically similar reagents in similar ways, one need not consider the detailed values of ΔN and $w_B^{eff}(\mathbf{r})$ for each reagent. In particular, the values of ΔN and $w_B^{eff}(\mathbf{r})$ for a simple “model reagent” will serve to indicate the reactivity of other chemically similar reagents. That is, a molecule’s reactivity can be revealed by measuring the response of a molecule to a model perturbation—a simple choice for ΔN and $w_B^{eff}(\mathbf{r})$ that nonetheless encapsulates the essential chemical information about the attacking reagent.

II.V. Regioselectivity

As a first example of the utility of the “perturbative perspective,” we consider the problem of predicting where a molecule undergoes electrophilic or nucleophilic attack. To obtain a simple model perturbation, consider that the “active site” of a nucleophile is generally negative, while the active site of an electrophile is usually positive. As a simple model, then, we can approximate the change in external potential due to electrophilic (nucleophilic) attack by a positive (negative) point charge. Substituting this into the Taylor series (Eq. (10)),

$$\begin{aligned} & E\left[v_A(\mathbf{r}) - \frac{q}{|\mathbf{r}-\mathbf{x}|}; N_A + \Delta N\right] - E[v_A(\mathbf{r}); N_A] \\ &= \Delta N \cdot \mu_A^\pm - q \int \frac{\rho_A(\mathbf{r})}{|\mathbf{r}-\mathbf{x}|} d\mathbf{r} \\ & \quad - q\Delta N \int \frac{f_A^\pm(\mathbf{r})}{|\mathbf{r}-\mathbf{x}|} d\mathbf{r} + \frac{q^2}{2} \iint \frac{P_A(\mathbf{r}, \mathbf{r}')}{|\mathbf{r}-\mathbf{x}||\mathbf{r}'-\mathbf{x}|} d\mathbf{r}d\mathbf{r}'. \end{aligned} \quad (16)$$

Equation (16) models the change in electronic energy; the change in total energy requires that one include the attraction/repulsion of the molecular nuclei for the point charge, thus

$$\begin{aligned} Y_A[q, \Delta N, \mathbf{x}] &\equiv E\left[v_A(\mathbf{r}) + \frac{q}{|\mathbf{r}-\mathbf{x}|}; N_A + \Delta N\right] - E[v_A(\mathbf{r}); N_A] + \sum_\alpha \frac{Z_\alpha q}{|\mathbf{R}_\alpha - \mathbf{x}|} \\ &\equiv \Delta N \cdot \mu_A^\pm + q\Phi_A(\mathbf{x}) - q\Delta N \int \frac{f_A^\pm(\mathbf{r})}{|\mathbf{r}-\mathbf{x}|} d\mathbf{r} + \frac{q^2}{2} \iint \frac{P_A(\mathbf{r}, \mathbf{r}')}{|\mathbf{r}-\mathbf{x}||\mathbf{r}'-\mathbf{x}|} d\mathbf{r}d\mathbf{r}'. \end{aligned} \quad (17)$$

where \mathbf{R}_α and Z_α denotes the position and atomic number of the α^{th} nucleus in molecule A . Here,

$$\Phi_A(\mathbf{x}) \equiv \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{R}_{\alpha} - \mathbf{x}|} - \int \frac{\rho_A(\mathbf{r})}{|\mathbf{r} - \mathbf{x}|} d\mathbf{x} \quad (18)$$

denotes the electrostatic potential. The quantities $Y_A(1,0,\mathbf{x})$ and $Y_A(-1,0,\mathbf{x})$ have been used as indicators of the Brönsted-Lowry acidity and basicity,⁶⁸ and our notation (Y is the Greek “Upsilon,” and not the Roman letter Y) is designed to reflect this relationship.

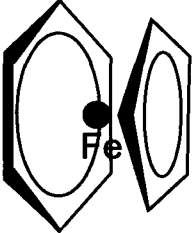
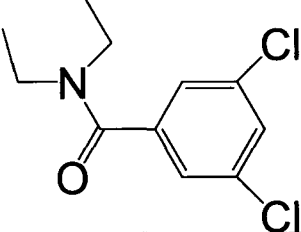
From Eq. (17), it is clear that the electrostatic potential can be used to describe regioselectivity⁶⁹ when the change in the number of electrons is small, so that $\Delta N \approx 0$. This amount of charge transfer can be approximated using the chemical potentials and hardnesses of the molecules, giving³³

$$|\Delta N| = \frac{|\mu_A - \mu_B|}{\eta_A + \eta_B} \quad (19)$$

Thus, for hard reagents we expect the effects of charge transfer to be relatively unimportant (the bond between the molecule and the attacking reagent tends to have a large amount of ionic character). The electrostatic potential, Eq. (18), is useful for describing such reactions, which are often said to be “charge controlled” or “electrostatically controlled.”⁷⁰

As an example of a charge-controlled reaction, we consider molecule **a** (see Table 1), which is a ferrocene molecule in which one of the cyclopentadienyl rings is replaced by benzene. Recalling that reactivity indicators are only rigorously valid when the conformational preference is apparent early on the reaction path, we plot the electrostatic potential on the van der Waals surface of the molecule in Figure 1.⁷¹ Electrophilic attack is clearly indicated on the cyclopentadienyl ring. This correct result is consistent with intuition: the cyclopentadiene ring is negatively charged because the metal center tends to donate an electron to the cyclopentadiene ring since the cyclopentadiene anion has $4(1)+2$ electrons and is thus aromatic.⁷² Donation of an electron from the cyclopentadiene to an attacking electrophile would result in the loss of aromaticity and consequent destabilization, so relatively little electron transfer occurs and the reaction is “charge controlled.” From an alternative perspective, the aromaticity of the cyclopentadienyl ring corresponds to a large hardness,^{73;74} which reduces the amount of electron transfer.

Table 1. Structures and orientations of the molecules in figures 1-4.

Structure Label	Structure
A	
B	
C	$\begin{array}{c} \text{HO} \quad \text{H}_2 \\ \quad \\ \text{H}_2\text{C} - \text{C} \\ \quad \\ \quad \quad \text{NH}_2 \end{array}$
D	NCS^-

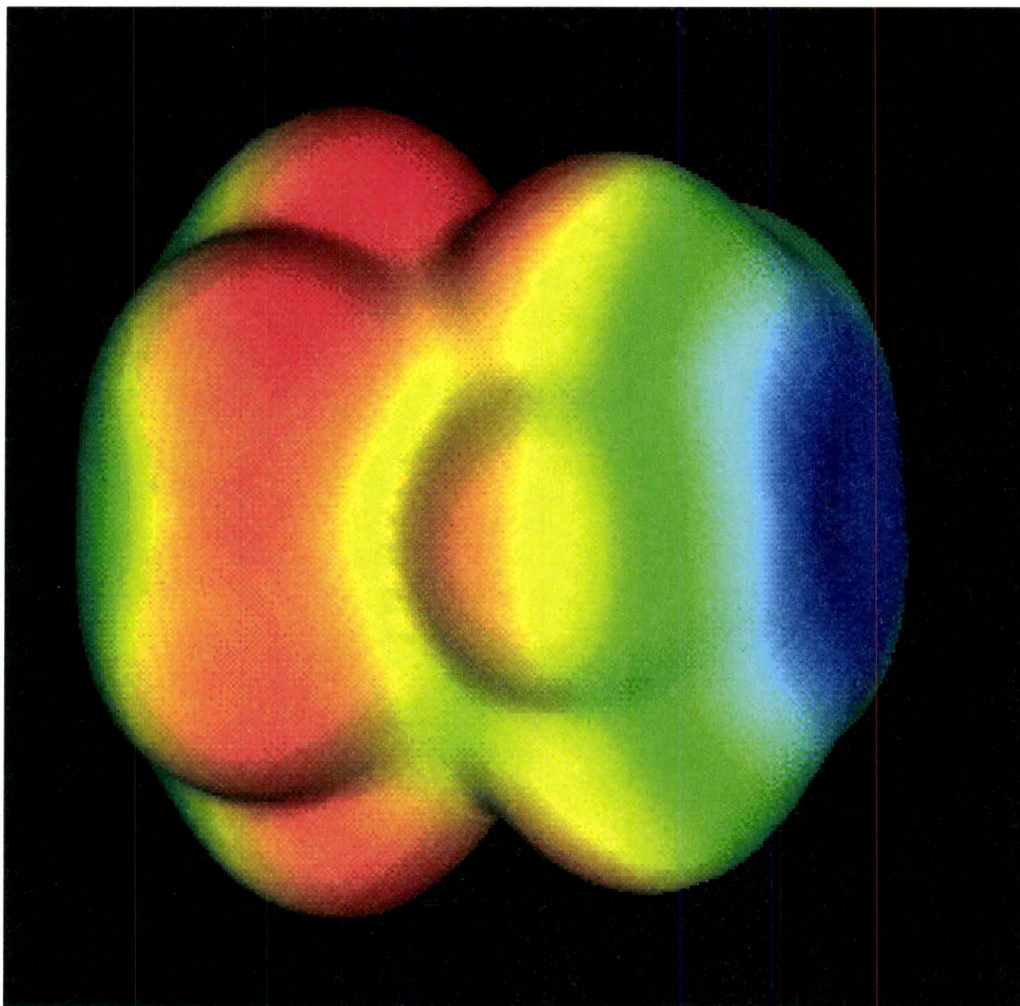


Figure 1. The electrostatic potential of **a**, plotted on the van der Waals surface of the molecule. Electrophilic attack occurs on the cyclopentadienyl ring.

When the charge of a reagent is small and the molecules being considered are soft, one expects that the dominant term in Eq. (17) is associated with the Fukui function and, more specifically, the Fukui potential,⁷⁵

$$v_f^\pm(\mathbf{r}) \equiv \int \frac{f_A^\pm(\mathbf{r})}{|\mathbf{r} - \mathbf{x}|} d\mathbf{r}. \quad (20)$$

The Fukui potential tends to be largest where the Fukui function is largest, and so the Fukui function provides a good indication of the susceptibility of a molecule to nucleophilic ($f^+(\mathbf{r})$) and electrophilic ($f^-(\mathbf{r})$) attack.^{30;76} (The appropriate choice of Fukui function is clear from the sign of ΔN ; $\Delta N > 0$ for nucleophilic attack and $\Delta N < 0$ for electrophilic attack.) The Fukui function is the density-functional analogue to the frontier molecular orbitals,³¹ and so when the dominant factors in determining regioselectivity are associated with the Fukui function a reaction is said to be “frontier controlled.”⁷⁰

As an example of a situation where the Fukui function is an appropriate measure, consider the lithiation of 3,5-dichloro-*N,N*-diethyl carboamide (entry **b** in Table 1).⁷⁷ The lithium atom is nominally neutral (though it is expected to have a slight positive charge) and is a soft reagent. We expect, then, that the lithiation will be accompanied by substantial charge transfer to **b** and that lithiation will occur at the site that is most able to accommodate additional electron density; in short, lithiation is expected to occur where $f^+(\mathbf{r})$ is the largest. Figure 2 plots $f^+(\mathbf{r})$ for **b** on the van der Waals surface of the molecule. The observed reactivity preference for the para position is somewhat surprising because carboxamides are powerful ortho directors. Lithiation at the para position, however, is experimentally observed. The present study reveals that the reactivity of **b** is in fact governed by electronic effects, rather than steric,⁷⁷ effects.

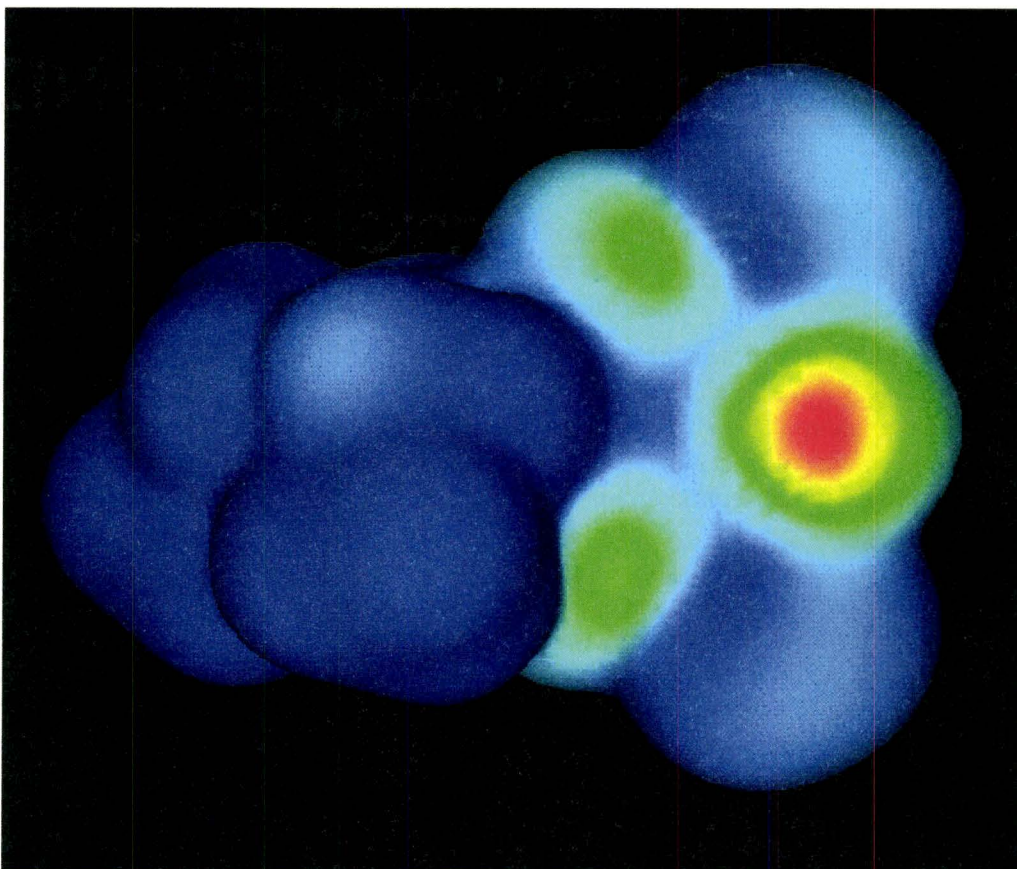
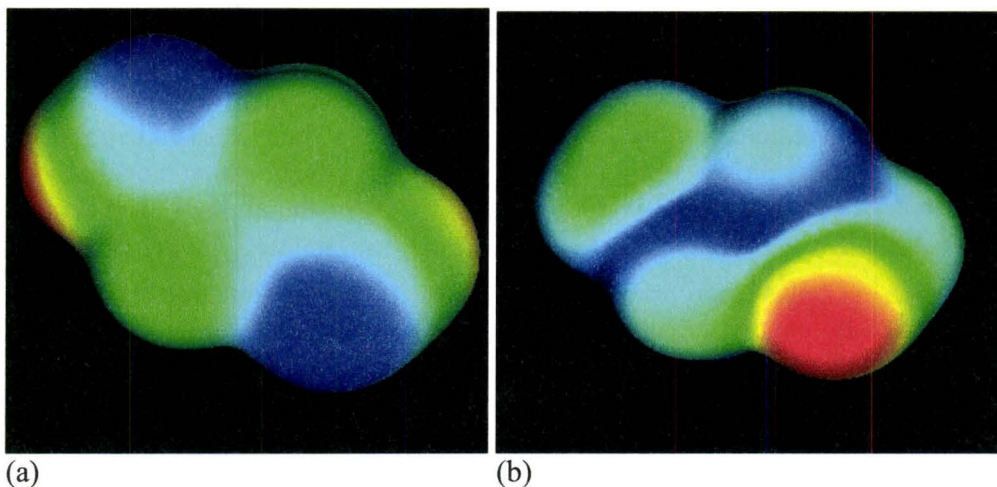


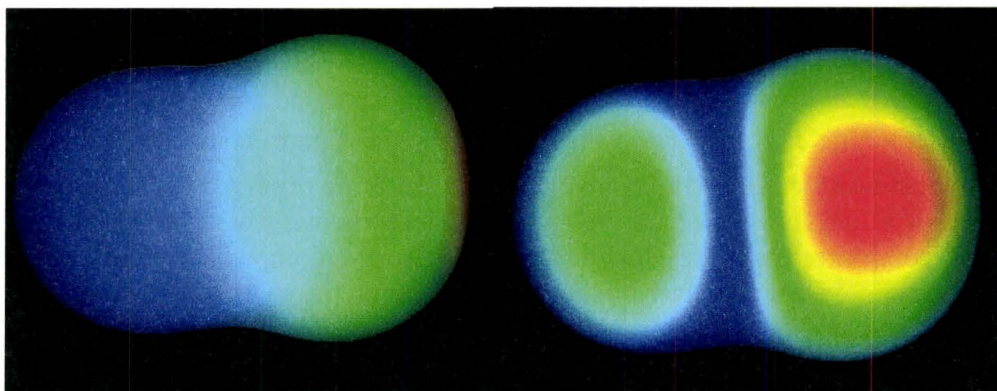
Figure 2. The Fukui function from above, $f^+(\mathbf{r})$, of **b**, plotted on the van der Waals surface of the molecule. Nucleophilic attack occurs para to the carboxamide group.

In many cases it is unclear whether the Fukui function or electrostatic potential terms in Eq. (17) will be dominant. Sometimes the result is surprising. Since the proton is the prototypical hard acid, one expects that protonation sites will usually be predicted by the electrostatic potential. The electrostatic potential for aminoethanol is plotted in Figure 3a. Based on the electrostatic potential, one cannot tell whether protonation of the alcohol (OH) or the amine (NH₂) will be preferred. (See Figure 3a.) The Fukui function from below, $f^-(\mathbf{r})$, (shown in Figure 3b) is also appropriate in this case (since protonation should be accompanied by charge transfer from aminoethanol to the proton). Referring to Figure 3b, the Fukui function shows a clear preference for the protonation of the amine, which is consistent with experimental observations.⁷⁸ In this case, then, the Fukui function serves as “tie-breaker.”



(a) (b)
Figure 3. (a) The electrostatic potential and (b) the Fukui function from below, $f^-(\mathbf{r})$, of the aminoethanol molecule (molecule **c** in Table 1). Protonation of aminoethanol occurs at the nitrogen atom. The color scheme is a spectral scheme where darker blue indicates more positive values and light red value indicate more negative values.

Matters are more subtle when the electrostatic potential and Fukui function predict different regioselectivities, as occurs for the thiocyanate anion (compound **d** in Table 1). In this case the electrostatic potential would predict that electrophilic attack occurs at the nitrogen atom (figure 4a) while $f^-(\mathbf{r})$ is largest at the sulfur center (figure 4b). Similar results are observed for many other ambidentate ligands. Based on the preceding analysis, we would predict that in “ionic” complexes binding occurs through the nitrogen atom, while in complexes with greater covalent character binding occurs through the sulfur atom. This is in fact observed. When the electrophile is a very hard Lewis acid (for example, the proton or Fe^{+3}), ΔN is relative small and so the term containing the Fukui function in Eq. (17) is relatively small. Consequently, SCN^- binds to hard Lewis acids through the nitrogen center. (Such compounds are called isothiocyanates.) Otherwise binding occurs through the sulfur center, forming the more common thiocyanate compounds. However, while binding is preferred at either the nitrogen or sulfur, in any reasonably large sample one is likely to find both isothiocyanates and thiocyanates present. Even for the proton (the prototypical hard Lewis acid) some binding through the sulfur occurs, though it is clear from the ratio of the acid dissociation constants ($\frac{K_a^{HSCN}}{K_a^{HNCS}} = 660$) that binding through the nitrogen is strongly preferred.



(a) (b)
Figure 4. (a) The electrostatic potential and (b) the Fukui function from below, $f^-(\mathbf{r})$, of thiocyanate (molecule **d** in Table 1). Reaction with very hard acids (e.g. H^+ , Fe^{+3}) occurs at the nitrogen atom to form an isothiocyanate compound. Softer Lewis acids attack at the sulfur atom. The color scheme is a spectral scheme where darker blue indicates more positive values and light red value indicate more negative values.

II.VI. Electrophilicity and the Quality of Leaving Groups

Often one is most interested in whether a reaction occurs, and not the specific product that is formed. In this case global reactivity indicators, which quantify the overall reactivity of a molecule, are of greater interest than the site reactivity indicators considered in the previous section. The two most important global reactivity indicators for predicting the susceptibility to and the type of reactions a reagent undergoes are probably the electronic chemical potential (electronegativity)²⁹ and the chemical hardness.³³ As a rule, a reagent with high electronic chemical potential, μ , is a good electron donor, while a reagent with a small μ is a good electron acceptor. The chemical hardness, η , is correlated to the stability of a reagent:^{45;54} the high reactivity of the alkali metals may be attributed to their small chemical hardness.

The chemical potential and hardness are generally considered intrinsic properties of a molecule. This allows one to “order” molecules by their hardness, for example, obtaining useful guides for predicting products of chemical reactions. Thus, when we quantify the chemical potential and hardness of a molecule, we should do so in terms of the properties of the isolated molecule. Using the quadratic model for this purpose leads to the definitions in Eqs. (14) and (15).

The chemical potential and chemical hardness are key indicators of the overall reactivity of the molecule, in general. When we wish to quantify a molecule’s susceptibility to a specific type of reaction, other indicators are appropriate. Since μ and η are the most fundamental descriptors of charge transfer in molecular reactivity, it is unsurprising that such indicators can usually be written as functions of μ and η .

For example, suppose one wishes to quantify the electrophilicity—the ability to accept electrons—of a reagent. (Such a quantity would be useful in describing, for example, the propensity of a reagent for attacking nucleophiles.) The chemical potential will not do as a measure of electrophilicity: while a molecule with low chemical potential is a good electrophile, an extremely hard molecule has reduced ability to accept electrons. Consequently, a measure of molecular electrophilicity will depend on both the chemical potential and the chemical hardness.

In keeping with the idea that intrinsic properties of a reagent should be defined using properties of the isolated molecule and in the spirit of the preceding section, we can model the electrophilicity with a model perturbation. For example, the electrophilicity of a molecule should be related to its ability to accept electrons from a “model” nucleophile, and the “perfect” nucleophile (that is, a reservoir of electrons with zero chemical potential and zero hardness) is the obvious (and perhaps the only non-arbitrary) choice. Defining the electrophilicity as the change in energy that occurs when a reagent is placed in contact with the perfect nucleophile (equivalently, the perfect electron donor) yields

$$\omega = \frac{\mu^2}{2\eta}. \quad (21)$$

This measure of electrophilicity was originally proposed by Maynard, Huang, Rice, and Covell⁷⁹ and justified, using the preceding argument, by Parr, Von Szentpaly, and Liu.⁸⁰ Note that a molecule that is a good electrophile is a molecule with a low (large negative) chemical potential and a small hardness, in accord with intuition. The nucleophilicity has been defined as $\nu = \frac{1}{\omega}$.⁸¹

Another property of a chemical reagent is its ability to serve as a leaving group in solvolysis or, equivalently, the first stage of an S_N1 or S_E1 reaction. For example, a good nucleofuge—a molecule with a high nucleofugality⁸²—will readily accept an electron from a system, leaving behind a nucleophile. Conversely, a good electrofuge—a molecule with a high electrofugality—will simultaneously dissociate from a substrate and donate an electron, leaving behind an electrophile. One might expect that the electrophilicity is a measure of the nucleofugality. However, a nucleofuge must take an *entire* electron with it upon dissociation, while an electrophile has the option of accepting only “a piece of an electron” from the reservoir. This suggests that the destabilization that occurs by *forcing* a molecule to accept an entire electron will be related to the nucleofugality. Thus, the difference in energy between the product (the anion) and the reactant (which we take to be reagent’s state in the presence of a perfect nucleophile),

$$\Delta E_{\text{nucleofuge}} \equiv \text{sgn}(\eta + \mu) \left(\frac{(\eta + \mu)^2}{2\eta} \right), \quad (22)$$

is a sort of “activation energy” for the reaction, and a small value of this “activation energy” will be associated with a high nucleofugality. This suggests that the nucleofugality can be described using the Boltzmann-type form,

$$\lambda_N \equiv e^{-\beta_N \Delta E_{\text{nucleofuge}}} \quad (23)$$

where, in order to obtain a scale that readily differentiates nucleofuges of widely varying quality, we have defined β_N so that the nucleofugality of the hydride ion is $\frac{1}{10}$.

Insofar as we have indices based on the difference in energy between the neutral reagent and the reagent in contact with a perfect nucleophile (electrophilicity) and the difference in energy between the anionic reagent and the reagent in contact with a perfect nucleophile (nucleofugality), we should complete this family of reactivity indicators by defining an index based on the difference in energy between the cation and the reagent in the presence of a perfect nucleophile,

$$\Delta E_{\text{electrofuge}} = \frac{(\mu-\eta)^2}{2\eta}. \quad (24)$$

Since $\Delta E_{\text{electrofuge}}$ measures the relative stability of the cation, a small value of $\Delta E_{\text{electrofuge}}$ should be associated with high electrofugality. This suggests that one define the electrofugality in analogy to Eq. (23), yielding

$$\lambda_E \equiv e^{-\beta_E \Delta E_{\text{electrofuge}}} \quad (25)$$

where β_E is chosen so that the electrofugality of the proton is $\frac{1}{10}$. This discussion of the electrofugality and nucleofugality is based on that of Ayers, Anderson, and Jawed.⁸³ Figure 5 summarizes the relationship between the electrophilicity, $\Delta E_{\text{nucleofuge}}$, and $\Delta E_{\text{electrofuge}}$.

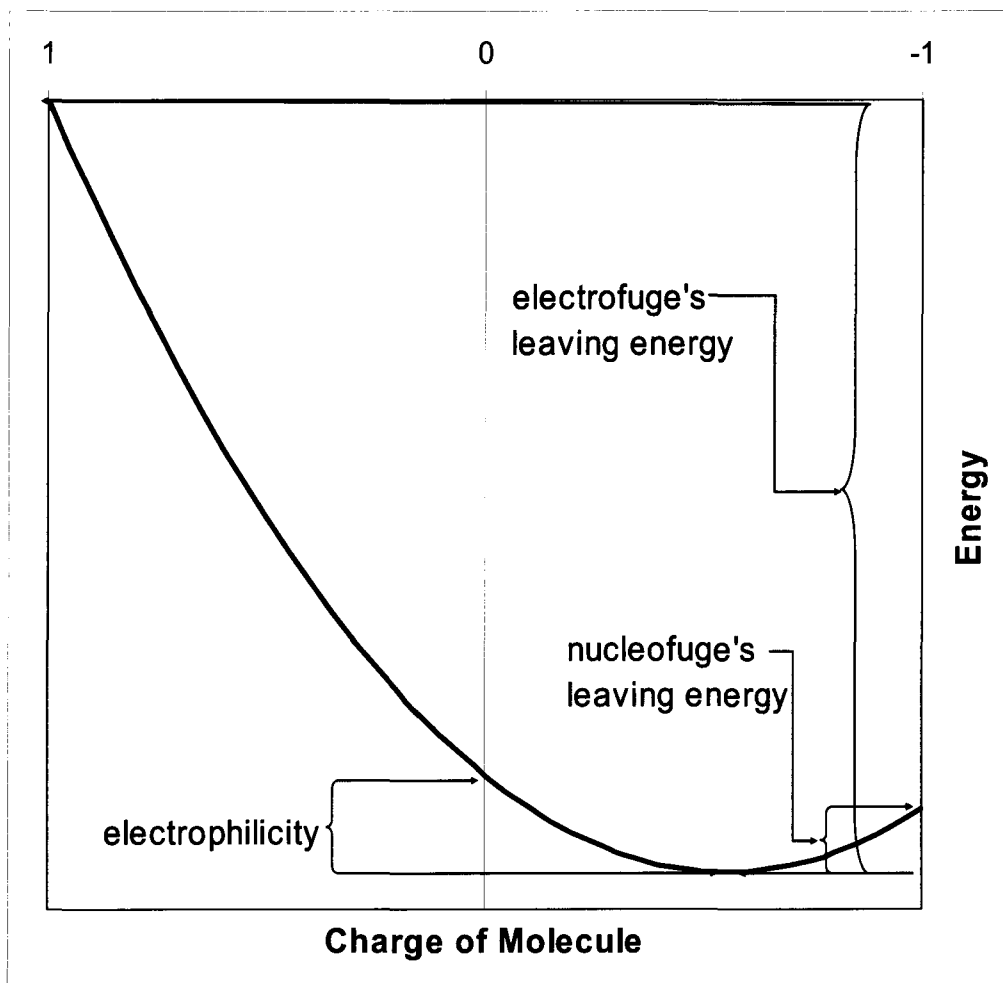


Figure 5. The quadratic model for the dependence of energy on the molecular charge (or, equivalently, the number of electrons), showing the relationship between the electrophilicity, $\Delta E_{\text{nucleofuge}}$, and $\Delta E_{\text{electrofuge}}$.

The appropriateness of the reaction with the “perfect nucleophile” as a model for the electrofugality is somewhat suspect, since an electrofuge leaves behind a good electrophile, and not a nucleophile. However, the energy of the reaction with a perfect electrophile (with $\mu = -\infty$ and zero hardness) is nominally infinite, so this is not a convenient choice of reference. Rather than make an arbitrary choice for the chemical potential of the reference electrophile, we opt to complete the family of indicators based on the reaction with a perfect nucleophile. The definition of the electrofugality is reasonable, however: the electrofugality is greatest when the ionization potential ($\mu - \eta = 3I - A \approx 3I$) is small and the hardness is large. Table 2 compares λ_E for several alkyl-mercury compounds (which may be used as leaving groups to generate carbanions).⁸⁴ The reproduction of the experimentally observed trend in electrofugality confirms the practical utility of λ_E .⁸⁵ The electrofugality could be used as an alternative measure of nucleophilicity.

Table 2. The electrofugality of alkyl-mercury compounds.⁸⁶

Compound	$\Delta E_{\text{electrofuge}}$ (eV)	electrofugality
H	15.60	0.10
HgCl	11.51	0.18
HgMe	8.03	0.31
HgEt	7.63	0.32
Hg-iPr	7.39	0.34
Hg-tBu	7.29	0.34

Support for the present definition of the nucleofugality may be obtained by comparing the nucleofugality to experimental measurements of the rate of solvolysis of halogenated 1-phenylethyl esters in 80% aqueous ethanol at 75°C.⁸⁷ If we assume that rate of solvolysis can be approximated with an Arrhenius-type expression,

$$k_{\text{solvolysis}} \propto e^{-\beta E_{\text{activation}}}, \quad (26)$$

then plotting $\log(k_{\text{solvolysis}})$ versus $\log(\lambda_N)$ will give a straight line if $E_{\text{activation}} \propto \Delta E_{\text{nucleofuge}}$. Plotting this data (cf. Figure 6) yields a line with $R^2 = .96$, which can be taken as experimental support of the hypothesized relationship between the activation energy and $\Delta E_{\text{nucleofuge}}$. Even stronger support comes from the slope of the line, which—if $\Delta E_{\text{nucleofuge}}$ is *equal* to the activation energy—should equal $\frac{\beta}{\beta_N}$. Using the slope of the line to approximate the temperature, we estimate that the experimental measurements were performed at 344 K, in remarkable agreement with the actual value of 348 K. Though the accuracy of this prediction should be considered fortuitous, this result testifies to the validity of the logic by which we derived the expression, Eq. (23), for the nucleofugality.

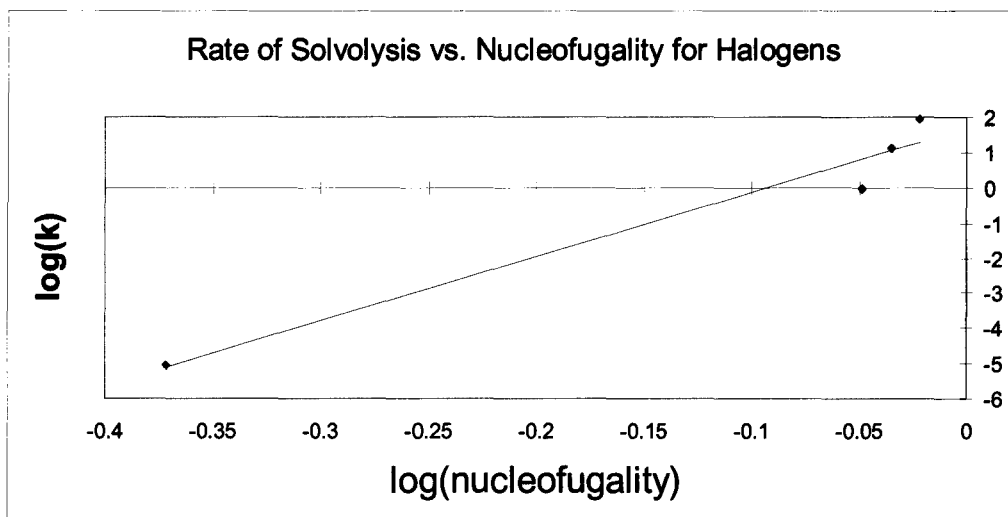


Figure 6. Comparison between the nucleofugality of the halides and rate of solvolysis of halogenated 1-phenylethyl esters in 80% aqueous ethanol at 75 C.⁸⁷ The value of correlation coefficient ($R^2 = .96$) is reasonable, especially since the empirical rates span seven orders of magnitude.

II.VII. Conclusion

It is a known chemical truth that molecules tend to have a set reactivity: a molecule has a given acidity or basicity and is subject to nucleophilic attack at certain sites and electrophilic attack at certain other sites; it is a good (or a bad) leaving group. These properties of the molecule—which are preserved (with rare exception) in a variety of reaction environments and with a large number of reaction partners—are described and predicted by the perturbative perspective: A given molecule tends to have a “universal” response to the approach of a positively (or negatively) charged electrophile (or nucleophile), regardless of the specific reagent containing the approaching reactive site. A given molecule tends to have a “universal” response in response to an electron donor or acceptor, regardless of the specific structure of that acceptor donor. Thus, by introducing model perturbations that capture the essence of an attacking species, we can predict the chemical reactivity of the molecule.

As particular examples, we have considered the electrostatic potential and Fukui function (which are relevant for regioselectivity) and the electrophilicity, nucleofugality, and electrofugality (relevant for describing the propensity of a substrate towards specific types of reaction). These specific examples show that the perturbative perspective is a powerful intuitive tool and, because of its direct contact with the important quantities from density-functional theory, represents a single theory that describes—both qualitatively and quantitatively—the reactivity of molecules.

Acknowledgement: The authors thank SHARCnet and the North Carolina Supercomputing Center for computational resources and NSERC, McMaster University, and the Canada Research Chairs for research support.

II.VIII. Appendix

We devote this appendix to a detailed exposition on the functional analytic Taylor series, concentrating on the zero-temperature result (Eq. (10)). Equation (10) models changes with respect to electron number *exactly* unless N passes an integer value not equal to N_A . (Mathematically, the radius of convergence is expressed in terms of the floor and ceiling functions, specifically the series converges if

$$\Delta N \in \left\{ \begin{array}{l} \left[\lfloor N_A \rfloor, \lceil N_A \rceil \right] \\ \left[N_A - 1, N_A + 1 \right] \end{array} \right. \quad \left. \begin{array}{l} N_A \text{ is not an integer} \\ N_A \text{ is an integer} \end{array} \right. \quad (27)$$

Equation (27) follows from the fact that the energy is a piecewise linear function of N_A with discontinuities at the integers. For changes in particle number that are larger than this, one may proceed by analytic continuation. First use (10) until N reaches the next integer past N_A ; compute ΔE^\pm for this change. Compute the coefficients in the Taylor series for this “intermediate” state; then, calculate the effect of further changes in the number of electrons in like manner. Such a procedure is exact.

When a system’s properties, Q , are modeled using the grand canonical ensemble with positive temperatures,

$$Q[v, \mu(N_A)] \equiv \sum_{N=0}^{\infty} \sum_{k=0}^{\infty} Q_k[v_A, N] e^{-\beta(E_k[v_A, N] - \mu N)} \quad (28)$$

the energy and its functional derivatives with respect to the external potential are analytic functions for $0 \leq N_A < \infty$. This follows directly from Eq. (28), where $Q_k[v_A, N]$ denotes the value of the property, Q , for the k^{th} -excited state of the N -electron system with external potential v_A , $\beta = \frac{1}{k_B T}$, and the chemical potential is chosen to reproduce the desired number of electrons. Even at $N_A = 0$, the derivatives with respect to N of the energy, density, polarizability kernel, etc. are all defined (provided we evaluate the derivatives from the electron-abundant side). This means that analytic continuation to negative electron number is technically possible, so that the Taylor expansion with respect to N_A converges for any choice of ΔN . (Of course it is only for $-N_A \leq \Delta N < \infty$ that this expansion is physically relevant.)

The more subtle convergence with respect to changes in the external potential is best understood by referring to the perturbation series for the Hamiltonian

$$\hat{H}(\lambda) = \hat{H}[v_A(\mathbf{r}); N_A] + \lambda \sum_{i=1}^{N_b} w_B^{\text{eff}}(\mathbf{r}). \quad (29)$$

The perturbation series for such operators have been explored mathematically, and most of the key results are reviewed in Kato’s monograph on the subject.⁸⁸ Using analytic continuation, one can evaluate the change in energy at any desired value of λ . Computing when a specific series ceases to converge and one must reevaluate the coefficients is more difficult, and the most powerful tools require information about the resolvent, $(\hat{H}(\lambda) - E)^{-1}$, that is frequently difficult to obtain. However, a general and helpful rule of thumb is that the series does not

converge to $\lambda = \tilde{\lambda}$ if there at some point in the interval $\lambda \in (-\tilde{\lambda}, \tilde{\lambda})$ there is a catastrophic change in either the character of the eigenstate (as from a bound state to a state embedded in the continuous spectrum) or if the perturbation is strong enough to cause “level crossing” so that the character of the ground state fundamentally changes.⁸⁹ The most intuitive and physically important case occurs when, for some λ , the state becomes unbound. In most other cases, the Taylor series converges, but may converge to an excited state if level crossing is not explicitly accounted for. Just as before, one can find the true change in the ground state energy using analytic continuation methods: exploit the power series up to its radius of convergence; reevaluate the functional derivatives in the Taylor series at the radius of convergence; then, resume progress toward the target system. One should note that, in general, at the radius of convergence there several states with the same energy, so degenerate perturbation theory must be used to evaluate the coefficients.

There are two problems with the preceding approach. The first is pragmatic: the coefficients of cubic and higher order terms in $w_B^{eff}(\mathbf{r})$ are difficult to evaluate computationally. These higher order terms are the “kernels” that measure the nonlinear response of a reagent to its environment and such “hyperpolarizabilities” are difficult to compute. (For example, evaluating the first hyperpolarizability kernel, $\frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}^*)\delta v(\mathbf{r}^*)}$, with perturbation theory requires an accurate second-order correction to the wave function.) For describing most molecules, the hyperpolarizability is much less important than polarizability, and higher-order coefficients are important only when the molecule is exposed to very strong electric fields. Unfortunately, due to the short distances involved, the electric fields to which reagents are exposed during molecule creation are very strong (electric fields are on the order of $10^{11} \frac{\text{C}}{\text{m}}$!), and the hyperpolarizabilities can be important. If one truncates the Taylor series after quadratic terms in $w_B^{eff}(\mathbf{r})$, one is essentially hoping that the hyperpolarizability does not affect the *qualitative* predictions of the model. Based on the observed utility of the quadratic truncation, this seems to be a good assumption. It is also in keeping with a general philosophical point: it is undesirable for a method that yields only qualitative information to require the same intense computational effort associated with highly-accurate quantitative predictions.

Apart from the problem of computing hyperpolarizability corrections, accurately estimating the radius of convergence of the series in the external potential is very difficult. The alternative is to use one of the functional analytic

analogues to “Taylor’s theorem with remainder” for the series in question. Analogous to the integral form of the remainder is the *exact* expression^{90;91}

$$\begin{aligned} & E\left[v_A(\mathbf{r}) + w_B^{\text{eff}}(\mathbf{r}); N_A + \Delta N\right] \\ &= E\left[v_A(\mathbf{r}); N_A\right] + \int_0^1 \left\{ \Delta N \cdot \mu_t + \int \rho_t(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} \right\} dt, \end{aligned} \quad (30)$$

where quantities with a subscript “ t ” are to be evaluated for the system with $N_A + t\Delta N$ electrons confined in the external potential $v_t(\mathbf{r}) = v_A(\mathbf{r}) + tw_B^{\text{eff}}(\mathbf{r})$. (At zero temperature one needs to use the appropriate one-sided chemical potentials, cf. Eq. (12).)

While Eq. (30) is useful for computational work, it is often instructive to manipulate the more intuitive forms,

$$E\left[v_A(\mathbf{r}) + w_B^{\text{eff}}(\mathbf{r}); N_A + \Delta N\right] = E\left[v_A(\mathbf{r}); N_A\right] + \overline{\mu}_t \cdot \Delta N + \overline{\int \rho_t(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r}} \quad (31)$$

where the “overbar” indicates that the coefficient is averaged over the change in the variable, e.g.

$$\overline{\rho_t(\mathbf{r})} \equiv \int_0^1 \rho_t(\mathbf{r}) dt. \quad (32)$$

When computing the polarizability kernel is unproblematic, it is advantageous to include the second order term in the Taylor series. The associated form is^{90;92}

$$\begin{aligned} E\left[v_A(\mathbf{r}) + w_B^{\text{eff}}(\mathbf{r}); N_A + \Delta N\right] &= E\left[v_A(\mathbf{r}); N_A\right] + \mu_A^\pm \Delta N + \int \rho_A(\mathbf{r}) w_B^{\text{eff}} d\mathbf{r} \\ &+ \int_0^1 (1-t) \left[\eta_t \Delta N^2 + 2 \int f_t^\pm(\mathbf{r}) \Delta N w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} \right. \\ &\quad \left. + \iint P_t(\mathbf{r}, \mathbf{r}') w_B^{\text{eff}}(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \right] dt \end{aligned} \quad (33)$$

This result can also be expressed as an average over the “range of perturbation,” yielding

$$\begin{aligned} E\left[v_A(\mathbf{r}) + w_B^{\text{eff}}(\mathbf{r}); N_A + \Delta N\right] &= E\left[v_A(\mathbf{r}); N_A\right] + \mu_A^\pm \Delta N + \int \rho_A(\mathbf{r}) w_B^{\text{eff}} d\mathbf{r} \\ &+ \Delta N^2 \overline{\eta_t} + 2\Delta N \overline{\int f_t^\pm(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r}} + \overline{\iint P_t(\mathbf{r}, \mathbf{r}') w_B^{\text{eff}}(\mathbf{r}) w_B^{\text{eff}}(\mathbf{r}') d\mathbf{r} d\mathbf{r}'} \\ &- \Delta N \overline{(t\eta_t \Delta N)} - 2\Delta N \overline{\int f_t^\pm(\mathbf{r}) t w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r}} - \overline{\iint w_B^{\text{eff}}(\mathbf{r}') P_t(\mathbf{r}, \mathbf{r}') t w_B^{\text{eff}}(\mathbf{r}) d\mathbf{r} d\mathbf{r}'}. \end{aligned} \quad (34)$$

Equations (33) and (34) have been written in a form that suggests both the zero-temperature generalization (set $\eta_t = 0$) and the finite-temperature result (in which case the \pm signs on the chemical potential and the Fukui function are unnecessary).

Equations (30)-(34) provide elegant methods for avoiding the truncation of perturbative expansions, but because they require integration over the entire range of the perturbative parameter, the cost of evaluating these expressions is dramatically (in principle, infinitely) more than would be required to evaluate the truncated Taylor series. In practice, the integrals over t will be approximated by efficient numerical quadratures, but a sequence of several quantum chemistry computations is still required. Since direct computation of change in energy requires only two calculations (one for the reactant and one for the product), one may well inquire whether there are any benefits to the perturbative ansatz introduced above. The primary benefits are conceptual. Each term in the Taylor series corresponds to a particular “response” of the system to a perturbation, the Taylor series represents not only the magnitude of the change in energy associated with a particular chemical reaction but, more importantly, reveals why the energy changes as it does.

Furthermore, if the change in the number of electrons is small enough, only the last three coefficients in the Taylor series in Eq. (33) depend upon the specific reaction. The leading coefficients in the perturbative treatment are independent of the nature of the reagent with which the molecule interacts, and thus are generally characteristic of the reactivity of the molecule in question, entirely independent of the specifics of any particular reaction. In this sense, the coefficients in the Taylor series are reactivity indicators of the “generic” reactivity of a molecular species. Furthermore, using the trapezoidal rule to approximate the integrals in Eq. (33), the dependence of the second order terms on the “target” of the perturbation series is removed and one recovers the truncated Taylor series, Eq. (8).

II.IX. References

- (1) Gillespie, R. J.; Popelier, P. L. A. *Chemical Bonding and Molecular Geometry*; Oxford: New York, 2001.
- (2) Wheland, G. *Resonance in Organic Chemistry*. 1955. New York, Wiley.
Ref Type: Serial (Book, Monograph)
- (3) Albright, T. A.; Burdett, J. K.; Whangbo, M. H. *Orbital Interactions in Chemistry*; Wiley-Interscience: New York, 1985.
- (4) Figgis, B. M.; Hitchman, M. A. *Ligand Field Theory and Its Applications*; Wiley-VCH: New York, 2000.
- (5) Woodward, R. B.; Hoffmann, R. *Angew. Chem., Int. Ed. Engl.* **1969**, *8*, 781-853.
- (6) Hoffmann, R.; Woodward, R. B. *Acc. Chem. Res.* **1968**, *1*, 17-22.
- (7) Hoffmann, R.; Woodward, R. B. *J. Am. Chem. Soc.* **1965**, *87*, 2046-2048.

- (8) Woodward, R. B.; Hoffmann, R. *J.Am.Chem.Soc.* **1965**, *87*, 2511-2513.
- (9) Woodward, R. B.; Hoffmann, R. *J.Am.Chem.Soc.* **1965**, *87*, 395-397.
- (10) Fujimoto, H.; Fukui, K. Intermolecular interactions and chemical reactivity; In *Chemical reactivity and reaction paths*; Klopman, G., ed. Wiley-Interscience: New York, 1974; pp 23-54.
- (11) Fukui, K. *Theory of orientation and stereoselection*; Berlin: Springer-Verlag, 1975.
- (12) Fukui, K. *Science* **1987**, *218*, 747-754.
- (13) Fukui, K.; Yonezawa, T.; Shingu, H. *J.Chem.Phys.* **1952**, *20*, 722-725.
- (14) Fukui, K.; Yonezawa, T.; Nagata, C. *J.Chem.Phys.* **1953**, *21*, 174-176.
- (15) Fukui, K.; Yonezawa, T.; Nagata, C. *Bull.Chem.Soc.Japan* **1954**, *27*, 423-427.
- (16) Ortiz, J. V. *J.Chem.Phys.* **1991**, *94*, 6064-6072.
- (17) Tong, W.; Morrison, R. C.; Day, O. W. *Int.J.Quantum Chem.* **1996**, *60*, 411-419.
- (18) Morrison, R. C.; Tong, W.; Day, O. W. *Int.J.Quantum Chem.* **1996**, *60*, 421-431.
- (19) Morrison, R. C.; Dixon, C. M.; Mizell, J. R. *Int.J.Quantum Chem.* **1994**, 309-314.
- (20) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford UP: New York, 1989.
- (21) Dreizler, R. M.; Gross, E. K. U. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*; Springer-Verlag: Berlin, 1990.
- (22) Hohenberg, P.; Kohn, W. *Phys.Rev.* **1964**, *136*, B864-B871.
- (23) Levy, M. *Proc.Natl.Acad.Sci.USA* **1979**, *76*, 6062-6065.
- (24) Lieb, E. H. *NATO ASI Ser., Ser.B* **1985**, *123*, 31-80.
- (25) Lieb, E. H. *Int.J.Quantum Chem.* **1983**, *24*, 243-277.
- (26) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem.Rev.* **2003**, *103*, 1793-1873.
- (27) Parr, R. G.; Yang, W. T. *Ann.Rev.Phys.Chem.* **1995**, *46*, 701-728.
- (28) Kohn, W.; Becke, A. D.; Parr, R. G. *J.Phys.Chem.* **1996**, *100*, 12974-12980.
- (29) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J.Chem.Phys.* **1978**, *68*, 3801-7.
- (30) Parr, R. G.; Yang, W. *J.Am.Chem.Soc.* **1984**, *106*, 4049-4050.
- (31) Yang, W.; Parr, R. G.; Pucci, R. *J.Chem.Phys.* **1984**, *81*, 2862-2863.
- (32) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (33) Parr, R. G.; Pearson, R. G. *J.Am.Chem.Soc.* **1983**, *105*, 7512-16.
- (34) Parr, R. G.; Von Szentpaly, L.; Liu, S. B. *J.Am.Chem.Soc.* **1999**, *121*, 1922-1924.
- (35) Nalewajski, R. F.; Parr, R. G. *J.Chem.Phys.* **1982**, *77*, 399-407.
- (36) Liu, S. B.; Parr, R. G. *Chem.Phys.Lett.* **1997**, *278*, 341-344.

- (37) De Proft, F.; Liu, S.; Parr, R. G. *J.Chem.Phys.* **1997**, *107*, 3000-3006.
- (38) Parr, R. G.; Liu, S. B. *Chem.Phys.Lett.* **1997**, *276*, 164-166.
- (39) Senet, P. *J.Chem.Phys.* **1997**, *107*, 2516-2524.
- (40) Senet, P. *J.Chem.Phys.* **1996**, *105*, 6471-6489.
- (41) Parr, R. G.; Von Szentpaly, L.; Liu, S. B. *J.Am.Chem.Soc.* **1999**, *121*, 1922-1924.
- (42) Ayers, P. W. *Theor.Chem.Acc.* **2001**, *106*, 271-279.
- (43) Ayers, P. W.; Morrison, R. C.; Roy, R. K. *J.Chem.Phys.* **2002**, *116*, 8731-8744.
- (44) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (45) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2000**, *122*, 2010-2018.
- (46) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (47) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (48) Ayers, P. W. *J.Chem.Phys.* **2000**, *113*, 10886-10898.
- (49) Sanderson, R. T. *Chemical Bonds and Bond Energy*; Academic: New York, 1976.
- (50) Sanderson, R. T. *Science* **1951**, *114*, 670-672.
- (51) York, D. M.; Yang, W. *J.Chem.Phys.* **1996**, *104*, 159-72.
- (52) Pearson, R. G. *J.Chem.Educ.* **1999**, *76*, 267-275.
- (53) Chattaraj, P. K. *Proc.Indian Natl.Sci.Acad., Part A* **1996**, *62*, 513-531.
- (54) Pearson, R. G. *J.Chem.Educ.* **1987**, *64*, 561-567.
- (55) Chattaraj, P. K.; Liu, G. H.; Parr, R. G. *Chem.Phys.Lett.* **1995**, *237*, 171-6.
- (56) Parr, R. G.; Chattaraj, P. K. *J.Am.Chem.Soc.* **1991**, *113*, 1854-5.
- (57) Chattaraj, P. K.; Lee, H.; Parr, R. G. *J.Am.Chem.Soc.* **1991**, *113*, 1855-6.
- (58) Pearson, R. G. *J.Am.Chem.Soc.* **1963**, *85*, 3533-3539.
- (59) Pearson, R. G. *Inorg.Chim.Acta* **1995**, *240*, 93-8.
- (60) Hellmann, H. *Einfuehrung in die Quantenchemie*; Deuticke: Leipzig, 1937.
- (61) Feynman, R. P. *Phys.Rev.* **1939**, *56*, 340-343.
- (62) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (63) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L., Jr. *Phys.Rev.Lett.* **1982**, *49*, 1691-1694.
- (64) Gyftopoulos, E. P.; Hatsopoulos, G. N. *Proc.Natl.Acad.Sci.USA* **1965**, *60*, 786-793.
- (65) Ayers, P. W. *Variational Principles for Describing Chemical Reactions*. 2001. University of North Carolina at Chapel Hill. 2001.

Ref Type: Thesis/Dissertation

- (66) Yang, W.; Zhang, Y.; Ayers, P. W. *Phys.Rev.Lett.* **2000**, *84*, 5172-5175.
- (67) Mulliken, R. S. *J.Chem.Phys.* **1934**, *2*, 782-793.
- (68) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (69) Politzer, P.; Truhlar, D. *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Plenum: New York, 1981.

- (70) Klopman, G. *J.Am.Chem.Soc.* **1968**, *90*, 223-234.
- (71) Flurchick, K.; Bartolotti, L. *Journal of Molecular Graphics* **1995**, *13*, 10-13.
- (72) Huheey, J. E.; Keiter, E. A.; Keiter, R. L. *Inorganic Chemistry: Principles of Structure and Reactivity*; HarperCollins: New York, 1993.
- (73) Zhou, Z.; Parr, R. G.; Garst, J. F. *Tetrahedron Lett.* **1988**, *29*, 4843-6.
- (74) Zhou, Z.; Parr, R. G. *J.Am.Chem.Soc.* **1989**, *111*, 7371-9.
- (75) Berkowitz, M. *J.Am.Chem.Soc.* **1987**, *109*, 4823-4825.
- (76) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (77) Demas, M.; Javadi, G. J.; Bradley, L. M.; Hunt, D. A. *Journal of Organic Chemistry* **2000**, *65*, 7201-7202.
- (78) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of Biochemistry, Upgrade Edition*; Wiley: New York, 2002.
- (79) Maynard, A. T.; Huang, M.; Rice, W. G.; Covell, D. G. *Proc.Natl.Acad.Sci.USA* **1998**, *95*, 11578-11583.
- (80) Parr, R. G.; Von Szentpaly, L.; Liu, S. B. *J.Am.Chem.Soc.* **1999**, *121*, 1922-1924.
- (81) Chattaraj, P. K.; Maiti, B. *J.Phys.Chem.A* **2001**, *105*, 169-183.
- (82) Stirling, C. J. M. *Acc.Chem.Res.* **1979**, *12*, 198-203.
- (83) Ayers, P. W.; Anderson, J. S. M.; Jawed, Z. *submitted* **2003**.
- (84) Ayers, P. W.; Anderson, J. S. M.; Jawed, Z. *submitted* **2003**.
- (85) March, J. *Advanced Organic Chemistry*; Wiley-Interscience: New York, 1992.
- (86) Ayers, P. W.; Anderson, J. S. M.; Jawed, Z. *submitted* **2003**.
- (87) Noyce, D. S.; Virgilio, J. A. *Journal of Organic Chemistry* **1972**, *37*, 2643.
- (88) Kato, Tosio. *Perturbation Theory for Linear Operators*. 2nd. 1980. Berlin, Springer-Verlag.

Ref Type: Serial (Book,Monograph)

- (89) Poles and branch points associated with complex values of the perturbation strength can also be problematic. Such perturbations, however, are difficult to interpret in any physical way.
- (90) van Leeuwen, R.; Baerends, E. J. *Phys.Rev.A* **1995**, *51*, 170-178.
- (91) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (92) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.

Chapter III

**Conceptual Density-Functional
Theory for General Chemical
Reactions, Including Those That Are
Neither Charge Nor Frontier-Orbital
Controlled**

**I. Theory and Derivation of a
General-Purpose Reactivity
Indicator^{*}**

^{*}The content of this chapter has been published: **J. S. M. Anderson**, J. Melin, P. W. Ayers
“Conceptual density-functional theory for general chemical reactions, including those that are
neither charge nor frontier-orbital controlled I. Theory and derivation of a general-purpose
reactivity indicator”; *J. Chem. Th. Comp.* **2007**, 3, 358-374.

III.I. Statement of the Problem

This chapter formulates a new general-purpose reactivity indicator that interpolates between the charge-controlled electron-transfer-controlled reactivity paradigms. The new indicator is designed to predict which sites in the molecule are the most reactive when the reaction is charge controlled, charge-transfer controlled, and any combination. To move between these two extremes the indicator has two parameters. These parameters are varied by the user based on the charge of an incoming reagent and the extent of electron transfer. The mathematical underpinnings of the general-purpose reactivity indicator allow us to solve some outstanding problems relate to the Hard/Soft Acid/Base principle. A side-effect of this analysis is the first rigorous explanation of the counterintuitive “minimum Fukui function” rule. In addition, a similarly counterintuitive “minimum electrostatic potential” rule is derived and explained.

III.II. Introduction

Many qualitative and semi-quantitative methods have been developed for predicting how and whether a reaction will take place. Perhaps the most popular method of prediction is frontier molecular orbital theory (FMO).¹ This method uses the shapes and symmetries of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) to indicate whether a reaction will occur. If the HOMO of the electron donor and the LUMO of the electron acceptor have the same shape and phase, then electron transfer from the HOMO of the first molecule to the LUMO of the second can occur, often forming a bond between the reagents. Kenichi Fukui and Roald Hoffmann shared the 1981 Nobel Prize in chemistry for precisely this: showing that the shape²⁻⁵ (Fukui) and phase⁶⁻¹⁰ (Hoffmann) of the HOMO and LUMO orbitals is indicative of chemical reactivity.

A primary limitation of the frontier molecular orbital theory approach is that it presupposes the validity of the orbital model and thus fails to incorporate the effects of electron correlation or orbital relaxation. This motivated the definition of a “Fukui function” in the context of density-functional theory (DFT), a function that encapsulates the essence of FMO¹¹⁻¹⁴ but, in principle, includes both electron correlation^{11;15} and orbital relaxation.^{12;16} The Fukui function from below, $f^-(\mathbf{r})$, is defined as the change in density that one observes when one goes from N to $N-1$ electrons (with the nuclear positions fixed); in simple molecular orbital theory this would give precisely the density of the HOMO

orbital. A similar function, $f^+(\mathbf{r})$, can be defined as the difference between the electron densities of the $N+1$ and N electron systems; this is analogous to the LUMO orbital density.

The Fukui function is labeled according to whether the system is acting as an electron acceptor or an electron donor. $f^+(\mathbf{r})$ says where an electron (received from a perfect electron donor) will add to the molecule. $f^-(\mathbf{r})$ says where an electron given to an electron acceptor (a perfect one, if you like) will come from. Electron donors tend to attack the molecule where $f^+(\mathbf{r})$ is large because this is where the molecule “wants electrons.” Electron acceptors tend to attack the molecule where $f^-(\mathbf{r})$ is large because this is where the molecule has electrons that it is “willing to give up.”^{11;17} This reasoning, of course, is only valid when the transition state lies early enough on the reaction path for the reacting fragments to still resemble the isolated reagents.¹⁸ The fact that conceptual DFT tends to work even when the transition state is not especially early may be attributed to the empirically observed utility of “hill-climbing” methods for locating transition states.^{19;20}

In 1989, Dewar²¹ listed several nucleophiles where FMO failed to describe electrophilic aromatic substitution, including isoquinoline, 10-hydroxy-10,9-borazarophenanthrene, and 10-methyl-10,9-borazarophenanthrene. We attempted to explain the reactivity of these molecules using the tools of conceptual DFT but it does not seem possible to describe the reactivity of these molecules without considering *both* electron-transfer effects *and* electrostatic interactions. This spurred us to develop a new reactivity indicator that combines the Fukui function and the electrostatic potential. The indicator we developed provides a general-purpose model for chemical reactivity. The purpose of this paper is to derive and discuss this reactivity indicator. The second paper in this series will discuss the application of the indicator to Dewar’s problematic molecules.²²

Before deriving the general-purpose reactivity indicator, we present a brief overview of FMO and conceptual DFT in section III. The new reactivity indicator is then derived, step-by-step, in section IV. Our model provides a unified picture of chemical reactivity and elucidates, among other things, the “minimum Fukui function rule.”^{23;24} Section V summarizes our findings.

III.III. Theoretical Background

A. Overview of Frontier Molecular Orbital Theory

Frontier Molecular Orbital theory (FMO) arises as a simplification of the treatment proposed by Coulson and Longuet-Higgins,²⁵⁻²⁸ who used second order perturbation theory to describe the interactions between the filled molecular orbitals of one reactant and the empty molecular orbitals of the other. The interaction energy between the fragments is then

$$E_{AB}^{FMO} = 2 \left(\sum_{i \in \mathbf{O}_A} \sum_{b \in \mathbf{U}_B} \frac{\left| \langle \phi_i^{(A)} | \hat{h}_{AB} | \phi_b^{(B)} \rangle \right|^2}{\varepsilon_i^{(A)} - \varepsilon_b^{(B)}} + \sum_{j \in \mathbf{O}_B} \sum_{a \in \mathbf{U}_A} \frac{\left| \langle \phi_j^{(B)} | \hat{h}_{AB} | \phi_a^{(A)} \rangle \right|^2}{\varepsilon_j^{(B)} - \varepsilon_a^{(A)}} \right). \quad (1)$$

Here, $\{\phi_i^{(A)}\}$ and $\{\phi_j^{(B)}\}$ are the molecular orbitals of fragments A and B ; $\{\varepsilon_i^{(A)}\}$ and $\{\varepsilon_j^{(B)}\}$ are their respective orbital energies, \mathbf{O}_A and \mathbf{O}_B are the sets of occupied molecular orbitals and \mathbf{U}_A and \mathbf{U}_B are the sets of unoccupied (virtual) orbitals in fragments A and B , respectively. \hat{h}_{AB} denotes the one-electron Hamiltonian for the “supermolecule” ($A+B$) (e.g., the Fock operator in Hartree-Fock or the Kohn-Sham Hamiltonian in DFT). The numerators in Eq. (1) are analogous to resonance integrals in Hückel theory.

In general, the most important terms in Eq. (1) are those with the smallest denominator. This suggests ignoring terms that do not depend on the frontier molecular orbitals, so that

$$E_{AB}^{FMO} \approx 2 \left(\frac{\left| \langle \phi_{HOMO}^{(A)} | \hat{h}_{AB} | \phi_{LUMO}^{(B)} \rangle \right|^2}{\varepsilon_{HOMO}^{(A)} - \varepsilon_{LUMO}^{(B)}} + \frac{\left| \langle \phi_{HOMO}^{(B)} | \hat{h}_{AB} | \phi_{LUMO}^{(A)} \rangle \right|^2}{\varepsilon_{HOMO}^{(B)} - \varepsilon_{LUMO}^{(A)}} \right). \quad (2)$$

If A (the Lewis acid) is the electron acceptor and B (the Lewis base) is the electron donor, then we expect that $\varepsilon_{LUMO}^{(A)}$ is small and $\varepsilon_{HOMO}^{(B)}$ is large. By this argument, the second term in Eq. (2) should be larger than the first term. Neglecting the first term gives

$$E_{AB}^{FMO} \approx 2 \left(\frac{\left| \langle \phi_{HOMO}^{(B)} | \hat{h}_{AB} | \phi_{LUMO}^{(A)} \rangle \right|^2}{\varepsilon_{HOMO}^{(B)} - \varepsilon_{LUMO}^{(A)}} \right). \quad (3)$$

This equation is the essential basis for the frontier molecular orbital theory. It is only valid when the neglected terms in Eq. (1) (and especially the neglected first term in Eq. (2)) are negligible.²¹ The numerator of Eq. (3) is a sort of “generalized resonance integral,” and so one can infer, by the usual arguments (e.g., from the justification of the Wolfsberg-Helmoltz approximation²⁹), that a large overlap between electron-donating and electron-accepting orbitals is favorable. This result can also be inferred from a Hölder inequality:

$$\begin{aligned} \left| \left\langle \phi_{HOMO}^{(B)} \left| \hat{h}_{AB} \right| \phi_{LUMO}^{(A)} \right\rangle \right|^2 &\leq \left\| \hat{h}_{AB} \right\|^2 \left\| \phi_{HOMO}^{(B)} \phi_{LUMO}^{(A)} \right\|^2 \\ &= \left\| \hat{h}_{AB} \right\|^2 \left(\int \left| \left(\phi_{HOMO}^{(B)}(\mathbf{r}) \right)^* \phi_{LUMO}^{(A)}(\mathbf{r}) \right| d\mathbf{r} \right)^2. \end{aligned} \quad (4)$$

B. The Fukui Function

As mentioned before, since the Fukui function contains similar information to the frontier molecular orbitals, it can be used to provide a DFT-based alternative to the standard rationalization of FMO theory. The Fukui function, $f(\mathbf{r})$, is defined as^{11;12;30}

$$f(\mathbf{r}) = \left[\frac{\delta\mu}{\delta v(\mathbf{r})} \right]_N = \left[\frac{\partial\rho(\mathbf{r})}{\partial N} \right]_v. \quad (5)$$

Here μ is the electronic chemical potential (equal to minus the electronegativity), $v(\mathbf{r})$ is the “external” potential due to the atomic nuclei, $\rho(\mathbf{r})$ is the electron density, and N is the number of electrons. The equality between the functional derivative of the chemical potential and the ordinary derivative of the electron density arises as a “Maxwell relation.”³¹

The slope of $\rho(\mathbf{r})$ as a function of N has discontinuities³²⁻³⁴ and thus the derivative must be evaluated from above and below (and averaged if necessary). This results in a Fukui function appropriate for describing nucleophilic attack,¹¹

$$f^+(\mathbf{r}) = \left[\frac{\partial\rho(\mathbf{r})}{\partial N} \right]_{v(\mathbf{r})}^+ \quad (6)$$

and a Fukui function appropriate for describing electrophilic attack,

$$f^-(\mathbf{r}) = \left[\frac{\partial\rho(\mathbf{r})}{\partial N} \right]_{v(\mathbf{r})}^-. \quad (7)$$

If one writes the electron density as a function of the Kohn-Sham orbitals, $\phi_i(\mathbf{r})$, and orbital occupation numbers, n_i ,

$$\rho(\mathbf{r}) = \sum_{i=1}^{\infty} n_i |\phi_i(\mathbf{r})|^2 \quad (8)$$

$$n_i = \begin{cases} 1 & i \leq \text{HOMO} \\ 0 & i \geq \text{LUMO} \end{cases} \quad (9)$$

then, from Eqs. (6) and (7), one has^{12;35}

$$f^+(\mathbf{r}) = |\phi_{\text{LUMO}}(\mathbf{r})|^2 + \sum_{i=1}^{\text{HOMO}} \left(\frac{\partial |\phi_i(\mathbf{r})|^2}{\partial N} \right)_{v(\mathbf{r})}^+ \quad (10)$$

$$f^-(\mathbf{r}) = |\phi_{\text{HOMO}}(\mathbf{r})|^2 + \sum_{i=1}^{\text{HOMO}} \left(\frac{\partial |\phi_i(\mathbf{r})|^2}{\partial N} \right)_{v(\mathbf{r})}^- \quad (11)$$

The link to frontier molecular orbital theory is obtained by neglecting the orbital relaxation terms, so that

$$f^+(\mathbf{r}) \approx |\phi_{\text{LUMO}}(\mathbf{r})|^2 \quad (12)$$

$$f^-(\mathbf{r}) \approx |\phi_{\text{HOMO}}(\mathbf{r})|^2 \quad (13)$$

C. Theoretical Description of Electrostatic and Electron-Transfer Effects

We have not yet provided a mathematical reason, comparable to Eq. (3), for interpreting the Fukui function as a reactivity indicator. To do this, consider how the energy of a molecule changes in response to an attacking electrophile. The attacking electrophile will take electrons from the molecule ($\Delta N_{\text{molecule}} < 0$). Additionally, the electrons that remain in the molecule will be perturbed by the presence of the reagent: that is, the electrons in the nucleophile will feel an additional external potential due to the electrons and nuclei of the electrophile. Combining these two effects, the change in the total energy of the molecule undergoing attack is

$$\begin{aligned} \Delta U = & \int \frac{\delta V_m[v]}{\delta v(\mathbf{r})} \Delta v(\mathbf{r}) d\mathbf{r} + \left(\frac{\partial E}{\partial N} \right)_{v(\mathbf{r})}^- \Delta N \\ & + \int \left(\frac{\delta E_v[v; N]}{\delta v(\mathbf{r})} \right)_N \Delta v(\mathbf{r}) d\mathbf{r} + \left[\int \left(\frac{\partial}{\partial N} \left(\frac{\delta E_v[v; N]}{\delta v(\mathbf{r})} \right) \right)_{v(\mathbf{r})}^- \Delta v(\mathbf{r}) d\mathbf{r} \right] \Delta N + \dots \end{aligned} \quad (14)$$

Here we are using U to denote the potential energy surface for the atomic nuclei. Within the Born-Oppenheimer approximation, $U = E + V_{nn}$; that is, U is sum of the electronic energy and the nuclear-nuclear repulsion energy.

The terms in Eq. (14) represent (a) the change in nuclear-nuclear repulsion energy, (b) the change in electronic energy due to electron transfer, (c) the change in electronic energy due to the change in external potential, and (d) the cross term linking electron transfer to changes in external potential. We will work within an exact formulation for the isolated system (e.g., the zero-temperature grand canonical ensemble³²), so second and higher-order derivatives with respect to the number of electrons vanish.^{32,36} We are neglecting terms including higher-order responses to the external potential even though the first term (which models the polarization of the system by the approaching electrophile) might be important. Polarization effects are commonly neglected in the DFT-based approach to chemical reactivity. This is mostly because it is difficult to compute the polarizability kernel;³⁷ the success of DFT-based reactivity methods even when the polarization term is neglected suggests that the approximation in Eq. (14) is often sufficient for qualitative considerations. This success may be rationalized by noting that the dominant contribution to polarization arises from induced dipoles on the atomic centers, and the resulting interaction is short-ranged and weak³⁸ compared to the interaction from atomic charges. As such, the polarization contribution will ordinarily be an order of magnitude smaller than the energetic contributions from the molecular electrostatic potential, so it is often negligible for qualitative and semiquantitative purposes. Further evidence for this assertion comes from the broad (but not universal!) success of classical molecular dynamics that do not include polarization.

Subject to the approximations inherent in Eq. (14), the change in the energy of the nucleophile due to the approaching electrophile is

$$\begin{aligned} \Delta U_{\text{nucleophile}} &= \left(-I_{\text{nucleophile}} \right) \Delta N \\ &\quad - \int \left(\sum_{\alpha \in \text{nucleophile}} \left(Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha}) \right) - \rho_{\text{nucleophile}}(\mathbf{r}) - \Delta N f_{\text{nucleophile}}^{-}(\mathbf{r}) \right) \Delta v(\mathbf{r}) d\mathbf{r} \end{aligned} \quad (15)$$

Here $I_{\text{nucleophile}}$ is the vertical ionization potential of the molecule, $\{Z_{\alpha}\}$ and $\{\mathbf{R}_{\alpha}\}$ denote the nuclear charges and their positions, and the summation is over all the atomic nuclei in the nucleophile. Since the first term in Eq. (15) depends only on ΔN and not on the position of the electrophile, the second term controls the regioselectivity of the reaction.

There is an exact model for $\Delta v(\mathbf{r})$ but it is not useful for computational purposes.^{39;40} For reactants that are far apart, exchange and correlation between electrons on different subsystems should be negligible, and so we can approximate $\Delta v(\mathbf{r})$ with minus the electrostatic potential of the attacking reagent.^{40;41} The resulting model for the change in external potential at the point \mathbf{r}_p is

$$\begin{aligned} \Delta v(\mathbf{r}_p) &\approx - \int \frac{\sum_{\beta \in \text{electrophile}} Z_{\beta} \delta(\mathbf{r} - \mathbf{R}_{\beta}) - (\rho_{\text{electrophile}}(\mathbf{r}) + (-\Delta N) f_{\text{electrophile}}^+(\mathbf{r}))}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \\ &= -\Phi_{\text{electrophile}}(\mathbf{r}_p) - \Delta N \int \frac{f_{\text{electrophile}}^+(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \end{aligned} \quad (16)$$

The first term in the second line is the electrostatic potential of the isolated electrophile and the second term is the correction to the electrostatic potential from electron transfer to the electrophile. In our model, $\Delta N < 0$ is the change in the number of electrons on the nucleophilic substrate, so the change in electron number for the electrophilic reagent is $-\Delta N$, which is greater than zero. This sign convention anticipates the second paper of this series,⁴² where we will use these results to describe where nucleophiles are most susceptible to electrophilic attack.

Just as the energy of the nucleophile is changed by the approaching electrophile, the energy of the electrophile is changed by the presence of the nucleophile. In analogy to Eq. (15), one has

$$\begin{aligned} \Delta U_{\text{electrophile}} &= (-A_{\text{electrophile}})(-\Delta N) \\ &- \int \left(\sum_{\beta \in \text{electrophile}} Z_{\beta} \delta(\mathbf{r} - \mathbf{R}_{\beta}) - \rho_{\text{electrophile}}(\mathbf{r}) - (-\Delta N) f_{\text{electrophile}}^+(\mathbf{r}) \right) \Delta v(\mathbf{r}) d\mathbf{r} \end{aligned} \quad (17)$$

Here the change in external potential is due to the electrons and nuclei in the nucleophile. Again, we approximate the change in external potential with the negative electrostatic potential of the nucleophile as corrected for electron transfer,

$$\begin{aligned} \Delta v(\mathbf{r}_p) &\approx - \int \frac{\sum_{\alpha \in \text{nucleophile}} Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha}) - (\rho_{\text{nucleophile}}(\mathbf{r}) + \Delta N f_{\text{nucleophile}}^-(\mathbf{r}))}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \\ &= -\Phi_{\text{nucleophile}}(\mathbf{r}_p) + \Delta N \int \frac{f_{\text{nucleophile}}^-(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \end{aligned} \quad (18)$$

The replacement of ΔN in Eqs. (15) and (16) with $-\Delta N$ in Eqs. (17) and (18) is dictated by charge conservation: electrons are transferred from the nucleophile to

the electrophile. Equations (16) and (18) lead to the identification of the electrostatic potential,

$$\Phi(\mathbf{r}_p) = \int \frac{\sum_{\beta \in \text{electrophile}} Z_{\beta} \delta(\mathbf{r} - \mathbf{R}_{\beta}) - \rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \quad (19)$$

and the Fukui potential

$$v^f(\mathbf{r}_p) = \int \frac{f(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \quad (20)$$

as key reactivity indicators. Clearly the Fukui potential is only relevant in cases where electron transfer is important. The Fukui potential also plays a key role in the reactivity model proposed by Berkowitz.⁴³

Combining Eqs. (15) and (17) and correcting for the double-counting of interactions gives a model for the interaction energy between the nucleophilic substrate and the electrophilic reagent:

$$\begin{aligned} U_{\text{int}} = & \left(A_{\text{electrophile}} - I_{\text{nucleophile}} \right) \Delta N \\ & + \int \left(\sum_{\alpha \in \text{nucleophile}} Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha}) - \rho_{\text{nucleophile}}(\mathbf{r}) \right) \Phi_{\text{electrophile}}(\mathbf{r}) d\mathbf{r} \\ & + \Delta N \int \left(f_{\text{electrophile}}^+(\mathbf{r}) \Phi_{\text{nucleophile}}(\mathbf{r}) - f_{\text{nucleophile}}^-(\mathbf{r}) \Phi_{\text{electrophile}}(\mathbf{r}) \right) d\mathbf{r} \\ & - (\Delta N)^2 \iint \frac{f_{\text{nucleophile}}^-(\mathbf{r}) f_{\text{electrophile}}^+(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \end{aligned} \quad (21)$$

The first term in this equation is a constant, and does not affect site selectivity. The next term reflects the electrostatic interactions between electrophiles and nucleophiles: because the active sites of electrophiles are usually positively charged while the reactive sites of nucleophiles are usually negatively charged, this term is usually negative. That is, electrostatic effects are usually attractive. This is in marked contrast to the terms in the third line of Eq. (21). Because $\Delta N < 0$, if $f_{\text{electrophile}}^+(\mathbf{r}) > 0$ at the reactive site of the electrophile and the reactive site of the nucleophile is negatively charged (so that $\Phi_{\text{nucleophile}}(\mathbf{r}) < 0$), then the first term in the third line of Eq. (21) is usually positive. Similarly, the second term in the third line of Eq. (21) is usually positive because $f_{\text{nucleophile}}^-(\mathbf{r}) > 0$ at the active site of the nucleophile and the active site of the electrophile is usually positively charged (so that $\Phi_{\text{electrophile}}(\mathbf{r}) > 0$). The third line of Eq. (21), then, reflects the fact that charge transfer from the nucleophile to the electrophile helps to equalize the charges of the reagents, which reduces the electrostatic attraction between them. The fourth line of Eq. (21) will generally be negative. The expression in the fourth line plays a key role in the theory of electron transfer

proposed by Berkowitz.⁴⁴ Of all the terms in Eq. (21), this is the term that most closely resembles the results from frontier molecular orbital theory. In fact, an expression very similar to Eq. (4) can be derived:

$$\begin{aligned}
 \iint \frac{f_{\text{nucleophile}}^-(\mathbf{r}) f_{\text{electrophile}}^+(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' &\leq \left\| \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right\| \left| \int f_{\text{nucleophile}}^-(\mathbf{r}) f_{\text{electrophile}}^+(\mathbf{r}) d\mathbf{r} \right| \\
 &\approx \left\| \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right\| \int |\phi_{\text{HOMO}}^{(B)}(\mathbf{r})|^2 |\phi_{\text{LUMO}}^{(A)}(\mathbf{r})|^2 d\mathbf{r} \\
 &\leq \left\| \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right\| \left| \int (\phi_{\text{HOMO}}^{(B)}(\mathbf{r}))^* \phi_{\text{LUMO}}^{(A)}(\mathbf{r}) d\mathbf{r} \right|^2 \\
 &\leq \left\| \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right\| \left(\int |(\phi_{\text{HOMO}}^{(B)}(\mathbf{r}))^* \phi_{\text{LUMO}}^{(A)}(\mathbf{r})| d\mathbf{r} \right)^2
 \end{aligned} \tag{22}$$

Reactants will approach each other in a way that minimizes the interaction energy. That is, the more negative the interaction energy, the stronger the attraction between reagents and the greater their susceptibility to reaction. Coulson and Longuet-Higgins used perturbation theory and molecular orbital theory to derive a formula for the interaction energy between two separated reagents; this led them to Eq. (1). Equation (21) is a just a density-functional theory inspired reformulation of perturbation theory about the separated reagent limit. Reactivity indicators based on these formulae might be unreliable when the transition state occurs late in the reaction, because in those cases the molecular geometry in the transition state may not resemble the isolated reagents.

Some readers may find it surprising that the electrostatic potential enters into Eq. (21) in such a natural way. Traditionally, the electrostatic potential has not been considered a reactivity index associated with conceptual DFT.⁴⁵ However, it is evident from the preceding analysis that whenever the change in the potential energy surface due to the electrostatic potential is addressed,⁴⁶

$$\int \frac{\delta U}{\delta v(\mathbf{r})} \delta v(\mathbf{r}) d\mathbf{r}, \tag{23}$$

the electrostatic potential enters into conceptual DFT in a very natural way. Evaluating expressions like Eq. (23) requires evaluating the functional derivative of the nuclear-nuclear repulsion energy with respect to changes in external potential. Evaluating that functional derivative is a nontrivial mathematical exercise; details can be found in Appendix A. This analysis helps provide a theoretical foundation for recent work relating “conventional” DFT-based reactivity indicators to the electrostatic potential.^{47;48}

III.IV. A General-Purpose Model for Chemical Reactivity

A. The Reactive Site Interaction Model

In general, the regioselective preferences of a nucleophile undergoing electrophilic attack are preserved across a broad range of electrophilic partners. Since the specific identity of the electrophile is not critical, we can replace the electrophile by a “model perturbation.”^{14;49;50} Presuming that the reactivity is dominated by the properties at the active site, we will represent the electrostatic potential of the electrophile as the electrostatic potential due to the charge on the reactive site

$$\Phi_{\text{electrophile}}(\mathbf{r}) \approx \frac{q_{\text{electrophile}}^{(0)}}{|\mathbf{r} - \mathbf{R}_{\text{electrophile}}|} \quad (24)$$

and, similarly, replace the Fukui function with its “condensed” value,^{51;52}

$$\begin{aligned} f_{\text{electrophile}}^+(\mathbf{r}) &\approx f_{\text{electrophile}}^{(+)} \delta(\mathbf{r} - \mathbf{R}_{\text{electrophile}}) \\ &= (q_{\text{electrophile}}^{(0)} - q_{\text{electrophile}}^{(+)}) \delta(\mathbf{r} - \mathbf{R}_{\text{electrophile}}) \end{aligned} \quad (25)$$

Here, $q_{\text{electrophile}}^{(0)}$ and $q_{\text{electrophile}}^{(+)}$ denote the effective charges on the reactive site of the electrophile and the electrophile with an additional electron. If the Fukui function is approximated by Eq. (25), then the equation for the Fukui potential becomes:

$$v_{\text{electrophile}}^{f^+}(\mathbf{r}) \approx \frac{f_{\text{electrophile}}^{(+)}}{|\mathbf{r} - \mathbf{R}_{\text{electrophile}}|} = \frac{q_{\text{electrophile}}^{(0)} - q_{\text{electrophile}}^{(+)}}{|\mathbf{r} - \mathbf{R}_{\text{electrophile}}|} \quad (26)$$

Models based on Eqs. (24)-(26) should be accurate for atomic cations (where $q_{\text{electrophile}}^{(0)} = f_{\text{electrophile}}^{(+)} = +1$) and reasonable for molecular electrophiles, though in that case it will be important that the “effective charge” and “effective condensed Fukui function” on the reactive site might not equal the atomic quantities, but might instead represent a partial sum over several atoms in the vicinity of the reactive site. With that caveat, this model should be reasonably realistic whenever the reagents are very far apart, so that the overlaps between their electron distributions and Fukui functions are negligible. Using these approximations, the interaction energy between a nucleophile and a model electrophile whose reactive site is (a) located at the point \mathbf{r}_p , (b) has charge $q_{\text{electrophile}}^{(0)}$, and (c) has condensed Fukui function $f_{\text{electrophile}}^{(+)}$ becomes

$$\begin{aligned}
 U_{\text{int}}(\mathbf{r}_p) \approx & \left(A_{\text{electrophile}} - I_{\text{nucleophile}} \right) \Delta N \\
 & + \left(q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} \right) \Phi_{\text{nucleophile}}(\mathbf{r}_p) \\
 & - \Delta N \left(q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} \right) v_{\text{nucleophile}}^{f^-}(\mathbf{r}_p).
 \end{aligned} \tag{27}$$

This “single reactive site interaction model” is easily generalized to cases where multiple reactive sites need to be considered: simply sum over the effective charges and effective condensed Fukui functions⁵¹ on all the relevant sites,

$$\begin{aligned}
 U_{\text{int}} \approx & \left(A_{\text{electrophile}} - I_{\text{nucleophile}} \right) \Delta N \\
 & + \sum_{\beta \in \text{electrophile}} \sum_{\alpha \in \text{nucleophile}} \frac{\left(q_{\beta}^{(0)} + \Delta N f_{\beta}^{(+)} \right) q_{\alpha}^{(0)}}{|\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|} \\
 & - \Delta N \sum_{\beta \in \text{electrophile}} \sum_{\alpha \in \text{nucleophile}} \frac{\left(q_{\beta}^{(0)} + \Delta N f_{\beta}^{(+)} \right) f_{\alpha}^{(-)}}{|\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|}
 \end{aligned} \tag{28}$$

Again, smaller (more negative) U_{int} represents greater attraction between reagents and indicates greater reactivity.

Insofar as we are modeling the attacking reagent with a point charge, these expressions for $U_{\text{int}}(\mathbf{r}_p)$ are similar to the indicator of Brønsted-Lowry acidity introduced in reference⁵³, though the Taylor series expansion in that work includes higher-order terms in the external potential and does not include the electron transfer contribution. The present model is not restricted to charges of unit magnitude, and in that sense is more like the single-interaction site point charge model that recently was used to elucidate the Hard/Soft Acid/Base principle.⁵⁴ More generally, models resembling this one are commonly encountered when the so-called perturbative perspective on conceptual density-functional theory is utilized.^{14;37;50;52;55;56} This sort of analysis, with its fundamental link to the Taylor series expansion to the energy and the use of “model perturbations” to define reactivity indicators, grew out of the work of De Proft, Liu, Nalewajski, Parr and Senet, among others.^{31;57-61}

B. Deriving A General-Purpose Reactivity Indicator for Nucleophiles

Equation (27) provides the basis for a general purpose reactivity indicator for nucleophiles. Note, first of all, that the key parameter,

$$\tilde{\kappa} \equiv q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} \tag{29}$$

modulates the electrostatic and Fukui function contributions to the interaction energy. Insert the definition of $\tilde{\kappa}$ into Eq. (27), and note that the first term in Eq. (27) does not depend on the position, \mathbf{r}_p . One obtains

$$\tilde{\Xi}_{\Delta N \leq 0}^{\tilde{\kappa}}(\mathbf{r}_p) \equiv \tilde{\kappa} \left(\Phi_{\text{nucleophile}}(\mathbf{r}_p) - \Delta N v_{\text{nucleophile}}^f(\mathbf{r}_p) \right), \quad (30)$$

which is a regioselectivity indicator for electrophilic attack. A nucleophile will be most reactive in places where $\tilde{\Xi}_{\Delta N \leq 0}^{\tilde{\kappa}}(\mathbf{r}_p)$ is small (most negative) because electrophilic attack at those sites is energetically favorable.

Equation (30) uses the reaction site interaction model to provide simple approximations to the complicated integral expressions in Eq. (21). But this is *too* simple. In particular, the reactive site interaction model is more accurate for some integrals than it is for others. The model typically underestimates the integrals because interactions between the asymptotic tails of the electron densities and the Fukui functions are neglected. Moreover, this error is largest for the integrals containing Fukui functions (which are concentrated on the periphery of the reagents). This implies that the reactive site interaction model is better at describing the interaction of the electrophile with the electrostatic potential (which contains the potential due to the electron density) than it is at describing the interaction of the electrophile with the Fukui potential. If we correct Eq. (30) for the errors associated with the reaction site interaction model, then we obtain an expression of the form,

$$\tilde{\Xi}_{\Delta N \leq 0}^{\tilde{\kappa}}(\mathbf{r}_p) \equiv (\tilde{\kappa} + \varepsilon_\rho) \Phi_{\text{nucleophile}}(\mathbf{r}_p) - (\tilde{\kappa} + \varepsilon_f) \Delta N v_{\text{nucleophile}}^f(\mathbf{r}_p). \quad (31)$$

Here ε_ρ corrects for the errors incurred by the reactive-site approximation for the electron density and ε_f corrects for the error incurred by the reactive-site approximation for the Fukui function. According to the preceding arguments,

$$\varepsilon_f < \varepsilon_\rho. \quad (32)$$

Usually both ε_f and ε_ρ are negative. Appendix B provides a full accounting of the error terms.

We cannot evaluate these error terms within the context of the reactive site interaction model. It is certainly true that the errors are small when the molecules are far apart and that the error “correction” terms dominate when the molecules are close together. When the molecules are in van-der-Waals contact, one suspects that the error terms make significant corrections to $\tilde{\kappa}$.

Our goal is to derive a *qualitative* indicator of chemical reactivity. Note that

- $\tilde{\kappa} \equiv q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)}$ has units of electric charge because the charge and the condensed Fukui function of the electrophile's reactive site both have units of electric charge.
- $\tilde{\kappa}$ is of order unity if one measures it in terms of the magnitude of charge on the electron, e . This is because e is the natural unit for expressing the charge and the condensed Fukui function of the electrophile's reactive site.
- We can eliminate the unknown error terms by introducing a new scale of electric charge, with the new unit of charge defined by

$$1 = \frac{\varepsilon_\rho - \varepsilon_f}{2}. \quad (33)$$

We then define

$$\kappa = \tilde{\kappa} + \varepsilon_\rho - 1. \quad (34)$$

This definition “sets the zero” of a scale for κ . Specifically, $\kappa = 0$ occurs when electrostatic and electron-transfer effects are perfectly balanced.

Using these relations, we can eliminate the unknown error terms from Eq. (31). This results in a general-purpose regioselectivity indicator for nucleophiles:

$$\Xi_{\Delta N \leq 0}^\kappa(\mathbf{r}_p) \equiv (\kappa + 1)\Phi_{\text{nucleophile}}(\mathbf{r}_p) - (\kappa - 1)\Delta N v_{\text{nucleophile}}^f(\mathbf{r}_p). \quad (35)$$

There is another way to think derive Eq. (35): starting with Eq. (31), choose the unit of energy and the zero of energy so that the unknown error terms are eliminated. (Notice, however, that energy scale one defines depends what position, \mathbf{r}_p , in the molecule is being considered.)

Equation (35) is our general-purpose reactivity indicator. Because $\Xi_{\Delta N \leq 0}^\kappa(\mathbf{r}_p)$ is a qualitative measure of the interaction energy of the nucleophile with an electrophile at the point \mathbf{r}_p , the nucleophile will be most reactive where $\Xi_{\Delta N \leq 0}^\kappa(\mathbf{r}_p)$ is most negative.

Equation (35) is the most conceptually transparent form for our indicator. However, for computational applications, we find it convenient to compute the indicator from the electrostatic potential of the nucleophile, $\Phi_{\text{nucleophile}}^{(0)}(\mathbf{r})$, and the electrostatic potential of the nucleophile with one electron removed, $\Phi_{\text{nucleophile}}^{(-)}(\mathbf{r})$:

$$\Xi_{\Delta N \leq 0}^\kappa(\mathbf{r}_p) = (1 + \kappa + \Delta N(\kappa - 1))\Phi_{\text{nucleophile}}^{(0)}(\mathbf{r}_p) - \Delta N(\kappa - 1)\Phi_{\text{nucleophile}}^{(-)}(\mathbf{r}_p). \quad (36)$$

Using this expression, our reactivity indicator can be evaluated using any popular quantum chemistry package. Remember that $\Xi_{\Delta N \leq 0}^\kappa(\mathbf{r}_p)$ models the interaction

energy, so the nucleophile is most reactive where $\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}_p)$ has the smallest (most negative) values.

The key to deriving Eq. (36) is to note that the Fukui potential (and the Fukui function) can be computed from the electrostatic potentials of the nucleophile, $\Phi_{\text{nucleophile}}^{(0)}(\mathbf{r})$ and the electrostatic potential of the nucleophile with one electron removed, $\Phi_{\text{nucleophile}}^{(-)}(\mathbf{r})$. Specifically,

$$\begin{aligned} v_{\text{nucleophile}}^{f^-}(\mathbf{r}) &\equiv \int \frac{f_{\text{nucleophile}}^-(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_p|} d\mathbf{r} = \int \frac{\rho_{\text{nucleophile}}^{(0)}(\mathbf{r}) - \rho_{\text{nucleophile}}^{(-)}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_p|} d\mathbf{r} \\ &= \Phi_{\text{nucleophile}}^{(-)}(\mathbf{r}_p) - \Phi_{\text{nucleophile}}^{(0)}(\mathbf{r}_p) \\ f_{\text{nucleophile}}^-(\mathbf{r}) &= \frac{1}{4\pi} \nabla^2 \left(\Phi_{\text{nucleophile}}^{(0)}(\mathbf{r}) - \Phi_{\text{nucleophile}}^{(-)}(\mathbf{r}) \right) \end{aligned} \quad (37)$$

Equation (35) or, equivalently, Eq. (36), is our general-purpose reactivity model for nucleophiles. Because of the way we have accommodated the error terms in Eq. (31), this is only a qualitative model for reactivity. One could derive other models with qualitatively similar behavior but somewhat different functional forms. We selected this form because it has an appealing symmetry and because it is easy to interpret (using Eq. (35)) and apply (using Eq. (36)).

$\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}_p)$ depends on two parameters, ΔN (measuring the extent of electron donation) and κ (measuring the relative importance of electrostatic effects and electron-transfer effects). $\kappa = 1$ corresponds to pure electrostatic control; $\kappa = -1$ corresponds to pure electron-transfer (or Fukui-function) control. $\kappa = 0$ corresponds to a perfect balance between electrostatic and electron-transfer control.

The change of units that accompanies the elimination of the error terms means that $\kappa \neq \tilde{\kappa}$. However, if the error terms are not too big, κ will be approximately proportional to $\tilde{\kappa}$. We denote this $\kappa \propto \tilde{\kappa}$. This insight gives us a working approximation for κ (cf. Eq. (29)), namely,

$$\kappa \propto \tilde{\kappa} \frac{q_{\text{electrophile}}^{(0)}}{|\Delta N| f_{\text{electrophile}}^{(+)}} \quad (38)$$

When the charge on the electrophile is very large, we still expect to see electrostatically controlled reactivity. Similarly, when the charge on the electrophile's reactive site is small compared with the extent of electron transfer to the reactive site ($q_{\text{electrophile}}^{(0)} \leq |\Delta N| f_{\text{electrophile}}^{(+)}$), the reaction is electron-transfer controlled.

Like $\tilde{\kappa}$, κ has units of electric charge and order of magnitude unity. Equation (38) is an approximate proportionality and it is useful for elucidating how a nucleophile's reactivity depends on the charge and condensed Fukui function of an electrophile's reactive site. However, Eq. (38) is not a quantitative formula for κ . Determining the value of κ that is most appropriate for a given reaction is sensitive to (a) the inherent errors in the reactive site interaction model and (b) molecular polarization and other effects that are neglected in this analysis. However, Eq. (38) should be sufficient to indicate whether κ is "large and positive", "small and positive", "almost zero", "small and negative, or "large and negative." Section III.F. contains a detailed discussion of each of these cases.

C. A General-Purpose Reactivity Indicator for Electrophiles

There is clearly an analogous indicator for electrophiles. Specifically, one has

$$\Xi_{\Delta N \geq 0}^{\kappa}(\mathbf{r}_p) \equiv -(\kappa + 1)\Phi_{\text{electrophile}}(\mathbf{r}_p) + \Delta N(\kappa - 1)v_{\text{electrophile}}^{f^+}(\mathbf{r}_p), \quad (39)$$

where the Fukui potential for the electrophile is given by

$$\begin{aligned} v_{\text{electrophile}}^{f^+}(\mathbf{r}) &\equiv \int \frac{f_{\text{electrophile}}^+(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} = \int \frac{\rho_{\text{electrophile}}^{(+)}(\mathbf{r}) - \rho_{\text{electrophile}}^{(0)}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \\ &= \Phi_{\text{electrophile}}^{(0)}(\mathbf{r}_p) - \Phi_{\text{electrophile}}^{(+)}(\mathbf{r}_p) \\ f_{\text{electrophile}}^+(\mathbf{r}) &= \frac{1}{4\pi} \nabla^2 \left(\Phi_{\text{electrophile}}^{(+)}(\mathbf{r}) - \Phi_{\text{electrophile}}^{(-)}(\mathbf{r}) \right) \end{aligned} \quad (40)$$

The appropriate values of κ are the same as before: $\kappa \geq 1$ (electrostatic control); $\kappa \leq -1$ (electron-transfer control), and $-1 < \kappa < 1$ (intermediate). Corresponding to Eq. (38), we can say that κ is approximately proportional to $-q_{\text{nucleophile}}^{(0)} - \Delta N f_{\text{nucleophile}}^{(-)}$. Here $q_{\text{nucleophile}}^{(0)}$ is the charge on the reactive site of the nucleophile, and is typically negative. $f_{\text{nucleophile}}^{(-)}$ is the condensed Fukui function from below at the reactive site of the nucleophile. It would be shocking to observe a negative value for $f_{\text{nucleophile}}^{(-)}$.^{52;62-66}

D. The Condensed General-Purpose Reactivity Indices

Because local reactivity indicators vary on a point-by-point basis, it is often convenient to "condense" their values to atomic sites.^{51;52} A condensed indicator related to $\Xi_{\Delta N}^{\kappa}(\mathbf{r}_p)$ follows directly from the fitting of the electrostatic potential to atomic charges. Specifically, the electrostatic potential can be expanded in an asymptotic series of atomic multipoles. Truncating the multipole expansion after

the monopoles yields an expression for the electrostatic potential in terms of atomic charges,

$$\begin{aligned}\Phi_{\text{nucleophile}}^{(0)}(\mathbf{r}) &\sim \sum_{\alpha \in \text{nucleophile}} \frac{q_{\text{nucleophile},\alpha}^{(0)}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \\ \Phi_{\text{nucleophile}}^{(-)}(\mathbf{r}) &\sim \sum_{\alpha \in \text{nucleophile}} \frac{q_{\text{nucleophile},\alpha}^{(-)}}{|\mathbf{r} - \mathbf{R}_{\alpha}|}\end{aligned}\quad (41)$$

These expressions are not very accurate close to the molecule. For example, these expressions do not reproduce the correct singularity at the atomic nuclei. However, the relevant values of a local reactivity indicator occur far from the molecule, on a “reactivity surface” that represents how closely two reagents can approach each other and still retain their separate identities. Not only do expressions like Eq. (41) suffice for this purpose, condensed expressions are actually preferable because they “average over” the irrelevant fine structure of the spatially varying indicators.⁵²

“Condensed” expressions for the Fukui potential are easily constructed from Eqs. (37), (40), and (41),

$$\begin{aligned}v_{\text{nucleophile}}^{f^-}(\mathbf{r}) &\sim \sum_{\alpha \in \text{nucleophile}} \frac{f_{\text{nucleophile},\alpha}^-}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \\ f_{\text{nucleophile},\alpha}^- &= q_{\text{nucleophile},\alpha}^{(-)} - q_{\text{nucleophile},\alpha}^{(0)}\end{aligned}\quad (42)$$

$$\begin{aligned}v_{\text{electrophile}}^{f^+}(\mathbf{r}) &\sim \sum_{\alpha \in \text{electrophile}} \frac{f_{\text{electrophile},\alpha}^+}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \\ f_{\text{electrophile},\alpha}^+ &= q_{\text{electrophile},\alpha}^{(0)} - q_{\text{electrophile},\alpha}^{(+)}\end{aligned}\quad (43)$$

f_{α}^{\pm} are called the condensed Fukui functions.⁵¹ We can now write asymptotic expressions for the reactivity indicators proposed in this paper. For nucleophiles,

$$\begin{aligned}\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}_p) &\sim \sum_{\alpha \in \text{nucleophile}} \frac{\Xi_{\Delta N \leq 0,\alpha}^{\kappa}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \\ \Xi_{\Delta N \leq 0,\alpha}^{\kappa} &= (\kappa + 1)q_{\text{nucleophile},\alpha}^{(0)} - \Delta N(\kappa - 1)f_{\text{nucleophile},\alpha}^- \\ &= (1 + \kappa + \Delta N(\kappa - 1))q_{\text{nucleophile},\alpha}^{(0)} - \Delta N(\kappa - 1)q_{\text{nucleophile},\alpha}^{(-)}.\end{aligned}\quad (44)$$

For electrophiles,

$$\begin{aligned}\Xi_{\Delta N \geq 0}^{\kappa}(\mathbf{r}_p) &\sim \sum_{\alpha \in \text{electrophile}} \frac{\Xi_{\Delta N \geq 0,\alpha}^{\kappa}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \\ \Xi_{\Delta N \geq 0,\alpha}^{\kappa} &= -(\kappa + 1)q_{\text{electrophile},\alpha}^{(0)} + \Delta N(\kappa - 1)f_{\text{electrophile},\alpha}^+ \\ &= (\Delta N(\kappa - 1) - \kappa - 1)q_{\text{electrophile},\alpha}^{(0)} - \Delta N(\kappa - 1)q_{\text{electrophile},\alpha}^{(+)}. \end{aligned}\quad (45)$$

Nucleophiles will be susceptible to electrophilic attack at the atomic sites where $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ is small, and ideally negative. Electrophiles will be susceptible to nucleophilic attack at the atomic sites where $\Xi_{\Delta N \geq 0, \alpha}^{\kappa}$ is small.

E. Single-Parameter Variants of the General-Purpose Reactivity Indicator

Because our general-purpose reactivity indicator, $\Xi_{\Delta N}^{\kappa}(\mathbf{r})$, depends on two parameters (κ and ΔN), it is sometimes difficult to visualize the wealth of information it contains. In some contexts, then, it would be convenient to use a variant of $\Xi_{\Delta N}^{\kappa}(\mathbf{r})$ that depended only on the amount of electron transfer, since that is easily computed using electronegativity equalization schemes. Since the appropriate value of κ varies depending on the amount of electron transfer, one could introduce a one-parameter model by setting

$$\kappa \sim \begin{cases} 1 + 2(\Delta N) & \Delta N \leq 0 \\ 1 - 2(\Delta N) & \Delta N \geq 0 \end{cases} \quad (46)$$

This model correctly predicts electrostatic control when $\Delta N \approx 0$ and electron-transfer control when $\Delta N \approx \pm 1$. Using Eq. (46), one obtains one-parameter models for the reactivity of nucleophiles

$$\Xi_{\Delta N \leq 0}^{1+2\Delta N}(\mathbf{r}_p) = 2 \left[(1 + \Delta N) \Phi_{\text{nucleophile}}(\mathbf{r}_p) - (\Delta N)^2 v_{\text{nucleophile}}^{f^-}(\mathbf{r}_p) \right] \quad (47)$$

and electrophiles

$$\Xi_{\Delta N \geq 0}^{1-2\Delta N}(\mathbf{r}_p) \equiv 2 \left[(\Delta N - 1) \Phi_{\text{electrophile}}(\mathbf{r}_p) - (\Delta N)^2 v_{\text{electrophile}}^{f^+}(\mathbf{r}_p) \right]. \quad (48)$$

While these single-parameter models are simpler than their two-parameter counterparts, our preliminary investigations indicate that in molecules with multiple reactive sites, the single parameter models sometimes fail to identify one or more reactive sites. For this reason, we will focus on the more general two-parameter models.

F. General-Purpose Reactivity Indicators: Discussion

We now discuss the interpretation of the general-purpose reactivity indicators for nucleophiles,

$$\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}) \equiv (\kappa + 1) \Phi_{\text{nucleophile}}(\mathbf{r}) - \Delta N (\kappa - 1) v_{\text{nucleophile}}^{f^-}(\mathbf{r}), \quad (49)$$

and electrophiles,

$$\Xi_{\Delta N \geq 0}^{\kappa}(\mathbf{r}) \equiv -(\kappa + 1) \Phi_{\text{electrophile}}(\mathbf{r}) + \Delta N (\kappa - 1) v_{\text{electrophile}}^{f^+}(\mathbf{r}). \quad (50)$$

These indicators are designed to reproduce the qualitative features of the molecular interaction energy expression that we derived in section II.C., namely,

$$\begin{aligned}
 U_{\text{int}} = & \left(A_{\text{electrophile}} - I_{\text{nucleophile}} \right) \Delta N \\
 & + \int \left(\sum_{\alpha \in \text{nucleophile}} Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha}) - \rho_{\text{nucleophile}}(\mathbf{r}) \right) \Phi_{\text{electrophile}}(\mathbf{r}) d\mathbf{r} \\
 & + \Delta N \int \left(f_{\text{electrophile}}^{+}(\mathbf{r}) \Phi_{\text{nucleophile}}(\mathbf{r}) - f_{\text{nucleophile}}^{-}(\mathbf{r}) \Phi_{\text{electrophile}}(\mathbf{r}) \right) d\mathbf{r} \quad (51) \\
 & - (\Delta N)^2 \iint \frac{f_{\text{nucleophile}}^{-}(\mathbf{r}) f_{\text{electrophile}}^{+}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'
 \end{aligned}$$

Negative values of the interaction energy, U_{int} , are associated with favorable interactions between the electrophile and the nucleophile. Similarly, molecules will be most highly reactive where $\Xi_{\Delta N}^{\kappa}(\mathbf{r})$ is the smallest.

In accord with a suggestion by Langenaeker, De Proft, and Geerlings,⁶⁷ each of our reactivity indicators is a linear combination of an appropriate reactivity index for hard-hard interactions (the electrostatic potential) and an appropriate reactivity index for soft-soft interactions (the Fukui function). The relative importance of these two contributions is controlled by the extent of charge transfer and the charge on the reactive site of the attacking reagent, as it should be. These are the key ingredients that make $\Xi_{\Delta N}^{\kappa}(\mathbf{r}_p)$ a true general-purpose reactivity indicator: it can model the extreme cases of electrostatic and electron-transfer control, but it can also model the “in between” cases. We will now explore each of these cases in more detail. In the interest of specificity and brevity, the following discussion is focused on the reactivity of nucleophiles. The results for electrophiles are broadly similar.

Case 1: Electrostatically Controlled Reactions; $\kappa \geq 1$

Electrostatically controlled (also called charge controlled) reactions are typified by:

- (a) The reagents have large charges (so $q_{\text{electrophile}}^{(0)} \gg 0$, $\Phi_{\text{nucleophile}}(\mathbf{r}_p) \ll 0$).
- (b) Reactivity does not necessarily occur in places where the Fukui functions are large. ($f_{\text{electrophile}}^{(+)}$ may not be very big.)
- (c) Charge transfer between reagents is minimal ($\Delta N \approx 0$).

Under these conditions, κ is greater than, or approximately equal to, one. We call the case $\kappa = 1$ “pure electrostatic control” and the case $\kappa > 1$ “strong electrostatic control.”

Case 1A: Pure Electrostatic Control; $\kappa = 1$.

When $\kappa = 1$,

$$\Xi_{\Delta N \leq 0}^{\kappa=1}(\mathbf{r}_p) = 2\Phi_{\text{nucleophile}}(\mathbf{r}_p) \quad (52)$$

Thus, for purely electrostatically controlled reactions, the appropriate reactivity indicator is the electrostatic potential. Furthermore, using the condensed version of the indicator,

$$\Xi_{\Delta N \leq 0, \alpha}^{\kappa=1} = 2q_{\text{nucleophile}, \alpha}^{(0)} \quad (53)$$

we infer that in the limit of pure electrostatic control, nucleophiles react at the most negatively charged atomic site. (For electrophiles, the most reactive sites are those where the electrostatic potential/atomic charge is most positive.)

Case 1B: Strong Electrostatic Control; $\kappa > 1$.

When $\kappa > 1$, the second term in Eq. (49) is usually very small because $\kappa - 1 < \kappa + 1$ and because the extent of electron transfer is very small ($\Delta N \approx 0$). For this reason, there can be no doubt that the most appropriate reactivity indicator for electrostatically controlled reactions is the electrostatic potential.^{24;68} What happens if there are two sites that are equivalent from an electrostatic perspective? Because $-\Delta N(\kappa - 1) > 0$ whenever $\kappa > 1$, $\Xi_{\Delta N \leq 0}^{\kappa > 1}$ will be smallest in those places where the Fukui potential is the smallest. Insofar as the Fukui potential is smallest in places where the Fukui function is also small, this suggests that: *for reactions that are strongly electrostatically controlled, if there are two reactive sites with the same electrostatic favorability, then the site with minimum Fukui function is the most favorable.* This rule can also be derived from the condensed version of our indicator. In that case: *For electrophilic attack occurring under conditions associated with strong electrostatic control of chemical reactivity, a nucleophile possessing two sites with equal negative charge will be most reactive at the site with smallest condensed Fukui function.* These rules might be called “Minimum Fukui Function Tie-Breaking Rules” because they indicate that if a reaction is strongly electrostatically controlled and there are two sites with equivalent electrostatic properties (and also other properties, notably including polarizability), then minimizing the Fukui function serves as a “tie-breaker” between the two equivalent sites.

In 1995, Li and Evans proposed that hard-hard interactions occur where the Fukui function is small.²³ This counterintuitive result started a debate in the literature.^{24;69-71} Our analysis establishes that it is true that sometimes having a small value for the Fukui function is favorable for reactivity, but the conditions under which this rule is valid are rather limited. There might be cases where the minimum Fukui function tie-breaker is decisive in determining the reactivity. However, the second term in Eq. (49) is usually *much smaller* than the first term, so the minimum Fukui function tie-breaker will *only* be operative when two hard reagents interact (so that $\Delta N \approx 0$), the electrophile is highly charged (so that

$q_{\text{electrophile}}^{(0)}$, and thus κ , is large), and the nucleophile has two reactive sites with *very similar* electrostatic profiles. Unless the possible reaction sites are very similar electrostatically, the weak dependence of $\Xi_{\Delta V \leq 0}^{\kappa > 1}$ on the Fukui function will not be decisive.

Even when it appears that the tie-breaking rule should apply, it may not. Because the dependence on the Fukui function is so weak, some of the interactions that were neglected in our model might be more important for determining the nucleophile's reactivity. The most important of these neglected interactions is probably the polarizability. Consider the Berkowitz-Parr formula for the polarizability kernel,^{37;72}

$$\left(\frac{\delta E}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} \right)_N = \left(\frac{\delta E}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} \right)_\mu + \left(\frac{\partial N}{\partial \mu} \right)_{v(\mathbf{r})} f(\mathbf{r}) f(\mathbf{r}'), \quad (54)$$

or the approximate formula for the polarizability kernel based on the Kohn-Sham noninteracting reference system^{37;73}

$$\begin{aligned} \left(\frac{\delta E}{\delta v(\mathbf{r}) \delta v(\mathbf{r}')} \right)_N &\approx \sum_i \sum_{j \neq i} \frac{n_j - n_i}{\epsilon_j - \epsilon_i} \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}') \\ &\approx \frac{-\phi_{\text{HOMO}}^*(\mathbf{r}) \phi_{\text{LUMO}}(\mathbf{r}) \phi_{\text{LUMO}}^*(\mathbf{r}') \phi_{\text{HOMO}}(\mathbf{r}')}{\epsilon_{\text{HOMO}} - \epsilon_{\text{LUMO}}} \\ &\quad - \frac{\phi_{\text{LUMO}}^*(\mathbf{r}) \phi_{\text{HOMO}}(\mathbf{r}) \phi_{\text{HOMO}}^*(\mathbf{r}') \phi_{\text{LUMO}}(\mathbf{r}')}{\epsilon_{\text{HOMO}} - \epsilon_{\text{LUMO}}} \end{aligned} \quad (55)$$

From these formulae, it is clear that a molecule tends to be most polarizable where the Fukui function (Eq. (54)) and the frontier orbitals (Eq. (55)) have a large amplitude. Given that strong electrostatic control will only occur for highly charged electrophiles, the second order response to the change in external potential might not be especially small. Consequently, polarization effects may well cancel out, or even reverse, the predictions that would be obtained by naïvely applying the minimum Fukui function tie-breaker.

The idea that the minimum Fukui function might correspond to the most reactive site goes against conventional wisdom, but it can be simply explained using the fundamental equation for the interaction energy, (51). Note that

- The first line of this equation is a constant, and does not influence regioselectivity.
- The second line of Eq. (51) models the electrostatic attraction between the electrophile and the nucleophile. If we are in a “tie-breaking” situation, this term is the same for electrophilic attack at both reactive sites. Thus, although the electrostatic interaction in the second line of this equation

makes the biggest contribution to the interaction energy, it doesn't determine the regioselectivity.

- Recall from our discussion of Eq. (51) that both terms on the third line are positive.
- The first term on the third line of (51) does not distinguish between electrostatically equivalent sites on the nucleophile.
- The term on the fourth line of Eq. (51) is negligible. If we are in a strongly electrostatically controlled regime the extent of electron transfer is small and so the $(\Delta N)^2$ term is entirely negligible. (The fourth term might be important if a “second tie-breaker” was needed, though.)

The remaining term in Eq. (51) must determine the regioselectivity. This term—the second term on the third line of Eq. (51)—can be rewritten as

$$-\Delta N \int f_{\text{nucleophile}}^-(\mathbf{r}) \Phi_{\text{electrophile}}(\mathbf{r}) d\mathbf{r} \quad (56)$$

Eq. (56) models how electron transfer quenches the electrostatic attraction between the nucleophile and the electrophile. The quenching arises because after the nucleophile donates electrons to the electrophile, the electron density on the nucleophile decreases by $\Delta N f_{\text{nucleophile}}^-(\mathbf{r})$. Because the reactive sites of the nucleophile are less negatively charged after the electron transfer, the electrostatic attraction between the nucleophile and the electrophile is weaker after electron transfer than it was before. If electron transfer quenches the electrostatic attraction at one interaction site more than the other, ($f_{\text{nucleophile}}^-(\mathbf{r}_1) > f_{\text{nucleophile}}^-(\mathbf{r}_2)$), then the second interaction site—the one with “minimum Fukui function”—will be the most favorable reactive site.

It should be stressed, again, that *the minimum Fukui function rule is a “tie-breaking” rule; it is not a general purpose reactivity rule.* The minimum Fukui function rule applies only when:

1. The reaction is strongly electrostatically controlled reactions. (The fourth term in Eq. (51) must be negligible!)
2. Multiple reactive sites are equivalent electrostatically. (The first term in Eq. (51) and the first term on the third line in Eq. (51) must fail to distinguish between the reactive sites.)
3. The Fukui-function term (the second term on the third line in Eq. (51)) is more important than the effects, like polarization, that are neglected by our model.

A minimum Fukui function tie-breaker is also operative in strongly electrostatically controlled nucleophilic attacks on electrophiles. In that case, it is the first term on the third line in Eq. (51) that determines the regioselectivity.

Case 2: Electron-transfer Controlled Reactions; $\kappa \leq -1$

Electron-transfer controlled (also called Fukui-function controlled and frontier-orbital controlled) reactions are typified by:

(a) The reactive sites have small charges. (So $q_{\text{electrophile}}^{(0)} \approx 0$ and $\Phi_{\text{nucleophile}}(\mathbf{r}_p) \approx 0$.)

(b) Reactivity occurs where the Fukui functions are large ($f_{\text{electrophile}}^{(+)} \gg 0$).

(c) Electron transfer between reagents is significant ($\Delta N \approx -1$).

Under these conditions, κ is less than, or approximately equal to, minus one. We call the case $\kappa = -1$ “pure electron-transfer control” and the case $\kappa < -1$ “strong electron-transfer control.”

Case 2A: Pure Electron-Transfer Control; $\kappa = -1$

When $\kappa = -1$,

$$\Xi_{\Delta N \leq 0}^{\kappa=-1}(\mathbf{r}_p) = 2\Delta N v_{\text{nucleophile}}^f(\mathbf{r}_p) \quad (57)$$

Thus, for pure electron-transfer controlled reactions, the appropriate reactivity indicator is the Fukui function’s potential. Because $\Delta N < 0$, the preferred reactive site is the location where the Fukui potential is the largest. Because the Fukui potential is usually large in the same places that the Fukui function is large, electron-transfer controlled reactions tend to occur where Fukui function is large.

For the condensed version of the general-purpose reactivity indicator,

$$\Xi_{\Delta N \leq 0, \alpha}^{\kappa=-1} = 2\Delta N f_{\text{nucleophile}, \alpha}^- \quad (58)$$

Based on this, we infer that for pure electron transfer control, the atomic site with the largest condensed Fukui function is the most reactive. The situation for electrophiles is essentially the same, but in that case it is the Fukui function from above, $f_{\text{electrophile}, \alpha}^+$, that is relevant.

Case 2B: Strong Electron-Transfer Control; $\kappa < -1$

When $\kappa < -1$, the magnitude of the first term in Eq. (49) is usually much smaller than the second term because $|\kappa - 1| > |\kappa + 1|$ and the nucleophile is not highly charged in electron-transfer controlled reactions ($q_{\text{nucleophile}, \alpha}^{(0)} \approx 0$). For this reason, there can be no doubt that the most appropriate reactivity indicator for electron-transfer controlled reactions is the Fukui potential or, alternatively, the Fukui function.^{24,68}

What happens if there are two sites that have equivalent values of the Fukui potential? Because $\kappa + 1 < 0$ whenever $\kappa < -1$, $\Xi_{\Delta N \leq 0}^{\kappa < -1}$ will be smallest in

those places where the electrostatic potential, $\Phi_{\text{nucleophile}}(\mathbf{r}_p)$, is the largest. That is, reactivity will be favored at the more positive (or less negative) reactive sites. This counterintuitive electrostatic potential tie-breaker is the analogue of the minimum Fukui function tie-breaker for strongly electron-transfer controlled reactions.

For nucleophiles, the “Electrostatic Potential Tie-Breaking Rule” states that: *if an electrophilic attack reaction on a nucleophile is strongly electron-transfer controlled and if there are two reactive sites with the same Fukui potential, then the site with the greatest electrostatic potential is the most favorable.* This principle can be restated in terms of condensed reactivity indicators: *for electrophilic attack occurring under conditions associated with strong electron-transfer control of chemical reactivity, a nucleophile possessing two sites with equal condensed Fukui function will be most reactive at the site with greatest charge.*

The analogous tie-breaking rule for electrophiles is: *If a nucleophilic attack reaction on a electrophile is strongly electron-transfer controlled and if there are two reactive sites with the same Fukui potential, then the site with the smallest electrostatic potential is the most favorable.* The condensed version of this rule is: *for nucleophilic attack occurring under conditions associated with strong electron-transfer control of chemical reactivity, an electrophile possessing two sites with equal condensed Fukui function will be most reactive at the site with least charge.*

Just as we did for the minimum Fukui function tie-breaker, we can elucidate the origins of the electrostatic potential tie-breaker by studying the fundamental equation for the interaction energy, (51). Note that

- The first line of this equation is a constant and does not influence regioselectivity.
- The fourth line of this equation models electron transfer from the nucleophile to the electrophile. Although this term usually dominates the regioselectivity, when we are in a “tie-breaking” situation, the Fukui potential of the nucleophile is the same at both reactive sites. Consequently, this term does not determine the regioselectivity.
- Because the nucleophile’s Fukui potential is similar at the two reactive sites, the second term on the third line does not contribute to the regioselectivity preference.
- Because we are in the strong electron-transfer control limit, neither the electrophile nor the nucleophile are highly charged. The second line of Eq. (51) models electrostatic effects, but if magnitudes of the atomic charges

on the electrophile and the atomic charges on the nucleophile are both small, the mutual attraction between these charges

$$E_{\text{int}}^{\text{electrostatic}} \approx \sum_{\alpha \in \text{nucleophile}} \sum_{\beta \in \text{electrophile}} \frac{q_{\text{nucleophile},\alpha} q_{\text{electrophile},\beta}}{|\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|} \quad (59)$$

should be very small.

The remaining term in Eq. (51),

$$\Delta N \int f_{\text{electrophile}}^+(\mathbf{r}) \Phi_{\text{nucleophile}}(\mathbf{r}) d\mathbf{r}, \quad (60)$$

must determine the regioselectivity. As electrons move from the nucleophile to the electrophile, the electron density on the electrophile increases by $-\Delta N f_{\text{electrophile}}^+(\mathbf{r})$. This causes the atomic charges on the electrophile to become less positive (or more negative) and decreases the electrostatic attraction between the nucleophile and the electrophile. (Since the electrophile was not highly charged to begin with, it is even conceivable that, after electron transfer, some key sites on the electrophile might be negatively charged. This corresponds to the extreme case where the electron transfer is so dramatic that the polarity of the nucleophile-electrophile bond is reversed.) Given a choice between two reactive sites with $\Phi_{\text{nucleophile}}(\mathbf{r}_1) > \Phi_{\text{nucleophile}}(\mathbf{r}_2)$, the most favorable site will be the first site, since this site is associated with a more favorable (or at least a less unfavorable) electrostatic interaction with the electrophile.

Case 3: Joint Electrostatic and Electron-Transfer Control, $-1 < \kappa < 1$

When κ is between minus one and plus one, a nucleophile will be most reactive at places where the electrostatic potential is negative and/or the Fukui potential is big. When κ is close to plus one or the extent of electron transfer is small ($\Delta N \approx 0$), reactivity preferences are predominately determined by the electrostatic potential. However, when two sites have similar electrostatic potentials, the most reactive site will be the site with the *largest* value of the Fukui potential. This should be contrasted with the extreme case of strong electrostatic control ($\kappa > 1$), where the site with the smallest value of the Fukui potential was the most favorable.

When κ is close to minus one and the extent of electron transfer is large ($\Delta N \approx -1$), reactivity preferences are predominately determined by the Fukui function. However, when two sites have similar Fukui functions, the site with the minimum electrostatic potential will be favored. This should be contrasted with the extreme case of strong electron-transfer control ($\kappa < -1$), where the site with maximum electrostatic potential would be favored.

In the intermediate regime, where $\kappa \approx 0$, the importance of electrostatic effects and electron-transfer effects are nearly balanced. In such cases, a reactive site that is negatively charged and has a reasonably positive value for the Fukui function might be favored over an uncharged site with a larger Fukui function. Similarly, such a site might be favored over an even more negatively charged site if that site was associated with a negligible value for the Fukui function.

Joint electrostatic and electron transfer control seems to be one of the most common situations in chemical reactivity. For example, one might expect that the protonation of aminoethanol, $(\text{H}_2\text{N})\text{H}_2\text{CCH}_2\text{OH}$ would be strongly electrostatically controlled: the proton is the prototypical hard acid⁷⁴ and aminoethanol is not an especially soft base. It is observed that the electrostatic potential around the oxygen atom and the nitrogen atom are about the same.⁵⁰ Assuming strong electrostatic control, one would then infer that protonation occurs on the Oxygen atom, since this is the site with minimum Fukui function. This is not the case; protonation occurs on the Nitrogen atom, where the Fukui function is the largest.⁵⁰ Thus, even in cases where strong electrostatic control might be expected, one frequently observes joint electrostatic and electron-transfer control.

It is interesting to notice what happens if one neglects the error-correction terms in Eq. (31) and returns to the oversimplified reactivity indicator in Eq. (30). In that model $\tilde{\kappa} > 0$ corresponds to strong electrostatic control and $\tilde{\kappa} < 0$ corresponds to strong electron-transfer control of reactivity. This shows that in a single reactive site interaction model, the possibility of joint electrostatic and electron-transfer control of reactivity arises *because* the point-charge representation of the electrophile provides a more accurate approximation to the electrostatic effects (involving the electrostatic potential of the nucleophile) than it does to the electron-transfer effects (involving the Fukui potential of the nucleophile). If this was not the case, then Eq. (32) would not be valid. If this equation was not valid, then every reaction would be either strongly electrostatically controlled or strongly electron-transfer controlled and a simple indicator like $\tilde{\Xi}_{\Delta N \leq 0}^{\tilde{\kappa}}(\mathbf{r}_p)$ (cf. Eq. (31)) would be qualitatively correct. However, joint electrostatic and electron-transfer control is commonly observed, while situations that require the minimum-Fukui function and electrostatic potential tie-breaker rules are uncommon. This observation provides experimental evidence for the validity of Eq. (32) and strongly supports the error analysis in section Appendix B.

G. Implications for the Local HSAB Principle

Our analysis provides a measure of support for the local Hard/Soft Acid/Base principle.⁷⁵⁻⁷⁸ The local Hard/Soft Acid/Base principle indicates that ambidentate ligands react with soft reagents in locations where the local softness is large and hard reagents in places where the local softness is small. (The local softness is just the global softness, S , times the Fukui function, $s(\mathbf{r}) = Sf(\mathbf{r})$.) Our analysis indicates that soft reagents ($\kappa \approx -1$) should react with the molecule in the places where the Fukui function (and thus the local softness) is the largest. For hard reagents, the Fukui function is not an important indicator, and reactions could occur where the Fukui function is small. Hard reagents might also react with the molecule in places where the Fukui function is large, however, because large values of the Fukui function are neither favorable nor unfavorable in electrostatically controlled reactions. The (probably extremely rare) exception is strongly electrostatic controlled ($\kappa > 1$) reactions of molecules that have multiple electrostatically favorable sites. In that case, a small value of the Fukui function (and thus the local softness) would be preferred over a large value.

The overall picture is entirely consistent with the work of Klopman:⁶⁸ electrostatic effects ($\Phi(\mathbf{r}_p)$ is the appropriate reactivity indicator) are typically dominant in reactions between hard reagents (where charge transfer is minimal); electron transfer effects ($f^-(\mathbf{r}_p)$ is the appropriate reactive indicator) are typically dominant in reactions between soft reagents (where substantial charge transfer occurs). For the intermediate cases, both effects are important.

III.V. Recapitulation

It seems desirable to review what we have accomplished. Starting from the Taylor expansion for the potential energy surface of interacting electrophiles and nucleophiles (Eq. (21)), we developed a “reactive site interaction model” based on the assumption that the interaction between the electrophile and the nucleophile is dominated by the interaction between their active sites. This led to a simplified indicator for the regioselectivity of the nucleophile (27). This model was then subjected to a detailed error analysis and parameterized to obtain our final indicators: Eq. (35) (for the regioselectivity of nucleophiles) and Eq. (39) (for the regioselectivity of electrophiles). These expressions could then be condensed into indicators for the reactivity of different atoms in the molecule, giving Eq. (44) (for nucleophiles) and Eq. (45) (for electrophiles).

All of our reactivity indicators represent models for the interaction energies between the electrophile and the nucleophile; because of this, highly reactive sites are associated with negative values of the reactivity indicator, which we denote $\Xi_{\Delta N}^{\kappa}$. This model interaction energy clearly depends on two parameters. The first parameter, ΔN , is the amount of electron transfer. ΔN could be computed from the chemical potential and the hardnesses of the reagents⁷⁹ or, alternatively, based on a quantum mechanical calculation of the product state (when the product of the chemical reaction is known). The second parameter, κ , quantifies whether the reaction is electrostatically controlled ($\kappa \geq 1$), electron-transfer controlled ($\kappa \leq -1$), or somewhere in between ($-1 < \kappa < 1$). The relative values of κ for different reagents can be compared using the approximate proportionalities

$$\kappa_{\Delta N \leq 0} \underset{\sim}{\propto} q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} \quad (61)$$

$$\kappa_{\Delta N \geq 0} \underset{\sim}{\propto} -\left(q_{\text{nucleophile}}^{(0)} + \Delta N f_{\text{nucleophile}}^{(-)} \right) \quad (62)$$

For highly charged electrophiles and small amounts of electron transfer, $\kappa \approx 1$ (electrostatic control). For weakly charged nucleophiles and significant electron transfer, $\kappa \approx -1$.

In most cases, the $\kappa \approx 1; |\Delta N| \approx 1$ case (electrostatic control and large amounts of electron transfer) is chemically irrelevant unless electrostatic effects are very, very strong. (This case could be important, for example, when the electrophile being reduced is a metal cation in a high oxidation state.) Similarly, the $\kappa \approx -1; \Delta N \approx 0$ case (electron-transfer control and negligible amounts of electron transfer) is chemically irrelevant unless electrostatic effects are weak. The $\kappa \approx 1; \Delta N \approx 0$ and $\kappa \approx -1; |\Delta N| \approx 1$ cases are very important limiting cases; they are associated with “classic” electrostatic and electron-transfer control, respectively. Most chemical reactions fall between those extremes; in these “jointly electrostatically and electron-transfer controlled” reactions the present indicator is preferable to existing approaches.

Using this reactivity indicator, we were able to gain some insight into appropriate indicators for different types of reactions. In the strong electrostatic control limit, $\kappa > 1$, it is observed that given two sites with similar electrostatic potential, the reactive site with the smallest Fukui function is favored. Similarly, in the strong electron-transfer control limit, $\kappa < 1$, it is observed that given two sites with similar Fukui potential, the reactive site with the greatest electrostatic potential (if the molecule is a nucleophile) or the least electrostatic potential (if the molecule is an electrophile) is most reactive. Both of these results are

counterintuitive, and it is reassuring that the cases of greatest chemical relevance, the electrostatic potential is the dominant indicator for electrostatically controlled reactions and the condensed Fukui function is the dominant indicator for electron-transfer controlled reactions. When the reactivity is between these two extremes, the most reactive site will be determined by a balance between the most favorable electrostatic potential and the most favorable Fukui potential, as one would expect.

When applying this model, it is important to the key assumptions that were made during its derivation:

- Reactive site interaction model. We assumed that the attacking reagent can be modeled as a point charge with a specified condensed Fukui function.
- Neglect of polarization and other terms from higher-order derivatives with respect to the external potential.

The first assumption is required for any reactivity indicator. Because we are seeking a qualitative reactivity indicator, it is imperative that our model depend only on the coarsest details of the attacking reagent. This is in keeping with experimental evidence: most molecules react at only one or two places, regardless of the choice of reagent. The “details” of the reagents cannot be very important for determining the reactivity of the molecule that is attacked.⁵⁰

The second assumption is merely pragmatic. We hope to incorporate polarization effects in our future work, but it is difficult to concoct a simple atom-condensed reactivity indicator that depends on a two-point quantity like the polarizability kernel. In addition, we believe that the most useful reactivity indicators are those that are easily evaluated using the output of standard quantum chemistry programs. The present reactivity indicator is easily computed from the atomic charges, which is a standard feature in quantum chemistry codes. By contrast, we do not know any simple way to extract a condensed polarizability kernel from the output file of a quantum chemistry program.

Acknowledgements:

Helpful discussions with Dr. David C. Thompson are acknowledged. NSERC, the Canada Research Chairs, and PREA provided funding for the Canadian authors. This research was performed when the second author visited McMaster University in the winter of 2005 and she wishes to thank the chemistry department at McMaster University for their hospitality.

III.VI. Appendix A. Derivation of the Electrostatic Potential Contribution

The key formula is the expression for the nuclear-nuclear repulsion energy in reference ⁸⁰

$$V_{nn}[v] = \frac{1}{32\pi^2} \iint_{\mathbf{r} \neq \mathbf{r}'} \frac{(\nabla_{\mathbf{r}}^2 v(\mathbf{r}))(\nabla_{\mathbf{r}'}^2 v(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (63)$$

We can derive $\frac{\delta V_{nn}[v]}{\delta v(\mathbf{r})}$ by finding the coefficient of $\delta v(\mathbf{r})$ in the expression $V_{nn}[v + \delta v] - V_{nn}[v]$.

$$\begin{aligned} V_{nn}[v + \delta v] - V_{nn}[v] &= \frac{1}{32\pi^2} \left(\iint_{\mathbf{r} \neq \mathbf{r}'} \frac{\nabla_{\mathbf{r}}^2 (v(\mathbf{r}) + \delta v(\mathbf{r})) \nabla_{\mathbf{r}'}^2 (v(\mathbf{r}') + \delta v(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right) \\ &\quad - \iint_{\mathbf{r} \neq \mathbf{r}'} \frac{\nabla_{\mathbf{r}}^2 v(\mathbf{r}) \nabla_{\mathbf{r}'}^2 v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\ &= \left\{ \frac{1}{32\pi^2} \left(\iint_{\mathbf{r} \neq \mathbf{r}'} \frac{\nabla_{\mathbf{r}}^2 \delta v(\mathbf{r}) \nabla_{\mathbf{r}'}^2 v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right) \right. \\ &\quad \left. + \iint_{\mathbf{r} \neq \mathbf{r}'} \frac{\nabla_{\mathbf{r}}^2 v(\mathbf{r}) \nabla_{\mathbf{r}'}^2 \delta v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right\} \\ &\quad + \left[\frac{1}{32\pi^2} \iint_{\mathbf{r} \neq \mathbf{r}'} \frac{\nabla_{\mathbf{r}}^2 \delta v(\mathbf{r}) \nabla_{\mathbf{r}'}^2 \delta v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right] \end{aligned} \quad (64)$$

The first term (in curly braces) represents the electrostatic interaction between the change in the external potential and the preexisting charge density. The second term (in square brackets) represents the self-repulsion energy of the perturbation. In an application such as ours, this term is neglected because it is already included in the energy expression for the attacking reagent. (If we included this term we would make a “double counting” error.) In any event, the term in square brackets is second order in the perturbing potential, and will not contribute to the first functional derivative. Neglecting this term and noting that the two integrals in the curly brace have identical values, Eq. (64) simplifies to

$$V_{nn}[v + \delta v] - V_{nn}[v] = \frac{1}{16\pi^2} \left(\iint_{\mathbf{r} \neq \mathbf{r}'} \frac{\nabla_{\mathbf{r}}^2 \delta v(\mathbf{r}) \nabla_{\mathbf{r}'}^2 v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right). \quad (65)$$

This equation is simplified by writing the functional variation as a nested integral

$$\delta_v [V_{nn}, \delta v] = \frac{1}{16\pi^2} \int \nabla_{\mathbf{r}'}^2 v(\mathbf{r}') \left(\int \frac{\nabla_{\mathbf{r}}^2 \delta v(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} \right) d\mathbf{r}' \quad (66)$$

and then simplifying the inner integral using Green’s theorem

$$\begin{aligned}
 \iiint \left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} \right) (\nabla_r^2 \delta v(\mathbf{r})) d\mathbf{r} &= \iiint \left(\nabla_r^2 \frac{1}{|\mathbf{r}-\mathbf{r}'|} \right) \delta v(\mathbf{r}) d\mathbf{r} \\
 &+ \iint \frac{1}{|\mathbf{r}-\mathbf{r}'|} (\nabla_r \delta v(\mathbf{r})) \cdot \hat{\mathbf{n}} da \quad (67) \\
 &- \iint \delta v(\mathbf{r}) \left(\nabla_r \frac{1}{|\mathbf{r}-\mathbf{r}'|} \right) \cdot \hat{\mathbf{n}} da.
 \end{aligned}$$

Our system is defined over all space, and so we choose the surface to be a sphere infinitely far from the origin. We assume that the change in charge density associated with the change in external potential, $\delta q(\mathbf{r}) = \frac{-1}{4\pi} \nabla_r^2 \delta v(\mathbf{r})$, is relatively localized, so that we can use the asymptotic form $\delta v(\mathbf{r}) \sim \frac{\langle \delta q \rangle}{r}$. With this assumption, the surface integral gives

$$\begin{aligned}
 \lim_{r \rightarrow \infty} \iint \frac{1}{|\mathbf{r}-\mathbf{r}'|} (\nabla \delta v(\mathbf{r})) \cdot \hat{\mathbf{n}} r^2 d\Omega &= \lim_{r \rightarrow \infty} \left(\frac{1}{r} \frac{\partial}{\partial r} \left(\frac{\langle \delta q \rangle}{r} \right) \right) (4\pi r^2) = 0 \\
 \lim_{r \rightarrow \infty} \iint \delta v(\mathbf{r}) \left(\nabla_r \frac{1}{|\mathbf{r}-\mathbf{r}'|} \right) \cdot \hat{\mathbf{n}} da &= \lim_{r \rightarrow \infty} \left(\frac{\langle \delta q \rangle}{r} \frac{\partial}{\partial r} \left(\frac{1}{r} \right) \right) (4\pi r^2) = 0
 \end{aligned} \quad (68)$$

Equation (68) shows that the surface terms vanish if $\delta v(\mathbf{r})$ falls off to zero at infinity at least as fast as r^{-1} . We simplify the volume integral in Eq. (67) using Poisson's equation for a point charge, obtaining

$$\begin{aligned}
 \iiint \left(\nabla_r^2 \frac{1}{|\mathbf{r}-\mathbf{r}'|} \right) \delta v(\mathbf{r}) d\mathbf{r} &= \iiint (-4\pi \delta(\mathbf{r}-\mathbf{r}')) \delta v(\mathbf{r}) d\mathbf{r} \quad (69) \\
 &= -4\pi \delta v(\mathbf{r}')
 \end{aligned}$$

Substitute Eqs. (68) and (69) into Eq. (67); then substitute that result into Eq. (65) to obtain

$$V_{nn}[v + \delta v] - V_{nn}[v] = \frac{-1}{4\pi} \left(\int \delta v(\mathbf{r}') \nabla_r^2 v(\mathbf{r}') d\mathbf{r}' \right). \quad (70)$$

The functional derivative is then

$$\begin{aligned}
 \frac{\delta V_{nn}[v]}{\delta v(\mathbf{r})} &= \frac{-1}{4\pi} \nabla_r^2 v(\mathbf{r}) \\
 &= -\rho_{nuc}(\mathbf{r})
 \end{aligned} \quad (71)$$

where $-\rho_{nuc}(\mathbf{r}) = -\sum_{\alpha} Z_{\alpha} \delta(\mathbf{r}-\mathbf{R}_{\alpha})$ is minus one times the nuclear charge density.

It is now easy to see how the electrostatic potential arises as a reactivity indicator in density-functional theory. Consider the change in total energy due to adding a small point charge, q , at the point \mathbf{r}_p :

$$\begin{aligned}\Delta(E + V_m) &= \int \left(\left(\frac{\delta E}{\delta v(\mathbf{r})} \right)_N + \frac{\delta V_m}{\delta v(\mathbf{r})} \right) \frac{q}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \\ &= q \int \frac{\rho(\mathbf{r}) - \sum_{\alpha} Z_{\alpha} \delta(\mathbf{r} - \mathbf{R}_{\alpha})}{|\mathbf{r} - \mathbf{r}_p|} d\mathbf{r} \\ &= -q\Phi(\mathbf{r}_p).\end{aligned}\tag{72}$$

III.VII. Appendix B. Error Analysis for the Reactive Site Interaction Model

This appendix performs an error analysis for the single reactive site interaction approximation to Eq. (21). The errors incurred by replacing the integrals in Eq. (21) with the point charge representations in Eq. (27) are conveniently summarized using the following expressions,

$$-\mathcal{E}_{(i)} \Phi_{\text{nucleophile}}(\mathbf{r}_p) \equiv \int \left(\sum_{\beta \in \text{electrophile}} Z_{\beta} \delta(\mathbf{r} - \mathbf{R}_{\beta}) - \rho_{\text{electrophile}}(\mathbf{r}) \right) \Phi_{\text{nucleophile}}(\mathbf{r}) d\mathbf{r} \tag{73}$$

$$\mathcal{E}_{(ii)} \Phi_{\text{nucleophile}}(\mathbf{r}_p) \equiv \int \left(f_{\text{electrophile}}^+(\mathbf{r}) - f_{\text{electrophile}}^{(+)} \delta(\mathbf{r} - \mathbf{r}_p) \right) \Phi_{\text{nucleophile}}(\mathbf{r}) d\mathbf{r} \tag{74}$$

$$-\mathcal{E}_{(iii)} v_{\text{nucleophile}}^{f^-}(\mathbf{r}_p) \equiv \int \left(\sum_{\beta \in \text{electrophile}} Z_{\beta} \delta(\mathbf{r} - \mathbf{R}_{\beta}) - \rho_{\text{electrophile}}(\mathbf{r}) \right) v_{\text{nucleophile}}^{f^-}(\mathbf{r}) d\mathbf{r} \tag{75}$$

$$\mathcal{E}_{(iv)} v_{\text{nucleophile}}^{f^-}(\mathbf{r}_p) \equiv \int \left(f_{\text{electrophile}}^+(\mathbf{r}) - f_{\text{electrophile}}^{(+)} \delta(\mathbf{r} - \mathbf{r}_p) \right) v_{\text{nucleophile}}^{f^-}(\mathbf{r}) d\mathbf{r} \tag{76}$$

We now rationalize these forms. It is important to recognize that each of these integrals is really a Coulomb integral associated with the attractive forces that electrons feel toward nuclei and the repulsive forces they feel towards other electrons. For example, Eq. (74) could be rewritten as

$$\mathcal{E}_{(ii)} \Phi_{\text{nucleophile}}(\mathbf{r}_p) \equiv \iint \frac{\left(\begin{array}{c} f_{\text{electrophile}}^+(\mathbf{r}) \\ -f_{\text{electrophile}}^{(+)} \delta(\mathbf{r} - \mathbf{r}_p) \end{array} \right) \left(\begin{array}{c} \sum_{\alpha \in \text{nucleophile}} Z_{\alpha} \delta(\mathbf{r}' - \mathbf{R}_{\alpha}) \\ -\rho_{\text{nucleophile}}(\mathbf{r}') \end{array} \right)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{77}$$

In an expression like this, it seems reasonable to assume that the largest error is associated with the approximation to the electron-electron repulsion-type term,

$$\begin{aligned}
 & - \iint \frac{f_{\text{electrophile}}^+(\mathbf{r}) \rho_{\text{nucleophile}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
 & \approx - \iint \frac{f_{\text{electrophile}}^+ \delta(\mathbf{r} - \mathbf{r}_p) \rho_{\text{nucleophile}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' ,
 \end{aligned} \tag{78}$$

instead of with the approximation to the electron-nuclear attraction-type term,

$$\begin{aligned}
 & \iint \frac{f_{\text{electrophile}}^+(\mathbf{r}) \left(\sum_{\alpha \in \text{nucleophile}} Z_{\alpha} \delta(\mathbf{r}' - \mathbf{R}_{\alpha}) \right)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
 & \approx \iint \frac{f_{\text{electrophile}}^+ \delta(\mathbf{r} - \mathbf{r}_p) \left(\sum_{\alpha \in \text{nucleophile}} Z_{\alpha} \delta(\mathbf{r}' - \mathbf{R}_{\alpha}) \right)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' .
 \end{aligned} \tag{79}$$

Because the integrands in Eq. (78) are very large when \mathbf{r} and \mathbf{r}' are close together, the point charge approximation does not provide a good representation for the interaction between the asymptotic tails of $\rho_{\text{nucleophile}}(\mathbf{r})$ and $f_{\text{electrophile}}^+(\mathbf{r})$. This suggests that: (1) the left hand side of Eq. (78) is less than the right hand side and (2) the error in Eq. (78) is significantly bigger than the error in Eq. (79). Extending this argument to the other integrals in Eq. (73)-(76)

- (a) In general, the left hand side in Eqs. (73) and (76) should be positive. This implies that, in general, $\varepsilon_{(i)} > 0$ and $\varepsilon_{(iv)} > 0$.
- (b) In general, the left hand side in Eq. (74) and (75) should be negative. This implies that, in general, $\varepsilon_{(ii)} > 0$ and $\varepsilon_{(iii)} > 0$.

The sign convention in Eqs. (73)-(76) is based on these observations. Recall that the electrostatic potential of a nucleophile is usually negative near a reactive site ($\Phi_{\text{nucleophile}}(\mathbf{r}_p) < 0$) while the potential due to the Fukui function is usually positive near a reactive site ($v_{\text{nucleophile}}^-(\mathbf{r}_p) > 0$). Referring back to (a) and (b), it is apparent that the sign convention in Eqs. (73)-(76) implies that $\varepsilon_{(i)} - \varepsilon_{(iv)}$ are usually positive constants. We will now characterize the size of these constants.

Examining Eqs. (73) and (74), it seems clear that the size of the error should be related to the magnitude of $\Phi_{\text{nucleophile}}(\mathbf{r})$ in the region where the electrophile is attacking. Similarly, in Eqs. (75) and (76), the error incurred by the point charge approximation should be small when $v_{\text{nucleophile}}^-(\mathbf{r})$ is small in the

region where the electrophile is attacking. On the other hand, if $v_{\text{nucleophile}}^{f^-}(\mathbf{r})$ is large at the reactive site, this suggests that $f_{\text{nucleophile}}^-(\mathbf{r})$ is also large near the reactive site, which suggests that the point charge approximation to the integrals in Eq. (75) will be poor. Hence:

- (c) The error in the left-hand sides of Eqs. (73) and (74) is roughly proportional to the magnitude of the electrostatic potential at the position where the electrophile attacks, $\Phi_{\text{nucleophile}}(\mathbf{r}_p)$.
- (d) The error in the left-hand sides of Eqs. (75) and (76) is roughly proportional to the magnitude of the Fukui potential at the position where the electrophile attacks, $v_{\text{nucleophile}}^{f^-}(\mathbf{r}_p)$.

By including an appropriate dependence on $\Phi_{\text{nucleophile}}(\mathbf{r}_p)$ or $v_{\text{nucleophile}}^{f^-}(\mathbf{r}_p)$ in the right-hand side of the defining equations, (73)-(76), we ensure that $\varepsilon_{(i)} - \varepsilon_{(iv)}$ depend only weakly on the relative magnitude of the electrostatic and Fukui potentials at the reactive sites. Our choice of definition, then, ensures that $\varepsilon_{(i)} - \varepsilon_{(iv)}$ measure the *intrinsic error* in the point charge approximations.

In the discussion surrounding Eq. (78), we pointed out that the primary error in the single reactive site approximation was associated with the interaction between the tails of electronic distributions centered on the electrophile and the nucleophile. This error is expected to be most severe in Eq. (76), because the Fukui functions (*I*) have slow asymptotic decays and (*2*) are concentrated on the “frontiers” of the reagents.

- (e) We expect $\varepsilon_{(iv)}$ to be relatively large. The error due to the point charge approximation can be rather large in this case because the interaction between the asymptotic tails of the Fukui functions might be significantly underestimated by the result from the point charge representation.

By contrast, the error in Eq. (73) might be relatively small, since the electron density is concentrated near the atomic nuclei.

- (f) We expect $\varepsilon_{(i)}$ to be the relatively small, because the interaction between the asymptotic tails of the electron densities should be only a small part of the total interaction between the electronic distributions.

The errors in Eqs. (74) and (75) should be intermediate.

- (g) We expect $\varepsilon_{(ii)}$ and $\varepsilon_{(iii)}$ to be relatively large compared to $\varepsilon_{(i)}$ because the Fukui function in these integrals is concentrated on the

frontiers of the molecule. However, we expect $\varepsilon_{(ii)}$ and $\varepsilon_{(iii)}$ to be smaller than $\varepsilon_{(iv)}$, because point charge approximation is least accurate in Eq. (76), where both the Fukui functions are involved.

Using results from (a), (b), (e), (f), and (g), we have the following ordering of relative errors.

$$0 < \varepsilon_{(i)} < \varepsilon_{(ii)} \approx \varepsilon_{(iii)} < \varepsilon_{(iv)} \quad (80)$$

It should be stressed that the preceding analysis is strictly qualitative. Exceptions to the ordering in Eq. (80) will occur. Henceforth, we will never rely upon the details of this analysis; the approximate ordering of errors in Eq. (80) is sufficient to establish our results.

As discussed in section III.A and III.B, the key parameters in the reactivity model are the number of electrons donated by the nucleophile to the electrophile, $\Delta N \leq 0$, and the quantity

$$\tilde{\kappa} = q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} \quad (81)$$

Inserting the expression for $\tilde{\kappa}$ into Eq. (27) gave the oversimplistic regioselectivity indicator in Eq. (30).

$\tilde{\kappa}$ measures the relative importance of electrostatic and electron-transfer effects. When $\tilde{\kappa}$ is significantly positive, the electrophile is highly charged and electron transfer is minimal; such reactions are expected to be electrostatically controlled. When $\tilde{\kappa}$ is significantly negative, then electron transfer to the electrophilic site is important, but the electrophile is not especially highly charged; such reactions are expected to be electron-transfer controlled.

As long as $\tilde{\kappa}$ is very different from zero, it is reasonable to neglect the errors due to the point-charge approximations to the integrals, because they should be small compared to $\tilde{\kappa}$. To address the case where $\tilde{\kappa} \approx 0$, insert the expressions for $\varepsilon_{(i)}$, $\varepsilon_{(ii)}$, and $\varepsilon_{(iii)}$, and $\varepsilon_{(iv)}$ into Eq. (21). This gives an “error corrected” version of the reactive-site interaction model in Eq. (30), namely,

$$\begin{aligned} \tilde{\Xi}_{\Delta N \leq 0}^{\tilde{\kappa}, \varepsilon_{(i)}, \varepsilon_{(ii)}}(\mathbf{r}_p) \equiv & \left(\tilde{\kappa} + \left(-\varepsilon_{(i)} + \varepsilon_{(ii)} \Delta N \right) \right) \Phi_{\text{nucleophile}}(\mathbf{r}_p) \\ & - \Delta N \left(\tilde{\kappa} + \left(-\varepsilon_{(iii)} + \varepsilon_{(iv)} \Delta N \right) \right) \nu_{\text{nucleophile}}^{\text{f}^-}(\mathbf{r}_p) \end{aligned} \quad (82)$$

Because $\varepsilon_{(i)} < \varepsilon_{(iii)}$, $\varepsilon_{(ii)} < \varepsilon_{(iv)}$, and $\Delta N \leq 0$,

$$-\varepsilon_{(iii)} + \varepsilon_{(iv)} \Delta N < -\varepsilon_{(i)} + \varepsilon_{(ii)} \Delta N. \quad (83)$$

Comparing Eqs. (31) and (82) allows us to make the identification

$$\varepsilon_p = -\varepsilon_{(i)} + \varepsilon_{(ii)} \Delta N \quad (84)$$

$$\varepsilon_f = -\varepsilon_{(iii)} + \varepsilon_{(iv)}\Delta N \quad (85)$$

Equation (83) then establishes that $\varepsilon_p > \varepsilon_f$, as stated in Eq. (32). The general purpose reactivity indicator then follows from the analysis in section III.B. In particular, because

$$\tilde{\kappa} - \varepsilon_{(i)} + \varepsilon_{(ii)}\Delta N > \tilde{\kappa} - \varepsilon_{(iii)} + \varepsilon_{(iv)}\Delta N, \quad (86)$$

the *qualitative* structure of the interaction energy model in Eq. (82) is recaptured by the simple expression

$$\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}_p) \equiv (\kappa + 1)\Phi_{\text{nucleophile}}(\mathbf{r}_p) - \Delta N(\kappa - 1)v_{\text{nucleophile}}^f(\mathbf{r}_p) \quad (87)$$

The new parameter, κ , is linearly related to the more fundamental $\tilde{\kappa}$ via the equation

$$\kappa = 2 \left(\frac{\tilde{\kappa} - \varepsilon_{(i)} + \varepsilon_{(ii)}\Delta N}{\varepsilon_{(iii)} - \varepsilon_{(i)} + (\varepsilon_{(ii)} - \varepsilon_{(iv)})\Delta N} \right) - 1 \quad (88)$$

The approximate proportionality reported in Eq. (38) is clearly accurate whenever $\tilde{\kappa}$ is large compared to the error terms. The motivation for this transformation of variables is that it gives a simple appealing qualitative picture, with $\kappa = +1$ and $\kappa = -1$ corresponding to pure electrostatic control and pure electron-transfer control, respectively.

III.VIII. References

- (1) Albright, T. A.; Burdett, J. K.; Whangbo, M. H. *Orbital Interactions in Chemistry*; Wiley-Interscience: New York, 1985.
- (2) Fukui, K. *Science* **1987**, *218*, 747-754.
- (3) Fukui, K.; Yonezawa, T.; Nagata, C. *J.Chem.Phys.* **1953**, *21*, 174-176.
- (4) Fukui, K.; Yonezawa, T.; Shingu, H. *J.Chem.Phys.* **1952**, *20*, 722-725.
- (5) Fukui, K.; Yonezawa, T.; Nagata, C. *Bull.Chem.Soc.Japan* **1954**, *27*, 423-427.
- (6) Woodward, R. B.; Hoffmann, R. *Angew.Chem., Int.Ed.Engl.* **1969**, *8*, 781-853.
- (7) Hoffmann, R.; Woodward, R. B. *Acc.Chem.Res.* **1968**, *1*, 17-22.
- (8) Hoffmann, R.; Woodward, R. B. *J.Am.Chem.Soc.* **1965**, *87*, 2046-2048.
- (9) Woodward, R. B.; Hoffmann, R. *J.Am.Chem.Soc.* **1965**, *87*, 2511-2513.
- (10) Woodward, R. B.; Hoffmann, R. *J.Am.Chem.Soc.* **1965**, *87*, 395-397.
- (11) Parr, R. G.; Yang, W. *J.Am.Chem.Soc.* **1984**, *106*, 4049-4050.
- (12) Yang, W.; Parr, R. G.; Pucci, R. *J.Chem.Phys.* **1984**, *81*, 2862-2863.
- (13) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.

- (14) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2000**, *122*, 2010-2018.
- (15) Melin, J.; Ayers, P. W.; Ortiz, J. V. *Journal of Chemical Sciences* **2005**, *117*, 387-400.
- (16) Bartolotti, L. J.; Ayers, P. W. *J.Phys.Chem.A* **2005**, *109*, 1146-1151.
- (17) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (18) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (19) Cerjan, C. J.; Miller, W. H. *J.Chem.Phys.* **1981**, *75*, 2800-2806.
- (20) Simons, J.; Jorgensen, P.; Taylor, H.; Ozment, J. *Journal of Physical Chemistry* **1983**, *87*, 2745-2753.
- (21) Dewar, M. J. S. *THEOCHEM* **1989**, *59*, 301-23.
- (22) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *Journal of Chemical Theory and Computation* **2006**, *submitted*.
- (23) Li, Y.; Evans, J. N. S. *J.Am.Chem.Soc.* **1995**, *117*, 7756-7759.
- (24) Melin, J.; Aparicio, F.; Subramanian, V.; Galvan, M.; Chattaraj, P. K. *J.Phys.Chem.A* **2004**, *108*, 2487-2491.
- (25) Coulson, C. A.; Longuet-Higgins, H. C. *Proc.R.Soc.London, A* **1947**, *192*, 39-60.
- (26) Coulson, C. A.; Longuet-Higgins, H. C. *Proc.R.Soc.London, A* **1947**, *192*, 16-32.
- (27) Coulson, C. A.; Longuet-Higgins, H. C. *Proc.R.Soc.London, A* **1948**, *193*, 447-456.
- (28) Coulson, C. A.; Longuet-Higgins, H. C. *Proc.R.Soc.London, A* **1948**, *193*, 456-464.
- (29) Wolfsberg, M.; Helmholz, L. *J.Chem.Phys.* **1952**, *20*, 837-843.
- (30) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (31) Nalewajski, R. F.; Parr, R. G. *J.Chem.Phys.* **1982**, *77*, 399-407.
- (32) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L., Jr. *Phys.Rev.Lett.* **1982**, *49*, 1691-1694.
- (33) Yang, W.; Zhang, Y.; Ayers, P. W. *Phys.Rev.Lett.* **2000**, *84*, 5172-5175.
- (34) Zhang, Y.; Yang, W. *Theor.Chem.Acc.* **2000**, *103*, 346-348.
- (35) Cohen, M. H.; Ganduglia-Pirovano, M. V. *J.Chem.Phys.* **1994**, *101*, 8988-97.
- (36) Yang, W.; Zhang, Y.; Ayers, P. W. *Phys.Rev.Lett.* **2000**, *84*, 5172-5175.
- (37) Ayers, P. W. *Theor.Chem.Acc.* **2001**, *106*, 271-279.
- (38) The contribution to molecular interaction energies from induced dipoles is normally less than a kcal/mol.
- (39) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (40) Ayers, P. W. *J.Chem.Phys.* **2000**, *113*, 10886-10898.
- (41) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (42) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *Journal of Chemical Theory and Computation* **2006**, *submitted*.
- (43) Berkowitz, M. *J.Am.Chem.Soc.* **1987**, *109*, 4823-4825.

- (44) Berkowitz, M. *J.Am.Chem.Soc.* **1987**, *109*, 4823-4825.
- (45) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem.Rev.* **2003**, *103*, 1793-1873.
- (46) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (47) Toro-Labbe, A.; Jaque, P.; Murray, J. S.; Politzer, P. *Chem.Phys.Lett.* **2005**, *407*, 143-146.
- (48) Toro-Labbe, A.; Gutierrez-Oliva, S.; Concha, M. C.; Murray, J. S.; Politzer, P. *J.Chem.Phys.* **2004**, *121*, 4570-4576.
- (49) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (50) Ayers, P. W.; Anderson, J. S. M.; Bartolotti, L. J. *Int.J.Quantum Chem.* **2005**, *101*, 520-534.
- (51) Yang, W.; Mortier, W. J. *J.Am.Chem.Soc.* **1986**, *108*, 5708-11.
- (52) Ayers, P. W.; Morrison, R. C.; Roy, R. K. *J.Chem.Phys.* **2002**, *116*, 8731-8744.
- (53) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (54) Ayers, P. W.; Parr, R. G.; Pearson, R. G. *J.Chem.Phys.* **2006**, *124*, 194107.
- (55) Ayers, P. W.; Anderson, J. S. M.; Rodriguez, J. I.; Jawed, Z. *Phys.Chem.Chem.Phys.* **2005**, *7*, 1918-1925.
- (56) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (57) Liu, S. B.; Parr, R. G. *Chem.Phys.Lett.* **1997**, *278*, 341-344.
- (58) De Proft, F.; Liu, S. B.; Parr, R. G. *J.Chem.Phys.* **1997**, *107*, 3000-3006.
- (59) Parr, R. G.; Liu, S. B. *Chem.Phys.Lett.* **1997**, *276*, 164-166.
- (60) Senet, P. *J.Chem.Phys.* **1997**, *107*, 2516-2524.
- (61) Senet, P. *J.Chem.Phys.* **1996**, *105*, 6471-6489.
- (62) Roy, R. K.; Pal, S.; Hirao, K. *J.Chem.Phys.* **1999**, *110*, 8236-8245.
- (63) Roy, R. K.; Hirao, K.; Pal, S. *J.Chem.Phys.* **2000**, *113*, 1372-1379.
- (64) Bultinck, P.; Carbo-Dorca, R. *Journal of Mathematical Chemistry* **2003**, *34*, 67-74.
- (65) Bultinck, P.; Carbo-Dorca, R.; Langenaeker, W. *J.Chem.Phys.* **2003**, *118*, 4349-4356.
- (66) Ayers, P. W. *Phys.Chem.Chem.Phys.* **2006**, *8*, 3387-3390.
- (67) Langenaeker, W.; Deproft, F.; Geerlings, P. *Journal of Physical Chemistry* **1995**, *99*, 6424-6431.
- (68) Klopman, G. *J.Am.Chem.Soc.* **1968**, *90*, 223-234.
- (69) Perez, P.; Simon-Manso, Y.; Aizman, A.; Fuentealba, P.; Contreras, R. *J.Am.Chem.Soc.* **2000**, *122*, 4756-4762.
- (70) Pal, S.; Chandrakumar, K. R. S. *J.Am.Chem.Soc.* **2000**, *122*, 4145-4153.
- (71) Nguyen, L. T.; Le, T. N.; De Proft, F.; Chandra, A. K.; Langenaeker, W.; Nguyen, M. T.; Geerlings, P. *J.Am.Chem.Soc.* **1999**, *121*, 5992-6001.
- (72) Berkowitz, M.; Parr, R. G. *J.Chem.Phys.* **1988**, *88*, 2554-2557.
- (73) Gross, E. K. U.; Kohn, W. *Phys.Rev.Lett.* **1985**, *55*, 2850-2852.
- (74) By some measures, the proton has infinite chemical hardness!

- (75) Gazquez, J. L.; Mendez, F. *Journal of Physical Chemistry* **1994**, *98*, 4591-4593.
- (76) Mendez, F.; Gazquez, J. L. *J.Am.Chem.Soc.* **1994**, *116*, 9298-9301.
- (77) Chattaraj, P. K. *J.Phys.Chem.A* **2001**, *105*, 511-513.
- (78) Geerlings, P.; De Proft, F. *Int.J.Quantum Chem.* **2000**, *80*, 227-235.
- (79) Parr, R. G.; Pearson, R. G. *J.Am.Chem.Soc.* **1983**, *105*, 7512-7516.
- (80) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.

Chapter IV

**Conceptual Density-Functional
Theory for General Chemical
Reactions, Including Those That Are
Neither Charge Nor Frontier-Orbital
Controlled**

**II. Application to Molecules Where
Frontier Molecular Orbital Theory
Fails***

*The content of this chapter has been published: **J. S. M. Anderson**, J. Melin, P. W. Ayers
“Conceptual density-functional theory for general chemical reactions, including those that
are neither charge nor frontier-orbital controlled II. Application to molecules where
frontier molecular orbital theory fails”; *J. Chem. Th. Comp.* **2007**, 3, 375-389.

IV.I. Statement of the Problem

This chapter presents the first chemical applications of the general-purpose reactivity indicator developed in the previous chapter. Particularly pernicious examples are chosen to illustrate the validity of the indicator. In particular, we focus on the so-called Dewar molecules, where the popular frontier-molecular theory approach gives incorrect results. A compact representation of the data that result from the two parameter general-purpose reactivity model called “reactivity transition tables” is introduced.

IV.II. Introduction

In the first paper in this series,¹ the authors derived a general-purpose reactivity indicator that is capable of describing not only electrostatically (or charge) controlled reactions and electron-transfer (or frontier-orbital) controlled reactions,² but also reactions that lie between these two extremes. This paper will apply this reactivity indicator to a particularly challenging set of molecules, where ordinary reactivity predictors have been observed to fail.

Before applying the reactivity indicator, we briefly summarize the results from the first paper in this series. The goal of the first paper was to derive a reactivity indicator that could truly be called a “general-purpose” reactivity indicator. That is, we sought a reactivity indicator that describes the full spectrum of chemical reactivity, from strong electrostatic control (minimum Fukui function is good), to joint electrostatic and electron-transfer control (maximum Fukui function is good), to strong electron-transfer control (maximum Fukui function is good, but maximum (for nucleophiles) or minimum (for electrophiles) electrostatic potential is also good). To achieve this goal, we used a perturbative expansion about the separated reagent limit to derive an expression for the interaction energy between an electrophile and a nucleophile. (See Eq. (21) in paper I.)

In order to derive a reactivity indicator, we introduced a single reactive-site interaction model for electrophiles and nucleophiles. In this model, the reactive site of the attacking electrophile/nucleophile is modeled with a point charge and a condensed Fukui function. Inserting this model into the expression for the interaction energy and performing a careful error analysis led to the desired indicators. One of our indicators is appropriate for predicting where an electrophile will attack a nucleophile,

$$\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}) \equiv (\kappa + 1) \Phi_{\text{nucleophile}}(\mathbf{r}) - \Delta N (\kappa - 1) v_{\text{nucleophile}}^{\text{r}}(\mathbf{r}). \quad (1)$$

The other indicator is appropriate for predicting where a nucleophile will attack an electrophile,

$$\Xi_{\Delta N \geq 0}^{\kappa}(\mathbf{r}) \equiv -(\kappa + 1)\Phi_{\text{electrophile}}(\mathbf{r}) + \Delta N(\kappa - 1)v_{\text{electrophile}}^{f^+}(\mathbf{r}) \quad (2)$$

Because these indicators model the interaction energy of the molecule, the molecule is most reactive in the places where $\Xi_{\Delta N}^{\kappa}(\mathbf{r})$ is most negative.

By simply varying the value of κ , the entire spectrum of chemical reactivity can be described, ranging from strong electrostatic control ($\kappa > 1$), to pure electrostatic control ($\kappa = 1$), to joint control by electrostatics and electron-transfer effects ($-1 < \kappa < 1$), to pure electron-transfer control ($\kappa = -1$), to strong electron transfer control ($\kappa < -1$). The value of κ that is appropriate for a particular reaction can be estimated using the approximate proportionalities:

$$\begin{aligned} \kappa &\propto \underset{\sim}{q}_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} && \text{for nucleophiles} \\ \kappa &\propto -\underset{\sim}{q}_{\text{nucleophile}}^{(0)} - \Delta N f_{\text{nucleophile}}^{(-)} && \text{for electrophiles} \end{aligned} \quad (3)$$

The constant of proportionality is positive and, based on our experience, has order of magnitude one.

The reactivity indicators in Eq. (1) and (2) also depend on the extent of electron transfer. The amount of electron transfer could be computed by minimizing the expression for the interaction energy directly, but the simple formula proposed by Parr and Pearson should be adequate for qualitative purposes,³

$$\begin{aligned} \Delta N_{\text{electrophile}} &\approx \frac{\mu_{\text{nucleophile}} - \mu_{\text{electrophile}}}{\eta_{\text{nucleophile}} + \eta_{\text{electrophile}}} \\ &\approx \frac{\left(I_{\text{nucleophile}} - I_{\text{electrophile}}\right) + \left(A_{\text{nucleophile}} - A_{\text{electrophile}}\right)}{2\left(I_{\text{nucleophile}} + I_{\text{electrophile}} - A_{\text{nucleophile}} - A_{\text{electrophile}}\right)} \end{aligned} \quad (4)$$

In this equation, μ denotes the electronic chemical potential⁴ and η denotes the chemical hardness.³ The second line of this equation approximates the chemical potential and chemical hardness using the vertical ionization potential (I) and vertical electron affinity (A) of the reagents. Because nucleophiles transfer electrons to electrophiles, we have chosen a sign convention where $\Delta N \leq 0$ for nucleophiles and $\Delta N \geq 0$ for electrophiles. The molecules we are studying in this paper are nucleophiles, so it is the $\Delta N \leq 0$ case that is of greatest interest here.

The general-purpose reactivity indicators are seen to have a dependence on the Fukui potential, $v^{f^{\pm}}(\mathbf{r})$ and the molecular electrostatic potential, $\Phi(\mathbf{r})$. The

electrostatic potential is essential for describing reactions that are electrostatically controlled ($\kappa = 1$); the Fukui potential is essential for describing reactions that are electron-transfer controlled ($\kappa = -1$). In all other cases, both the electrostatic potential and the Fukui potential play a role in determining a molecule's regioselectivity preferences.

It is useful to approximate the electrostatic potential and the Fukui potentials using atomic charges,^{5,6}

$$\Phi^{(0)}(\mathbf{r}) \sim \sum_{\alpha} \frac{q_{\alpha}^{(0)}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \quad (5)$$

$$v^{f^{-}}(\mathbf{r}) \sim \sum_{\alpha} \frac{f_{\alpha}^{-}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \quad (6)$$

$$v^{f^{+}}(\mathbf{r}) \sim \sum_{\alpha} \frac{f_{\alpha}^{+}}{|\mathbf{r} - \mathbf{R}_{\alpha}|}$$

where $q_{\alpha}^{(0)}$ denotes the atomic charges on the reagent. f_{α}^{+} and f_{α}^{-} are the condensed Fukui functions. Using these results, one can derive a condensed version of the indicators in Eqs. (1) and (2)

$$\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}_p) \sim \sum_{\alpha \in \text{nucleophile}} \frac{\Xi_{\Delta N \leq 0, \alpha}^{\kappa}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \quad (7)$$

$$\Xi_{\Delta N \leq 0, \alpha}^{\kappa} = (\kappa + 1)q_{\text{nucleophile}, \alpha}^{(0)} - \Delta N(\kappa - 1)f_{\text{nucleophile}, \alpha}^{-}$$

$$\Xi_{\Delta N \geq 0}^{\kappa}(\mathbf{r}_p) \sim \sum_{\alpha \in \text{electrophile}} \frac{\Xi_{\Delta N \geq 0, \alpha}^{\kappa}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \quad (8)$$

$$\Xi_{\Delta N \geq 0, \alpha}^{\kappa} = -(\kappa + 1)q_{\text{electrophile}, \alpha}^{(0)} + \Delta N(\kappa - 1)f_{\text{electrophile}, \alpha}^{+}$$

For special cases, this reactivity indicator recovers the reactivity patterns that would be predicted based on the Fukui function ($\kappa = -1$) or the electrostatic potential ($\kappa = 1$) alone. The new indicator has no value, then, unless it supersedes the description of chemical reactivity that is possible using these reactivity indicators in isolation. This suggests that the new indicator might be useful for studying molecules where frontier molecular orbital theory (FMO) would be expected to work but has been observed to fail. The regioselectivity of electrophilic aromatic substitution reactions is usually well-described using both FMO and density-functional theory analogues to frontier molecular orbital theory like the Fukui function.⁷⁻⁹ This is not always true, however: Dewar showed that FMO fails to describe electrophilic aromatic substitution in isoquinoline, 10-hydroxy-10,9-borazarophenanthrene, and 10-methyl-10,9-borazarophenanthrene. The question arises: does the Fukui function,¹⁰⁻¹² which extends the frontier molecular orbital theory but is nonetheless motivated by FMO ideas, fail in the

same way? In section V, it is observed that while the Fukui function does seem to work better than FMO for these molecules, it still fails to adequately describe their reactivity. (That is, the Fukui function fails “in the same way” as FMO, but not as badly.)

The Fukui function is intimately linked to the idea of electron transfer, so it is an appropriate indicator for “electron-transfer controlled” (also called frontier-orbital controlled and Fukui-function controlled) reactions. When hard reagents interact, electron transfer is either limited or occurs late in the chemical reaction profile; such reactions are usually called “electrostatically controlled” or “charge controlled.”² The electrostatic potential is a more appropriate indicator than the Fukui function in these cases.¹³ Since the Fukui function fails to adequately describe the reactivity of these molecules, perhaps the electrostatic potential will suffice. The results in section V show that the electrostatic potential does not describe the reactivity of these molecules either.

These results suggest that the electrophilic substitution on borazarophenanthrenes represent a difficult, but otherwise suitable, test for the general-purpose reactivity indicator, $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$, that we derived and discussed in the first paper of this series. Indeed, the theoretical developments in the first paper were motivated by our inability to describe these molecules using ordinary reactivity indicators. Section V contains the main results; we observe that $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ does an excellent job of describing a variety of electrophilic aromatic substitution reactions on borazarophenanthrenes. Before presenting our results, however, we need to state our computational methods.

IV.III. Computational Methods

To demonstrate the power of these indicators, we decided to analyze molecules where Dewar found contradictions to frontier molecular orbital effects.¹⁴ In what follows, all calculations were conducted using Gaussian03¹⁵ and the B3LYP¹⁶⁻¹⁸ functional with the 6-31++G* basis set.¹⁹ Figures were generated using GaussView 3.0. The atomic charges used to compute condensed reactivity indicators were obtained from four different methods: the Mulliken population analysis²⁰⁻²³ and natural population analysis (NPA)²⁴⁻²⁶ approaches to partitioning the density matrix and the Merz-Singh-Kollman^{27;28} (MSK) and CHelpG²⁹ (CHG) methods for fitting the electrostatic potential.

IV.IV. Results and Discussion

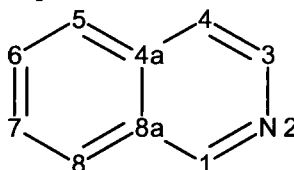
A. Overview

We will explore three of the molecules (isoquinoline, 10-hydroxy-10,9-bozazarophenanthrene, and 10-methyl-10,9-bozazarophenanthrene) Dewar gave as examples where frontier molecular orbital theory (FMO) fails to adequately describe regioselectivity.¹⁴ (Dewar gave a fourth example, nitrobenzene. This molecule has been extensively studied using reactivity indicators associated with conceptual DFT and will not be revisited here. More information on nitrobenzene can be found, for example, in the paper of Langenaeker, Demel, and Geerlings.⁷)

For each molecule, we will first present the experimentally observed reactivity preferences. Then we will present the reactivity preferences predicted by FMO (which predicts that electrophilic attack occurs where the magnitude of the highest occupied molecular orbital is largest), electrostatic considerations (which predicts that electrophilic attack occurs where the electrostatic potential is most negative), and the Fukui function (which predicts that electrophilic attack occurs where $f^-(\mathbf{r})$ is the largest). When no single method can predict the observed reactivity, we will examine the more general index, $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$, which combines information from the electrostatic potential and the Fukui function.

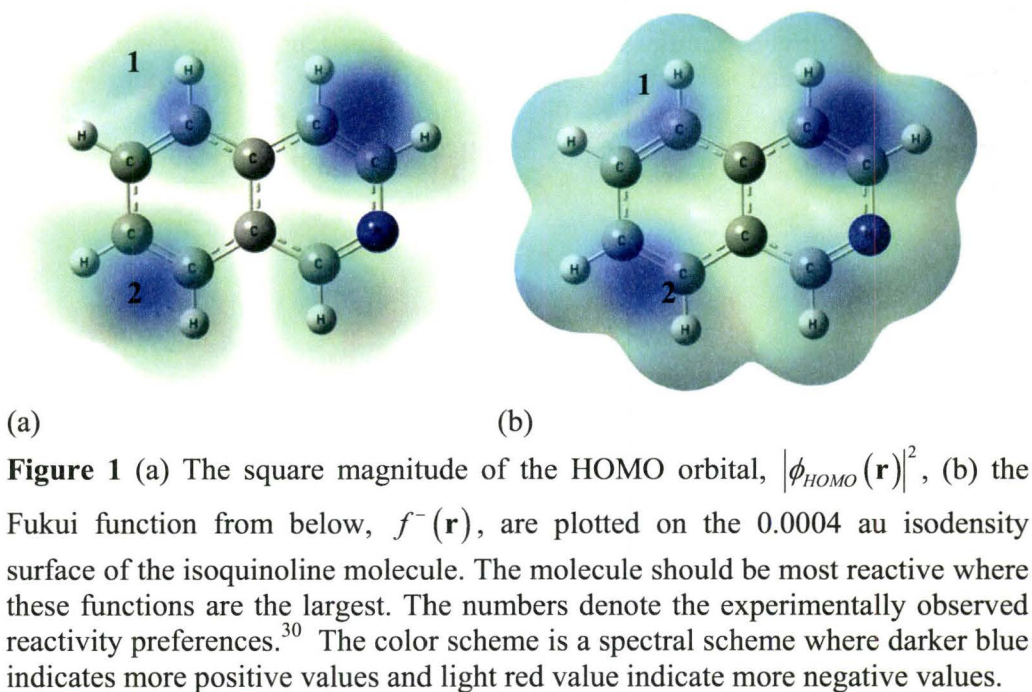
B. Isoquinoline

Experimental studies on Isoquinoline,³⁰



have shown that the most reactive site of this molecule is carbon 5, with secondary reactivity at carbon 8. Products from electrophilic substitution on carbon 4 were not found so this site is believed to be unreactive. Figure 1a reports the value of the highest occupied molecular orbital density, $|\phi_{HOMO}(\mathbf{r})|^2$ on the van der Waals surface of the molecule. We model the van der Waals surface with the $\rho(\mathbf{r}) = .0004$ isodensity surface; this models the reactive surface of the molecule

and is an appropriate indicator of site selectivity.^{31;32;33} In accord with Dewar's FMO analysis, an electrophilic attack is predicted to occur on the double bonds between carbon 3 and 4, then carbon 7 and 8, and finally carbon 5 and 6. This reactivity order is the opposite of the experimental products: carbon 5 is the primary reaction site; carbon 8 is the secondary reaction site, and carbon 4 is unreactive.



Sometimes orbital relaxation effects are important for describing chemical reactivity, and so the Fukui function (which includes orbital relaxation effects)¹¹ sometimes performs better than $|\phi_{HOMO}(\mathbf{r})|^2$ for predicting sites of electrophilic attack. Orbital relaxation effects have been shown to be important in electrophilic aromatic substitution⁷ and electrophilic attack on double bonds (in organic molecules)³¹ and multiple bonds (in inorganic complexes with metal-metal bonds).³² It seems plausible, then, that the Fukui function will locate the appropriate sites for electrophilic attack on isoquinoline. To investigate this possibility, we plotted the value of the Fukui function on the molecular van der Waals surface (see Figure 1b). While the Fukui function and $|\phi_{HOMO}(\mathbf{r})|^2$ are quantitatively different, they are qualitatively similar: the Fukui function at carbon 4 is slightly larger than that at carbon 8, which is significantly larger than the Fukui function at the dominant reaction site (carbon 5). It is interesting that this trend is altered somewhat if condensed Fukui functions are used. With condensed Fukui functions and ChelpG charges, we have $f_{C5} = 0.229$, $f_{C4} = 0.193$, and $f_{C8} = 0.165$. While carbon 4 is still predicted to be reactive, at least the first-choice reaction site is now identified correctly. To the best of our knowledge, this is the first molecule where condensed Fukui functions perform decisively better than Fukui functions plotted on a reactive surface. In this case, the Fukui function on carbon 5 is concentrated near the nucleus and thus, while the Fukui potential (which depends on the total size of the Fukui function in the vicinity of the atom) will be large in the region of carbon 5, the amplitude of the Fukui function (which decays exponentially, rather than as $\frac{1}{r}$) has decreased almost to zero on the reactive surface. Thus chastened, we will henceforth focus on condensed reactivity indicators.

For reactions between hard molecules, the appropriate indicator of site selectivity is the electrostatic potential. This is clear already from the work of Klopman and Berkowitz,^{2,34} but is also a topic of recent emphasis in the conceptual DFT literature.^{13,35,36} It follows very clearly from our analysis also, since the extent of electron transfer

$$\Delta N_{\text{nucleophile}} \approx \frac{\mu_{\text{electrophile}} - \mu_{\text{nucleophile}}}{\eta_{\text{electrophile}} + \eta_{\text{nucleophile}}} \quad (9)$$

will be small when the hardness of the reagents is large. When $|\Delta N|$ is small, the dominant contribution to both $\Xi_{\Delta N \leq 0}^{\kappa}(\mathbf{r}_p)$ and $\Xi_{\Delta N \geq 0}^{\kappa}(\mathbf{r}_p)$ is from the molecular electrostatic potential. While isoquinoline is not especially hard, insofar as the Fukui function has failed to successfully describe its reactivity, it seems desirable to explore the electrostatic potential. This is done by computing the atomic

charges, which represent the “condensed” electrostatic potential (cf. Eq. (5)).³⁷ The charges from several different population analysis schemes are plotted in Figure 2a. The Mulliken charges are manifestly unreliable, as is to be expected for a basis set including diffuse functions. Henceforth we will not report the results from Mulliken population analysis. The other charge schemes reported in Figure 2a are more reasonable, with the two methods of electrostatic potential fitting (Merz-Singh-Kollmann (MSK) and CHelpG (CHG)) giving similar results. The natural population analysis (NPA) scheme is less similar, which may also be due to this method’s stronger dependence on the basis set or due to the fact that NPA, unlike MSK and CHG, is based on a population analysis of the density matrix, and not on fitting the electrostatic potential.

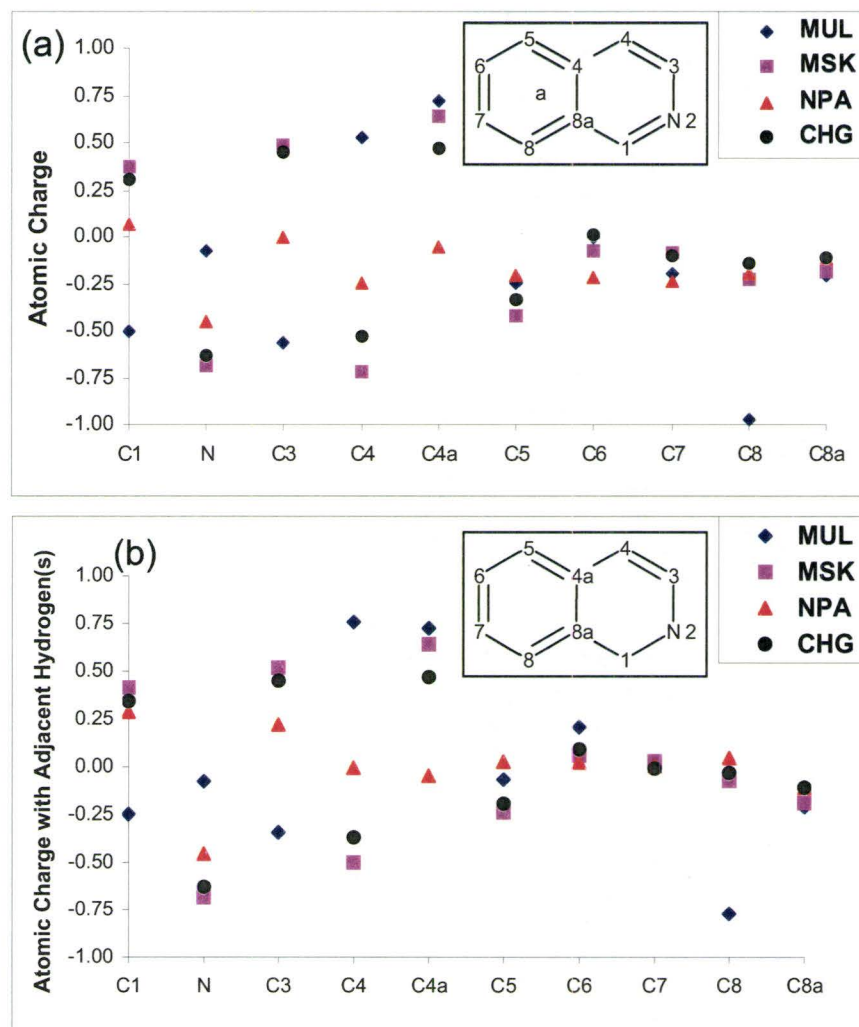


Figure 2 The atomic charges (condensed electrostatic potential) of isoquinoline. (a) The atomic charges on the indicated atoms. (b) The atomic charges on the indicated atoms plus the charges of hydrogen atoms bonded to those atoms. The atomic numbering scheme is included as an inset. Experimentally, C5 is most reactive, followed by C8. C4 is unreactive.

Recall that the reactive site of strong electrophiles is usually positively charged. Electrophilic attack thus tends to occur at the most negatively charged sites of the nucleophile. Looking at Figure 2a, it is clear that the most negative sites in the molecule are the nitrogen atom and carbon 4. Carbon 4, however, is unreactive. The charges on carbon 5 (the most reactive position) and carbon 8 (the second most reactive position) are not especially small.

Sometimes one argues that one should add the charges of hydrogen atoms into the charges of the atoms they are bonded to. This seems especially useful in electrostatic fitting procedures: because the carbon-hydrogen bond is short, it can be difficult to determine how to partition a reactive carbon's charge between the carbon atom and the adjacent hydrogen. More generally, the regioselectivity of a molecule is usually determined by interactions that occur when the molecular substrate and the attacking reagent are in van-der-Waals contact. Because the separation between the molecules is much larger than the length of a carbon-hydrogen bond, from the perspective of the attacking reagent, $-CH_n$ groups appear as a single point charge. Based on this reasoning, it is preferable to consider "functional group" charges that are computed by adding the charges of hydrogen atoms to the charges of the adjacent "heavy" atom. This data is reported in Figure 2b. Unfortunately, this does not alter the fact that carbon 4 is predicted to be reactive, while the molecular sites that actually are reactive are predicted to be relatively unreactive.

Further thought about the chemistry of isoquinoline rectifies these unsuccessful predictions. Isoquinoline is relatively basic, $pK_b = 8.6$. Unsurprisingly, this is similar to the pK_b of pyridine, which is 8.7. Experimentally, it is difficult to perform an electrophilic substitution reaction on pyridine: electrophiles are Lewis acids, and so the solutions used for electrophilic aromatic substitutions are also acidic.³⁸ The experimental studies of Dewar were carried out in a mixture of nitric and sulfuric acids!³⁰ Even when less extreme conditions are used, the pH of those solutions is almost always less than 5.3, and in that environment pyridine is protonated. Clearly, performing an electrophilic attack on a protonated molecule will be difficult! Isoquinoline, which is marginally less basic than pyridine, is also expected to be in its protonated state when it undergoes electrophilic aromatic substitution. Based on this logic, we performed calculations on protonated isoquinoline.³⁹ The atomic charges are reported in Figures 3a (raw charges) and 3b (summed with adjacent hydrogens). The condensed Fukui functions are reported in figures 4a (computed from atomic charges) and 4b (computed from atomic charges summed with adjacent hydrogens). Based on the charges, we would predict that the reaction occurs at carbon 5, with carbon 8 and carbon 4 having similar reactivity. Based on the Fukui function, we would predict that carbon 5 and carbon 8 are both highly

reactive (carbon 5 perhaps slightly more so), and that carbon 4 is not very reactive. This agrees with experiment: the substantial negative charge on carbon 5, coupled with its significant Fukui function, makes this site highly susceptible to electrophilic attack. carbon 8 is also favorable electrostatically and based on its Fukui function, but its less negative charge is associated with slightly diminished reactivity. Carbon 4 is not significantly reactive compared to carbon 5 or carbon 8.

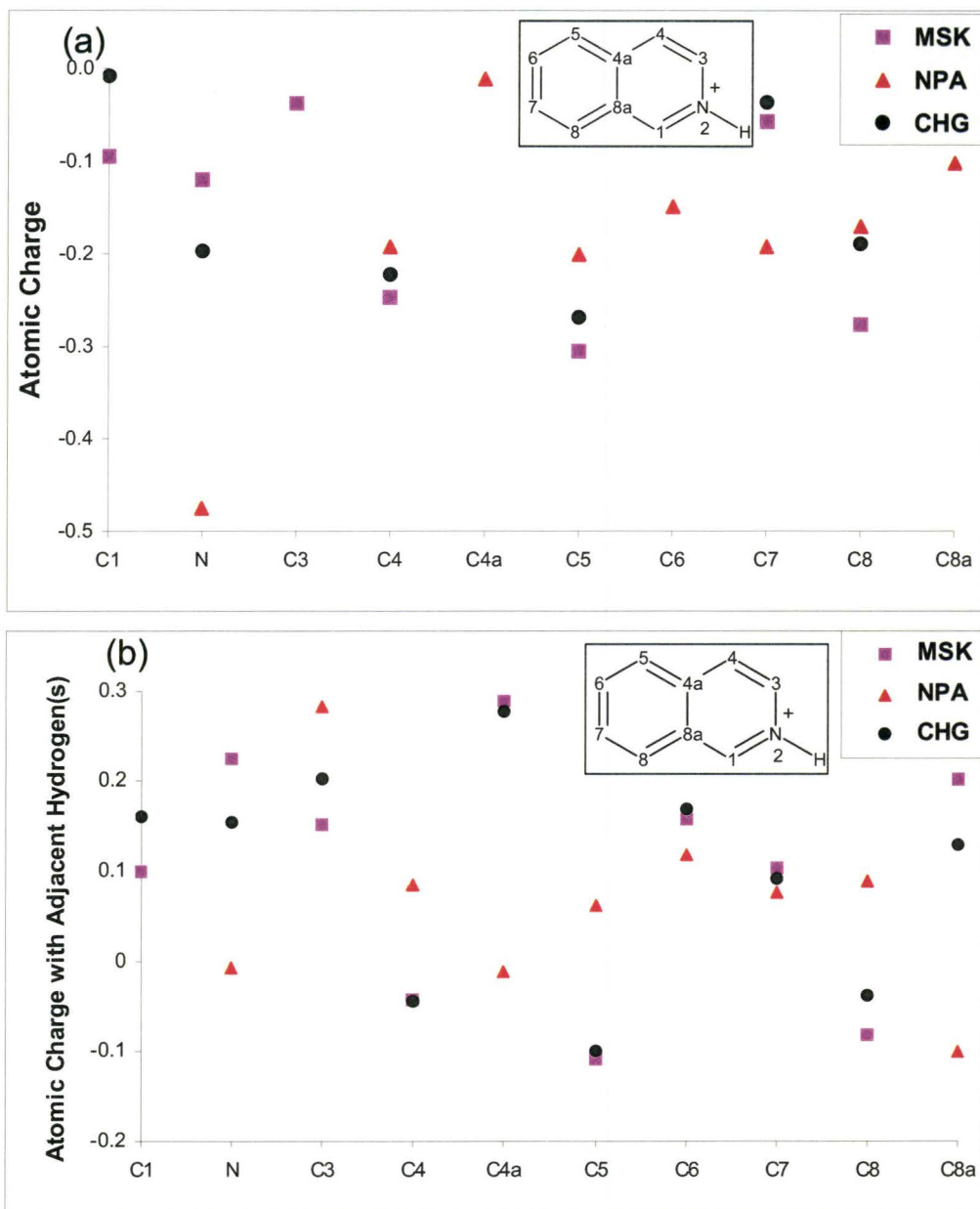


Figure 3 The atomic charges (condensed electrostatic potential) of *protonated* isoquinoline.

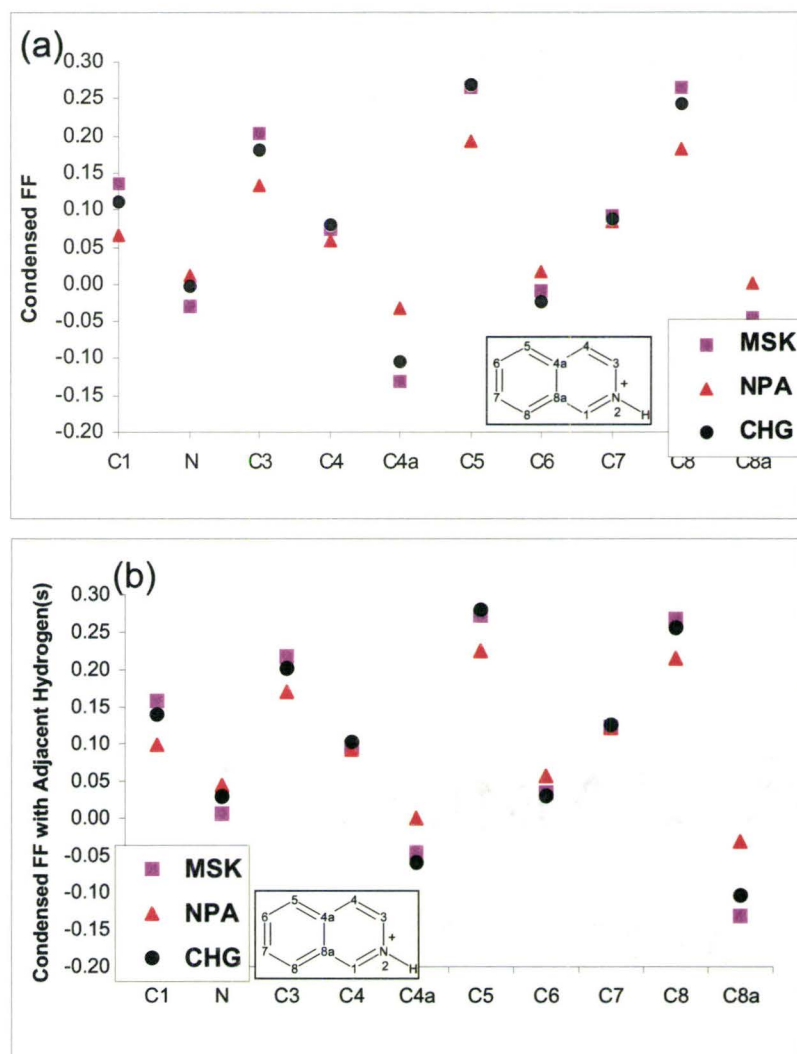


Figure 4 The condensed Fukui function of protonated isoquinoline. In (a), we compute the Fukui function from the difference of atomic charges. In (b), we add to each indicated atom the condensed Fukui functions of the hydrogen atoms bonded to it. The atomic numbering scheme is included as an inset. Experimentally, C5 is most reactive, followed by C8. C4 is unreactive.

Perhaps protonating isoquinoline is too strong an assumption. One could, after all, carry out an electrophilic substitution in an aprotic solvent. To explore this possibility, we considered the complex of pyridine with the sodium cation, which is an extremely weak Lewis acid. (See figure 5.) In this case, carbon 4 remains the most negatively charged reactive site. However, the Fukui function is highly concentrated on carbon 5 and, to a lesser extent, carbon 8. Since isoquinoline is not especially hard, it seems reasonable to infer that under any reasonable set of experimental conditions, the “ion paired isoquinoline” that is subject to chemical reaction will react first at carbon 5, with a secondary product associated with reaction at carbon 8.

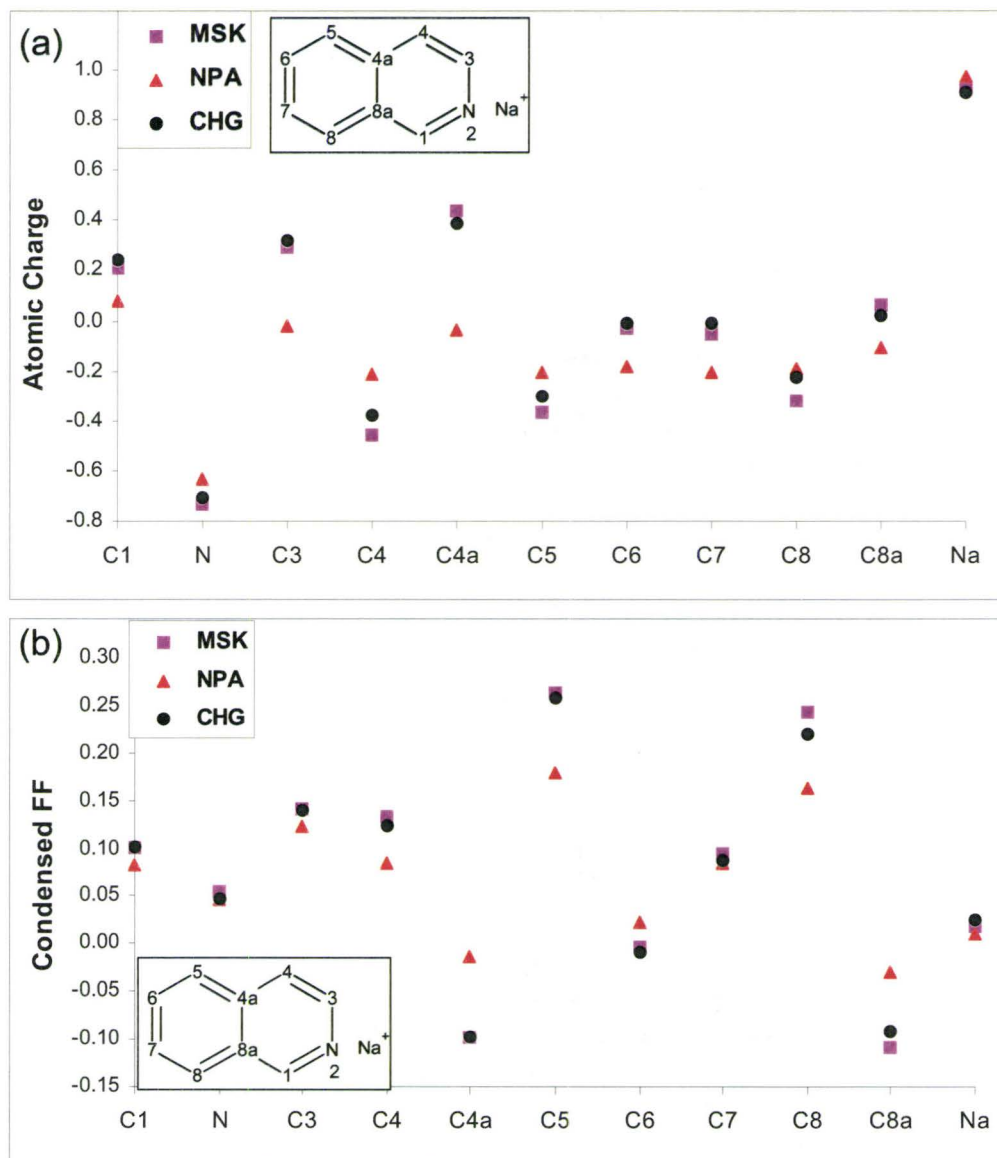


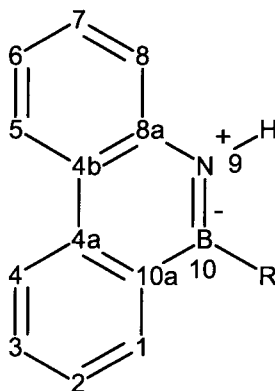
Figure 5 The (a) atomic charges and (b) condensed Fukui functions of isoquinoline with a sodium “spectator cation.” Qualitatively, the plot with hydrogens summed into adjacent carbons is very similar.

We could apply our general-purpose reactivity indicator to isoquinoline, but the picture that emerges is quite boring since the electrostatically favored and electron-transfer favored sites on protonated isoquinoline are the same. Instead, we will apply our reactivity indicator to a much more challenging problem: electrophilic aromatic substitution on 10,9-borazarophenanthrenes.

C. 10-R-10,9-borazarophenanthrene

1. Summary of Experimental Observations

In addition to isoquinoline, Dewar also pointed out two, more challenging, molecules where experimental results were not in accord with frontier molecular orbital theory. Experimental studies of 10-R-10,9-borazarophenanthrene (R=OH or R=CH₃),



indicate that carbon 8 and carbon 6 are both susceptible to electrophilic attack.⁴⁰⁻⁴² Chlorination of 10-methyl-10,9-borazarophenanthrene favors carbon 8 over carbon 6,⁴² while chlorination of 10-hydroxy-10,9-borazarophenanthrene gives the disubstituted product corresponding to reaction at both carbon 6 and carbon 8.⁴¹ Nitration produces a mixture of the products associated with reaction at carbon 6 and carbon 8.⁴² It seems reasonable to infer that carbon 6 and carbon 8 are the most reactive positions, with carbon 8 being slightly more reactive, at least for some electrophiles. If one increases the temperature and the amount of reagent, then one can add another chlorine to 10-hydroxy-10,9-borazarophenanthrene; forming the 2,6,8 trisubstituted product.⁴¹ Carbon 2 is essentially unreactive with respect to nitration and bromination, however.^{41;42} We infer, then, that the experimental reactivity profile can be summarized as

$$C8 \gtrsim C6 \gg C2. \quad (10)$$

Theoretical electronic-structure treatments suggest that carbon 4 and carbon 2 should have similar reactivity. It is known, however, that steric congestion significantly reduces the rate of reaction at carbon 4.⁴³

One might expect that, like isoquinoline, 10-*R*-10,9-borazarophenanthrene would be protonated. This is not the case. First of all, note that the nitrogen atom in 10,9-bozazarophenanthrene already has four bonds, which reduces its susceptibility to protonation. Second, the presence of the adjacent boron atom (which accepts electrons from the *p* orbital of the *sp*² hybridized nitrogen atom) reduces the basicity of the nitrogen. Protonation does occur, but only when the compound is heated in concentrated sulfuric acid.^{40;44} (Upon protonation in that environment, the central ring breaks and boron is lost.) Under any reasonable set of experimental conditions, then, 10-*R*-10,9-bozazarophenanthrene is not protonated.

2. Frontier Molecular Orbitals, Condensed Fukui Functions, and Atomic Charges

Frontier molecular orbital theory does not predict the reactivity in these molecules. According to FMO, carbon 2 is slightly more reactive than carbon 4, which is slightly more reactive than the bond between carbon 6 and carbon 7.¹⁴ (Both carbon 6 and carbon 7 have substantial contributions from the HOMO orbital.) There is very little frontier molecular orbital density on carbon 8.

The condensed Fukui functions for 10-hydroxy-10,9-bozaroarphenanthrene (R=OH) are reported in Figures 6b and 7b. The condensed Fukui function gives results that are more in line with experiment than frontier molecular orbital theory. In particular, the Fukui function predicts that carbon 6 is the most reactive position in the molecule. Carbon 2 and carbon 4b are the next most reactive positions, followed by carbon 4. The Fukui function predicts that carbon 8 is essentially unreactive. The predicted reactivity at carbon 4b demonstrates a recurrent feature in electrophilic polyaromatic substitution reactions. Qualitative reactivity indicators often predict ipso addition of electrophiles to polyaromatic compounds but, in many cases, ipso addition is a mechanistic dead end: although it is sometimes the most favorable orientation for the reactants, the barrier separating the ipso reactive intermediate from stable product molecules is very high. For example, because there are no hydrogen atoms at carbon 4b, electrophilic aromatic substitution cannot occur. Consequently, addition at carbon 4b requires a loss of aromaticity.

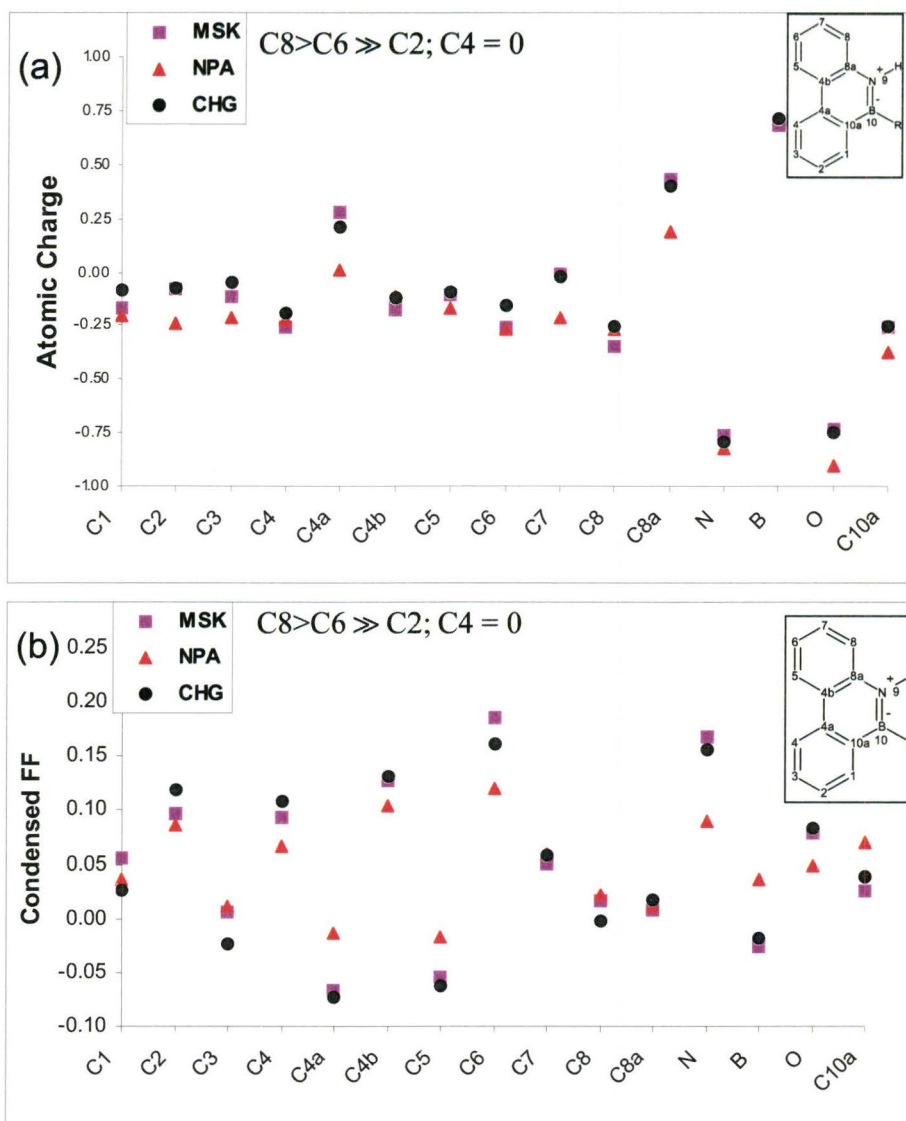


Figure 6 The (a) atomic charges and (b) condensed Fukui functions of 10-hydroxy-10,9-benzo[ar]phenanthrene. The atomic numbering scheme is included as an inset. Experimentally, C8 is most reactive, followed by C6 and then C2. Other sites are unreactive.

The atomic charges are reported in Figures 6a and 7a. Though carbon 10a, nitrogen, and oxygen are all negatively charged, these sites are not susceptible to electrophilic aromatic substitution. While a positively charged electrophile might form an “ion pair” association complex with one of these sites, no further reaction at these sites is possible under ordinary conditions. For carbon 10a, there is no hydrogen atom to serve as a leaving group. While there is a hydrogen bound to the nitrogen atom, bonds between hydrogen and electron-deficient sp^2 hybridized nitrogen atoms are very strong, so this hydrogen atom is a poor leaving group.⁴⁵ Oxygen-hydrogen bonds are generally stronger than carbon-hydrogen bonds, so the hydrogen atom bonded to oxygen is also a poor leaving group. Consequently, we expect the chemistry at C10a, nitrogen, and oxygen is limited to the formation and dissociation of weak “ion pairing” type interactions.

Among sites that are susceptible to electrophilic aromatic substitution, carbon 8 is the most negatively charged, with carbon 6 and carbon 4 having somewhat smaller, but still substantial, negative charges. Carbon 2 is also negatively charged.

We predict, then, that carbon 8 and carbon 6 are the most highly reactive sites based on electrostatic effects and electron-transfer effects, respectively. Carbon 2 and carbon 4 represent trade-offs: these molecular sites have reasonably large Fukui functions and reasonably negative charges. If the reactivity of 10-hydroxy-10,9-bozaroarphenanthrene were strongly electrostatically controlled, then carbon 6 should be less reactive because large values for the Fukui function are unfavorable in that situation. Similarly, if the reactivity were strongly electron-transfer controlled, then carbon 8 should be less reactive because small values of the electrostatic potential are unfavorable in that situation. Both carbon 8 and carbon 6 are observed to be reactive, so the reactivity must be jointly controlled by electrostatic and electron-transfer effects.

Because Merz-Singh-Kollman (MSK) and ChelpG (CHG) charges are both designed to reproduce the electrostatic potential, these two choices of charges should resemble one another and, hopefully, also the charges derived from natural population analysis (NPA). Examining Figure 7, it seems that the agreement between the different population analysis schemes is slightly better if the charges on the heavy atoms are combined with the charges on their adjacent hydrogen atoms. We will base the rest of our analysis on the plots with the hydrogen atoms summed in; due to the similarity between the results in Figures 6 and 7, this will not affect our results very much. Indeed, the main trends in the values of the charges and the condensed Fukui functions seem to be reproduced no matter which population analysis scheme is being employed. This is

reassuring, because there is no *a priori* reason to assert the superiority of any one of the population analysis schemes.⁴⁶

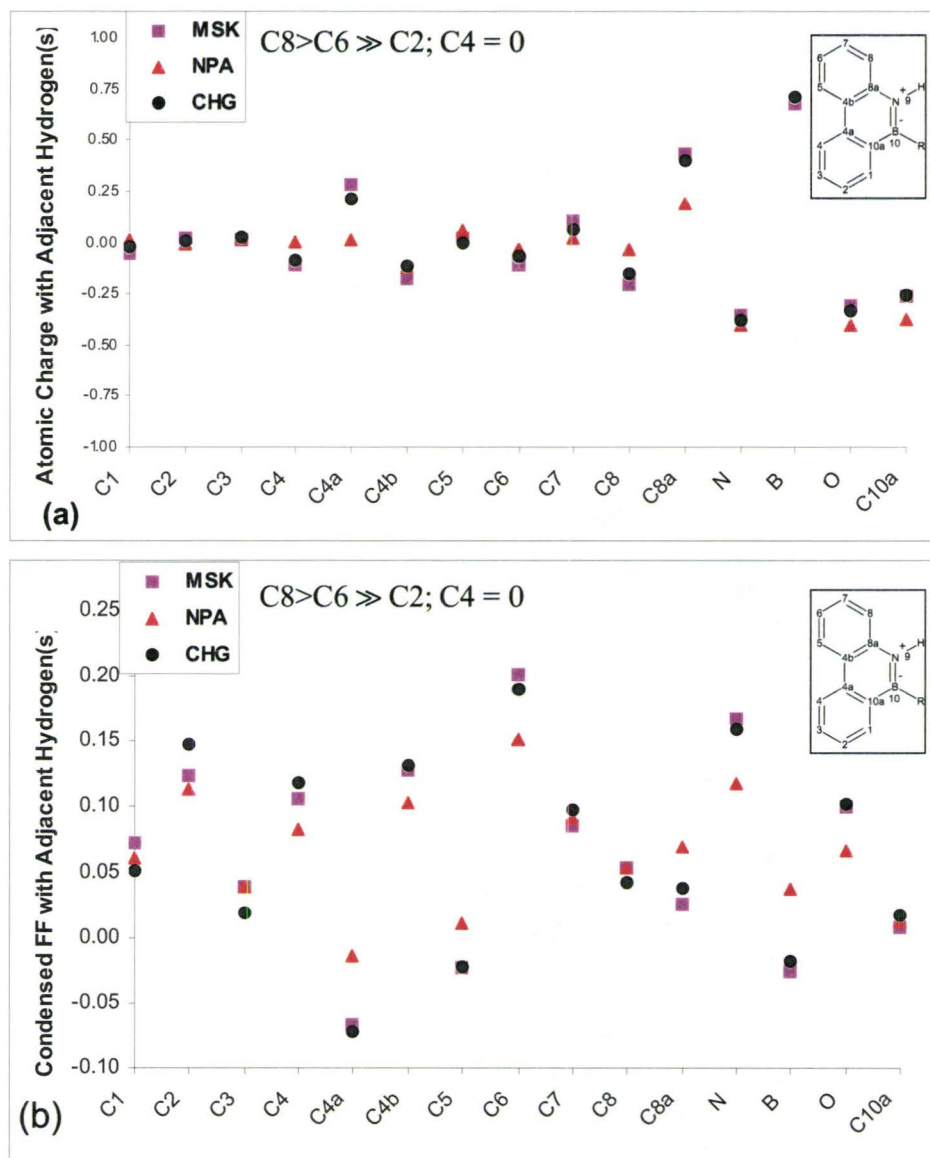


Figure 7 The (a) “hydrogen summed” atomic charges and (b) “hydrogen summed” condensed Fukui functions of 10-hydroxy-10,9-bozaroarphenanthrene. Unlike Figure 6, the reactivity indicators for hydrogen atoms have been added to the reactivity indicators of the heavy atoms to which they are bonded.

3. Application of the Condensed General-Purpose Reactivity Indicator

Because 10-hydroxy-10,9-bozazarophenanthrene is a case of joint electrostatic and electron-transfer control, it seems to be a suitable test for the general-purpose reactivity indicator developed in the previous paper.¹ It is difficult to distill the immense amount of information in this indicator,

$$\Xi_{\Delta N \leq 0, \alpha}^{\kappa} = (\kappa + 1)q_{\text{nucleophile}, \alpha}^{(0)} - \Delta N(\kappa - 1)f_{\text{nucleophile}, \alpha}^{-} \quad (11)$$

into an easily digestible form, however. Recall that $\Xi_{\Delta N, \alpha}^{\kappa}$ depends on two parameters: κ (which measures the relative importance of electrostatic control ($\kappa \approx 1$) and electron-transfer control ($\kappa \approx -1$)) and ΔN , which measures the amount of electron transfer. $\Xi_{\Delta N, \alpha}^{\kappa}$ is a bivariate function for every atom in the system! What we really want to know is which site is most reactive, and how (and whether) the choice of reaction site changes due to changing electrophilic reagents and reaction conditions. We have found that the key information about the reactivity of the molecule can be summarized using what we term “reactivity transition tables.” (See Tables 1-4.) To make a reactivity transition table, one starts by computing the value of $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ for every atom in the molecule and for the entire chemically relevant range of choices for the amount of electron transfer ($-1 \leq \Delta N \leq 0$) and the extent of electrostatic/electron-transfer control (κ). As established in the previous section, electrophilic attack on 10-hydroxy-10,9-bozaroarphenanthrene is jointly controlled by electrostatic and electron-transfer effects, so we consider only $-1 \leq \kappa \leq 1$. In constructing the reactivity transition tables, we restrict ourselves to carbons that are susceptible to electrophilic aromatic substitution.

One determines which atom is most reactive by locating the atom with the smallest value of $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$. We insert this value in the “first choice” reactivity transition table and then color-code the cell so that it is clear which carbon is reacting. The second most reactive atom is the one with the second smallest value for $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$. We insert this value in the “second choice” reactivity transition table and color-code the cells to indicate the second most reactive carbon. If the difference between the values of $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ for the first and second choices is relatively small, then, one expects that both possible product molecules will form. If the difference between the “first choice” and “second choice” values of $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ is relatively large, vigorous reaction conditions may be required to form the secondary product.

Reactivity transition tables contain both qualitative and quantitative data on reactivity. At a quantitative level, the “first choice” and “second choice” reactivity transition tables give information about the relative favorability of the primary and secondary product. At a qualitative level, the reactivity transition table can be read as a “phase diagram” for chemical reactivity. Examine Table 1a. When the nucleophile reacts with a very hard electrophile (so that the reaction is mostly electrostatically controlled and $\kappa \approx 1$), reactions occur at carbon 8. As the electrophile becomes softer and the extent of electron transfer increases, carbon 6 becomes the preferred site for reactivity. In the “transition region” between carbon 8 and carbon 6, one would expect a mixture of products. In this way, a reactivity transition table contains information about how to choose the electrophile so that a desired product is formed.

Reactivity transition tables for different types of population analyses are provided in Table 1 (NPA), Table 2 (MSK), and Table 3 (CHelpG). In each case, carbon 8 is the most favorable site when electrostatic effects are dominant ($\kappa \approx 1$, $\Delta N \approx 0$) and carbon 6 is the most favorable site when electron-transfer effects are dominant ($\kappa \approx -1$, $\Delta N \approx -1$). Experimental results indicate that chlorination and nitration of 10-hydroxy-10,9-bozaroarphenanthrene occurs on carbon 8 and carbon 6, with a small preference for carbon 8. In these reactions, the electrophile is reasonably hard, and so the reaction is probably more electrostatically controlled than it is electron-transfer controlled ($0 < \kappa < 1$). Electron transfer to the electrophile will be important, but incomplete ($\Delta N \approx -.5$). In Table 2a and 3a (based on CHelpG and MSK population analysis, with the hydrogenic contributions summed into the adjacent heavy atoms), these experimental conditions place one near the transition region between carbon 8 and carbon 6, but in regions where carbon 8 would be predicted to be slightly more reactive. In Table 1a, (based on natural population analysis), this places one in the region where carbon 6 is most favorable, though one is still reasonably close to the region where carbon 8 would be favored.

Table 1 Reactivity Transition Tables for 10-hydroxy-10,9-benzoarphenanthrene using natural population analysis for the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice: the minimum values of $\Xi_{\Delta N \leq 0}^{\kappa}$ denotes where the molecule is most reactive. (b) Second-Choice: the second smallest values for $\Xi_{\Delta N \leq 0}^{\kappa}$, denoting the second most reactive carbon.

(a)	κ										
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0
-1.0	-0.072	-0.090	-0.113	-0.137	-0.161	-0.185	-0.208	-0.232	-0.256	-0.280	-0.303
-0.9	-0.072	-0.086	-0.107	-0.128	-0.149	-0.169	-0.190	-0.211	-0.232	-0.252	-0.273
-0.8	-0.072	-0.083	-0.101	-0.119	-0.137	-0.154	-0.172	-0.190	-0.207	-0.225	-0.243
-0.7	-0.072	-0.080	-0.095	-0.110	-0.124	-0.139	-0.154	-0.168	-0.183	-0.198	-0.212
-0.6	-0.072	-0.077	-0.089	-0.101	-0.112	-0.124	-0.136	-0.147	-0.159	-0.170	-0.182
-0.5	-0.072	-0.074	-0.083	-0.092	-0.100	-0.109	-0.117	-0.126	-0.135	-0.143	-0.152
-0.4	-0.072	-0.071	-0.077	-0.082	-0.088	-0.094	-0.099	-0.105	-0.110	-0.116	-0.121
-0.3	-0.072	-0.068	-0.071	-0.073	-0.076	-0.078	-0.081	-0.083	-0.086	-0.088	-0.091
-0.2	-0.072	-0.067	-0.065	-0.064	-0.064	-0.063	-0.063	-0.062	-0.062	-0.061	-0.061
-0.1	-0.072	-0.066	-0.060	-0.055	-0.052	-0.048	-0.045	-0.041	-0.037	-0.034	-0.030
0.0	-0.072	-0.065	-0.058	-0.051	-0.043	-0.036	-0.029	-0.022	-0.014	-0.007	
	Carbon 2		Carbon 4			Carbon 6			Carbon 8		

(b)	κ										
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0
-1.0	-0.066	-0.076	-0.079	-0.083	-0.095	-0.117	-0.139	-0.161	-0.184	-0.206	-0.228
-0.9	-0.066	-0.075	-0.077	-0.080	-0.086	-0.106	-0.125	-0.145	-0.165	-0.185	-0.205
-0.8	-0.066	-0.074	-0.075	-0.076	-0.078	-0.094	-0.112	-0.129	-0.147	-0.165	-0.182
-0.7	-0.066	-0.073	-0.073	-0.073	-0.073	-0.083	-0.098	-0.113	-0.129	-0.144	-0.160
-0.6	-0.066	-0.071	-0.071	-0.070	-0.069	-0.071	-0.084	-0.098	-0.111	-0.124	-0.137
-0.5	-0.066	-0.070	-0.069	-0.067	-0.065	-0.063	-0.071	-0.082	-0.092	-0.103	-0.114
-0.4	-0.066	-0.069	-0.066	-0.063	-0.061	-0.058	-0.057	-0.066	-0.074	-0.083	-0.091
-0.3	-0.066	-0.068	-0.064	-0.060	-0.056	-0.052	-0.048	-0.050	-0.056	-0.062	-0.068
-0.2	-0.066	-0.065	-0.062	-0.057	-0.052	-0.047	-0.042	-0.037	-0.038	-0.042	-0.046
-0.1	-0.066	-0.062	-0.059	-0.054	-0.048	-0.041	-0.035	-0.029	-0.023	-0.021	-0.023
0.0	-0.066	-0.059	-0.053	-0.046	-0.039	-0.033	-0.026	-0.020	-0.013	-0.007	
	Carbon 2		Carbon 4			Carbon 6			Carbon 8		

Table 2 Reactivity Transition Tables for 10-hydroxy-10,9-bozaroarphenanthrene using CHelpG population analysis with the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice: the minimum values of $\Xi_{\Delta N \leq 0}^{\kappa}$ denotes where the molecule is most reactive. (b) Second-Choice: the second smallest values for $\Xi_{\Delta N \leq 0}^{\kappa}$, denoting the second most reactive carbon.

(a)	κ											
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0	
-1.0	-0.312	-0.289	-0.267	-0.244	-0.233	-0.257	-0.282	-0.307	-0.331	-0.356	-0.381	
-0.9	-0.312	-0.289	-0.265	-0.241	-0.218	-0.238	-0.259	-0.280	-0.301	-0.322	-0.343	
-0.8	-0.312	-0.288	-0.263	-0.239	-0.214	-0.219	-0.236	-0.253	-0.270	-0.288	-0.305	
-0.7	-0.312	-0.287	-0.262	-0.236	-0.211	-0.200	-0.214	-0.227	-0.240	-0.253	-0.266	
-0.6	-0.312	-0.286	-0.260	-0.234	-0.208	-0.181	-0.191	-0.200	-0.210	-0.219	-0.228	
-0.5	-0.312	-0.285	-0.258	-0.231	-0.204	-0.177	-0.168	-0.173	-0.179	-0.185	-0.190	
-0.4	-0.312	-0.284	-0.257	-0.229	-0.201	-0.173	-0.145	-0.147	-0.149	-0.150	-0.152	
-0.3	-0.312	-0.284	-0.255	-0.226	-0.197	-0.169	-0.140	-0.120	-0.118	-0.116	-0.114	
-0.2	-0.312	-0.283	-0.253	-0.224	-0.194	-0.165	-0.135	-0.105	-0.088	-0.082	-0.076	
-0.1	-0.312	-0.282	-0.251	-0.221	-0.191	-0.160	-0.130	-0.100	-0.069	-0.048	-0.038	
0.0	-0.312	-0.281	-0.250	-0.219	-0.187	-0.156	-0.125	-0.094	-0.062	-0.031		
	Carbon 2		Carbon 4			Carbon 6			Carbon 8			

(b)	κ											
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0	
-1.0	-0.174	-0.180	-0.187	-0.208	-0.221	-0.206	-0.212	-0.219	-0.234	-0.264	-0.294	
-0.9	-0.174	-0.178	-0.182	-0.197	-0.218	-0.194	-0.198	-0.202	-0.211	-0.237	-0.264	
-0.8	-0.174	-0.176	-0.177	-0.185	-0.202	-0.190	-0.184	-0.185	-0.187	-0.211	-0.235	
-0.7	-0.174	-0.173	-0.173	-0.174	-0.187	-0.186	-0.170	-0.169	-0.168	-0.185	-0.206	
-0.6	-0.174	-0.171	-0.168	-0.165	-0.172	-0.181	-0.155	-0.152	-0.149	-0.158	-0.176	
-0.5	-0.174	-0.169	-0.163	-0.158	-0.157	-0.162	-0.150	-0.135	-0.130	-0.132	-0.147	
-0.4	-0.174	-0.166	-0.158	-0.150	-0.143	-0.143	-0.145	-0.119	-0.111	-0.105	-0.117	
-0.3	-0.174	-0.164	-0.154	-0.143	-0.133	-0.124	-0.122	-0.111	-0.092	-0.082	-0.088	
-0.2	-0.174	-0.161	-0.149	-0.136	-0.123	-0.111	-0.099	-0.094	-0.076	-0.060	-0.059	
-0.1	-0.174	-0.159	-0.144	-0.129	-0.114	-0.099	-0.084	-0.069	-0.057	-0.039	-0.029	
0.0	-0.174	-0.157	-0.139	-0.122	-0.104	-0.087	-0.070	-0.052	-0.035	-0.017		
	Carbon 2		Carbon 4			Carbon 6			Carbon 8			

Table 3 Reactivity Transition Tables for 10-hydroxy-10,9-benzo[ar]phenanthrene using Merz-Singh-Kollman population analysis with the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice: the minimum values of $\Xi_{\Delta N \leq 0}^{\kappa}$ denotes where the molecule is most reactive. (b) Second-Choice: the second smallest values for $\Xi_{\Delta N \leq 0}^{\kappa}$, denoting the second most reactive carbon.

(a)	κ										
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0
-1.0	-0.697	-0.631	-0.565	-0.499	-0.453	-0.439	-0.426	-0.412	-0.398	-0.385	-0.371
-0.9	-0.697	-0.631	-0.564	-0.498	-0.438	-0.421	-0.404	-0.386	-0.369	-0.351	-0.334
-0.8	-0.697	-0.630	-0.563	-0.497	-0.430	-0.402	-0.381	-0.360	-0.339	-0.318	-0.297
-0.7	-0.697	-0.630	-0.563	-0.496	-0.428	-0.384	-0.359	-0.334	-0.309	-0.284	-0.260
-0.6	-0.697	-0.630	-0.562	-0.494	-0.427	-0.365	-0.337	-0.308	-0.280	-0.251	-0.223
-0.5	-0.697	-0.629	-0.561	-0.493	-0.425	-0.358	-0.314	-0.282	-0.250	-0.218	-0.185
-0.4	-0.697	-0.629	-0.561	-0.492	-0.424	-0.356	-0.292	-0.256	-0.220	-0.184	-0.148
-0.3	-0.697	-0.629	-0.560	-0.491	-0.423	-0.354	-0.285	-0.230	-0.191	-0.151	-0.111
-0.2	-0.697	-0.628	-0.559	-0.490	-0.421	-0.352	-0.283	-0.214	-0.161	-0.118	-0.074
-0.1	-0.697	-0.628	-0.558	-0.489	-0.420	-0.350	-0.281	-0.212	-0.142	-0.084	-0.037
0.0	-0.697	-0.627	-0.558	-0.488	-0.418	-0.349	-0.279	-0.209	-0.139	-0.070	
	Carbon 2		Carbon 4			Carbon 6			Carbon 8		

(b)	κ										
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0
-1.0	-0.513	-0.494	-0.481	-0.467	-0.433	-0.366	-0.318	-0.286	-0.254	-0.221	-0.194
-0.9	-0.513	-0.491	-0.473	-0.456	-0.431	-0.365	-0.307	-0.273	-0.238	-0.204	-0.175
-0.8	-0.513	-0.487	-0.466	-0.445	-0.424	-0.363	-0.296	-0.260	-0.223	-0.187	-0.156
-0.7	-0.513	-0.483	-0.458	-0.433	-0.409	-0.361	-0.294	-0.246	-0.208	-0.170	-0.136
-0.6	-0.513	-0.479	-0.451	-0.422	-0.394	-0.359	-0.292	-0.233	-0.193	-0.153	-0.117
-0.5	-0.513	-0.476	-0.443	-0.411	-0.379	-0.347	-0.290	-0.222	-0.178	-0.136	-0.097
-0.4	-0.513	-0.472	-0.436	-0.400	-0.364	-0.328	-0.287	-0.219	-0.163	-0.119	-0.078
-0.3	-0.513	-0.468	-0.429	-0.389	-0.349	-0.310	-0.270	-0.217	-0.148	-0.102	-0.058
-0.2	-0.513	-0.466	-0.421	-0.378	-0.334	-0.291	-0.248	-0.204	-0.145	-0.085	-0.039
-0.1	-0.513	-0.464	-0.414	-0.367	-0.320	-0.273	-0.225	-0.178	-0.131	-0.073	-0.019
0.0	-0.513	-0.462	-0.411	-0.359	-0.308	-0.257	-0.205	-0.154	-0.103	-0.051	
	Carbon 2		Carbon 4			Carbon 6			Carbon 8		

Examining the second-choice reactivity transition diagrams in Tables 1b (NPA), 2b (ChelpG), and 3b (MSK), one finds that carbon 4 and carbon 2 are also predicted to be reactive. Carbon 4 would be predicted to be more reactive than carbon 2 except when electron-transfer effects are dominant ($\kappa \approx -1$). Experimentally, carbon 4 is much less reactive than one would expect based on electronic-structure considerations alone. Dewar noted this effect as early as 1956, and suggested that the abnormally low reactivity of carbon 4 is due to steric effects, rather than electronic structure considerations.⁴³ Our indicator does not include any information about steric hindrance, so it is not surprising that it also overestimates the reactivity of carbon 4. Examining the values of $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ where carbon 4 emerges as the “second choice” reactivity site, we observe that there is a large gap between the predicted reactivity of carbon 4 and the reactivity of the most reactive site.⁴⁷ This indicates that whenever carbon 4 is the “second choice” reactive site, the “first choice” reactive site is much more reactive. When the second-choice reactive site is much less favorable than the first-choice reactive site, one expects that the secondary product will be a very small percentage of the total yield, and thus difficult to isolate and characterize.

The overall picture that emerges is fairly convincing: carbon 6 and carbon 8 are most reactive, with the population analysis schemes based on electrostatic fitting successfully predicting that carbon 8 should be slightly more reactive. Carbon 2 is predicted to be much less reactive, which agrees with the experimental observation that reactions at carbon 2 only occur under very special conditions: chlorination at high temperatures with excess chlorine.⁴¹

Our analysis also indicates that under the appropriate conditions, carbon 6 should be more reactive than carbon 8. Indeed, Friedel-Crafts acetylation of 10-hydroxy-10,9-benzophenanthrene occurs predominately at carbon 6, with the diacetylated product corresponding to reaction at both carbon 6 and carbon 8 as a secondary product.⁴⁴ The monoacetylated product corresponding to reaction at carbon 8 is not observed. Dewar explained this by hypothesizing that the nitrogen atom is complexed by the AlCl_3 catalyst, which might make carbon 8 sterically inaccessible.^{44,48} Our analysis suggests another possibility, however. The mechanism of Friedel-Crafts acetylation involves the addition of the resonance-stabilized acetyl carbocation (CH_3CO^+) to the aromatic ring. Because this is a carbocation, we expect that $\Delta N \approx -1$. Moreover, this carbocation is resonance stabilized and, additionally, might be somewhat stabilized by complexation of the catalyst. As such, the charge on the electrophilic carbon atom is relatively small and, moreover, we expect that the acetyl carbocation is relatively soft. Accordingly, we expect that the reaction will be mostly electron-transfer controlled, $-1 < \kappa < 0$. Referring to Tables 1a through 3a, one finds that under

these conditions, carbon 6 is much more reactive than any of the other molecular sites.

To this point we have focused on hydroxylated 10,9-bozazarophenanthrene, ($R=OH$) rather than on the methylated compound. For the hydroxylated compound, the various population analysis schemes gave qualitatively similar results. This was not true for $R=CH_3$, and summing the charges of the hydrogen atoms into the adjacent heavy atoms did not substantially improve the agreement between the various population analysis schemes. Indeed, even though the MSK and CHelpG charges are both based on fitting the electrostatic potential, the results from these population analysis schemes were significantly different for $R=CH_3$. This can be contrasted with the favorable results in Figure 7a, where the different electrostatic potential fitting methods gave substantially similar results. Given the unreliability of our charges when $R=CH_3$, we have chosen to focus our discussion on the hydroxylated compound. Nonetheless, the main results for $R=CH_3$ are broadly similar: carbon 6 and carbon 8 are ordinarily most reactive, with carbon 4 and carbon 2 being less reactive. In many cases, carbon 4 is more reactive than carbon 2, and steric effects need to be invoked to describe the lack of reactivity at this site.⁴³ An exception occurs for the natural population analysis scheme, which is reported in Tables 4a and 4b. In that case, carbon 8, carbon 6, and carbon 2 are all reasonably reactive, though carbon 8 and carbon 6 are likely to be the most reactive sites for chlorination and nitration.⁴⁹

Table 4 Reactivity Transition Tables for 10-methyl-10,9-bozaroarphenanthrene using natural population analysis with the charges and Fukui functions on hydrogens summed into the adjacent heavy atoms. (a) First-Choice: the minimum values of $\Xi_{\Delta N \leq 0}^{\kappa}$ denotes where the molecule is most reactive. (b) Second-Choice: the second smallest values for $\Xi_{\Delta N \leq 0}^{\kappa}$, denoting the second most reactive carbon.

(a)	κ										
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0
-1.0	-0.067	-0.066	-0.074	-0.094	-0.115	-0.135	-0.159	-0.183	-0.207	-0.232	-0.256
-0.9	-0.067	-0.066	-0.069	-0.087	-0.105	-0.123	-0.143	-0.165	-0.187	-0.209	-0.231
-0.8	-0.067	-0.065	-0.065	-0.080	-0.096	-0.112	-0.128	-0.147	-0.166	-0.186	-0.205
-0.7	-0.067	-0.064	-0.062	-0.073	-0.086	-0.100	-0.113	-0.129	-0.146	-0.163	-0.179
-0.6	-0.067	-0.064	-0.061	-0.066	-0.077	-0.088	-0.099	-0.111	-0.125	-0.140	-0.154
-0.5	-0.067	-0.063	-0.059	-0.059	-0.067	-0.076	-0.084	-0.093	-0.105	-0.117	-0.128
-0.4	-0.067	-0.063	-0.058	-0.054	-0.058	-0.064	-0.070	-0.076	-0.084	-0.093	-0.102
-0.3	-0.067	-0.062	-0.057	-0.052	-0.048	-0.052	-0.056	-0.060	-0.064	-0.070	-0.077
-0.2	-0.067	-0.061	-0.056	-0.050	-0.045	-0.040	-0.042	-0.043	-0.045	-0.047	-0.051
-0.1	-0.067	-0.061	-0.055	-0.049	-0.043	-0.036	-0.030	-0.027	-0.026	-0.025	-0.026
0.0	-0.067	-0.060	-0.054	-0.047	-0.040	-0.033	-0.027	-0.020	-0.013	-0.007	
	Carbon 2		Carbon 4			Carbon 6			Carbon 8		

(b)	κ										
ΔN	1.0	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1.0
-1.0	-0.033	-0.054	-0.065	-0.086	-0.110	-0.134	-0.156	-0.176	-0.196	-0.217	-0.237
-0.9	-0.033	-0.051	-0.064	-0.078	-0.100	-0.122	-0.141	-0.159	-0.177	-0.195	-0.213
-0.8	-0.033	-0.049	-0.063	-0.070	-0.089	-0.109	-0.127	-0.143	-0.158	-0.174	-0.190
-0.7	-0.033	-0.047	-0.060	-0.063	-0.079	-0.096	-0.113	-0.126	-0.140	-0.153	-0.166
-0.6	-0.033	-0.044	-0.055	-0.057	-0.069	-0.083	-0.097	-0.110	-0.121	-0.131	-0.142
-0.5	-0.033	-0.042	-0.050	-0.056	-0.059	-0.070	-0.082	-0.093	-0.102	-0.110	-0.119
-0.4	-0.033	-0.039	-0.046	-0.052	-0.050	-0.057	-0.066	-0.075	-0.083	-0.089	-0.095
-0.3	-0.033	-0.037	-0.041	-0.045	-0.047	-0.045	-0.051	-0.058	-0.064	-0.067	-0.071
-0.2	-0.033	-0.035	-0.036	-0.038	-0.039	-0.039	-0.036	-0.040	-0.043	-0.046	-0.047
-0.1	-0.033	-0.032	-0.031	-0.030	-0.029	-0.029	-0.028	-0.024	-0.023	-0.024	-0.024
0.0	-0.033	-0.030	-0.027	-0.023	-0.020	-0.017	-0.013	-0.010	-0.007	-0.003	
	Carbon 2		Carbon 4			Carbon 6			Carbon 8		

IV.V. Summary

To explore the validity of our methods, we studied three molecules where frontier molecular orbital theory fails to predict the correct reactivity. This analysis underscores, among other things, the importance of chemical reasoning when applying reactivity indicators. For example, while our results for isoquinoline were in stark disagreement with experiment, isoquinoline is protonated under the experimental conditions. Adding a proton to isoquinoline brought our predictions into agreement with the experimental results. Similarly, when exploring electrophilic aromatic substitution on 10-*R*-10,9-borazarophenanthrenes, it was important to remember that “ipso” attack—attack on carbon atoms that do not have any adjacent hydrogen atoms—is a mechanistic dead end since there is no leaving group that the electrophile can “substitute” for.

The complex reactivity of 10-*R*-10,9-borazarophenanthrenes provided an ideal situation for testing the general-purpose reactivity indicator derived in the first paper of this series. Experimentally, chlorination, bromination, and nitration of 10-*R*-10,9-bozazarophenanthrenes occurs primarily on carbons 6 and 8 (with carbon 8 slightly favored). At higher temperatures, chlorination of 10-hydroxy-10,9-bozazarophenanthrene also gives some of the trichlorinated product, with the additional reaction occurring at carbon 2. Friedel-Crafts acetylation of 10-*R*-10,9-bozazarophenanthrenes occurs primarily on carbon 6. Using the condensed version of our general-purpose reactivity indicator, $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$, we were able to explain these results: carbon 6 and carbon 8 are the most reactive sites, with carbon 8 favored for hard electrophiles; carbon 2 is significantly less reactive than carbon 6 and carbon 8, but is predicted more reactive than any other site except the sterically hindered carbon 4; because the acetyl carbocation is resonance-stabilized and a very good electron acceptor, the most reactive site should be carbon 6. To obtain these results, we used reactivity transition tables, which list the value of $\Xi_{\Delta N \leq 0, \alpha}^{\kappa}$ at the most reactive (Tables 1a, 2a, 3a, and 4a) and second most-reactive (Tables 1b, 2b, 3b, and 4b) sites. The entries in the table are then color-coded according to identity of the most reactive site. For molecules with multiple reactive sites, reactivity transition tables provide a useful way of predicting how the regioselectivity of the molecule depends on the characteristics of the attacking reagent.

We are encouraged by these results. The new indicator, $\Xi_{\Delta N}^{\kappa}$, coincides with conventional conceptual DFT reactivity indicators whenever they work, but also provides theoretical insight and computational results that can be used to clarify situations where conventional reactivity indicators fail. Our future efforts

in this area will focus on extending these results to other types of chemical reactions and refining the present indicator to account for the polarization of reactive sites.

Acknowledgements:

Helpful discussions with Dr. David C. Thompson and Dr. Cherif Matta are acknowledged. NSERC, the Canada Research Chairs, and PREA provided funding for the Canadian authors. This research was performed when the second author visited McMaster University in the winter of 2005 and she wishes to thank the chemistry department at McMaster University for their hospitality.

IV.VI. References

- (1) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *Journal of Chemical Theory and Computation* **2007**, *3*, 358-374.
- (2) Klopman, G. *J.Am.Chem.Soc.* **1968**, *90*, 223-234.
- (3) Parr, R. G.; Pearson, R. G. *J.Am.Chem.Soc.* **1983**, *105*, 7512-7516.
- (4) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J.Chem.Phys.* **1978**, *68*, 3801-3807.
- (5) Yang, W.; Mortier, W. J. *J.Am.Chem.Soc.* **1986**, *108*, 5708-11.
- (6) Ayers, P. W.; Morrison, R. C.; Roy, R. K. *J.Chem.Phys.* **2002**, *116*, 8731-8744.
- (7) Langenaeker, W.; Demel, K.; Geerlings, P. *THEOCHEM* **1991**, *80*, 329-342.
- (8) Meneses, L.; Tiznado, W.; Contreras, R.; Fuentealba, P. *Chem.Phys.Lett.* **2004**, *383*, 181-187.
- (9) Martinez, A.; Vazquez, M. V.; Carreon-Macedo, J. L.; Sansores, L. E.; Salcedo, R. *Tetrahedron* **2003**, *59*, 6415-6422.
- (10) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (11) Yang, W.; Parr, R. G.; Pucci, R. *J.Chem.Phys.* **1984**, *81*, 2862-2863.
- (12) Parr, R. G.; Yang, W. *J.Am.Chem.Soc.* **1984**, *106*, 4049-4050.
- (13) Melin, J.; Aparicio, F.; Subramanian, V.; Galvan, M.; Chattaraj, P. K. *J.Phys.Chem.A* **2004**, *108*, 2487-2491.
- (14) Dewar, M. J. S. *THEOCHEM* **1989**, *59*, 301-23.
- (15) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Peersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.;

- Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvetti, O.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham M.A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03, Revision C.02*; Gaussian Inc.: Wallingford, CT, 2004.
- (16) Becke, A. D. *Phys.Rev.A* **1988**, *38*, 3098-3100.
 - (17) Becke, A. D. *J.Chem.Phys.* **1993**, *98*, 5648-5652.
 - (18) Lee, C.; Yang, W.; Parr, R. G. *Phys.Rev.B* **1988**, *37*, 785-789.
 - (19) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J.Chem.Phys.* **1980**, *72*, 650-654.
 - (20) Mulliken, R. S. *J.Chem.Phys.* **1955**, *23*, 1833.
 - (21) Mulliken, R. S. *J.Chem.Phys.* **1955**, *23*, 1841.
 - (22) Mulliken, R. S. *J.Chem.Phys.* **1955**, *23*, 2343.
 - (23) Mulliken, R. S. *J.Chem.Phys.* **1955**, *23*, 2338.
 - (24) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem.Rev.* **1988**, *88*, 899-926.
 - (25) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J.Chem.Phys.* **1985**, *83*, 735-746.
 - (26) Reed, A. E.; Weinhold, F. *J.Chem.Phys.* **1983**, *78*, 4066-4073.
 - (27) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J.Comp.Chem.* **1990**, *11*, 431-439.
 - (28) Singh, U. C.; Kollman, P. A. *J.Comp.Chem.* **1984**, *5*, 129-145.
 - (29) Breneman, C. M.; Wiberg, K. B. *J.Comp.Chem.* **1990**, *11*, 361-373.
 - (30) Dewar, M. J. S.; Maitlis, P. M. *Journal of the Chemical Society* **1957**, 2521-2528.
 - (31) Flurchick, K.; Bartolotti, L. *Journal of Molecular Graphics* **1995**, *13*, 10-13.
 - (32) Bartolotti, L. J.; Ayers, P. W. *J.Phys.Chem.A* **2005**, *109*, 1146-1151.
 - (33) We have confirmed that the qualitative features of these plots do not alter when the isodensity surface is moved closer to the molecule. To the eye, plots generated using the .001 and .002 isodensity surfaces look the same as the ones in Figure 1.
 - (34) Berkowitz, M. *J.Am.Chem.Soc.* **1987**, *109*, 4823-4825.
 - (35) Chattaraj, P. K. *J.Phys.Chem.A* **2001**, *105*, 511-513.
 - (36) Hocquet, A.; Toro-Labbe, A.; Chermette, H. *THEOCHEM* **2004**, *686*, 213-218.
 - (37) Because of the slow asymptotic decay of the electrostatic potential and the Fukui potential, plots like Figure 1 do not provide useful information.

- (38) Miller, B. *Advanced organic chemistry: Reactions and mechanisms*; Prentice-Hall: Upper Saddle River, N. J., 1998.
- (39) The reader might notice that atomic charges the nitrogen atom and carbon 4 have similar negative charges in the electrostatic charge fitting schemes (Figure 2a). Note, however, that once one considers the hydrogen-summed charges (Figure 2b), protonation on the nitrogen is clearly preferred. Protonation of carbon 4 would be associated with electrophilic aromatic substitution of hydrogen for hydrogen. So even if carbon 4 were protonated, the concentration of the tetrahedral reactive intermediate would be very, very low. It follows that the case where the molecule is protonated at carbon 4 is unlikely to contribute to the overall chemical reactivity in any significant way. Protonation of carbon 4 could be studied by deuterium exchange experiments, but we are unaware of experimental studies of this type.
- (40) Dewar, M. J. S. *Progress in Boron Chemistry* **1964**, *1*, 235-263.
- (41) Dewar, M. J. S.; Kubba, V. P. *Journal of Organic Chemistry* **1960**, *25*, 1722-1724.
- (42) Dewar, M. J. S.; Kubra, V. P. *Tetrahedron* **1959**, *7*, 213-222.
- (43) Dewar, M. J. S.; Warford, E. W. T. *Journal of the Chemical Society* **1956**, 3570-3572.
- (44) Dewar, M. J. S.; Kubba, V. P. *J. Am. Chem. Soc.* **1961**, *83*, 1757-1760.
- (45) To the extent that electrophilic attack on the nitrogen atom does occur, it will be associated with molecular decomposition (just like protonation on this site). Though Dewar's experiments often produced the featured products in high yield, some of the impurities might be due to the decomposition of the starting reagent induced by electrophilic attack on the nitrogen atom. It is impossible to quantify the importance of this process, so we will focus on the reaction pathways for which there is incontrovertible experimental evidence.
- (46) One might argue that since our model is based on changes in the interaction potential, charges from electrostatic potential fitting are the more natural choice. While we share this view, it is also true that if one is far enough from the molecule, the charges based on natural population analysis (or any other density-matrix population analysis method) reproduce the electrostatic potential.
- (47) This is especially true in Tables 2 and 3, but not as true for Table 1.
- (48) We attempted to add a sodium cation near the nitrogen atom to mimic the complexation. If there is a potential energy well in this area, it seems to be very shallow: in our calculations the cation always either dissociated or migrated to the top of one of the aromatic rings.
- (49) Again, naive application of the reactivity indicator sometimes indicates that ipso attack is favored. The carbon of the methyl group is also

predicted to be rather reactive, but electrophilic substitution at this carbon can not occur because of the absence of the electrofuge. In the text and in the tables, we focus our treatment on the ring carbons that are subject to substitution, carbons 1-8.

Chapter V

Predicting the Reactivity of Ambidentate Nucleophiles and Electrophiles Using a Single, General-Purpose, Reactivity Indicator*

*The content of this chapter has been published: **J. S. M. Anderson**, P. W. Ayers “Predicting the reactivity of ambidentate nucleophiles and electrophiles using a single, general purpose, reactivity indicator”; *Phys. Chem. Chem. Phys.* **2007**, *9*, 2371-2378.

V.I. Statement of the Problem

Chapter III derived a new general-purpose reactivity indicator and chapter IV applied the indicator to some difficult cases. In this chapter a larger number of molecules are presented. These molecules are particularly well-suited for illustrating the validity of the general-purpose reactivity indicator because they have multiple reactive sites, with different reactive sites preferring to bond to different substrates. Such molecules are said to be ambidentate. The molecules treated here have one site that prefers to react with hard reagents and another that prefers to react with soft reagents. Normal reactivity indicators will predict that one of the active sites is always preferred, but the general purpose reactivity indicator was specifically designed to discern between these two types of reactivity. It succeeds.

V.II. Background

Electronic structure theory is the art of transforming molecular structure information into information about chemistry and, especially, chemical reactivity. The last few decades have seen immense improvements in this quest, especially when one considers the improvements in quantitative accuracy for modeling chemical reactions.^{1,2} In contrast, conceptual advances have lagged somewhat, mostly because the most accurate *ab initio* techniques for achieving quantitative accuracy are not amenable to direct interpretation. An exception is density-functional theory (DFT). DFT is exact in principle but, as stressed by Parr and his proselytes, it provides a simple and direct qualitative picture of what drives chemical reactions.³⁻⁸ Importantly, the picture does not change as the quality of the underlying density-functional method improves from inaccurate local functionals, to more accurate hybrid functionals, to extremely accurate approaches inspired by “wavefunction based” *ab initio* methods.⁹⁻¹⁴ Unlike simple molecular orbital models, density-functional reactivity theory includes electron correlation and orbital relaxation effects, both of which can be significant.¹⁵ Density-functional reactivity theory has proved very successful for describing known principles of chemical reactivity (e.g., electronegativity equalization¹⁶ and the Hard/Soft Acid/Base Principle¹⁷⁻²²) and also for the development of new principles (e.g., the maximum hardness principle²³⁻²⁵).

Recently, the authors proposed new reactivity indicators, one appropriate for predicting reactivity of nucleophiles and one appropriate for predicting the reactivity of electrophiles, that were designed to not only model conventional “charge controlled” reactions (where electrostatic interactions are dominant) and

“frontier controlled” reactions (where electron-transfer is dominant),^{20;26} but also intermediate cases.²⁷ Our first papers considered only a few simple, but problematic, molecules.²⁸ Based on the encouraging results, it occurred to us that we should test the indicators more thoroughly. To do this, we decided to explore ambidentate molecules: molecules with multiple reactive sites where it is known that reactivity “switches” from one site to another based on the properties of the attacking reagent. Very commonly one reactive site is associated with highly charged (or “hard”) reagents while the other site is associated with highly polarizable (or “soft”) reagents. (In fact, one of the key innovations in ref. ²⁷ was the ability to rigorously explain the local Hard/Soft Acid/Base phenomenon that Pearson proposed long ago.^{17;21;22,})

The “general-purpose reactivity indicator,” Ξ , that we proposed is based on what is termed the “perturbative perspective on chemical reactivity.”^{7;25;29} The basic idea is that most molecules are only reactive in one or two sites. For example, nitrogen-containing bases tend to react on the lone pair of the nitrogen regardless of the structure of the remainder of the molecule and regardless of the detailed nature of the acidic reagent that is being considered. This tells us that one can obtain a qualitative description of chemical reactivity by formulating a simplified model for the reagent and then studying how this model perturbs the molecule of interest. If the perturbation lowers the energy of the molecule (or at least does not increase the energy very much) then the reaction is favorable. Otherwise the reaction is unfavorable.

The reagent perturbs the molecule under scrutiny in two ways: it changes the number of electrons in the molecule (ΔN) and it changes the external potential felt by the electrons in the molecule ($\Delta v(\mathbf{r})$). The change in energy can then be computed using density-functional perturbation theory,

$$\begin{aligned} \Delta U = & \int \frac{\delta V_m[v]}{\delta v(\mathbf{r})} \Delta v(\mathbf{r}) d\mathbf{r} + \left(\frac{\partial E}{\partial N} \right)_{v(\mathbf{r})}^- \Delta N \\ & + \int \left(\frac{\delta E[v; N]}{\delta v(\mathbf{r})} \right)_N \Delta v(\mathbf{r}) d\mathbf{r} \\ & + \left[\int \left(\frac{\partial}{\partial N} \left(\frac{\delta E[v; N]}{\delta v(\mathbf{r})} \right) \right)_{v(\mathbf{r})}^- \Delta v(\mathbf{r}) d\mathbf{r} \right] \Delta N + \dots \end{aligned} \quad (1)$$

Here U represents the total energy of the molecule within the Born-Oppenheimer approximation, i.e. U is the sum of the total electronic energy and nuclear-nuclear repulsion energy, $U = E + V_m$. It is, of course, impossible to consider the entire expansion and so the expansion needs to be truncated or, alternatively,

treated using the functional-analytic generalization of Taylor's theorem with remainder.^{7,30} The general-purpose reactivity indicator is one of the simplest conceivable reactivity indicators based on Eq. (1): it is derived by truncating the expansion after the first-order terms in the external potential. (Higher order contributions from ΔN are identically zero in exact *ab initio* quantum mechanics.^{9,31-33,34}) As such, only the terms that are explicitly shown in Eq. (1) are considered. The reactive site of the attacking reagent is then modeled using a condensed reactive-site only model,^{20;21;27} where the reactive site is modeled as a point charge with a given condensed Fukui function.³⁵⁻³⁹ After some rather lengthy analysis,²⁷ one arrives at indicators for the reactivity of each site in the molecule with respect to electrophilic attack on a nucleophile ($\Delta N \leq 0$)

$$\Xi_{\Delta N \leq 0, \alpha}^{\kappa} = (\kappa + 1)q_{\text{nucleophile}, \alpha}^{(0)} - \Delta N(\kappa - 1)f_{\text{nucleophile}, \alpha}^{-} \quad (2)$$

and nucleophilic attack on an electrophile ($\Delta N \geq 0$):

$$\Xi_{\Delta N \geq 0, \alpha}^{\kappa} = -(\kappa + 1)q_{\text{electrophile}, \alpha}^{(0)} + \Delta N(\kappa - 1)f_{\text{electrophile}, \alpha}^{+} \quad (3)$$

Here q_{α}^{0} denotes the atomic charges on the nucleophile/electrophile that is under study and

$$f_{\text{nucleophile}, \alpha}^{-} = q_{\text{nucleophile}, \alpha}^{+} - q_{\text{nucleophile}, \alpha}^{0} \quad (4)$$

and

$$f_{\text{electrophile}, \alpha}^{+} = q_{\text{electrophile}, \alpha}^{0} - q_{\text{electrophile}, \alpha}^{-} \quad (5)$$

denote the condensed Fukui functions for electron removal from the nucleophile and electron addition to the electrophile.^{38;39} $q_{\text{nucleophile}, \alpha}^{+}$ denotes the atomic charges on the nucleophile with one electron removed (which is typically a cation, and typically choosing vertical ionization). $q_{\text{electrophile}, \alpha}^{-}$ denotes the atomic charges on the electrophile with one electron added (which is typically an anion, and typically choosing vertical electron attachment).⁴⁰ The general-purpose reactivity indicator, $\Xi_{\Delta N}^{\kappa}$, is a two-parameter model, that depends on the amount of electron transfer ΔN and the parameter κ . The parameter κ encapsulates the key information about the electronic structure of the attacking reagent. When one goes through the mathematics, one discovers that the qualitative preferences of an electrophilic reagent for a particular site in a nucleophile ($\Delta N \leq 0$) are modulated by a quantity that is approximately proportional to the sum of the charge on the reactive site electrophile and the amount of electron transfer to the electrophile's reactive site:

$$\kappa \sim q_{\text{electrophile}}^{(0)} + \Delta N f_{\text{electrophile}}^{(+)} \quad (6)$$

Similarly, the reactivity of a nucleophilic reagent attacking an electrophile ($\Delta N \geq 0$) is modulated by

$$\kappa \sim -q_{\text{nucleophile}}^{(0)} - \Delta N f_{\text{nucleophile}}^{(-)} \quad (7)$$

In most (but not all) cases, κ is between one and minus one. When $\kappa \approx 1$, the charge on the reactive site of the attacking reagent is much larger than the extent of electron transfer to/from the attacking reagent, and the reaction is electrostatically controlled, hard reaction. When $\kappa \approx -1$, the charge on the reactive site of the attacking reagent is much smaller than the extent of electron transfer, and the reaction is electron-transfer (or Fukui-function) controlled, soft reaction. Many reactions are somewhere between these two extremes.^{27;28}

The $\Xi_{\Delta N}^{\kappa}$ reactivity indicator is preferable to other indicators whenever (a) a reaction is neither electron-transfer nor electrostatically controlled and (b) when different types of reagents react with the molecule at different reactive sites. The latter phenomenon is associated with ambidentate reactivity, wherein a molecule can react at either of two reactive sites, and which site is most reactive depends on the nature of the attacking reagent. Our hypothesis is that the $\Xi_{\Delta N}^{\kappa}$ will elucidate the reactivity of ambidentate molecules. The goal of this paper is to demonstrate that this hypothesis is correct.

V.III. Computational Methodology

All of the calculations presented were performed using Gaussian 03 using the B3LYP exchange-correlation energy functional^{41;42} using the 6-31++G* basis set.⁴³ The calculations on SeCN⁻ were also compared with the LanL2DZ basis set, which treats the core electrons with an effective core potential. The pseudopotential calculation gave similar results to the 6-31++G* basis set, indicating that relativistic considerations were not determinative for the reactivity of this molecule.

For each molecule, the geometry was first optimized for the molecular substrate and the geometry was then held fixed when the charges of the $(N+1)$ -electron system (for electrophilic molecules) and the $(N-1)$ -electron system (for nucleophilic molecules) was computed. The calculations on the nucleophiles (including the geometry optimization) included a solvent model, the PCM model.⁴⁴ (The solvent model is especially important for stabilizing the anions.) The electrophiles considered are neutral molecules so solvent is inessential. As such, no solvent model was used for the electrophiles. The condensed general-purpose reactivity indicator is derived with electrostatics in mind and as such a population scheme that uses electrostatic fitting is likely most appropriate. In our study we presented a detailed analysis including the Mulliken (MPA), Mertz-

Singh-Kollman (MSK), natural (NPA), and cHelpG (CHG) schemes. We did the same studies here but, for brevity, we will only present the CHG results (the other schemes give qualitatively similar results). CHG has the reputation for being a reliable population scheme,⁴⁵ and it gave good results in our previous calculations using this reactivity indicator.²⁸ Also, when applicable, the charges on the hydrogen atoms were summed into the heavy atoms they are bonded to, this is often done since distinguishing how much charge is on the hydrogen atom and how much is on the adjacent heavy atom is often not clear due to the short XH bond length. For qualitative arguments including the hydrogens bounded to the reactive atom in the reactive site is permissible since one can interpret the value of $\Xi_{\Delta N, \alpha}^{\kappa}$ assigned to heavy atoms as representing the reactivity of XH_n functional groups (X = heavy atom, n=0,1,2,...).

We will present our data using the “reactivity transition tables” introduced in our previous paper. Reactivity transition tables are a concise and visually appealing way to represent the most reactive site in a molecule organized by (1) the extent of electron transfer *to* the molecule being studied (ΔN) and (2) the reactivity preferences of the attacking reagent (i.e. hard or soft, with $\kappa = 1$ representing charge control (hard reagents) and $\kappa = -1$ indicating electron-transfer control (soft reagents)). For each reasonable choice of κ and ΔN , the value of $\Xi_{\Delta N, \alpha}^{\kappa}$ was computed for each non-hydrogen atom (or XH_n functional group) and the most reactive atom is identified. The value of $\Xi_{\Delta N, \alpha}^{\kappa}$ at the most reactive atom is recorded in the appropriate cell of the reactivity transition table and identity of the atom is indicated by color-coding the cell of the table. In this work, we have introduced the convention that, if known, the product that experiments predict for electrostatically (i.e. charge) controlled reactions will be colored “dark red” and the product that experiments predict for electron-transfer (i.e. frontier) controlled reactions will be colored “sky blue.”⁴⁶ If there is a third reactive site associated with “intermediate” conditions this would be assigned a “violet” color. Reactive sites that are theoretically predicted but not experimentally observed will be assigned different colors. A “legend” for the colors is presented at the bottom of the table. One advantage of this “standardized color scheme” is that it allows one to, at a glance, identify to what extent a reactivity transition table agrees with the experimentally predicted results.

We remind the reader that since $\Xi_{\Delta N, \alpha}^{\kappa}$ is based on a model for the interaction energy between the molecule and an approaching reagent, the most reactive atoms/functional groups are those with the smallest (most negative or least positive) values of the $\Xi_{\Delta N, \alpha}^{\kappa}$.

V.IV. Results

A. Ambidentate Nucleophiles

The nucleophiles studied illustrate the efficacy of the condensed general-purpose reactivity indicator. In all four cases presented, the cases of SCN^- , SeCN^- , NO_2^- , and SO_3^{2-} , reactivity transition tables 1-4, the condensed indicator correctly indicated the hard (electrostatic controlled or charge controlled) and soft (charge transfer or Fukui controlled) sites of the molecule.

1. SCN^-

The condensed general-purpose reactivity indicator correctly predicts the sulfur atom to be the soft site and the nitrogen atom to be the hard site. (See Table 1.) Moreover, if one then looks at the second choice (not shown) (i.e. the atom with the second smallest value of $\Xi_{\Delta N, \alpha}^*$), the carbon atom is never predicted as the reactive site. (This agrees with the experimental result (the carbon atom is unreactive) but disagrees with some of the prior results in the chemical reactivity literature, which would have predicted that carbon (and not nitrogen) would have been the reactive site for a hard reagent, because the carbon atom has the smallest value of the Fukui function.⁴⁷ More information about the (rather limited) utility of the minimum Fukui function rule can be found in the work of Melin and her collaborators.^{27;48})

Table 1 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \leq 0}^{\kappa}$, for electrophilic attack on SCN⁻ using cHelpG charges.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
-1	-1.59843	-1.50432	-1.41878	-1.43619	-1.45359	-1.471	-1.48841	-1.50581	-1.52322	-1.54063	-1.55803
-0.9	-1.59843	-1.49775	-1.39706	-1.38945	-1.39127	-1.3931	-1.39493	-1.39675	-1.39858	-1.4004	-1.40223
-0.8	-1.59843	-1.49117	-1.38392	-1.34271	-1.32895	-1.3152	-1.30144	-1.28769	-1.27394	-1.26018	-1.24643
-0.7	-1.59843	-1.4846	-1.37077	-1.29596	-1.26663	-1.2373	-1.20796	-1.17863	-1.14929	-1.11996	-1.09062
-0.6	-1.59843	-1.47803	-1.35762	-1.24922	-1.20431	-1.15939	-1.11448	-1.06956	-1.02465	-0.97974	-0.93482
-0.5	-1.59843	-1.47145	-1.34448	-1.2175	-1.14199	-1.08149	-1.021	-0.9605	-0.90001	-0.83951	-0.77902
-0.4	-1.59843	-1.46488	-1.33133	-1.19778	-1.07967	-1.00359	-0.92752	-0.85144	-0.77536	-0.69929	-0.62321
-0.3	-1.59843	-1.45831	-1.31818	-1.17806	-1.03794	-0.92569	-0.83403	-0.74238	-0.65072	-0.55907	-0.46741
-0.2	-1.59843	-1.45174	-1.30504	-1.15834	-1.01164	-0.86495	-0.74055	-0.63332	-0.52608	-0.41884	-0.31161
-0.1	-1.59843	-1.44516	-1.29189	-1.13862	-0.98535	-0.83208	-0.67881	-0.52554	-0.40144	-0.27862	-0.1558
0	-1.59843	-1.43859	-1.27875	-1.1189	-0.95906	-0.79922	-0.63937	-0.47953	-0.31969	-0.15984	
Nitrogen						Sulfur					

2. SeCN⁻

Similar to SCN⁻ the condensed general-purpose reactivity indicator correctly predicts that the selenium atom is site where soft acids react and the nitrogen atom is the site where hard acids react. (See Table 2.) Just as in the SCN⁻ case, carbon is always predicted to be the least reactive atom, in accord with experimental facts. Notice how similar the reactivity transition table for this molecule is to the reactivity transition table for SCN⁻.

Table 2 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \leq 0}^{\kappa}$, for electrophilic attack on SeCN⁻ using cHelpG charges.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
-1	-1.19038	-1.23139	-1.22358	-1.13382	-1.04406	-0.9543	-0.86454	-0.77478	-0.68502	-0.59526	-0.5055
-0.9	-1.19038	-1.21538	-1.21347	-1.11866	-1.02384	-0.92902	-0.83421	-0.7394	-0.64458	-0.54977	-0.45495
-0.8	-1.19038	-1.19938	-1.20336	-1.10349	-1.00362	-0.90375	-0.80388	-0.70401	-0.60414	-0.50427	-0.4044
-0.7	-1.19038	-1.18337	-1.17637	-1.08832	-0.9834	-0.87848	-0.77355	-0.66863	-0.5637	-0.45878	-0.35385
-0.6	-1.19038	-1.16737	-1.14436	-1.07316	-0.96318	-0.8532	-0.74322	-0.63324	-0.52326	-0.41328	-0.3033
-0.5	-1.19038	-1.15136	-1.11235	-1.058	-0.94296	-0.82793	-0.71289	-0.59786	-0.48282	-0.36779	-0.25275
-0.4	-1.19038	-1.13536	-1.08034	-1.02532	-0.92274	-0.80265	-0.68256	-0.56247	-0.44238	-0.32229	-0.2022
-0.3	-1.19038	-1.11935	-1.04833	-0.97731	-0.90252	-0.77738	-0.65223	-0.52709	-0.40194	-0.2768	-0.15165
-0.2	-1.19038	-1.10335	-1.01632	-0.92929	-0.84227	-0.7521	-0.6219	-0.4917	-0.3615	-0.2313	-0.1011
-0.1	-1.19038	-1.08735	-0.98431	-0.88128	-0.77825	-0.67521	-0.57218	-0.45632	-0.32106	-0.1858	-0.05055
0	-1.19038	-1.07134	-0.9523	-0.83326	-0.71423	-0.59519	-0.47615	-0.35711	-0.23808	-0.11904	
Nitrogen						Selenium					

3. NO_2^-

The condensed general-purpose reactivity indicator correctly identifies nitrogen to be the soft site and oxygen to be the hard site of the nitrite anion. (See Table 3.) Based on Table 3, it seems unlikely that the nitrogen atom will be reactive except for extremely soft acids. Certainly the overall trend is correct, since protonation (the proton is the prototypical hard acid) of NO_2^- occurs on an oxygen atom.

Table 3 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \leq 0}^{\kappa}$, for electrophilic attack on NO_2^- using cHeloG charges.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
-1	-0.94812	-0.91822	-0.88832	-0.85842	-0.82852	-0.79862	-0.76872	-0.73883	-0.70893	-0.67903	-0.70177
-0.9	-0.94812	-0.91173	-0.87534	-0.83895	-0.80256	-0.76617	-0.72978	-0.69339	-0.657	-0.62061	-0.63159
-0.8	-0.94812	-0.90523	-0.86235	-0.81947	-0.77659	-0.73371	-0.69083	-0.64795	-0.60507	-0.56219	-0.56141
-0.7	-0.94812	-0.89874	-0.84937	-0.8	-0.75063	-0.70125	-0.65188	-0.60251	-0.55314	-0.50376	-0.49124
-0.6	-0.94812	-0.89225	-0.83639	-0.78052	-0.72466	-0.6688	-0.61293	-0.55707	-0.50121	-0.44534	-0.42106
-0.5	-0.94812	-0.88576	-0.82341	-0.76105	-0.6987	-0.63634	-0.57399	-0.51163	-0.44928	-0.38692	-0.35088
-0.4	-0.94812	-0.87927	-0.81042	-0.74158	-0.67273	-0.60388	-0.53504	-0.46619	-0.39734	-0.3285	-0.28071
-0.3	-0.94812	-0.87278	-0.79744	-0.7221	-0.64677	-0.57143	-0.49609	-0.42075	-0.34541	-0.27008	-0.21053
-0.2	-0.94812	-0.86629	-0.78446	-0.70263	-0.6208	-0.53897	-0.45714	-0.37531	-0.29348	-0.21165	-0.14035
-0.1	-0.94812	-0.8598	-0.77148	-0.68316	-0.59483	-0.50651	-0.41819	-0.32987	-0.24155	-0.15323	-0.07018
0	-0.94812	-0.8533	-0.75849	-0.66368	-0.56887	-0.47406	-0.37925	-0.28443	-0.18962	-0.09481	
Oxygen											Nitrogen

4. SO_3^{2-}

The condensed general-purpose reactivity indicator correctly identifies sulfur to be the soft site and oxygen to be the hard site of the sulfite anion, see Table 4. Notice that the similarity in the reactivity transition tables of sulfite and nitrite reflect their chemical similarity.

Table 4 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \leq 0}^{\kappa}$, for electrophilic attack on SO_3^{2-} using cHelpG charges.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
-1	-1.53177	-1.4162	-1.30064	-1.18508	-1.06952	-0.95396	-0.8384	-0.72284	-0.60727	-0.72577	-0.87237
-0.9	-1.53177	-1.41244	-1.29312	-1.1738	-1.05447	-0.93515	-0.81583	-0.69651	-0.57718	-0.64726	-0.78513
-0.8	-1.53177	-1.40868	-1.2856	-1.16251	-1.03943	-0.91634	-0.79326	-0.67017	-0.54709	-0.56875	-0.6979
-0.7	-1.53177	-1.40492	-1.27807	-1.15123	-1.02438	-0.89754	-0.77069	-0.64384	-0.517	-0.49023	-0.61066
-0.6	-1.53177	-1.40116	-1.27055	-1.13994	-1.00934	-0.87873	-0.74812	-0.61751	-0.48691	-0.41172	-0.52342
-0.5	-1.53177	-1.3974	-1.26303	-1.12866	-0.99429	-0.85992	-0.72555	-0.59118	-0.45681	-0.33321	-0.43619
-0.4	-1.53177	-1.39364	-1.2555	-1.11737	-0.97924	-0.84111	-0.70298	-0.56485	-0.42672	-0.28859	-0.34895
-0.3	-1.53177	-1.38987	-1.24798	-1.10609	-0.9642	-0.82231	-0.68041	-0.53852	-0.39663	-0.25474	-0.26171
-0.2	-1.53177	-1.38611	-1.24046	-1.09481	-0.94915	-0.8035	-0.65784	-0.51219	-0.36654	-0.22088	-0.17447
-0.1	-1.53177	-1.38235	-1.23294	-1.08352	-0.93411	-0.78469	-0.63528	-0.48586	-0.33645	-0.18703	-0.08724
0	-1.53177	-1.37859	-1.22541	-1.07224	-0.91906	-0.76588	-0.61271	-0.45953	-0.30635	-0.15318	
Oxygen											
											Sulfur

B. Ambidentate Electrophiles

Relatively few ambidentate electrophiles are known and even the simplest ambidentate electrophiles are relatively large and complex molecules. Because of this, it is often difficult to discern whether the “switch” in reactivity of ambidentate electrophiles is due to switching between electrostatic and electron-transfer control or due to steric interactions with soft nucleophiles (because soft nucleophiles tend to be larger than hard nucleophiles). Even though our general-purpose reactivity indicator does not model steric effects, it seems to be reasonably effective for predicting the reactivity of ambidentate electrophiles. The electrophiles examined include dimethyl carbonate,⁴⁹ *N*-methyl-*N*-nitrosotoluene-*p*-sulfonamide (MNTS),⁵⁰ and 1-chloro-2,4,6-trinitrobenzene (CNB).⁵¹

1. Dimethyl Carbonate

The case of dimethyl carbonate, I, shows the greatest efficacy of this indicator. The results clearly show that the condensed general-purpose reactivity indicator correctly predicts the methyl groups to be the soft sites and the carbonyl carbon to be the hard site. (See Table 5.)

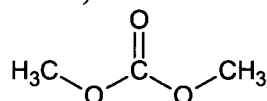
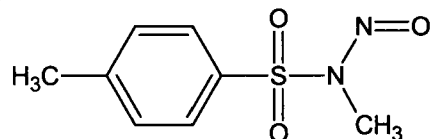


Table 5 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \geq 0}^{\kappa}$, for nucleophilic attack on dimethyl carbonate using cHelfG charges.

$\Delta N \backslash \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
0	-2.16137	-1.94523	-1.72909	-1.51296	-1.29682	-1.08068	-0.86455	-0.64841	-0.43227	-0.21614	
0.1	-2.16137	-1.93849	-1.71562	-1.49274	-1.26987	-1.047	-0.82412	-0.60125	-0.37837	-0.1555	-0.02419
0.2	-2.16137	-1.93175	-1.70214	-1.47253	-1.24292	-1.01331	-0.7837	-0.55409	-0.32448	-0.09786	-0.04837
0.3	-2.16137	-1.92502	-1.68867	-1.45232	-1.21597	-0.97962	-0.74328	-0.50693	-0.27058	-0.11963	-0.07256
0.4	-2.16137	-1.91828	-1.67519	-1.43211	-1.18902	-0.94594	-0.70285	-0.45977	-0.21668	-0.14139	-0.09674
0.5	-2.16137	-1.91154	-1.66172	-1.4119	-1.16207	-0.91225	-0.66243	-0.4126	-0.2054	-0.16316	-0.12093
0.6	-2.16137	-1.90481	-1.64825	-1.39168	-1.13512	-0.87856	-0.622	-0.36544	-0.22475	-0.18493	-0.14511
0.7	-2.16137	-1.89807	-1.63477	-1.37147	-1.10818	-0.84488	-0.58158	-0.31828	-0.24409	-0.20669	-0.1693
0.8	-2.16137	-1.89133	-1.6213	-1.35126	-1.08123	-0.81119	-0.54116	-0.29842	-0.26344	-0.22846	-0.19348
0.9	-2.16137	-1.88459	-1.60782	-1.33105	-1.05428	-0.7775	-0.50073	-0.31535	-0.28279	-0.25023	-0.21767
1	-2.16137	-1.87786	-1.59435	-1.31084	-1.02733	-0.74382	-0.46031	-0.33228	-0.30214	-0.27199	-0.24185
Carbonyl Carbon											
						Methyl Carbon					

2. *N*-methyl-*N*-nitrosotoluene-*p*-sulfonamide (MNTS)

Dimethyl carbonate is an ideal test of the $\Xi_{\Delta N}^{\chi}$ indicator because it is a relatively small molecule and steric effects are rather negligible. This is not true of *N*-methyl-*N*-nitrosotoluene-*p*-sulfonamide (MNTS), II, but the reactivity indicator still correctly assigns the sulfur atom to be the hard site. The experimental evidence⁵⁰ indicates most of the reactivity is seen at the sulfur atom. This observation is consistent with our results since the reactivity transition table is dominated by the sulfur. Notice that the general-purpose reactivity indicator correctly predicts that the sulfur atom can react as a soft-base (e.g., thiocyanate, II.A.1) or as a hard acid (e.g., MNTS). In MNTS, the sulfur atom has a formal charge of +4; atoms in high oxidation states are hard acids.^{20;21;52}



The indicator also successfully predicts that the nitroso group, NO, is the soft site. However, the indicator fails to discern that it is the nitrogen atom, and not the oxygen atom, of the nitroso group that is most reactive. (Instead, it predicts that both atoms are somewhat reactive towards soft reagents, with the oxygen atom being the most reactive atom and nitrogen either the second (NPA population analysis) or third (CHG) most reactive atom.⁵³) This could indicate that a nucleophile attacks the oxygen atom directly and then migrates to the nitrogen atom or indicate that the oxygen attack is a “mechanistic dead end.”⁵⁴) This could also be due to the difficulty of partitioning the electron density between the nitrogen and oxygen atoms (since the NO bond is rather short).

Table 6 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \geq 0}^{\kappa}$, for nucleophilic attack on MNTS using cHelpG charges.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
0	-1.42896	-1.28607	-1.14317	-1.00027	-0.85738	-0.71448	-0.57159	-0.42869	-0.28579	-0.1429	
0.1	-1.42896	-1.28848	-1.14799	-1.0075	-0.86701	-0.72653	-0.58604	-0.44555	-0.30507	-0.16458	-0.03538
0.2	-1.42896	-1.29089	-1.15281	-1.01473	-0.87665	-0.73857	-0.60049	-0.46242	-0.32434	-0.18626	-0.07076
0.3	-1.42896	-1.29329	-1.15763	-1.02196	-0.88629	-0.75062	-0.61495	-0.47928	-0.34361	-0.20794	-0.10614
0.4	-1.42896	-1.2957	-1.16244	-1.02918	-0.89592	-0.76266	-0.6294	-0.49614	-0.36288	-0.22962	-0.14152
0.5	-1.42896	-1.29811	-1.16726	-1.03641	-0.90556	-0.77471	-0.64386	-0.51301	-0.38216	-0.2513	-0.1769
0.6	-1.42896	-1.30052	-1.17208	-1.04364	-0.9152	-0.78675	-0.65831	-0.52987	-0.40143	-0.27299	-0.21228
0.7	-1.42896	-1.30293	-1.1769	-1.05087	-0.92483	-0.7988	-0.67277	-0.54673	-0.4207	-0.29467	-0.24766
0.8	-1.42896	-1.30534	-1.18172	-1.05809	-0.93447	-0.81084	-0.68722	-0.5636	-0.43997	-0.31635	-0.28304
0.9	-1.42896	-1.30775	-1.18653	-1.06532	-0.9441	-0.82289	-0.70167	-0.58046	-0.45925	-0.33803	-0.31842
1	-1.42896	-1.31016	-1.19135	-1.07255	-0.95374	-0.83493	-0.71613	-0.59732	-0.47852	-0.35971	-0.3538
Sulfur						Oxygen					

3. 1-chloro-2,4,6-trinitrobenzene (CNB).

The third electrophile considered, CNB, III, was suggested by Boga and coworkers⁵¹ to be an ambidentate molecule, but they only observed reactivity at the carbon bonded to chlorine and no reactivity at the *meta* carbons. They further suggested that the *meta* carbons were likely the kinetic product of the reaction but were unstable. The general-purpose reactivity indicator when comparing only the *meta* carbons and the carbon bonded to the chlorine (with the hydrogens summed into the *meta* carbons) one observes the *meta* carbons are the most reactive site. (See Table 7.) This confirms our suspicion that the general-purpose reactivity indicator—like other reactivity indicators based on perturbation theory about the isolated reagents—provide predictions for the “kinetic products” of a reaction and not the “thermodynamic products.” Notice that in this case the experimental reactivity predictions were not easily classified as “electrostatic” or “electron-transfer” control (this is why this table does not use the red/purple/blue color scheme), yet the general-purpose reactivity indicator is still consistent with the interpretation of the experimentalists.

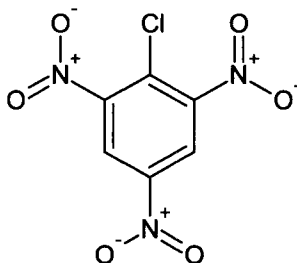


Table 7 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \geq 0}^{\kappa}$, for nucleophilic attack on CNB using cHelpG charges. Only the meta carbon and the carbon bonded to chlorine are included in this table.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
0	-0.02975	-0.02678	-0.0238	-0.02083	-0.01785	-0.01488	-0.0119	-0.00893	-0.00595	-0.00298	-0.01461
0.1	-0.02975	-0.02823	-0.02672	-0.0252	-0.02368	-0.02216	-0.02064	-0.01912	-0.01761	-0.0161	-0.01461
0.2	-0.02975	-0.02969	-0.02963	-0.02957	-0.02951	-0.02944	-0.02938	-0.02932	-0.02928	-0.02925	-0.02923
0.3	-0.02975	-0.03115	-0.03254	-0.03394	-0.03533	-0.03673	-0.03812	-0.03953	-0.04097	-0.0424	-0.04384
0.4	-0.02975	-0.0326	-0.03546	-0.03831	-0.04116	-0.04401	-0.04687	-0.04976	-0.05266	-0.05556	-0.05846
0.5	-0.02975	-0.03406	-0.03837	-0.04268	-0.04699	-0.0513	-0.05563	-0.05999	-0.06435	-0.06871	-0.07307
0.6	-0.02975	-0.03552	-0.04128	-0.04705	-0.05282	-0.05858	-0.0644	-0.07022	-0.07604	-0.08186	-0.08769
0.7	-0.02975	-0.03697	-0.0442	-0.05142	-0.05864	-0.06588	-0.07317	-0.08045	-0.08773	-0.09502	-0.1023
0.8	-0.02975	-0.03843	-0.04711	-0.05579	-0.06447	-0.07319	-0.08193	-0.09068	-0.09942	-0.10817	-0.11691
0.9	-0.02975	-0.03989	-0.05003	-0.06016	-0.0703	-0.0805	-0.0907	-0.10091	-0.11112	-0.12132	-0.13153
1	-0.02975	-0.04135	-0.05294	-0.06453	-0.07614	-0.0878	-0.09947	-0.11114	-0.12281	-0.13447	-0.14614
<i>Meta</i>											

If one included the *ortho* and *para* carbons in the analysis also, one obtains Table 8. This would indicate that the electrostatically controlled reaction site is the *para* carbon, which is the most positively charged carbon. While nucleophilic attack on the *para* site can occur, attack at this site does not lead to product (because the hydride anion is a very poor leaving group). This is an example of a “mechanistic dead end”: sometimes the most favorable “first step” for a reaction (which is what DFT-based reactivity indicators identify) cannot lead to a product because reaction barrier for the second step (here, the leaving of the hydride group) is too large. This is a seemingly universal characteristic of simple reactivity indicators that are based on the isolated reagents and it is commonly observed in electrophilic aromatic substitution also, because electrophilic attack at an *ipso* site is often favorable, but the reaction barrier to *ipso* substitution is very large and so the *ipso* product is rarely observed.

Table 8 Condensed general-purpose reactivity indicator, $\Xi_{\Delta N \geq 0}^{\kappa}$, for nucleophilic attack on CNB using cHelpG charges. All of the carbon atoms are included in this table.

$\Delta N \setminus \kappa$	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
0	-0.27847	-0.25062	-0.22278	-0.19493	-0.16708	-0.13923	-0.11139	-0.08354	-0.05569	-0.02785	
0.1	-0.27847	-0.25064	-0.22282	-0.19499	-0.16717	-0.13934	-0.11152	-0.08369	-0.05587	-0.02804	-0.01461
0.2	-0.27847	-0.25067	-0.22286	-0.19506	-0.16726	-0.13945	-0.11165	-0.08385	-0.05604	-0.02925	-0.02923
0.3	-0.27847	-0.25069	-0.22291	-0.19512	-0.16734	-0.13956	-0.11178	-0.084	-0.05622	-0.0424	-0.04384
0.4	-0.27847	-0.25071	-0.22295	-0.19519	-0.16743	-0.13967	-0.11191	-0.08415	-0.05639	-0.05556	-0.05846
0.5	-0.27847	-0.25073	-0.22299	-0.19526	-0.16752	-0.13978	-0.11204	-0.0843	-0.06435	-0.06871	-0.07307
0.6	-0.27847	-0.25075	-0.22304	-0.19532	-0.1676	-0.13989	-0.11217	-0.08445	-0.07604	-0.08186	-0.08769
0.7	-0.27847	-0.25078	-0.22308	-0.19539	-0.16769	-0.14	-0.1123	-0.08461	-0.08773	-0.09502	-0.1023
0.8	-0.27847	-0.2508	-0.22312	-0.19545	-0.16778	-0.1401	-0.11243	-0.09068	-0.09942	-0.10817	-0.11691
0.9	-0.27847	-0.25082	-0.22317	-0.19552	-0.16786	-0.14021	-0.11256	-0.10091	-0.11112	-0.12132	-0.13153
1	-0.27847	-0.25084	-0.22321	-0.19558	-0.16795	-0.14032	-0.11269	-0.11114	-0.12281	-0.13447	-0.14614
<i>Para</i>						<i>Meta</i>					

In some electrophilic and nucleophilic substitution reactions, whether a given site of attack leads to product can be ascertained by evaluating the quality of the leaving group using, for example, any of the several recently proposed reactivity indicators for leaving group ability.^{7;55-57}

V.V. Discussion

In collaboration with Junia Melin, we recently proposed a general-purpose reactivity indicator that is capable of describing both charge controlled (hard-hard interactions) and electron-transfer controlled (soft-soft interactions) reactions, as well as intermediate cases.^{27;28} The purpose of this paper is to test whether the general-purpose reactivity indicator can successfully predict the reactivity of ambidentate electrophiles and nucleophiles. The first part of this paper (§II.A) establishes that the reactivity indicator works very well at predicting where electrophiles attack ambidentate nucleophiles. This result is somewhat unsurprising, since in our previous (albeit very limited) tests, the indicator seemed to successfully predict electrophilic attack on aromatic compounds.²⁸ The second part of this paper (§II.B) marks the first applications of the general-purpose reactivity indicator to nucleophilic attack. It is reassuring that the reactivity indicator also works well at predicting where nucleophiles attack ambidentate electrophiles.

It is interesting to observe that this reactivity indicator successfully models Hard/Soft-Acid/Base (HSAB) behavior without using the Parr-Pearson value for the hardness ($\eta = I - A$). The hardnesses, η^+ and η^- , are zero in exact *ab initio* theory, so it is reassuring that we are able to recover the types of results one associates with the HSAB principle using a model based on $\eta^+ = \eta^- = 0$.

Based on our previous computational study of this reactivity indicator,²⁸ the mathematical formalism underlying the indicator,²⁷ and this paper, we can issue some guidance for subsequent applications of this reactivity tool.

- This reactivity indicator does not contain information about steric hindrance. Sterically congested sites may not be as reactive as the indicator would lead one to expect.
- This reactivity indicator shows the most favorable site for initial attack. Sometimes this site might be a mechanistic dead end (i.e., the second step in the reaction might be unfavorable); sometimes the reagent might attack at one site and then migrate to another site. Also, the site predicted by the reactivity indicator will usually correspond to the kinetic product; the thermodynamic product will sometimes be different.

- This reactivity indicator is most effective for small, simple molecules, where the electronic structure of the reagent is dominant over steric effects, molecular rearrangements, and other “molecule geometry” factors.
- It is recommended to consider several different population analysis schemes; in cases where several different (but reasonable) population analysis schemes give different results, one must be very cautious when drawing conclusions from the data.

Overall, we believe that the general-purpose reactivity indicator is an effective reactivity tool, which should be added to the toolbox of chemists working in the theory of chemical reactivity. It is far from the ultimate goal of the qualitative density functional theory of chemical reactivity—it is an interpretative, rather than a predictive, tool—but among reactivity indicators that capture the key physics of chemical reactivity, the proposed general-purpose reactivity indicator is one of the simplest and most easily computable formulae that has been conceived. It is hoped that further studies will lead to greater insight into the physical foundations of this indicator, greater understanding of its range of chemical applicability, and further improvements to its accuracy and utility.

Acknowledgements:

Funding from NSERC, the Canada Research Chairs, and PREA is acknowledged. James Anderson acknowledges support from an OGSST graduate fellowship. Junia Melin and David C. Thompson have made important contributions to our work using the general-purpose reactivity indicator and are warmly acknowledged for these contributions.

V.VI. References

- (1) Head-Gordon, M. *Journal of Physical Chemistry* **1996**, *100*, 13213-13225.
- (2) Pople, J. A. *Rev.Mod.Phys.* **1999**, *71*, 1267-1274.
- (3) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem.Rev.* **2003**, *103*, 1793-1873.
- (4) Parr, R. G.; Yang, W. T. *Ann.Rev.Phys.Chem.* **1995**, *46*, 701-728.
- (5) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford UP: New York, 1989.
- (6) Kohn, W.; Becke, A. D.; Parr, R. G. *J.Phys.Chem.* **1996**, *100*, 12974-12980.

- (7) Ayers, P. W.; Anderson, J. S. M.; Bartolotti, L. J. *Int.J.Quantum Chem.* **2005**, *101*, 520-534.
- (8) Chermette, H. *J.Comp.Chem.* **1999**, *20*, 129-154.
- (9) Melin, J.; Ayers, P. W.; Ortiz, J. V. *Journal of Chemical Sciences* **2005**, *117*, 387-400.
- (10) Ayers, P. W.; Melin, J. *Theor.Chem.Acc.* **2006**, (accepted).
- (11) Grabowski, I.; Hirata, S.; Ivanov, S.; Bartlett, R. J. *J.Chem.Phys.* **2002**, *116*, 4415-4425.
- (12) Gorling, A.; Levy, M. *Phys.Rev.A* **1994**, *50*, 196-204.
- (13) Mori-Sanchez, P.; Wu, Q.; Yang, W. T. *J.Chem.Phys.* **2005**, *123*, 062204.
- (14) Ayers, P. W. *Theor.Chem.Acc.* **2001**, *106*, 271-279.
- (15) Bartolotti, L. J.; Ayers, P. W. *J.Phys.Chem.A* **2005**, *109*, 1146-1151.
- (16) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J.Chem.Phys.* **1978**, *68*, 3801-3807.
- (17) Pearson, R. G. *J.Am.Chem.Soc.* **1963**, *85*, 3533-3539.
- (18) Chattaraj, P. K.; Lee, H.; Parr, R. G. *J.Am.Chem.Soc.* **1991**, *113*, 1855-6.
- (19) Ayers, P. W. *J.Chem.Phys.* **2005**, *122*, 141102.
- (20) Ayers, P. W.; Parr, R. G.; Pearson, R. G. *J.Chem.Phys.* **2006**, *124*, 194107.
- (21) Ayers, P. W. *Faraday Discussions* **2007**, *135*, 161-190.
- (22) Mendez, F.; Gazquez, J. L. *J.Am.Chem.Soc.* **1994**, *116*, 9298-9301.
- (23) Pearson, R. G. *J.Chem.Educ.* **1987**, *64*, 561-567.
- (24) Parr, R. G.; Chattaraj, P. K. *J.Am.Chem.Soc.* **1991**, *113*, 1854-5.
- (25) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2000**, *122*, 2010-2018.
- (26) Klopman, G. *J.Am.Chem.Soc.* **1968**, *90*, 223-234.

- (27) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *Journal of Chemical Theory and Computation* **2007**, (accepted).
- (28) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *Journal of Chemical Theory and Computation* **2007**, (accepted).
- (29) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (30) Ayers, P. W.; Parr, R. G. *J.Am.Chem.Soc.* **2001**, *123*, 2007-2017.
- (31) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L., Jr. *Phys.Rev.Lett.* **1982**, *49*, 1691-1694.
- (32) Yang, W.; Zhang, Y.; Ayers, P. W. *Phys.Rev.Lett.* **2000**, *84*, 5172-5175.
- (33) Ayers, P. W.; Melin, J. *Theor.Chem.Acc.* **2006**, (accepted).
- (34) P. W. Ayers, *J. Math. Chem.* (accepted).
- (35) Yang, W.; Parr, R. G.; Pucci, R. *J.Chem.Phys.* **1984**, *81*, 2862-2863.
- (36) Parr, R. G.; Yang, W. *J.Am.Chem.Soc.* **1984**, *106*, 4049-4050.
- (37) Ayers, P. W.; Levy, M. *Theor.Chem.Acc.* **2000**, *103*, 353-360.
- (38) Yang, W.; Mortier, W. J. *J.Am.Chem.Soc.* **1986**, *108*, 5708-11.
- (39) Ayers, P. W.; Morrison, R. C.; Roy, R. K. *J.Chem.Phys.* **2002**, *116*, 8731-8744.
- (40) The reader is warned that, somewhat illogically, many researchers use "q" to represent the electronic population, rather than the charge. With this alternative convention the definitions in Eqs. (4) and (5) are changed by a multiplicative factor of -1.
- (41) Becke, A. D. *J.Chem.Phys.* **1993**, *98*, 1372-1377.
- (42) Lee, C.; Yang, W.; Parr, R. G. *Phys.Rev.B* **1988**, *37*, 785-789.
- (43) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Peersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.;

Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvetti, O.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham M.A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03, Revision C.02*; Gaussian Inc.: Wallingford, CT, 2004.

- (44) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J.Chem.Phys.* **2002**, *117*, 43-54.
- (45) Breneman, C. M.; Wiberg, K. B. *J.Comp.Chem.* **1990**, *11*, 361-373.
- (46) These are the "Microsoft" names for the colors we have chosen to use.
- (47) Li, Y.; Evans, J. N. S. *J.Am.Chem.Soc.* **1995**, *117*, 7756-7759.
- (48) Melin, J.; Aparicio, F.; Subramanian, V.; Galvan, M.; Chattaraj, P. K. *J.Phys.Chem.A* **2004**, *108*, 2487-2491.
- (49) Tundo, P.; Rossi, L.; Loris, A. *Journal of Organic Chemistry* **2005**, *70*, 2219-2224.
- (50) Leis, J. R.; Pena, M. E.; Rios, A. M. *Journal of the Chemical Society-Perkin Transactions 2* **1995**, 587-593.
- (51) Boga, C.; Bonamartini, A. C.; Forlani, L.; Modarelli, V.; Righi, L.; Sgarabotto, P.; Todesco, P. E. *European Journal of Organic Chemistry* **2001**, 1175-1182.
- (52) Pearson, R. G. *Science* **1966**, *151*, 172-177.
- (53) Here we are considering raw population analysis without summing the hydrogens into the neighboring carbons. If one sums hydrogens into the neighboring carbons, the carbons in the "ortho" position become more reactive than the nitrogen atom in the nitroso group.

- (54) A similar effect was observed in our first paper, where we noted that "ipso" attack on an aromatic ring is sometimes predicted. However, because there is no leaving group at such a carbon, the second step in the reaction has a very high energy. A similar effect might be operative here.
- (55) Ayers, P. W.; Anderson, J. S. M.; Rodriguez, J. I.; Jawed, Z. *Phys.Chem.Chem.Phys.* **2005**, *7*, 1918-1925.
- (56) Guerra, D.; Castillo, R.; Andres, J.; Fuentealba, P.; Aizman, A.; Contreras, R. *Chem.Phys.Lett.* **2006**, *424*, 437-442.
- (57) Jaramillo, P.; Domingo, L. R.; Perez, P. *Chem.Phys.Lett.* **2006**, *420*, 95-99.

Chapter VI

Relativistic Quantum Theory of Atoms in Molecules: Results for the ZORA Hamiltonian*

*The content of this chapter is in preparation to be submitted: **J. S. M. Anderson, P. W. Ayers**
“Relativistic Quantum Theory of Atoms in Molecules: Results for the ZORA Hamiltonian”;
special issue of *J. Chem. Phys.* for the REHE conference 2010

VI.I. Statement of the Problem

This chapter extends the quantum theory of atoms in molecules (QTAIM) to include relativistic corrections. Specifically, this chapter uses the quantum mechanical Hamiltonian at the scalar-relativistic zeroth-order regular approximation (ZORA) level theory. This level of theory is approximate, but it is usually adequate for chemical accuracy. The only change from non-relativistic QTAIM is the kinetic energy term. The quantum theory of atoms in molecules can be extended to include relativity at this level of theory. Interestingly, this is possible despite the fact atomic partitioning of molecules is sensitive to the form of the kinetic energy. Moreover, the mathematical form of the atoms is preserved in this theory. This is the first successful relativistic generalisation of QTAIM.

VI.II. Introduction

Predicting how a molecule changes in response to the insertion, deletion, or substitution of atoms or functional groups is of the utmost importance to chemists, theoretical and experimental alike. Such predictions are critical to chemists who want to design molecules with particular properties. It is a fundamental precept of chemistry that atoms and functional groups have transferable behaviour, and one theoretical approach that explains this observation is the quantum theory of atoms in molecules (QTAIM).¹⁻⁶ In QTAIM the value of a physical property of an entire system (usually a molecule) is obtained by adding the value of that property for all the atoms in the system. Under the assumption of perfect transferability, this allows one to—via a simple process of addition and subtraction—see how changes in atoms and functional groups affect molecular properties. (Perfect transferability is impossible,⁷⁻¹¹ however atoms and functional groups are commonly highly transferable.⁷)

When considering molecules that include heavy atoms, relativistic effects cannot be ignored. The most common way to address relativistic effects is to modify the nonrelativistic Hamiltonian in a way that approximately models the effects of relativity. This is the approach of the popular Zeroth-Order Regular Approximation (ZORA).¹²⁻¹⁹ (Although more sophisticated two-component and four-component methods are becoming increasingly prevalent,²⁰⁻²⁹ addressing those models requires an altogether different mathematical treatment, which we shall not pursue here.) The goal of this paper is to explore the structure of QTAIM when the normal non-relativistic Hamiltonian is replaced by a Hamiltonian with relativistic corrections. Here we will focus on the ZORA

Hamiltonian in its scalar-relativistic form. The addition of spin-orbit effects is more complicated and will be deferred to a later paper.

As espoused by Bader and his coworkers, QTAIM is motivated by and mathematically developed from a generalisation of the quantum mechanical principle of stationary action to a subsystem.^{1,4-6} This mathematical development has recently been the subject of some controversy,³⁰⁻³⁷ but we will overlook that controversy here. Our main goal is to extend the usual argumentation of QTAIM to the ZORA Hamiltonian; doing so establishes the relativistic AIM which is most nearly analogous to the ordinary, nonrelativistic AIM. (In this paper, we will use AIM to refer to the atom in a molecule that is associated with QTAIM and its relativistic generalisations and evade the deeper epistemological issues.^{7,38-40})

We are aware of only one previous attempt to mathematically generalise QTAIM to relativistic considerations. Cioslowski has considered relativistic QTAIM using a four-component relativistic formulation.⁴¹ In that work, a certain arbitrariness in the definition of the relativistic Lagrangian is noted. As noted by several workers and emphasized by Cohen, the arbitrariness in the Lagrangian arises even in the nonrelativistic formulation. Although this arbitrariness does not cause any problems for entire systems, it means that local (and subsystem) properties cannot be uniquely defined.⁴²⁻⁴⁷ Cioslowski explored the implications of this arbitrariness in the context of Hamilton's action principle (and not the principle of stationary action that subsumes it) and the arbitrariness makes it difficult to unambiguously identify an atom in a molecule. Our goal is more pragmatic than Cioslowski's (most calculations of molecules containing heavy atoms are performed using ZORA and Kohn-Sham density-functional theory) and less ambitious (since ZORA is only an approximation to the fully relativistic four-component form). One advantage of focussing on the ZORA Hamiltonian is that it is easy to identify the Lagrangians that are most nearly analogous to the nonrelativistic Lagrangians commonly employed in QTAIM.

The construction of the relativistic AIM is performed by constructing relativistic Lagrangians appropriate to the ZORA Hamiltonians and then applying the same subsystem least-action principle used for nonrelativistic QTAIM.^{1,4-6} The subsequent mathematical simplifications will be treated meticulously in subsequent work; this paper provides only the level of mathematical detail that is conventional in QTAIM. Our treatment supersedes the ordinary QTAIM argumentation, which can be recovered in the nonrelativistic ($c \rightarrow \infty$) limit of our equations. In the nonrelativistic QTAIM, the zero-flux atomic surfaces arise naturally from the second derivative terms in the kinetic energy operator and the observation that the potential energy operator is multiplicative. (The electron-

electron repulsion and electron-nuclear attraction do not depend on the electronic momentum.) We find it surprising, then, that when one modifies the kinetic-energy operator using the scalar-relativistic ZORA Hamiltonian in Kohn-Sham density-functional theory, the surfaces binding the atoms are still zero-flux surfaces of the electron density.

In the next section, we present background information on QTAIM and the ZORA Hamiltonian. In section V, we present the QTAIM analysis for the scalar-relativistic ZORA Hamiltonian (SR-ZORA). Section VI contains concluding remarks. Most of the mathematical justifications are deferred to the appendix.

VI.III. Background

A. The Atom in a Molecule

Matter is composed of atoms. The nature of chemistry, as embodied in the periodic table, is that molecules can be described in terms of their atomic constituents. Individual atoms are not anonymous constituents of molecules. Instead, each atom in a molecule contributes to the molecule's overall spectroscopic, thermodynamic, and kinetic properties in a well-defined, transferable way. An atomic leaving group is a simple example of this: when taught S_N1 reactions we learn that chlorine, bromine and iodine are exceptional leaving groups. Functional groups are simply characteristic arrangements of atoms that produce distinctive molecular properties. (For example, tosylate groups are excellent leaving groups.) These two observations motivate the postulate that within molecules there exist atoms (functional groups), and that these atoms (functional groups) have distinct and assignable properties. Furthermore, summing the properties of all the atoms in a molecule furnishes the properties of the entire molecule.

The quantum theory of atoms in molecules (QTAIM) is a theoretical formulation that captures these chemical principles. The mathematical foundations of QTAIM can be found in the literature^{1-6,48} and will not be repeated here. As was already noted by Cioslowski,⁴¹ if one attempts to build an AIM that encompasses relativistic quantum mechanics that is in the spirit of Bader's prior work,^{1,4-6} then the treatment must revert to the standard, nonrelativistic, version of QTAIM in the limit as $c \rightarrow \infty$. Just as in the nonrelativistic version of QTAIM, the relativistic quantum theory of atoms in molecules (R-QTAIM) hinges on

generalising the principle of stationary action for the total system to an open quantum subsystem. This defines the relativistic atom in a molecule (RAIM).

One advantage of QTAIM is that it is essentially axiomatic: one starts by requiring certain criteria that one wishes atoms in molecules to meet and then one derives the definition of an atom in a molecule using those criteria. While the underlying criteria can be debated, once they are adopted there is limited freedom to “customise” the atom. We believe the criteria proposed by Bader and Nguyen-Dang are reasonable, and so we will adopt them here.⁴

- *The definition of an atom must define all of its average properties. Moreover, these definitions must reduce to the quantum mechanical definition of the properties of the isolated atom.* This very pragmatic requirement is essential for chemistry. The first statement indicates merely that the atom is fully defined; if some of the properties of atoms were not defined, then QTAIM would not be very useful! The second statement ensures that QTAIM can describe the dissociation of molecules and the properties (e.g., atomisation energies) associated with that process.
- *Atomic properties must be additive. When the values of each atomic property are summed over all constituent atoms, they should give the average property for the molecule.* This requirement is necessary if one is to reproduce the experimentally observed transferability of atomic properties.⁷ While there is no *a priori* reason not to allow “bond” contributions (based on two atoms) to molecular properties (and even 3-atom, 4-atom, etc. contributions could also be considered), certainly the exhaustive atomic-level partitioning of molecular properties is the simplest.
- *The definition of an atom should be time-dependent.* Time-dependent phenomena are important to chemistry and an ideal theory of atoms in molecules will be able to address them. At present, however, truly time-dependent studies (e.g., analysis of molecular wavepacket propagation) using QTAIM are rarely (if ever) performed.
- *An atom should be defined as a region in real space.*⁴⁹ Many population analysis schemes are based on Hilbert-space partitioning,⁵⁰⁻⁵⁸ but insofar as matter exists in real space, it seems more concrete (and, arguably, more desirable) to define atoms in real space also. While one can define “fuzzy” overlapping atoms in real space,^{59,60} it is simpler to define atoms as regions in space. When atoms are associated with regions in space, atomic properties can be constructed using projection operators on the wave function, which makes a link to the Hilbert-space alternative.⁶¹

These precepts are the motivation for Bader’s approach to defining a quantum theory of atom in a molecule.^{1-6,48} Provided some technical mathematical

assumptions (notably the choice of Lagrangian and certain “Occam’s razor” simplifications) are made, this approach leads inexorably to the zero-flux surfaces that are ubiquitous in QTAIM.

VI.IV. The ZORA Hamiltonian

The zeroth-order regular approximation (ZORA) is widely used in relativistic molecular electronic structure calculations, especially in the context of Kohn-Sham density-functional theory.^{13,19} The ZORA approach captures most of the relativistic effects and, in particular, provides a reasonably accurate description of both valence and sub-valence electrons in atomic and molecular calculations.¹³

ZORA was first proposed by Chang *et al.*,¹² and by Heully *et al.*¹⁴ It follows from the Foldy-Wouthuysen¹⁸ transformation of the Dirac equation. Van Lenthe^{15,16} showed that ZORA is variationally stable and demonstrated that it contains similar relativistic corrections to the Pauli Hamiltonian. The ZORA Hamiltonian has also been shown to contain every term in the Breit-Pauli Hamiltonian to second order.^{13,62} However, ZORA does not have the problematic delta-function singularities at the nuclei that the Breit-Pauli Hamiltonian does.¹⁵ These advantages make this ZORA tractable, and applications of ZORA to a broad range of chemical phenomena illustrate its accuracy.¹⁵

The ZORA equation is

$$\left[(\vec{\sigma} \cdot \vec{p}) \frac{c^2}{2mc^2 - V} (\vec{\sigma} \cdot \vec{p}) + V \right] \Psi = i\hbar \frac{\partial \Psi}{\partial t} \quad (1)$$

where $\vec{\sigma}$ is the Pauli spin vector, \vec{p} is the three dimensional momentum operator, m is the electron mass, c is the speed of light, V is the potential energy, Ψ is the wave function, and t is time. Applying the identity

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = \vec{A} \cdot \vec{B} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B}) \quad (2)$$

to Eq. (1) allows one to divide the Hamiltonian into two parts, a spin orbit part

$$i\vec{\sigma} \cdot \left(\vec{p} \times \frac{c^2}{2mc^2 - V} \vec{p} \right) \quad (3)$$

and what is termed the scalar-relativistic (SR) contribution

$$\vec{p} \cdot \frac{c^2}{2mc^2 - V} \vec{p}. \quad (4)$$

This gives the physically intuitive form of the ZORA equations:

$$\begin{aligned} \left[\begin{array}{l} \vec{p} \cdot \frac{c^2}{2mc^2 - V} \vec{p} \\ + i\vec{\sigma} \cdot \left(\vec{p} \times \frac{c^2}{2mc^2 - V} \vec{p} \right) + V \end{array} \right] \Psi &= i\hbar \frac{\partial \Psi}{\partial t} \\ \left[\begin{array}{l} \vec{p} \cdot \frac{c^2}{2mc^2 - V} \vec{p} \\ + i\vec{\sigma} \cdot \left(\vec{p} \times \frac{c^2}{2mc^2 - V} \vec{p} \right) + V \end{array} \right] \Psi^* &= -i\hbar \frac{\partial \Psi^*}{\partial t} \end{aligned} \quad (5)$$

Neglecting the spin-orbit term results in the SR-ZORA equation,

$$\left[\vec{p} \cdot \frac{c^2}{2mc^2 - V} \vec{p} + V \right] \Psi = i\hbar \frac{\partial \Psi}{\partial t} \quad (6)$$

The SR-ZORA equation can be rewritten in a more explicit form

$$\begin{aligned} \left[-\hbar^2 \vec{\nabla} \cdot \frac{c^2}{2mc^2 - V} \vec{\nabla} + V \right] \Psi &= i\hbar \frac{\partial \Psi}{\partial t} \\ \left[-\hbar^2 \vec{\nabla} \cdot \frac{c^2}{2mc^2 - V} \vec{\nabla} + V \right] \Psi^* &= -i\hbar \frac{\partial \Psi^*}{\partial t} \end{aligned} \quad (7)$$

Turning the time-dependent Schrödinger equation into the SR-ZORA equation simply entails modifying the kinetic energy operator. From Eqs. (7), we identify the SR-ZORA Hamiltonian as

$$\hat{\mathcal{H}}^{SR} = -\hbar^2 \Psi^* \vec{\nabla} \cdot \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} \Psi + V \Psi^* \Psi \quad (8)$$

ZORA is used almost exclusively within the Kohn-Sham DFT.^{63,64} The potential-energy term in the denominator of the ZORA Hamiltonian causes difficulties for traditional, wave function based, *ab initio* methods. (*Ab initio* schemes for implementing the ZORA equations do exist, however.⁶⁵⁻⁶⁸) Henceforth we will assume that the potential in the denominator of the ZORA equation is a local one-electron potential (e.g., the Kohn-Sham potential, $v_{KS}(\mathbf{r})$).

VI.V. The SR-ZORA Extension of QTAIM

Extending the QTAIM to the SR-ZORA Hamiltonian closely follows the analysis in nonrelativistic QTAIM. The detailed motivation is presented in the appendices, but the main concepts will be sketched here. A detailed treatment of applying the Schwinger principle of stationary action is deferred to a following paper. Based on Eq. (8), a suitable Lagrangian, \mathcal{L}^{SR} , is defined through

$$\mathcal{L}^{SR} = \left[\frac{i\hbar}{2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t} \right) - \left(\hbar^2 \left[\frac{c^2}{2mc^2 - v_{KS}} \right] \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi + V \Psi^* \Psi \right) \right]. \quad (9)$$

Using identities from vector calculus and the time-dependent SR-ZORA equations, Eq. (5), the Lagrangian simplifies to

$$\mathcal{L}^{SR} = -\frac{\hbar^2}{2} \vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right]. \quad (10)$$

Upon integration over all but one electron, this Lagrangian resembles the Laplacian of the electron density, hence we will call this the SR-ZORA Laplacian. Notice that in the limit as $c \rightarrow \infty$, this becomes

$$\mathcal{L}^{QAIM} = -\frac{\hbar^2}{4m} \nabla^2 (\Psi^* \Psi), \quad (11)$$

which is the nonrelativistic Lagrangian used to derive QAIM. This indicates that the choice of Lagrangian in Eq. (10) is an appropriate choice for the SR-ZORA analogue to the standard nonrelativistic QAIM.

Just as in nonrelativistic QAIM, the atom is defined by requiring the variation of the total Lagrangian over the atomic regions to be zero:

$$0 = \delta \left(\int_{\Omega} d\vec{r} \int d\vec{\tau} \frac{-\hbar^2}{2} \vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right] \right) \quad (12)$$

or, applying the divergence theorem,

$$0 = \delta \left(\frac{-\hbar^2}{2} \oint_{\partial\Omega} d\vec{S} \cdot \int d\vec{\tau} \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right) \quad (13)$$

Here we are denoting integration over the $N-1$ electrons not operated upon by the Lagrangian operator by $\int d\vec{\tau}$; integration over these coordinates gives the local Lagrangian density,

$$\mathcal{L}^{SR}(\vec{r}) = \int d\vec{\tau} \frac{-\hbar^2}{2} \vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right]. \quad (14)$$

The outer integral in Eq. (12), denotes the integration of the local Lagrangian density over the atomic region, $\int_{\Omega} d\vec{r} \mathcal{L}^{SR}(\vec{r})$. In equation (13) the surface integral is denoted in the obvious way, with $\oint_{\partial\Omega} d\vec{S} \cdot = \oint_{\partial\Omega} da \hat{n} \cdot$ and \hat{n} the outward unit normal to the surface.

It follows from Eq. (13) that a sufficient (but not necessary!) condition for stationary action is that

$$0 = \left[\frac{c^2}{2mc^2 - v_{KS}(\vec{r})} \right] \vec{\nabla} \rho(\vec{r}) \cdot \hat{n} \quad (15)$$

on the atomic surface. Experience shows that for neutral atoms and molecules, the Kohn-Sham potential is always negative. (Regardless, it seems very unlikely that the Kohn-Sham potential would ever be greater than $2mc^2 = 37558$ Hartree and the assumption that $2mc^2 - v_{KS}(\vec{r}) > 0$ is inherent in the derivation of the ZORA.)

Since the coefficient $\vec{\nabla} \rho(\vec{r}) \cdot \hat{n}$ is always positive, the surface bounding the atoms in SR-ZORA is exactly the same as the one that defines the AIM the in the original nonrelativistic QTAIM formulation.

To obtain the final expression of the stationary action principle in its most convenient form for, e.g., evaluating atomic properties, one takes the total variation of the action, $\delta \mathcal{W}_{\Omega,12}^{SR}[\Psi, \Omega]$, and simplifies, obtaining

$$\delta \mathcal{W}_{\Omega,12}^{SR}[\Psi, \Omega] = \frac{\varepsilon}{2} \int_1^2 dt \left\{ \frac{i}{\hbar} \langle [\hat{\mathcal{H}}^{SR}, \hat{\mathcal{F}}] \rangle_{\Omega} + cc \right\}, \quad (16)$$

This is (except for the difference in Hamiltonian), exactly the same as the result in the nonrelativistic case.^{1,4-6} The derivation of Eq. (16) is subtle and lengthy, but it works precisely the same way as in normal QTAIM (cf. ref. ¹), except for the replacement of the classical Hamiltonian and classical Laplacian by their ZORA counterparts.

VI.VI. Conclusion

Replacing the kinetic-energy operator in the nonrelativistic Hamiltonian with the scalar-relativistic ZORA kinetic-energy operator (1) does not change the definition of an atom in a molecule and (2) does not change the formulation of the quantum theory of atoms in molecules (QTAIM) espoused by Bader in any significant way. This result is perhaps surprising, since the derivations associated with QTAIM critically depend on the form of the kinetic energy operator.

This result provides evidence for the general utility of density-based partitioning schemes in general, and QTAIM in particular, across the periodic table. This is a reassuring result for chemists who wish to use the powerful conceptual tools in QTAIM to clarify structure and bonding in molecules and materials containing heavy atoms. The tools are not only extensible to those

substances with heavy atoms but, at least at the level of SR-ZORA, they go through unchanged.

Motivated by this favourable result, we have recently studied the effects of the spin-orbit term and the more general ZORA (cf. Eq. (1)). The results in those cases will be published separately,⁶⁹ but it may be mentioned that QTAIM carries over intact to those more general forms, albeit sometimes with (slight) modifications to the surfaces bounding the atoms. How (and whether) QTAIM can be extended to two-component or four-component relativistic equations without encountering the problems previously identified by Cioslowski is a more daunting question, and one for which presently we have no decisive answer.

Acknowledgements

We would like thank Dr. Richard F. W. Bader and Dr. David C. Thompson for helpful questions and discussions. Funding was provided by an NSERC graduate fellowship (JSMA) and by the Canada Research Chairs and an NSERC Discovery Grant (PWA).

VI.VII. Appendices

A. Defining the SR-ZORA Lagrangian

The purpose of this appendix is to show that the chosen form of the Lagrangian (Eq. (9)) is reasonable.

In order to derive a relativistic QTAIM, one must first propose an appropriate form for the quantum-mechanical Lagrangian. Because our goal is to derive the relativistic counterpart of QTAIM, we choose a Lagrangian that is directly analogous to the one used by Bader and his coworkers, specifically

$$\mathcal{L}^{SR} = \frac{i\hbar}{2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t} \right) - \left(\hbar^2 \left[\frac{c^2}{2mc^2 - v_{KS}} \right] \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi + V \Psi^* \Psi \right). \quad (17)$$

(Equation (17) is directly relevant to the special case of a one-electron molecule. The general N -electron case is encompassed by considering Ψ and Ψ^* as field operators. One could show the multi-electron character explicitly by summing the kinetic energy term over all electrons, but we will not do this since it complicates the notation.)

We now want to show the link between the form of the kinetic energy in Eq. (17) and the form that arises in the SR-ZORA Hamiltonian (cf. Eq. (8)),

$$-\frac{\hbar^2}{2} \Psi^* \vec{\nabla} \cdot \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} \Psi + cc . \quad (18)$$

To this end, we apply the product rule for the divergence,

$$A \vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (A \vec{B}) - \vec{\nabla} A \cdot \vec{B} , \quad (19)$$

to expression (18) (using $A = \Psi^*$ and $\vec{B} = \frac{c^2}{2mc^2 - V} \vec{\nabla} \Psi$). This gives

$$\begin{aligned} -\frac{\hbar^2}{2} \Psi^* \vec{\nabla} \cdot \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} \Psi + cc = \\ \left[-\frac{\hbar^2}{2} \vec{\nabla} \cdot \left(\frac{c^2}{2mc^2 - v_{KS}} \Psi^* \vec{\nabla} \Psi \right) \right. \\ \left. + \frac{\hbar^2}{2} \left[\frac{c^2}{2mc^2 - v_{KS}} \right] \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi \right] + cc \end{aligned} \quad (20)$$

The kinetic energy term in our proposed Lagrangian, (17), appears when we explicitly show the terms from the complex conjugate,

$$\begin{aligned} -\frac{\hbar^2}{2} \Psi^* \vec{\nabla} \cdot \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} \Psi + cc = \\ \left[-\frac{\hbar^2}{2} \vec{\nabla} \cdot \left(\frac{c^2}{2mc^2 - v_{KS}} \right) (\Psi^* \vec{\nabla} \Psi + \Psi \vec{\nabla} \Psi^*) \right. \\ \left. + \hbar^2 \left[\frac{c^2}{2mc^2 - v_{KS}} \right] \vec{\nabla} \Psi \cdot \vec{\nabla} \Psi^* \right] . \end{aligned} \quad (21)$$

Just as in nonrelativistic QTAIM,^{1,4-6} however, there is an additional “non-kinetic energy” term (the first term in Eq. (21)). The additional term is simplified by using the product rule for the gradient,

$$\vec{\nabla} (AB) = B \vec{\nabla} A + A \vec{\nabla} B \quad (22)$$

to obtain

$$\begin{aligned} -\frac{\hbar^2}{2} \Psi^* \vec{\nabla} \cdot \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} \Psi + cc = \\ \left[-\frac{\hbar^2}{2} \vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right] \right. \\ \left. + \hbar^2 \left[\frac{c^2}{2mc^2 - v_{KS}} \right] \vec{\nabla} \Psi \cdot \vec{\nabla} \Psi^* \right] \end{aligned} \quad (23)$$

As $c \rightarrow \infty$, $\frac{c^2}{2mc^2 - V} \rightarrow \frac{1}{2m}$. Substituting this result into Eq. (23) recovers the analogous non-relativistic formula, namely

$$-\frac{\hbar^2}{4m} \nabla^2 (\Psi^* \Psi) + \frac{\hbar^2}{2m} \vec{\nabla} \Psi^* \cdot \vec{\nabla} \Psi. \quad (24)$$

As expected, Eq. (24) is simply the sum of a Laplacian term and the nonrelativistic kinetic-energy contribution to the Lagrangian. The analogy between Eqs. (23) and (24) motivates us to identify

$$\vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right] \quad (25)$$

as the SR-ZORA Laplacian. The reasonability of our choice for the SR-ZORA Lagrangian is confirmed by the recovery of the analogous nonrelativistic expressions in the appropriate $c \rightarrow \infty$ limit.

Identity (23) relates the kinetic-energy term in the Lagrangian to the kinetic-energy term that enters the time-dependent SR-ZORA equations (cf. Eq.(7)). Inserting identity (23) and the time-dependent SR-ZORA equations, into the SR-ZORA Lagrangian (cf. Eq. (17)) gives:

$$\begin{aligned} \mathcal{L}^{SR} &= \frac{i\hbar}{2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - c c \right) - \frac{\hbar^2}{2} \vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right] \\ &+ \left(\frac{\hbar^2}{2} \Psi^* \vec{\nabla} \cdot \frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} \Psi + c c \right) - V \Psi^* \Psi \\ &= -\frac{\hbar^2}{2} \vec{\nabla} \cdot \left[\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right] \end{aligned} \quad (26)$$

So the SR-ZORA Lagrangian reduces to the SR-ZORA Laplacian. This is directly analogous to the relativistic case because⁵⁴

$$\lim_{c \rightarrow \infty} \mathcal{L}^{SR} = -\frac{\hbar^2}{4m} \nabla^2 (\Psi^* \Psi). \quad (27)$$

B. Motivating the Zero-Flux Condition

The purpose of this appendix is to show that the zero-flux condition is a reasonable choice for the bounding surfaces of atoms. This argument is based on the absence of any “best” choice for the atomic Lagrangian, which suggests that AIM should be defined so that several “reasonable” choices for the Lagrangian give identical results.

Bader and Nguyen-Dang⁴ define the atomic Lagrangian by integrating total Lagrangian (Eq. (17)) over an atomic region. For the electron on which the kinetic/potential energy operators apply, this entails integration over the atomic volume, Ω ; the other $N-1$ electrons over all of space,

$$\mathcal{L}_{\Omega}^{SR} [\Psi, t] = \int_{\Omega} d\vec{r} \int d\vec{\tau} \mathcal{L}^{SR}. \quad (28)$$

Bader *et al.* also define an alternative, Hamiltonian-based, formulation of the atomic Lagrangian,

$$\tilde{\mathcal{L}}_{\Omega}^{SR} [\Psi, t] = \int_{\Omega} d\vec{r} \int d\vec{\tau} \left\{ \frac{i\hbar}{2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t} \right) - \frac{1}{2} \left(\Psi^* \hat{\mathcal{H}}^{SR} \Psi + \Psi \hat{\mathcal{H}}^{SR} \Psi^* \right) \right\} \quad (29)$$

It is difficult to argue for the superiority of one of these Lagrangians over the other, so it seems reasonable to require that the two Lagrangians give identical formulations of atoms in molecules. Referring back to identity (23), it is clear that as

$$\mathcal{L}_{\Omega}^{SR} [\Psi, t] = \tilde{\mathcal{L}}_{\Omega}^{SR} [\Psi, t] - \frac{\hbar^2}{2} \int_{\Omega} d\vec{r} \int d\vec{\tau} \vec{\nabla} \cdot \left(\frac{c^2}{2mc^2 - v_{KS}} \vec{\nabla} (\Psi^* \Psi) \right) \quad (30)$$

The two formulations for the atomic Lagrangian are identical when

$$0 = -\frac{\hbar^2}{2} \int_{\Omega} d\vec{r} \vec{\nabla} \cdot \left(\frac{c^2}{2mc^2 - v_{KS}(\vec{r})} \vec{\nabla} \rho(\vec{r}) \right) \quad (31)$$

According to the divergence theorem, a sufficient (but not necessary) condition for this to be achieved is that

$$0 = \frac{c^2}{2mc^2 - v_{KS}(\vec{r})} \vec{\nabla} \rho(\vec{r}) \cdot \hat{n} \quad (32)$$

which leads to the “zero flux” condition that the gradient of the electron density is orthogonal to the normal to the atomic surface.

VI.VIII. References

- 1 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*. (Clarendon, Oxford, 1990).
- 2 R. F. W. Bader and P. M. Beddall, *J. Chem. Phys.* **56** (7), 3320 (1972).
- 3 R. F. W. Bader, Y. Tal, S. G. Anderson, and T. T. Nguyen-Dang, *Isr.J.Chem.* **19**, 8 (1980).

- 4 R. F. W. Bader and T. T. Nguyendang, *Adv. Quantum Chem.* **14**, 63 (1981).
- 5 R. F. W. Bader, S. Srebrenik, and T. T. Nguyendang, *J. Chem. Phys.* **68** (8), 3680 (1978).
- 6 S. Srebrenik, R. F. W. Bader, and T. T. Nguyendang, *J. Chem. Phys.* **68** (8), 3667 (1978).
- 7 C. F. Matta and R. F. W. Bader, *J. Phys. Chem. A* **110**, 6365 (2006).
- 8 P. G. Mezey, *Mol. Phys.* **96** (2), 169 (1999).
- 9 R. F. W. Bader and P. Becker, *Chem. Phys. Lett.* **148**, 452 (1988).
- 10 J. Riess and W. Munch, *Theor. Chim. Act.* **58** (4), 295 (1981).
- 11 P. W. Ayers, *J. Chem. Phys.* **113**, 10886 (2000).
- 12 C. Chang, M. Pelissier, and P. Durand, *Phys. Scr.* **34** (5), 394 (1986).
- 13 M. Filatov and D. Cremer, *Mol. Phys.* **101** (14), 2295 (2003).
- 14 J. L. Heully, I. Lindgren, E. Lindroth, S. Lundqvist, and A. M. Martenssonpendrill, *J. Phys. B* **19**, 2799 (1986).
- 15 E. Vanlenthe, E. J. Baerends, and J. G. Snijders, *J. Chem. Phys.* **99** (6), 4597 (1993).
- 16 E. Vanlenthe, E. J. Baerends, and J. G. Snijders, *J. Chem. Phys.* **101** (11), 9783 (1994).
- 17 E. vanLenthe, R. vanLeeuwen, E. J. Baerends, and J. G. Snijders, *Int. J. Quantum Chem.* **57** (3), 281 (1996).
- 18 L. L. Foldy and S. A. Wouthuysen, *Physical Review* **78**, 29 (1950).
- 19 G. T. Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. Van Gisbergen, J. G. Snijders, and T. Ziegler, *J. Comput. Chem.* **22**, 931 (2001).
- 20 A. H. Macdonald and S. H. Vosko, *Journal of Physics C* **12**, 2977 (1979).
- 21 W. J. Liu, F. Wang, and L. M. Li, *Journal of Theoretical & Computational Chemistry* **2** (2), 257 (2003).
- 22 T. Yanai, H. Iikura, T. Nakajima, Y. Ishikawa, and K. Hirao, *J. Chem. Phys.* **115** (18), 8267 (2001).
- 23 M. Mayer, S. Kruger, and N. Rosch, *J. Chem. Phys.* **115** (10), 4411 (2001).
- 24 A. Willetts, L. Gagliardi, A. G. Ioannou, A. M. Simper, C. K. Skylaris, S. Spencer, and N. C. Handy, *Int. Rev. Phys. Chem.* **19**, 327 (2000).
- 25 S. Varga, B. Fricke, H. Nakamatsu, T. Mukoyama, J. Anton, D. Geschke, A. Heitmann, E. Engel, and T. Bastug, *J. Chem. Phys.* **112** (8), 3499 (2000).
- 26 H. M. Quiney, H. Skaane, and I. P. Grant, *Advances in Quantum Chemistry, Vol 32: Quantum Systems in Chemistry and Physics, Pt Ii* **32**, 1 (1999).

- 27 W. J. Liu, G. Y. Hong, D. D. Dai, L. M. Li, and M. Dolg, *Theor. Chem. Acc.* **96** (2), 75 (1997).
- 28 E. Engel and R. M. Dreizler, *Density Functional Theory II* **181**, 1 (1996).
- 29 K. G. Dyall, I. P. Grant, C. T. Johnson, F. A. Parpia, and E. P. Plummer, *Comput. Phys. Commun.* **55** (3), 425 (1989).
- 30 P. Cassam-Chenai and D. Jayatilaka, *Theor. Chem. Acc.* **105**, 213 (2001).
- 31 R. F. W. Bader, *Theor. Chem. Acc.* **107**, 381 (2002).
- 32 P. Cassam-Chenai and D. Jayatilaka, *Theor. Chem. Acc.* **107** (6), 383 (2002).
- 33 P. Cassam-Chenai, *J. Math. Chem.* **31**, 145 (2002).
- 34 L. Delle Site, *Theor. Chem. Acc.* **107**, 378 (2002).
- 35 E. S. Kryachko, *Theor. Chem. Acc.* **107**, 375 (2002).
- 36 J. R. Mohallem, *Theor. Chem. Acc.* **107**, 372 (2002).
- 37 R. F. W. Bader, *Theor. Chem. Acc.* **105**, 276 (2001).
- 38 R. G. Parr, P. W. Ayers, and R. F. Nalewajski, *J. Phys. Chem. A* **109**, 3957 (2005).
- 39 S. Shahbazian and M. Zahedi, *Foundations of Chemistry* **8**, 37 (2006).
- 40 S. Shahbazian and M. Zahedi, *Foundations of Chemistry* **9**, 85 (2007).
- 41 J. Cioslowski and J. Karwowski, in *Fundamentals of molecular similarity* (Kluwer, Dordrecht, 2001), pp. 101.
- 42 L. Cohen, *J. Chem. Phys.* **70**, 788 (1979).
- 43 L. Cohen, *J. Chem. Phys.* **80**, 4277 (1984).
- 44 L. Cohen and C. Lee, *Found. of Phys.* **17**, 561 (1987).
- 45 L. Cohen, *Phys. Lett. A* **212**, 315 (1996).
- 46 P. W. Ayers, R. G. Parr, and A. Nagy, *Int. J. Quantum Chem.* **90**, 309 (2002).
- 47 J. S. M. Anderson, P. W. Ayers, and J. I. R. Hernandez, *J. Phys. Chem. A* **114**, 8884 (2010).
- 48 R. F. W. Bader, *J. Chem. Phys.* **73**, 2871 (1980).
- 49 R. F. W. Bader, *Int. J. Quantum Chem.* **49** (3), 299 (1994).
- 50 R. S. Mulliken, *J. Chem. Phys.* **23**, 1833 (1955).
- 51 R. S. Mulliken, *J. Chem. Phys.* **23**, 1841 (1955).
- 52 R. S. Mulliken, *J. Chem. Phys.* **23**, 2343 (1955).
- 53 R. S. Mulliken, *J. Chem. Phys.* **23**, 2338 (1955).
- 54 R. F. Nalewajski and O. Sikora, *J. Phys. Chem. A* **104**, 5638 (2000).
- 55 A. E. Reed, L. A. Curtiss, and F. Weinhold, *Chem. Rev.* **88** (6), 899 (1988).
- 56 A. E. Reed, R. B. Weinstock, and F. Weinhold, *J. Chem. Phys.* **83** (2), 735 (1985).
- 57 A. E. Reed and F. Weinhold, *J. Chem. Phys.* **78** (6), 4066 (1983).
- 58 K. Morokuma, *J. Chem. Phys.* **55**, 1236 (1971).

- 59 F. L. Hirshfeld, *Theor. Chim. Act.* **44**, 129 (1977).
- 60 P. Salvador and I. Mayer, *J. Chem. Phys.* **120** (11), 5046 (2004).
- 61 P. Bultinck, M. Rafat, R. Ponec, B. Van Gheluwe, R. Carbó-Dorca, and P. Popelier, *J. Phys. Chem. A* **110**, 7642 (2006).
- 62 E. van Lenthe, R. van Leeuwen, E. J. Baerends, and J. G. Snijders, *Int. J. Quantum Chem.* **57** (3), 281 (1996).
- 63 W. Kohn and L. J. Sham, *Phys.Rev.* **140**, A1133 (1965).
- 64 R. G. Parr and W. T. Yang, *Annu. Rev. Phys. Chem.* **46**, 701 (1995).
- 65 S. Faas, J. G. Snijders, and J. H. van Lenthe, in *Quantum Systems in Chemistry and Physics, Vol 1 - Basic Problems and Model Systems*, edited by A. HernandezLaguna, J. Maruani, R. McWeeny, and S. Wilson (2000), Vol. 2, pp. 251.
- 66 W. Klopper, J. H. van Lenthe, and A. C. Hennen, *J. Chem. Phys.* **113**, 9957 (2000).
- 67 S. Faas, J. H. Van Lenthe, and J. G. Snijders, *Mol. Phys.* **98**, 1467 (2000).
- 68 S. Faas, J. H. van Lenthe, A. C. Hennen, and J. G. Snijders, *J. Chem. Phys.* **113**, 4052 (2000).
- 69 J. S. M. Anderson and P. W. Ayers, in preparation.

Chapter VII

How Ambiguous Is the Local Kinetic Energy?*

*The content of this chapter has been published: **J. S. M. Anderson**, P. W. Ayers, J. I. Rodriguez Hernandez “How Ambiguous is the Local Kinetic Energy?”; *J. Phys. Chem. A* **2010**, *114*, 8884-8895 (invited issue in honour of Klaus Ruedenberg Festschrift).

VII.I. Statement of the Problem

As demonstrated in Chapter VI, the changing the form of the quantum-mechanical kinetic energy operator need not change the underlying quantum theory of atoms in molecules (QTAIM). This chapter shows that this is not always true. To do this, all forms of the kinetic energy that are consistent with the quantum-classical correspondence principle are considered. While many of these forms seem obvious to those experienced with quantum mechanics, some of the forms are very surprising. All the obvious forms of the kinetic energy give the same value when evaluated over the atoms defined in QTAIM. An infinite number of nonobvious forms give consistent atomic kinetic energies as well. However, there are also an infinite number of exotic forms for the kinetic energy that give different values of the atomic kinetic energies in QTAIM. One can define different proper open quantum subsystems based on these exotic forms of the kinetic energy operator, but those subsystems generally do not represent atoms in molecules.

VII.II. Introduction

Upon formation of a chemical bond, electrons from the bonding atoms delocalize, reducing the kinetic energy and stabilizing the system. The critical role of the kinetic energy in chemical bonding may be attributed to Hellmann,¹ but the quantitative and thorough justification of this perspective is usually attributed to Ruedenberg's classic 1962 paper "The Physical Nature of the Chemical Bond."² (A decade later, Goddard and Wilson elaborated upon this work^{3,4} and Kutzelnigg wrote a classic paper from the same viewpoint entitled "The Physical Mechanism of the Chemical Bond"⁵) Of course, no interpretation of chemical bonding based on energy decomposition analysis is unassailable,⁶ but Ruedenberg's perspective is considered to be reasonable and elucidating not only by the majority of theoretical chemists, but by undergraduate physical chemistry students.

Early on, it was realized that *local* changes in the kinetic energy would be similarly useful in interpreting chemical bonding. In follow-up papers to ref. ², Feinberg, Ruedenberg, and Mehler utilize the local kinetic energy,⁷⁻⁹

$$t_+(\mathbf{r}) = \left\langle \sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) \frac{\nabla_i \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \cdot \nabla_i \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)}{2} \right\rangle, \quad (1)$$

noting that because this quantity "contains only positive contributions from all volume elements, it is particularly well suited to elucidate the relationship between wavefunction rearrangements and the kinetic energy."⁸ They recognize,

however, that the alternative “Schrödinger form”

$$t_{Sch}(\mathbf{r}) = \left\langle \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \sum_{i=1}^N \left(-\frac{1}{2} \delta(\mathbf{r}_i - \mathbf{r}) \nabla_i^2 \right) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \right\rangle \quad (2)$$

is also acceptable.⁹ So is any appropriate linear combination between the two,

$$t_{\alpha}(\mathbf{r}) = \alpha t_{+}(\mathbf{r}) + (1 - \alpha) t_{Sch}(\mathbf{r}), \quad (3)$$

with the choice $\alpha = 1/2$ being the preferred definition of several authors.¹⁰⁻¹² Does Eq. (3) capture the full generality of permissible local kinetic energies? In section IV we argue that it does not.¹³ It seems, however, that all of the *explicit* definitions of the local kinetic energy that have been proposed in the literature fit into the form of Eq. (3).^{6,10,14-24}

The key role of the kinetic energy in the formation of the covalent bond has been recognized and used repeatedly since the mid-1970s, and it is impossible to summarize more than a tiny segment of the literature. The utility of plotting the *local* kinetic energy (in this case, $t_{+}(\mathbf{r})$) is clearly demonstrated in the recent paper of Ruedenberg and Schmidt,²⁵ these authors also give a survey of the literature. Tachibana and his coworkers have also considered the role of the local kinetic energy (using $t_{Sch}(\mathbf{r})$) in chemical bonding.^{26,27} Because the regions, far from an atom or molecule, where $t_{Sch}(\mathbf{r})$ is negative can be labeled “classically forbidden,” Tachibana *et al.* associate chemical bonding with the elimination of classically forbidden barriers between reagents. As the two reagents approach the volumes surrounded by their zero-contours of $t_{Sch}(\mathbf{r})$ eventually unite, indicating that sharing of electrons between the two reagents is classically allowed. This is considered a signature of covalent bonding. The picture is very similar to that of Ruedenberg and coworkers: it is the ability of electrons to delocalize between fragments (lowering the kinetic energy relative to a hypothetical reference system in which electrons are not allowed to be delocalized) that causes the formation of a covalent bond to be energetically favorable.^{25,28-31} More recently, Nalewajski, emphasizing the link between kinetic energy and information,³² has stressed the role of the kinetic energy as the driving force of chemical bonding.³³

Many other researchers have used some variant of the local kinetic energy to elucidate bonding. For example, the key ingredient in the electron localization function is the difference between $t_{+}(\mathbf{r})$ and the equivalent expression for noninteracting bosons; this quantity is then divided by a “reference” to render it dimensionless.³⁴⁻³⁷ Various local covariance measures are extremely closely related to the electron localization function and the local kinetic energy.³⁸ The localized orbital locator³⁹ of Schmier and Becke is precisely $t_{+}(\mathbf{r})$.⁴⁰⁻⁴³ The local temperature/nighness measure associated with Parr and his coworkers is proportional to $t(\mathbf{r})/\rho(\mathbf{r})$, where

$$\rho(\mathbf{r}) = \left\langle \Psi^*(\mathbf{r}_1, \dots, \mathbf{r}_N) \left(\sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) \right) \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \right\rangle \quad (4)$$

is the electron density and $t(\mathbf{r})$ is usually taken to be either $t_+(\mathbf{r})$ or $t_{\alpha=1/2}(\mathbf{r})$ (from Eq. (3)).^{10,19,44} The kinetic energy density is the second momentum moment, and other researchers have based their investigations more directly on the momentum density or a local momentum density (based, for example, on the Wigner distribution^{45,46}). Pioneering work along these lines was performed by Rozendaal and Baerends;⁴⁷ the parity function of Schmider and Becke is closely related.^{39,48} Another local momentum density model has been recently explored by Bohórquez and Boyd.⁴⁹

The ambiguity in the local kinetic energy can be considered to be inherited from the ambiguity in the local electronic stress tensor,⁵⁰⁻⁵²

$$t(\mathbf{r}) = -\frac{1}{2} \text{Tr}[\tilde{\sigma}(\mathbf{r})]. \quad (5)$$

Equation (5) is the typical definition of the electronic stress tensor in the chemistry literature; the physics literature sometimes uses a slightly different definition, differing by multiplicative factors and additional terms associated with external fields. If one defines the one-electron reduced density matrix in the usual way using either explicit integration or field operators,

$$\begin{aligned} \gamma(\mathbf{r}, \mathbf{r}') &= N \iint \dots \int \Psi^*(\mathbf{r}', \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r}_2 \dots d\mathbf{r}_N \\ &= \langle \Psi | \hat{\psi}^\dagger(\mathbf{r}') \hat{\psi}(\mathbf{r}) | \Psi \rangle \end{aligned} \quad (6)$$

then an infinite family of stress tensors, with elements given by

$$\left[\tilde{\sigma}_{\alpha, \beta}(\mathbf{r}) \right]_{ij} = -\frac{1}{2} \left[\begin{array}{c} \alpha \left(\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r_j} \right) \\ -(1-\alpha) \left(\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r'_j} \right) \\ + \delta_{ij} \beta \nabla^2 \rho(\mathbf{r}) \end{array} \right]_{\mathbf{r}'=\mathbf{r}} \quad (7)$$

can be proposed. Here $r_1 = x$, $r_2 = y$, $r_3 = z$ is shorthand for the Cartesian coordinates, δ_{ij} is the Kronecker delta function, α and β are real numbers, and $\rho(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r})$ is the electron density. All of the stress tensors in common use in the literature belong to this family,⁵²⁻⁵⁷ including the popular Schrödinger-Pauli-Epstein form $\tilde{\sigma}_{1/2,0}(\mathbf{r})$.⁵⁰⁻⁵² The stress tensor in Eq. (7) is defined so that it is consistent with the local kinetic energy in Eq. (3),

$$t_\alpha(\mathbf{r}) = -\frac{1}{2} \text{Tr}[\tilde{\sigma}_{\alpha,0}(\mathbf{r})] \quad (8)$$

More generally,

$$t_{\alpha+3\beta}(\mathbf{r}) = -\frac{1}{2} \text{Tr}[\tilde{\sigma}_{\alpha,\beta}(\mathbf{r})] \quad (9)$$

The idea of using the stress tensor to elucidate chemical bonding originates in the work of Epstein⁵² and was popularized through the work of Bader, Deb, and Parr.^{54,58-65} Within the chemistry community, most of the recent work has arisen from the groups of Tachibana and Jenkins.^{26,27,66-73} Physicists have recently explored the concept also.^{55,56,74-82} For the chemist, however, it is the work of Tachibana and Jenkins, who have stressed how the eigenvalues and eigenvectors of the stress tensor can be used to understand the strength, directionality, and deformability of chemical bonding patterns, whose work is the most elucidating. Tachibana's group, in particular, has stressed the characteristic "spindle structure" that the stress tensor takes in covalent chemical bonds.^{66,67} The importance of this characteristic "spindle structure" appears in the recent analysis, by Ichikawa and Tachibana, of the prototypical covalent chemical bond (the hydrogen molecule-ion H_2^+).⁶⁹ (It is worth noting that the physicists had studied the stress tensor of H_2^+ even earlier.⁷⁹)

The form of the local kinetic energy also plays a key role in Bader's quantum theory of atoms in molecules (QTAIM).^{20,83,84} The particular form of the atom in QTAIM may be justified by two different approaches. The simplest approach is to require that the expectation value of the kinetic energy in an AIM is well-defined, so that the virial theorem can be used to infer the energy of the AIM.^{20,21,85,86} In Bader's formulation, the AIM are bounded by zero-flux surfaces of the electron density, that is, the normal vector to the surface is orthogonal to the gradient of the electron density. With this definition, the integral of the Laplacian of the electron density over an atomic region is zero

$$\int_{\Omega} \nabla^2 \rho(\mathbf{r}) d\mathbf{r} = \oint_{\partial\Omega} \nabla \rho(\mathbf{r}) \cdot \mathbf{n} da = 0. \quad (10)$$

It is worth noting that the members of the family of local kinetic energies defined by Eq. (3) differ by just such a term, and therefore give precisely the same expectation value of the kinetic energy for AIM bound by zero-flux surfaces. In particular, we may rewrite Eq. (3) as

$$t_{\alpha}(\mathbf{r}) = t_{+}(\mathbf{r}) - \left(\frac{1-\alpha}{4} \right) \nabla^2 \rho(\mathbf{r}). \quad (11)$$

The alternative method for deriving the quantum theory of atoms in molecules is based on a stationary-action argument.^{16,20,87-90} In that approach, one must define a quantum mechanical Lagrangian, which in turn requires making a choice for the kinetic energy operator. Typically one uses the forms implied by

either Eq. (1) or Eq. (2), but, as recently shown by Nasertayoob and Shahbazian, there is much larger family of quantum mechanical Lagrangians that lead back to the conventional QTAIM definition of the atom.⁹⁰ However, if one chose a different form for the local kinetic energy, then one might obtain a different criterion for defining the proper open quantum subsystems.⁹¹ The definition of the AIM, once again, is entangled with the unfortunate arbitrariness of the local kinetic energy in quantum mechanics.

At this stage we hope that the reader is convinced that the local kinetic energy, the stress tensor, and related quantities are useful—and perhaps even essential—for studying the origins, properties, and presence of the covalent bonds between atoms in chemical systems. But what should be made of the inherent ambiguity in the definition of the local kinetic energy and, similarly, the local stress tensor? It seems dangerous to base one’s understanding of chemical bonding on such shaky ground. The goal of this paper is to characterize the ambiguity in the local kinetic energy in a more precise way than has heretofore appeared in the literature. Since the ambiguity in the local kinetic energy can be deduced from the ambiguity in the local stress tensor (cf. Eq. (5)), we will first summarize the literature on ambiguity in the stress tensor. In particular, we will show that Eq. (7) encapsulates the family of stress tensors that arises in the work of Godfrey⁷⁷ and Rogers and Rappe.⁸¹ Based on these considerations, the local kinetic energy is ambiguous by at most an additive multiple of the Laplacian of the electron density. However, a more general formulation, based on quasiprobability distribution functions, is also consistent with classical reasoning. This extends the family of allowed local kinetic energies beyond the “Laplacian family” in Eq. (11). This approach is presented in section IV, where some examples of these “strange” local kinetic energies are presented. Other “strange” forms of the local kinetic energy arise naturally from the virial theorem in density-functional theory (DFT). The DFT-based approach is sketched in section V. Section VI concludes our analysis and discusses the implications of our findings. Readers who do not wish to scrutinize the mathematical details may safely skip to the last subsections in III, IV, and V, where a capsule-summary of the mathematical analysis, along with a discussion of practical implications, is provided.

VII.III. Local Kinetic Energy from the Local Electronic Stress Tensor

A. Forms of the Local Electronic Stress Tensor

Within the physics literature, what chemists call the local electronic stress tensor is usually called the momentum flux density. The ambiguity in the momentum flux density is well-known, and it is usually analyzed by first proposing an explicit general form for the flux density, and then deriving constraints on that form using physical arguments. Probably the most lucid discussion of the key issue is by Godfrey.⁷⁷ Adding a divergence-free quantity to the stress tensor does not change any of the physical properties of the system. Therefore, the following alteration of the stress tensor may be made without any real effect,

$$\bar{\sigma}(\mathbf{r}) \rightarrow \bar{\sigma}(\mathbf{r}) + \vec{G}(\mathbf{r}) \quad (12)$$

$$0 = \nabla \cdot \vec{G}(\mathbf{r}) \quad (13)$$

The arbitrariness of $\vec{G}(\mathbf{r})$ is analogous to the freedom to choose a convenient gauge in classical electrodynamics.

The divergence criterion is not very restrictive, of course, so Godfrey *assumes* that the “reasonable” choices for $\vec{G}(\mathbf{r})$ will contain the second, and only the second, derivatives of the wavefunction. As he notes,⁷⁷ “the classical quantity is a real sum of single-particle dynamical variables, and $\bar{\sigma}_{ij}(\mathbf{r})$ should reflect this by being a Hermitian single-particle operator, involving only the field operators and their spatial derivatives.” This leads Godfrey to the form

$$\vec{G}(\mathbf{r}) = \left[\left(A \left(\frac{\partial^2}{\partial r_i \partial r_j} + \frac{\partial^2}{\partial r_i' \partial r_j'} \right) + B \left(\frac{\partial^2}{\partial r_i \partial r_j'} + \frac{\partial^2}{\partial r_i' \partial r_j} \right) \right) + \delta_{ij} \left(C \sum_{k=1}^3 \frac{\partial^2}{\partial r_k \partial r_k'} + D \sum_{k=1}^3 \left(\frac{\partial^2}{\partial r_k^2} + \frac{\partial^2}{\partial (r_k')^2} \right) \right) \right] \gamma(\mathbf{r}, \mathbf{r}') \Big|_{\mathbf{r}'=\mathbf{r}} \quad (14)$$

where A , B , C , and D are arbitrary real constants. The requirement that the divergence vanish (Eq. (13)) reveals that, in effect, only one of these quantities may be varied independently. Ensuring that the stress tensor responds sensibly upon Galilean transformations of the wavefunction, $\Psi(\boldsymbol{\tau}) \rightarrow \Psi(\boldsymbol{\tau}) e^{i\mathbf{P} \cdot \boldsymbol{\tau}}$ further specifies that A be real. Godfrey then obtains the form⁷⁷

$$\tilde{G}(\mathbf{r}) = A \left(\frac{\partial^2}{\partial r_i \partial r_j} - \delta_{ij} \left(\sum_{k=1}^3 \frac{\partial^2}{\partial r_k^2} \right) \right) \gamma(\mathbf{r}, \mathbf{r}) \quad (15)$$

This can be rewritten in terms of the electron density, $\rho(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r})$, as

$$\tilde{G}(\mathbf{r}) = A \left(\frac{\partial^2}{\partial r_i \partial r_j} - \delta_{ij} \left(\sum_{k=1}^3 \frac{\partial^2}{\partial r_k^2} \right) \right) \rho(\mathbf{r}). \quad (16)$$

Godfrey proposes that Eq. (15) can be added to the Schrödinger-Pauli-Epstein stress-tensor form, $\tilde{\sigma}_{1/2,0}(\mathbf{r})$, from Eq. (7) to obtain a family of “reasonable” electronic stress tensors,

$$\tilde{\sigma}_A^{(G)}(\mathbf{r}) = -\frac{1}{4} \left[\begin{array}{c} \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r_j} - \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r_j} - \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} \\ -4A \left(\frac{\partial^2 \rho(\mathbf{r})}{\partial r_i \partial r_j} - \delta_{ij} \nabla^2 \rho(\mathbf{r}) \right) \end{array} \right]_{\mathbf{r}'=\mathbf{r}} \quad (17)$$

Rogers and Rappe obtain a similar conclusion by a more complicated argument. The idea behind their analysis is that the stress tensor should behave appropriately in the limit of a flat, generally Euclidean, space.⁸¹ To analyze this limit, they *assume* that all “reasonable” stress tensors in a general curved space can be written in the form:

$$\tilde{\sigma}(\mathbf{r}) = \frac{1}{2} \sqrt{|\tilde{g}|} \left[\left(\begin{array}{c} \left[\tilde{g}^{-1} \right]_{ij} \frac{\partial^2}{\partial r_i \partial r'_j} + bR \\ + c \sum_{k,l} \left(\left[\tilde{R}^{-1} \right]_{kl} \frac{\partial^2}{\partial r_k \partial r'_l} \right) \end{array} \right) \gamma(\mathbf{r}, \mathbf{r}') \right]_{\mathbf{r}'=\mathbf{r}} + \frac{1}{2} aR \sqrt{|g|} \quad (18)$$

Here g_{ij} denotes an elements of the metric tensor, $|g|$ is the determinant of \tilde{g} , and $[\tilde{g}^{-1}]_{ij}$ is an element of the inverse of the metric tensor. R is the scalar curvature of the space and \tilde{R} is the Ricci curvature. It may be argued that a is zero because otherwise the stress tensor of the vacuum is nonzero. The constant c is set to zero because otherwise the field theory would not be renormalizable. Finally they work through the math and obtain, for the Euclidean metric, a form of the ambiguity that coincides with Godfrey’s,⁸¹

$$-b \left(\frac{\partial^2}{\partial r_i \partial r_j} - \delta_{ij} \left(\sum_{k=1}^3 \frac{\partial^2}{\partial r_k^2} \right) \right) \rho(\mathbf{r}) \quad (19)$$

Rogers and Rappe propose adding this to the “positive-definite” stress tensor, $\tilde{\sigma}_{1,0}(\mathbf{r})$, obtaining an infinite family of stress tensors with the form:

$$\bar{\sigma}_b^{(RR)}(\mathbf{r}) = -\frac{1}{2} \left[\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r_j} + 2b \left(\frac{\partial^2 \rho(\mathbf{r})}{\partial r_i \partial r_j} - \delta_{ij} \nabla^2 \rho(\mathbf{r}) \right) \right]_{\mathbf{r}'=\mathbf{r}} \quad (20)$$

The Rogers-Rappe and Godfrey families do not have any common members, but they are both members of the two-parameter superfamily defined by Eq. (7). For example, the “positive definite” stress tensor ($\bar{\sigma}_0^{(RR)}(\mathbf{r}) = \bar{\sigma}_{0,0}(\mathbf{r})$) is not a member of the Godfrey family. Conversely, the Schrödinger-Pauli-Epstein stress tensor ($\bar{\sigma}_0^{(G)}(\mathbf{r}) = \bar{\sigma}_{1/2,0}(\mathbf{r})$) is not a member of the Rogers-Rappe family.

To clarify these differences, use the expansion of the one-electron reduced density matrix in terms of natural orbitals,

$$\gamma(\mathbf{r}, \mathbf{r}') = \sum_{p=1}^{\infty} n_p \phi_p^*(\mathbf{r}') \phi_p(\mathbf{r}), \quad (21)$$

to derive

$$\begin{aligned} \frac{\partial^2 \rho(\mathbf{r})}{\partial r_i \partial r_j} &= \sum_{p=1}^{\infty} n_p \frac{\partial^2 (\phi_p^*(\mathbf{r}') \phi_p(\mathbf{r}))}{\partial r_i \partial r_j} \\ &= \left[\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r_j} \right]_{\mathbf{r}'=\mathbf{r}}. \end{aligned} \quad (22)$$

Using this identity, the Godfrey and Rogers-Rappe families can then be rewritten as

$$\begin{aligned} \bar{\sigma}_A^{(G)}(\mathbf{r}) &= \frac{-1}{2} \left[\begin{aligned} & \left(\frac{1}{2} - 2A \right) \left(\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r_j} \right) \\ & + \left(\frac{1}{2} + 2A \right) \left(-\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r_j} - \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r'_j} \right) \\ & + \delta_{ij} (2A) \nabla^2 \rho(\mathbf{r}) \end{aligned} \right]_{\mathbf{r}'=\mathbf{r}} \\ &= \bar{\sigma}_{\frac{1}{2}-2A, 2A}(\mathbf{r}) \end{aligned} \quad (23)$$

and

$$\begin{aligned} \bar{\sigma}_b^{(RR)}(\mathbf{r}) &= \frac{-1}{2} \left[\begin{array}{l} (1+2b) \left(\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r'_j} + \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r_j} \right) \\ -2b \left(-\frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r_i \partial r_j} - \frac{\partial^2 \gamma(\mathbf{r}, \mathbf{r}')}{\partial r'_i \partial r'_j} \right) \\ -\delta_{ij} (2b) \nabla^2 \rho(\mathbf{r}) \end{array} \right]_{\mathbf{r}'=\mathbf{r}} \\ &= \bar{\sigma}_{1+2b, -2b}(\mathbf{r}) \end{aligned} \quad (24)$$

respectively. They arise as special cases of Eq. (7), which leads us to conclude that any stress tensor with the general form of Eq. (7) is reasonable.

B. Implications for the Ambiguity in the Local Kinetic Energy

By first *assuming* a general form for the local stress tensor, $\bar{\sigma}(\mathbf{r})$, and then requiring, based on physical arguments, that $\bar{\sigma}(\mathbf{r})$ be invariant to various coordinate transformations, Godfrey⁷⁷ and Rogers and Rappe⁸¹ derived two distinct families of stress tensors, Eqs. (17) and (20), respectively. It does not seem to have been appreciated that even though Godfrey and Rogers-Rappe characterize the essential ambiguity within the stress tensor in the same way (cf. Eqs. (16) and (19)), the families of stress tensors they define do not have any common members. Nor does it seem to be appreciated that the Rogers-Rappe family does not include the popular Schrödinger-Pauli-Epstein form that is, by far, the most commonly used form in the literature. Both the Godfrey and Rogers-Rappe families of electronic stress tensors are elements of the extended family of stress tensors we defined in Eq. (7) (cf. Eqs. (23) and (24)).

The local kinetic energy arises is simply minus one-half the trace of the electronic stress tensor, Eqs. (5). It is important to notice that even though the Godfrey and Rogers-Rappe families of stress tensors are different, the family of possible local kinetic energies they define are exactly the same,

$$\begin{aligned} t_A^{(G)}(\mathbf{r}) &= -\frac{1}{2} \text{Tr} \left[\bar{\sigma}_A^{(G)}(\mathbf{r}) \right] = t_+(\mathbf{r}) + \left(A - \frac{1}{2} \right) \nabla^2 \rho(\mathbf{r}) \\ &= t_{\frac{1}{2}+4A}(\mathbf{r}) \end{aligned} \quad (25)$$

$$\begin{aligned} t_A^{(RR)}(\mathbf{r}) &= -\frac{1}{2} \text{Tr} \left[\bar{\sigma}_A^{(RR)}(\mathbf{r}) \right] = t_+(\mathbf{r}) - b \nabla^2 \rho(\mathbf{r}) \\ &= t_{1-4b}(\mathbf{r}). \end{aligned} \quad (26)$$

That is, either the Godfrey or the Rogers-Rappe forms of the stress tensor suffice to generate the Laplacian family of local kinetic energies, Eq. (11). The extended

family of stress tensors we defined, Eq. (7), also generates the Laplacian family of local kinetic energies, Eq. (9).

Based on the analysis of the local electronic stress tensor in the physics literature, then, we might suppose that every reasonable definition for the local kinetic energy is an element of the Laplacian family, $t_\alpha(\mathbf{r})$. In the next section, however, we will argue that there are other forms for the local kinetic energy that are also consistent with physical intuition.

VII.IV. Local Kinetic Energy from the Quasiprobability Distributions

A. Forms of the Quasiprobability Distribution Function

The motivation for defining a local kinetic energy is that it allows us to extend our classical understanding to quantum systems. In a classical system, it is possible to observe both the momentum and the position of a particle, so the local kinetic energy has the simple expression,

$$t^{(cl)}(\mathbf{r}) = \iint \cdots \int \left(\sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2m_i} \right) F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{r}_1 d\mathbf{p}_1 \dots d\mathbf{r}_N d\mathbf{p}_N \quad (27)$$

where $F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N)$ is the phase-space probability distribution function representing the probability of observing the particles at the specified positions with the specified momenta. Any quantum-mechanical definition of the local kinetic energy that is consistent with the classical limit ($\hbar \rightarrow 0$) may be considered reasonable. Unfortunately, this specification is not very restrictive.

As first noted by Wigner, it is possible to use the wavefunction to derive a quasiprobability distribution function, $F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N)$. This quasiprobability distribution function obviously must be normalized,

$$1 = \iint \cdots \int F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{r}_1 d\mathbf{p}_1 \dots d\mathbf{r}_N d\mathbf{p}_N \quad (28)$$

and must reduce to the position-space and momentum-space distribution functions that are implied by the underlying wavefunction,

$$|\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)|^2 = \iint \cdots \int F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{p}_1 \dots d\mathbf{p}_N \quad (29)$$

$$|\Phi(\mathbf{p}_1, \dots, \mathbf{p}_N)|^2 = \iint \cdots \int F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{r}_1 \dots d\mathbf{r}_N \quad (30)$$

Of course, since the wavefunction depends on only $3N$ variables, but the quasiprobability distribution function depends on $6N$ variables, there are infinitely

many quasiprobability distribution functions that are consistent with constraints (28)-(30).

It is an exercise in Fourier analysis to demonstrate that the complete family of quasiprobability distributions that are consistent with Eqs. (28)-(30) is characterized by the formula⁹²

$$\begin{aligned}
 F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) = & \\
 \left(\frac{1}{2\pi}\right)^{6N} \iiint & \left[\prod_{n=1}^N \exp(-i\boldsymbol{\tau}_n \cdot \mathbf{p}_n) \exp(-i\boldsymbol{\theta}_n \cdot [\mathbf{r}_n - \mathbf{u}_n]) \right] \\
 & \times f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N) \Psi^* \left(\mathbf{u}_1 - \frac{\hbar}{2} \boldsymbol{\tau}_1, \dots, \mathbf{u}_N - \frac{\hbar}{2} \boldsymbol{\tau}_N \right) \\
 & \times \Psi \left(\mathbf{u}_1 + \frac{\hbar}{2} \boldsymbol{\tau}_1, \dots, \mathbf{u}_N + \frac{\hbar}{2} \boldsymbol{\tau}_N \right) \Big] d\mathbf{u}_1 d\boldsymbol{\theta}_1 d\boldsymbol{\tau}_1 \dots d\mathbf{u}_N d\boldsymbol{\theta}_N d\boldsymbol{\tau}_N
 \end{aligned} \tag{31}$$

with

$$\begin{aligned}
 1 &= f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1 = \mathbf{0}, \dots, \boldsymbol{\tau}_N = \mathbf{0}) \\
 1 &= f(\boldsymbol{\theta}_1 = \mathbf{0}, \dots, \boldsymbol{\theta}_N = \mathbf{0}; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N)
 \end{aligned} \tag{32}$$

$$f^*(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N) = f(-\boldsymbol{\theta}_1, \dots, -\boldsymbol{\theta}_N; -\boldsymbol{\tau}_1, \dots, -\boldsymbol{\tau}_N) \tag{33}$$

The last requirement ensures that the quasiprobability distribution function is real-valued. Every different choice of f defines a classical-quantum correspondence that is mathematically acceptable.⁹³ For example, the Wigner-Weyl correspondence is obtained by choosing^{45,94}

$$f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N) = 1 \tag{34}$$

and the Margenau-Hill correspondence arises from⁹⁵

$$f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N) = \sum_{n=1}^N \cos\left(\frac{\hbar}{2} \boldsymbol{\theta}_n \cdot \boldsymbol{\tau}_n\right). \tag{35}$$

There are innumerable many alternatives, with no obvious way to establish a preference between the various different specific realizations of the quantum-classical correspondence principle.

Using Eq. (31), we can derive all of the forms for the local kinetic energy that are consistent with classical reasoning. However, there are certain additional restrictions that we might reasonably impose. For example, because the classical local kinetic energy can be determined from the momentum of a single particle, it is reasonable to require the quantum-mechanical kinetic energy to be a one-particle property. This means that f cannot couple the $\boldsymbol{\theta}$'s and $\boldsymbol{\tau}$'s associated with different electrons together, suggesting the form

$$f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N) = \sum_{n=1}^N f_1(\boldsymbol{\theta}_n, \boldsymbol{\tau}_n). \tag{36}$$

This *assumption* severely culls the family of permissible quasiprobability distributions. This is also very convenient because it allows us to treat the one-electron reduced quasiprobability distribution function, which we can obtain from the one-electron reduced density matrix as,

$$F_1(\mathbf{r}, \mathbf{p}) = \left(\frac{1}{2\pi} \right)^6 \iiint e^{-i\tau \cdot \mathbf{p}} e^{-i\theta(\mathbf{r}-\mathbf{u})} f_1(\theta, \tau) \gamma(\mathbf{u} + \frac{\hbar}{2} \tau, \mathbf{u} - \frac{\hbar}{2} \tau) d\mathbf{u} d\theta d\tau \quad (37)$$

For convenience, we often use the natural orbital expansion for the density matrix (cf. Eq. (21)), so that Eq. (37) can be written in a form that resembles the one-electron problem,

$$F_1(\mathbf{r}, \mathbf{p}) = \sum_{p=1}^{\infty} n_p \left[\left(\frac{1}{2\pi} \right)^6 \iiint e^{-i\tau \cdot \mathbf{p}} e^{-i\theta(\mathbf{r}-\mathbf{u})} f_1(\theta, \tau) \phi_p^*(\mathbf{u} - \frac{\hbar}{2} \tau) \phi_p(\mathbf{u} + \frac{\hbar}{2} \tau) d\mathbf{u} d\theta d\tau \right]. \quad (38)$$

Because we will not use the N -electron quasiprobability distributions in the remainder of the paper, henceforth we will drop the subscripts on $F_1(\mathbf{r}, \mathbf{p})$ and $f_1(\theta, \tau)$.

B. Local Kinetic Energy from Quasiprobability Distribution Functions

Using Eq. (38), the local kinetic energy has the form:

$$t_{f(\theta, \tau)}(\mathbf{r}) = \sum_{p=1}^{\infty} n_p \left[\begin{aligned} & -\frac{\hbar^2}{8m} (\phi_p^*(\mathbf{r}) \nabla^2 \phi_p(\mathbf{r}) - \nabla \phi_p^*(\mathbf{r}) \cdot \nabla \phi_p(\mathbf{r})) + c.c. \\ & - \left(\frac{1}{2\pi} \right)^3 \iint e^{-i\theta(\mathbf{r}-\mathbf{u})} \left(\frac{1}{2m} |\phi_p(\mathbf{u})|^2 \right) [\nabla_{\tau}^2 f(\theta, \tau)]_{\tau=0} d\mathbf{u} d\theta \\ & - \left(\frac{1}{2\pi} \right)^3 \iint e^{-i\theta(\mathbf{r}-\mathbf{u})} [\nabla_{\tau} f(\theta, \tau)]_{\tau=0} \cdot i \mathbf{j}_p(\mathbf{u}) d\mathbf{u} d\theta \end{aligned} \right] \quad (39)$$

where

$$\mathbf{j}_p(\mathbf{r}) = \frac{\hbar}{2mi} (\phi_p^*(\mathbf{r}) \nabla \phi_p(\mathbf{r}) - \phi_p(\mathbf{r}) \nabla \phi_p^*(\mathbf{r})) \quad (40)$$

is the orbital current density and $c.c.$ denotes the complex conjugate. In this paper we restrict ourselves to stationary states, where the current density is zero.

Using Eq. (3),

$$t_{f(\theta, \tau)}(\mathbf{r}) = t_{\alpha=\frac{1}{2}}(\mathbf{r}) + \sum_{p=1}^{\infty} n_p \left[- \left(\frac{1}{2\pi} \right)^3 \iint e^{-i\theta(\mathbf{r}-\mathbf{u})} \left(\frac{1}{2m} |\phi_p(\mathbf{u})|^2 \right) [\nabla_{\tau}^2 f(\theta, \tau)]_{\tau=0} d\mathbf{u} d\theta \right] \quad (41)$$

Because the classical local kinetic energy is the sum of the kinetic energies in the component directions, $t_C(\mathbf{r}) = (p_x^2(\mathbf{r}) + p_y^2(\mathbf{r}) + p_z^2(\mathbf{r}))/2m$, it is reasonable to *assume* that $f(\boldsymbol{\theta}, \boldsymbol{\tau})$ (a) does not couple the different Cartesian directions and (b) treats all three Cartesian directions equivalently. I.e.,

$$f(\boldsymbol{\theta}, \boldsymbol{\tau}) = g(\theta_x, \tau_x) + g(\theta_y, \tau_y) + g(\theta_z, \tau_z) \quad (42)$$

Furthermore, to preserve the inherent symmetry between the momentum-space and the position-space description, it is reasonable to *assume* that

$$g(\theta_x, \tau_x) = g(\tau_x, \theta_x). \quad (43)$$

Finally, we will assume that $f(\boldsymbol{\theta}, \boldsymbol{\tau})$ is a normal function, and not a functional of the wavefunction (or natural orbital). This means that the quasiprobability distribution function will not always be nonnegative, complicating its classical interpretation. (As shown by Wigner^{46,96} and Cohen and Zaporovanny (via constructive proof),^{97,98} the quasiprobability distribution function can be forced to be nonnegative only when $f(\mathbf{r}, \mathbf{p})$ is wavefunction-dependent.) The (very complicated) analysis that is required when these assumptions are relaxed will be considered in a future paper. Some results pertaining to the use of nonnegative quasiprobability distribution functions to evaluate the local kinetic energy⁹⁹ and local operators in general¹⁰⁰ can be found in the work of Cohen.

With the assumption in Eq. (42), the local kinetic energy can be reexpressed as

$$t_{g(\boldsymbol{\theta}, \boldsymbol{\tau})}(\mathbf{r}) = t_{\alpha=\frac{1}{2}}(\mathbf{r}) + \sum_{p=1}^{\infty} n_p \left[-\sum_{k=1}^3 \left(\frac{1}{2\pi} \right)^3 \iint e^{-i\boldsymbol{\theta} \cdot (\mathbf{r}-\mathbf{u})} \left(\frac{1}{2m} |\phi_p(\mathbf{u})|^2 \right) \left[\frac{\partial^2 g(\theta_k, \tau_k)}{\partial \tau_k^2} \right]_{\tau=0} d\mathbf{u} d\boldsymbol{\theta} \right] \quad (44)$$

The second term vanishes in the special case where g is a constant or a linear function of $\boldsymbol{\tau}$.

Even with the assumptions we have made to obtain the simplified form in Eq. (44), there are a vast range of different forms for the local kinetic energy. Suppose, for simplicity, that the function g is analytic in θ , so that it can be represented by a Taylor series with the form

$$g(\theta, \tau) = \frac{1}{3N} + \sum_{l=1}^{\infty} h_l(i\tau)(i\theta)^l \quad (45)$$

with

$$h_l(0) = 0 \quad l = 1, 2, \dots \quad (46)$$

The dependence on $i = \sqrt{-1}$ in this expression is suggested by condition (33). The constant term in Eq. (45) and the constraint in Eq. (46) are required by the “initial

condition” in Eq. (32). Using the relationship between the Fourier transform and derivatives,

$$\left(\frac{1}{2\pi}\right) \int e^{-i\theta(r-u)} (i\theta)^l d\theta = (-1)^l \delta^{(l)}(r-u) \quad (47)$$

and the formula for the derivatives of the Dirac delta function,

$$\int G(u) \delta^{(l)}(r-u) du = \frac{d^l G(r)}{dr^l} \quad (48)$$

gives

$$\begin{aligned} t_{g(\theta,\tau)}(\mathbf{r}) &= t_{\alpha=\frac{1}{2}}(\mathbf{r}) + \sum_{p=1}^{\infty} n_p \left[\sum_{k=1}^3 \left(\frac{1}{2m} \sum_{l=1}^{\infty} (-1)^{l-1} h_l''(0) \sum_{k=1}^3 \frac{\partial^l |\phi_p(\mathbf{r})|^2}{\partial r_k^l} \right) \right] \\ &= t_{\alpha=\frac{1}{2}}(\mathbf{r}) + \sum_{l=1}^{\infty} \frac{(-1)^{l-1} h_l''(0)}{2m} \left(\frac{\partial^l}{\partial x^l} + \frac{\partial^l}{\partial y^l} + \frac{\partial^l}{\partial z^l} \right) \rho(\mathbf{r}) \end{aligned} \quad (49)$$

Obviously this family of local kinetic energy functionals is much larger than the “Laplacian family” from Eq. (11). The Laplacian family is regained if $h_l''(0) = 0$ for all $l \neq 2$. Otherwise this form is more general than the Laplacian family.

Still, there are very many choices for $g(\theta,\tau)$ that are associated with the Laplacian family of local kinetic energies. For example, consider the multiplicative form where g is any sufficiently integrable function depending only on the product $\theta h(\tau)$

$$g(\theta,\tau) = g(\theta h(\tau)), \quad (50)$$

where $h(\tau)$ is an analytic function of the form¹⁰¹

$$h(\tau) = \sum_{l=0}^{\infty} a_l \tau^l \quad a_2 = 0 \quad (51)$$

Using the chain rule for second derivatives,

$$\begin{aligned} &\left[\frac{\partial^2 g(\theta h(\tau))}{\partial \tau^2} \right]_{\tau=0} \\ &= \theta^2 \left[\frac{\partial^2 g(\theta h(\tau))}{\partial (\theta h(\tau))^2} \left(\frac{\partial h(\tau)}{\partial \tau} \right)^2 \right]_{\tau=0} + \theta \left[\frac{\partial g(\theta h(\tau))}{\partial (\theta h(\tau))} \frac{\partial^2 h(\tau)}{\partial \tau^2} \right]_{\tau=0} \end{aligned} \quad (52)$$

The second term vanishes because of Eq. (51), and so we may conclude that any g with the form given by Eqs. (50)-(51) belongs to the Laplacian family. Note that if one requires that g be symmetric in θ and τ (Eq. (43)), then the only choice for h

is $h(\tau) = \tau$. That the general form $g(\theta\tau)$ recovers the Laplacian family is a key result from Cohen's work on the local kinetic energy; this is a particularly important observation because all of the commonly used quasiprobability distributions are of this form.²²

The requirement that g be symmetric in θ and τ is not enough, however, to ensure that the local kinetic energy belongs to the Laplacian family. Suppose that g is well-behaved enough to be expanded in a Taylor series. Then all of the requirements (Eqs. (32), (33), and (43)) can be fulfilled by any g whose expansion looks like

$$g(\theta, \tau) = 1 + \sum_{k=1}^{\infty} \sum_{l=1}^k i^{k+l} a_{kl} (\theta^k \tau^l + \tau^k \theta^l) \quad a_{kl} \in \mathbb{R} \quad (53)$$

The form of the local kinetic energy is then determined by the derivative

$$\left[\frac{\partial^2 g(\theta, \tau)}{\partial \tau^2} \right]_{\tau=0} = \sum_{k=1}^{\infty} \sum_{l=1}^k 2i^{k+l} a_{kl} (\delta_{l2} \theta^k + \delta_{k2} \theta^l) \quad (54)$$

It is only the case $k=l$ that gives the Laplacian family. That is, the only analytic functions, $g(\theta, \tau)$, that generate a quasiprobability distribution functions that (i) satisfy the constraints (32) and (33), (ii) are consistent with the assumptions (36), (42), and (43), and (iii) always give a local kinetic energy with the Laplacian form have the special multiplicative form considered by Cohen, $g(\theta\tau)$. If $g(\theta, \tau) \neq g(\theta\tau)$, yet the key assumptions and constraints are still satisfied, then the local kinetic energy will not have the Laplacian form because, if any of the $k \neq l$ terms do not vanish, then there will be a contribution to the local kinetic energy from a term of the form $2i^{k+2} a_{k2} \theta^k$, with $k \neq 2$. This produces a non-Laplacian term in the local kinetic energy expression.

Probably the simplest example of a "reasonable" non-Laplacian local kinetic energy arises from choosing

$$g(\theta, \tau) = 1 + i^3 (\theta^2 \tau + \tau^2 \theta) \quad (55)$$

gives

$$t_{(55)}(\mathbf{r}) = t_{\alpha=\frac{1}{2}}(\mathbf{r}) + \frac{1}{m} (\mathbf{1} \cdot \nabla \rho(\mathbf{r})) \quad (56)$$

where $\mathbf{1}$ is the constant vector containing one's. A more general, but related, example is:

$$g(\theta, \tau) = 1 + i^{(l+2)} (\theta^2 \tau^l + \tau^2 \theta^l) \quad (57)$$

which gives:

$$t_{(57)}(\mathbf{r}) = t_{\alpha=\frac{1}{2}}(\mathbf{r}) + \frac{\left((-1)^{l-1} - \delta_{l2}\right)}{m} \left(\nabla \cdot \left[\frac{\partial^{l-1} \rho(\mathbf{r})}{\partial x^{l-1}}, \frac{\partial^{l-1} \rho(\mathbf{r})}{\partial y^{l-1}}, \frac{\partial^{l-1} \rho(\mathbf{r})}{\partial z^{l-1}} \right] \right) \quad (58)$$

This is a non-Laplacian form of the local kinetic energy except when $l=2$.

There are many, more complicated, examples. One that is particular interesting is

$$g(\theta, \tau) = 1 + (\cos(a\theta) - 1)(\cos(a\tau) - 1), \quad (59)$$

which causes finite-difference approximations to $\nabla^2 \rho(\mathbf{r})$ to enter into the local kinetic energy,

$$t_{(59)}(\mathbf{r}) = t_{\alpha=\frac{1}{2}}(\mathbf{r}) - \frac{a^4}{4m} \left(\begin{array}{l} \frac{\rho(x+a, y, z) - 2\rho(x, y, z) + \rho(x-a, y, z)}{a^2} \\ + \frac{\rho(x, y+a, z) - 2\rho(x, y, z) + \rho(x, y-a, z)}{a^2} \\ + \frac{\rho(x, y, z+a) - 2\rho(x, y, z) + \rho(x, y, z-a)}{a^2} \end{array} \right) \quad (60)$$

C. Generalizations: The Multiplicative Alternative and the Electronic Stress Tensor

Instead of the additive forms in Eqs. (36) and (42), one can use the multiplicative form,

$$f(\theta_1 \dots \theta_N; \tau_1 \dots \tau_N) = \prod_{n=1}^N g(\theta_{n,x}, \tau_{n,x}) g(\theta_{n,y}, \tau_{n,y}) g(\theta_{n,z}, \tau_{n,z}) \quad (61)$$

or a sum of terms with this form. Because of the conditions on $g(\theta, \tau)$, however,

$$\begin{aligned} g(0, 1) &= g(1, 0) = 1 \\ g^*(\theta, \tau) &= g(-\theta, -\tau) \end{aligned} \quad (62)$$

the analysis in the previous section still holds, and the same restrictions on $g(\theta, \tau)$ apply.

Although we have not discussed the stress tensor in any detail, it is worth noting that the same considerations apply there. Only the form of the classical operator changes, giving (instead of Eq. (27)),

$$\begin{aligned} \bar{\sigma}^{(cl)}(\mathbf{r}) \\ = - \iint \dots \iint \left(\sum_{i=1}^N \delta(\mathbf{r}_i - \mathbf{r}) [\mathbf{p}_i \otimes \mathbf{p}_i] \right) F(\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) d\mathbf{r}_1 d\mathbf{p}_1 \dots d\mathbf{r}_N d\mathbf{p}_N \end{aligned} \quad (63)$$

D. Implications for the Ambiguity in the Local Kinetic Energy

To the extent that the purpose of the local kinetic energy in quantum mechanics is to facilitate the use of “classical” intuition to understand quantum phenomenon, then any form of the local kinetic energy that arises from a quasiprobability distribution function is acceptable. Even so, in the preceding analysis, we opted to make additional assumptions ((36), (42), and (43)) so that the forms of the local kinetic energy that we derived would have a simpler, more intuitive form. This led (cf. Eq. (49)) to the following family of local kinetic energies,

$$t_{\alpha, \{a_l\}}(\mathbf{r}) = t_{\alpha}(\mathbf{r}) + \sum_{\substack{l=1 \\ l \neq 2}}^{\infty} a_l \left(\frac{\partial^l \rho(\mathbf{r})}{\partial x^l} + \frac{\partial^l \rho(\mathbf{r})}{\partial y^l} + \frac{\partial^l \rho(\mathbf{r})}{\partial z^l} \right) \quad (64)$$

Every quasiprobability distribution function that arises from an analytic function, $f(\theta_1, \dots, \theta_N; \tau_1, \dots, \tau_N)$ (cf. Eq. (31)), and satisfies assumptions (33), (36), (42), and (43) leads to a local kinetic energy of this form. However, when

$$f(\theta_1 \dots \theta_N; \tau_1 \dots \tau_N) = \sum_{n=1}^N \left(g(\theta_{n,x}, \tau_{n,x}) + g(\theta_{n,y}, \tau_{n,y}) + g(\theta_{n,z}, \tau_{n,z}) \right) \quad (65)$$

or

$$f(\theta_1 \dots \theta_N; \tau_1 \dots \tau_N) = \prod_{n=1}^N g(\theta_{n,x}, \tau_{n,x}) g(\theta_{n,y}, \tau_{n,y}) g(\theta_{n,z}, \tau_{n,z}) \quad (66)$$

and $g(\theta, \tau)$ has the multiplicative form $g(\theta h(\tau))$, then the summation in Eq. (64) vanishes and the Laplacian family of kinetic energies is retained. This slightly extends a result of Cohen, who showed that the Laplacian family arises when $g(\theta, \tau) = g(\theta\tau)$. We showed that the converse is also true: subject to our restrictions on the form of g , every analytic function g that gives rise to a local kinetic energy in the Laplacian family has the simple multiplicative form $g(\theta\tau)$. Other forms of g (e.g., Eqs. (55), (57), and (59)) give non-Laplacian forms for the local kinetic energy.

There are other interesting forms of the local kinetic energy. For example, if one replaces Eq. (65) or Eq. (66) with the more general forms¹⁰²

$$f(\theta_1 \dots \theta_N; \tau_1 \dots \tau_N) = \sum_{n=1}^N f_1(\theta_n, \tau_n) \quad (67)$$

or

$$f(\theta_1 \dots \theta_N; \tau_1 \dots \tau_N) = \prod_{n=1}^N f_1(\theta_n, \tau_n), \quad (68)$$

then one can derive quasiprobability distributions corresponding to the following family of local kinetic energies,

$$t_{\alpha, \{b_l\}, \{c_m\}}(\mathbf{r}) = t_\alpha(\mathbf{r}) + \sum_{l=0}^{\infty} b_l (\mathbf{1} \cdot \nabla) (\nabla \cdot \nabla)^l \rho(\mathbf{r}) + \sum_{m=2}^{\infty} c_m (\nabla \cdot \nabla)^m \rho(\mathbf{r}) \quad (69)$$

Based on our analysis of quasiprobability distributions, the non-Laplacian forms of the local kinetic energy density cannot be excluded. Such forms are clearly consistent with classical reasoning (as they originate from a permissible quasiprobability distribution function). While one may argue that the Laplacian forms of the local kinetic energy are somehow “simpler” than the others, this relies upon the seemingly arbitrary decision that the $l=2$ terms in Eq. (49) are “simpler” than the $l=1$ terms. One may argue that the second-derivative terms are preferable because second derivatives already appear in $t_{1/2}(\mathbf{r})$ or because g has a particular simple form in this case, but these are at best “Occam’s razor” arguments (although they are appealing). There does not seem to be a physical or intuitive justification for excluding a non-Laplacian local kinetic energy like the one in Eqs. (60).

The reader might suppose, at this point, that every “reasonable” definition for the local kinetic energy is an element of the family, Eq. (64), (or alternatively Eq. (69)) that can be derived from the simple classes of quasiprobability distributions we have considered here. However, there is another family of local kinetic energies, commonly used in the density-functional theory, that is not derived from any simple quasiprobability distribution function.

VII.V. Analysis from Virial Theorems for Density Functionals

A. Local Kinetic Energy from DFT Virial Theorems

In density-functional theory, one may evaluate the non-interacting kinetic energy (i.e., the kinetic energy of the Kohn-Sham determinant) using the virial theorem,¹⁰³⁻¹⁰⁵

$$T_s[\rho] = \frac{1}{2} \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla v_s(\mathbf{r}) d\mathbf{r}, \quad (70)$$

where $v_s(\mathbf{r})$ denotes the Kohn-Sham potential.¹⁰⁶ Even though Eq. (70) is often used in both formal and practical studies of DFT,¹⁰⁷⁻¹¹⁴ the form

$$t_s(\mathbf{r}) = \frac{1}{2} \rho(\mathbf{r}) \mathbf{r} \cdot \nabla v_s(\mathbf{r}) \quad (71)$$

can be rejected as a permissible form for the local kinetic energy because it is not invariant to translation of the origin of the coordinate system. Integration by parts in Eq. (70) leads to the alternative form,

$$t_s(\mathbf{r}) = -\frac{1}{2}(v_s(\mathbf{r})\mathbf{r} \cdot \nabla \rho(\mathbf{r}) + 3v_s(\mathbf{r})\rho(\mathbf{r})) \quad (72)$$

but this form is not translationally invariant either.

To preserve translational invariance, Sim *et al.* use a standard trick,¹¹⁵ and derive a property density that has the same integral as the problematic term when integrated over all space.¹¹⁶ I.e., define a local noninteracting kinetic energy function, $t_s(\mathbf{r})$ such that

$$\int t_s(\mathbf{r}) d\mathbf{r} = \frac{1}{2} \int \mathbf{r} \cdot \rho(\mathbf{r}) \nabla v_s(\mathbf{r}) d\mathbf{r}. \quad (73)$$

Integration by parts reveals that

$$\begin{aligned} & \iiint t_s(x, y, z) dx dy dz \\ &= \frac{1}{3} \left[\iint x \frac{\partial t_s(x, y, z)}{\partial x} dy dz + \iint y \frac{\partial t_s(x, y, z)}{\partial y} dx dz + \iint z \frac{\partial t_s(x, y, z)}{\partial z} dx dy \right]_{-\infty}^{\infty} \\ & - \frac{1}{3} \left(\iiint \left(x \frac{\partial t_s(x, y, z)}{\partial x} + y \frac{\partial t_s(x, y, z)}{\partial y} + z \frac{\partial t_s(x, y, z)}{\partial z} \right) dx dy dz \right) \end{aligned} \quad (74)$$

Assuming that $t_s(\mathbf{r})$ decays rapidly asymptotically, this expression simplifies to

$$\iiint t_s(x, y, z) dx dy dz = \int \mathbf{r} \cdot \frac{-1}{3} \nabla t_s(\mathbf{r}) d\mathbf{r} \quad (75)$$

This allows us to write

$$\frac{1}{2} \rho(\mathbf{r}) \nabla v_s(\mathbf{r}) = \frac{-1}{3} \nabla t_s(\mathbf{r}) + \nabla \times \mathbf{a}(\mathbf{r}), \quad (76)$$

where $\mathbf{a}(\mathbf{r})$ is an arbitrary vector because

$$\int \mathbf{r} \cdot \nabla \times \mathbf{a}(\mathbf{r}) d\mathbf{r} = - \int \mathbf{a}(\mathbf{r}) \cdot \nabla \times \mathbf{r} d\mathbf{r} = \int 0 d\mathbf{r} = 0. \quad (77)$$

Now we need to determine $t_s(\mathbf{r})$. To do this, take the divergence of both sides,¹¹⁶

$$\begin{aligned} \frac{1}{2} \nabla \cdot \rho(\mathbf{r}) \nabla v_s(\mathbf{r}) &= \frac{-1}{3} \nabla \cdot \nabla t_s(\mathbf{r}) + \nabla \cdot \nabla \times \mathbf{a}(\mathbf{r}) \\ &= \frac{-1}{3} \nabla^2 t_s(\mathbf{r}) \end{aligned} \quad (78)$$

which gives a Poisson equation for $t_s(\mathbf{r})$ with the solution

$$t_s(\mathbf{r}) = \frac{3}{8\pi} \int \frac{\nabla' \cdot \rho(\mathbf{r}') \nabla' v_s(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (79)$$

Significantly, the same expression for the local noninteracting kinetic energy is obtained if one removes the translational dependence starting from Eq. (72), instead of Eq. (71).

There are several ways to extend this analysis to the total noninteracting kinetic energy. One can use one of the virial formulae for the correlation-kinetic energy,^{104,117}

$$T_c = - \int [\rho(\mathbf{r})\mathbf{r} \cdot \nabla v_{xc}(\mathbf{r}) + \varepsilon_{xc}(\mathbf{r})] d\mathbf{r} \quad (80)$$

$$T_c = \int [-\varepsilon_{xc}(\mathbf{r}) + 3\rho(\mathbf{r})v_{xc}(\mathbf{r}) + v_{xc}(\mathbf{r})(\mathbf{r} \cdot \nabla \rho(\mathbf{r}))] d\mathbf{r}, \quad (81)$$

where $v_{xc}(\mathbf{r})$ and $\varepsilon_{xc}(\mathbf{r})$ are the exchange-correlation potential and the exchange-correlation energy density, respectively. Although Eq. (81) has been used for computing the correlation-kinetic energy over subspaces, it is not translationally invariant either.¹¹⁸ This can be fixed using the same method as before, giving the form

$$t_{\varepsilon_{xc}}(\mathbf{r}) = \frac{3}{4\pi} \int \frac{\nabla' \cdot \rho(\mathbf{r}') \nabla' (\frac{1}{2}v_s(\mathbf{r}') - v_{xc}(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - \varepsilon_{xc}(\mathbf{r}) \quad (82)$$

Since there is also an ambiguity in the exchange-correlation energy density, however,^{115,119} it is more appealing to instead use the coordinate-scaling identities^{115,120}

$$\tilde{v}_{xc}(\mathbf{r}) = \int_0^1 \frac{v_{xc} \left[\alpha^3 \rho(\alpha x, \alpha y, \alpha z); \frac{x}{\alpha}, \frac{y}{\alpha}, \frac{z}{\alpha} \right]}{\alpha} d\alpha \quad (83)$$

$$T_c = - \int \rho(\mathbf{r})\mathbf{r} \cdot \nabla (v_{xc}(\mathbf{r}) - \tilde{v}_{xc}(\mathbf{r})) d\mathbf{r} \quad (84)$$

to obtain

$$t_{\text{DFT}}(\mathbf{r}) = \frac{3}{4\pi} \int \frac{\nabla' \cdot \rho(\mathbf{r}') \nabla' (\frac{1}{2}v_s(\mathbf{r}') + \tilde{v}_{xc}(\mathbf{r}') - v_{xc}(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (85)$$

B. Implications for the Ambiguity in the Local Kinetic Energy

Some of the most powerful tools for computing, and analyzing, the kinetic energy in density-functional theory are the virial expressions, (70), (80), (81), and (84), which allow us to compute the kinetic energy by integrating over the various potentials that appear in DFT. These expressions are not suitable for deriving a local kinetic energy because the resulting forms are not translationally invariant (cf. Eqs. (71) and (72)). Sim *et al.* showed how to remove this ambiguity, expressing the local noninteracting kinetic energy as the solution to a Poisson Equation (Eq. (78))¹¹⁶ which we then (trivially) solved, giving an explicit expression for the local noninteracting kinetic energy, Eq. (79). The same argument allows us to obtain explicit expressions for the total local kinetic energy, either Eq. (82) or (better, but more difficult) Eq. (85).

From the viewpoint of DFT, the local noninteracting kinetic energy expression (Eq. (79)) and the local total kinetic energy expressions (Eqs. (82) and (85)) are very reasonable expressions for the local kinetic energy. However, $t_s(\mathbf{r})$ is definitely not of the Laplacian form (Eq. (11)), and it is readily apparent that, for approximate exchange-correlation functionals, the local total kinetic energy expressions are not of the Laplacian form either.¹²¹ Moreover, the DFT-based $t(\mathbf{r})$ correspond to local kinetic energies might not arise quasiprobability distributions either. Based on the special cases derived by Sim *et al.*, if there is a quasiprobability distribution associated with DFT-based $t(\mathbf{r})$, then it corresponds to a wavefunction-dependent form of $f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N)$. Certainly the DFT-based local kinetic energies do not arise from the simple quasiprobability distributions that arise from Eqs. (65) and (66). However, although the authors cannot prove it, they postulate that the DFT-based $t(\mathbf{r})$ are consistent with a quasiprobability distribution function formulation, albeit an exceedingly complicated one in which $f(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N; \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_N)$ is an extremely complicated functional of the wavefunction.¹²²

VII.VI. Discussion

How ambiguous is the local kinetic energy? On the one hand, it is very ambiguous: even the most restrictive forms that we derived (based on invariance properties of the electronic stress tensor, section III) gives rise to an infinite family of local kinetic energies, namely those with the Laplacian form,

$$t_\alpha(\mathbf{r}) = t_+(\mathbf{r}) + \left(\frac{\alpha-1}{4}\right) \nabla^2 \rho(\mathbf{r}), \quad (86)$$

where $t_+(\mathbf{r})$ is the positive-definite local kinetic energy (Eq. (1)) preferred by many authors.^{8,19} One may argue, however, that any local kinetic energy that arises from a quasiprobability distribution function is physically reasonable because it entirely consistent with classical intuition. But even with additional assumptions (Eqs. (33), (36), (42), and (43)), the family of local kinetic energies that arises from quasiprobability distributions is much larger than the Laplacian family, cf. Eq. (64). However, there are eminently reasonable expressions for the local kinetic energy that arise in density-functional theory (section V) that do not belong to this family of local kinetic energies, and which might not be derivable at all from a quasiprobability distribution.

The authors cannot think of any physical, or intuitive justification for excluding local kinetic energies from outside the Laplacian family although one may, following Godfrey,⁷⁷ simply *assume* that the local kinetic energy should contain only second derivatives of the wavefunction, and then discard all other

forms, including most of those derived from quasiprobability distributions and all of those derived from DFT. This Occam's razor argument is appealing at a practical level: the conceptual utility of various members of the Laplacian family (especially the Schrödinger form, $t_{\text{Sch}}(\mathbf{r}) = t_0(\mathbf{r})$, the positive-definite form, $t_+(\mathbf{r}) = t_1(\mathbf{r})$, and the average between them, $t_{1/2}(\mathbf{r})$) is well-established.^{8,9,20,25,26,39-43,55,83} The general utility of the Laplacian family of local kinetic energies for chemical interpretation may be attributed to the fact that $t_+(\mathbf{r})$ and $\nabla^2\rho(r)$ both contain, by themselves, information about electron pairing and chemical bonding. For example, $t_+(\mathbf{r})$ is the key ingredient in the localized orbital locator³⁹ and the nighness¹⁹. The Laplacian of the electron density is a commonly used identifier of electron shells in the quantum theory of atoms in molecules (QTAIM), and its utility in valence-electron-pair-repulsion theory (VSEPR) is well-established.^{20,58,83,123-132} It is unsurprising, then, that linear combinations of $t_+(\mathbf{r})$ and $\nabla^2\rho(r)$ are useful for interpreting chemical bonding patterns. By contrast, the conceptual utility of more general forms of the local kinetic energy is at best unproven, and at worst exceedingly unlikely. (Consider, for example, how uninformative the local kinetic energy in Eq. (60) will be when the parameter a is large.)

At an intuitive or physical level, there is nothing that requires that $t(\mathbf{r})$ should be conceptually useful, and so the more general forms that arise from quasiprobability distributions or DFT-based arguments cannot be immediately discarded. While we recommend using Eq. (86) and the associated family of local electronic stress tensors, Eq. (7), for conceptual work, it must be kept in mind that the selection of these forms is based on *ad hoc assumptions* about the form of the electronic stress tensor (section III) or the quasiprobability distribution function (section IV).

The wealth of local kinetic energy forms proves particularly problematic in QTAIM. One derivation of QTAIM is based on the precept that the kinetic energy of an atom in a molecule (AIM) should be the same, no matter what form of local kinetic energy one uses.^{20,21,85,86} By choosing AIM to be bounded by "zero-flux" surfaces of the electron density, the kinetic energy of an AIM is the same for every local kinetic energy of the Laplacian form (cf. Eq. (10)). In general, however, *different* AIM kinetic energies are obtained if one uses the local kinetic energy derived from a more general quasiprobability distribution (e.g., Eqs. (56), (58), or (60)) function or the local kinetic energy derived based on the virial theorem in DFT. If one chose a different family of local kinetic energies (based, for example, on a different subclass of the local kinetic energies defined in Eq. (64)), one would require a different definition of molecular subsystems in

order to maintain the property that the subsystem kinetic energy is the same for every density in the subclass. For example, the small family of local kinetic energies defined by

$$t_a^{(l)}(\mathbf{r}) = t_+(\mathbf{r}) + a \left(\nabla \cdot \left[\frac{\partial^{l-1} \rho(\mathbf{r})}{\partial x^{l-1}}, \frac{\partial^{l-1} \rho(\mathbf{r})}{\partial y^{l-1}}, \frac{\partial^{l-1} \rho(\mathbf{r})}{\partial z^{l-1}} \right] \right) \quad (87)$$

give rise to subsystems that are bounded by surfaces for which

$$0 = \mathbf{n} \cdot \left[\frac{\partial^{l-1} \rho(\mathbf{r})}{\partial x^{l-1}}, \frac{\partial^{l-1} \rho(\mathbf{r})}{\partial y^{l-1}}, \frac{\partial^{l-1} \rho(\mathbf{r})}{\partial z^{l-1}} \right] \quad (88)$$

where \mathbf{n} denotes the normal vector to the surface. ($l = 2$ is the conventional case.) Similarly, using a different family of local kinetic energies (this one based on Eq. (69)),

$$t_c^{(m)}(\mathbf{r}) = t_+(\mathbf{r}) + c (\nabla \cdot \nabla)^m \rho(\mathbf{r}) \quad (89)$$

gives rise to subsystems that are bounded by surfaces for which

$$0 = \mathbf{n} \cdot \nabla (\nabla \cdot \nabla)^{m-1} \rho(\mathbf{r}) \quad (90)$$

In the $m = 2$ case, this gives rise to subsystems that are bounded by zero-flux surfaces of the Laplacian of the electron density. Regardless of how one partitions the system, it seems that the kinetic energy of an atom in a molecule is not uniquely defined in quantum mechanics: for any choice of subsystem partitioning, one can always find two quasiprobability distribution functions that give different values for the regional kinetic energy.

An alternative pathway to deriving form of the open quantum subsystems in QTAIM relies upon the principle of stationary action.^{16,20,87-90} This approach has been often criticized on subtle mathematical grounds,¹³³⁻¹³⁵ but the proof can be considered valid at a level of mathematical rigor that most chemists will find reasonable.^{90,136-140} From the present perspective, the problem is that the ambiguity in the local kinetic energy induces a corresponding arbitrariness in the form of the Lagrangian.⁹¹ Different forms for the Lagrangian give rise to different forms for the quantum subsystems. (Indeed, even after one chooses a form for the Lagrangian, one can still choose different quantum subsystems by enforcing only the “net zero flux” condition.^{16,141,142,143,144}) One can argue, if one wishes, that choosing a Lagrangian like the following (based on Eq. (56))

$$\mathcal{L} = \frac{i\hbar}{2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t} \right) - \left(\frac{\hbar^2}{2m} (\bar{\nabla} \Psi^* \cdot \bar{\nabla} \Psi + \mathbf{1} \cdot \nabla \Psi^* \Psi) + V \Psi^* \Psi \right) \quad (91)$$

is “unfair,” but this Lagrangian is physically and intuitively reasonable insofar as it corresponds to a quasiprobability distribution function, and thus is consistent with the classical-quantum correspondence principle.

One implication of this analysis is that there are an infinite number of different QTAIMs, each with different appropriate conditions for the boundaries of the quantum subsystems, associated with different choices of the local kinetic energy. Most of these alternatives to the “conventional QTAIM” are probably not useful because they will not give atom-like regions, but some choices also define atomic subsystems. For example, if the value of the coefficients $\{a_i\}$ in Eqs. (64) are sufficiently small, the proper open quantum subsystems will differ only slightly from the conventional QTAIM regions. The more general kinetic energy formulation here also provides a way to extend the mathematics of QTAIM to non-atomic subsystems. A particularly interesting example is the $m=2$ case in Eq. (89), which provides a justification for treating the topological basins of the Laplacian as quantum open systems.

Acknowledgement: The authors wish to acknowledge the friendship and inspiration of Klaus Ruedenberg. Funding was provided by NSERC (Discovery grant to PWA, graduate fellowship to JSMA) and the Canada Research Chairs. Constructive comments and criticism from Profs. Shant Shahbazian, Patrick Bultinck, and W. H. Eugen Schwarz are acknowledged.

VII.VII. References

- (1) Hellmann, H. *Z. Phys.* **1933**, *35*, 180.
- (2) Ruedenberg, K. *Reviews of Modern Physics* **1962**, *34*, 326.
- (3) Goddard, W. A.; Wilson, C. W. *Theor. Chim. Act.* **1972**, *26*, 195.
- (4) Goddard, W. A.; Wilson, C. W. *Theor. Chim. Act.* **1972**, *26*, 211.
- (5) Kutzelnigg, W. *Angew. Chem. Int. Ed.* **1973**, *12*, 546.
- (6) Bader, R. F. W.; Preston, H. J. T. *Int. J. Quantum Chem.* **1969**, *3*, 327.
- (7) Feinberg, M. J.; Ruedenberg, K. *J. Chem. Phys.* **1971**, *55*, 5804.
- (8) Feinberg, M. J.; Ruedenberg, K. *J. Chem. Phys.* **1971**, *54*, 1495.
- (9) Feinberg, M. J.; Ruedenberg, K.; Mehler, E. L. *Adv. Quantum Chem.* **1970**, *5*, 27.
- (10) Ghosh, S. K.; Berkowitz, M.; Parr, R. G. *Proc. Natl. Acad. Sci.* **1984**, *81*, 8028.
- (11) Lombard, R. J.; Mas, D.; Moszkowski, S. A. *Journal of Physics G-Nuclear and Particle Physics* **1991**, *17*, 455.

(12) Lombard, R. J.; Moszkowski, S. A. *Nuovo Cimento Della Societa Italiana Di Fisica B-General Physics Relativity Astronomy and Mathematical Physics and Methods* **1994**, *109*, 1291.

(13) Equations (1) and (2) are related by partial integration, subject only to the vanishing of a surface term. It is intuitively clear that there are many other forms for the local kinetic energy that also reducing to Eq. (1) by partial integration, but which have more complicated forms for the vanishing surface term.

(14) Ziff, R. M.; Uhlenbeck, G. E.; Kac, M. *Physics Reports-Review Section of Physics Letters* **1977**, *32*, 169.

(15) Tal, Y.; Bader, R. F. W. *Int. J. Quantum Chem.* **1978**, 153.

(16) Bader, R. F. W.; Nguyendang, T. T. *Adv. Quantum Chem.* **1981**, *14*, 63.

(17) Yang, Z. Z.; Liu, S. B.; Wang, Y. A. *Chem. Phys. Lett.* **1996**, *258*, 30.

(18) Yang, Z. Z.; Wang, Y. A.; Liu, S. B. *Science in China Series B-Chemistry* **1998**, *41*, 174.

(19) Ayers, P. W.; Parr, R. G.; Nagy, A. *Int. J. Quantum Chem.* **2002**, *90*, 309.

(20) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Clarendon: Oxford, 1990.

(21) Bader, R. F. W.; Beddall, P. M. *J. Chem. Phys.* **1972**, *56*, 3320.

(22) Cohen, L. *J. Chem. Phys.* **1979**, *70*, 788.

(23) Real, C. C.; Muga, J. G.; Brouard, S. *American Journal of Physics* **1997**, *65*, 157.

(24) Muga, J. G.; Seidel, D.; Hegerfeldt, G. C. *J. Chem. Phys.* **2005**, *122*, 154106.

(25) Ruedenberg, K.; Schmidt, M. W. *J. Phys. Chem. A* **2009**, *113*, 1954.

(26) Tachibana, A. *J. Chem. Phys.* **2001**, *115*, 3497.

(27) Szarek, P.; Sueda, Y.; Tachibana, A. *J. Chem. Phys.* **2008**, *129*, 094102.

(28) Bitter, T.; Ruedenberg, K.; Schwarz, W. H. E. *J. Comput. Chem.* **2007**, *28*, 411.

(29) Ruedenberg, K.; Schmidt, M. W. *J. Comput. Chem.* **2007**, *28*, 391.

(30) Wang, S. G.; Qiu, Y. X.; Schwarz, W. H. E. *Chemistry-a European Journal* **2009**, *15*, 6032.

(31) Shu-Guang Wang, Yi-Xiang Qiu, W. H. Eugen Schwarz, submitted.

(32) Sears, S. B.; Parr, R. G.; Dinur, U. *Isr.J.Chem.* **1980**, *19*, 165.

(33) Nalewajski, R. F. *J. Math. Chem.* **2010**, *47*, 667.

- (34) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397.
- (35) Savin, A.; Becke, A. D.; Flad, J.; Nesper, R.; Preuss, H.; Vonschnering, H. G. *Angew. Chem.* **1991**, *30*, 409.
- (36) Savin, A.; Nesper, R.; Wengert, S.; Fassler, T. F. *Angew. Chem.* **1997**, *36*, 1809.
- (37) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683.
- (38) Ayers, P. W. *J. Chem. Sci.* **2005**, *117*, 441.
- (39) Schmider, H. L.; Becke, A. D. *Theochem-Journal of Molecular Structure* **2000**, *527*, 51.
- (40) Jacobsen, H. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **2008**, *86*, 695.
- (41) Jacobsen, H. *J. Comput. Chem.* **2009**, *30*, 1093.
- (42) Jacobsen, H. *Dalton Transactions* **2009**, 4252.
- (43) Jacobsen, H. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **2009**, *87*, 965.
- (44) Chattaraj, P. K.; Chamorro, E.; Fuentealba, P. *Chem. Phys. Lett.* **1999**, *314*, 114.
- (45) Wigner, E. *Phys. Rev.* **1932**, *40*, 749.
- (46) Wigner, E. P.; Yourgrau, W.; van der Merwe, A. In *Perspectives in Quantum Theory*; MIT: Cambridge, 1971, p 25.
- (47) Rozendaal, A.; Baerends, E. J. *Chem. Phys.* **1985**, *95*, 57.
- (48) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **2002**, *116*, 3184.
- (49) Bohorquez, H. J.; Boyd, R. J. *J. Chem. Phys.* **2008**, *129*, 024110.
- (50) Pauli, W. *Handbuch der physik*; Springer: Berlin, 1958.
- (51) Schrödinger, E. *Ann. Phys.* **1927**, *387*, 265.
- (52) Epstein, S. T. *J. Chem. Phys.* **1975**, *63*, 3573.
- (53) Holas, A.; March, N. H. *Phys. Rev. A* **1995**, *51*, 2040.
- (54) Bader, R. F. W. *J. Chem. Phys.* **1980**, *73*, 2871.
- (55) Tao, J. M.; Vignale, G.; Tokatly, I. V. *Phys. Rev. Lett.* **2008**, *100*, 206405.
- (56) Grafenstein, J.; Ziesche, P. *Phys. Rev. B* **1996**, *53*, 7143.
- (57) Morante, S.; Rossi, G. C.; Testa, M. *J. Chem. Phys.* **2006**, *125*, 034101.
- (58) Bader, R. F. W.; Essen, H. *J. Chem. Phys.* **1984**, *80*, 1943.
- (59) Hernandez-Trujillo, J.; Cortes-Guzman, F.; Fang, D. C.; Bader, R. F. W. *Faraday Discuss.* **2007**, *135*, 79.
- (60) Bader, R. F. W.; Austen, M. A. *J. Chem. Phys.* **1997**, *107*, 4271.
- (61) Bader, R. F. W.; Fang, D. C. *Journal of Chemical Theory and Computation* **2005**, *1*, 403.
- (62) Deb, B. M.; Bamzai, A. S. *Mol. Phys.* **1978**, *35*, 1349.
- (63) Deb, B. M.; Bamzai, A. S. *Mol. Phys.* **1979**, *38*, 2069.

- (64) Deb, B. M.; Ghosh, S. K. *Journal of Physics B-Atomic Molecular and Optical Physics* **1979**, *12*, 3857.
- (65) Bartolotti, L. J.; Parr, R. G. *J. Chem. Phys.* **1980**, *72*, 1593.
- (66) Tachibana, A. *Int. J. Quantum Chem.* **2004**, *100*, 981.
- (67) Tachibana, A. *J. Mol. Model.* **2005**, *11*, 301.
- (68) Szarek, P.; Tachibana, A. *J. Mol. Model.* **2007**, *13*, 651.
- (69) Ichikawa, K.; Tachibana, A. *Phys. Rev. A* **2009**, *80*, 062507.
- (70) Jenkins, S.; Heggie, M. I. *Journal of Physics-Condensed Matter* **2000**, *12*, 10325.
- (71) Jenkins, S.; Morrison, I. *Journal of Physics-Condensed Matter* **2001**, *13*, 9207.
- (72) Jenkins, S.; Kirk, S. R.; Cote, A. S.; Ross, D. K.; Morrison, I. *Can. J. Phys.* **2003**, *81*, 225.
- (73) Ayers, P. W.; Jenkins, S. *J. Chem. Phys.* **2009**, *130*, 154104.
- (74) Ziesche, P.; Lehmann, D. *Physica Status Solidi B-Basic Research* **1987**, *139*, 467.
- (75) Tokatly, I. V. *Phys. Rev. B* **2005**, *71*, 165104.
- (76) Tao, J. M.; Vignale, G.; Tokatly, I. V. *Phys. Rev. B* **2007**, *76*, 195126.
- (77) Godfrey, M. J. *Phys. Rev. B* **1988**, *37*, 10176.
- (78) Grafenstein, J.; Ziesche, P. *Phys. Rev. B* **1990**, *41*, 3245.
- (79) Godfrey, M. J. *Journal of Physics B-Atomic Molecular and Optical Physics* **1990**, *23*, 2427.
- (80) Ramprasad, R. *Journal of Physics-Condensed Matter* **2002**, *14*, 5497.
- (81) Rogers, C. L.; Rappe, A. M. *Phys. Rev. B* **2002**, *65*, 224117.
- (82) Maranganti, R.; Sharma, P.; Wheeler, L. *Journal of Aerospace Engineering* **2007**, *20*, 22.
- (83) Popelier, P. L. A. *Atoms in Molecules: An Introduction*; Pearson: Harlow, 2000.
- (84) Matta, C. F.; Bader, R. F. W. *J. Phys. Chem. A* **2006**, *110*, 6365.
- (85) Bader, R. F. W. *J. Phys. Chem. A* **2007**, *111*, 7966.
- (86) Nasertayoob, P.; Shahbazian, S. *Int. J. Quantum Chem.* **2009**, *109*, 726.
- (87) Bader, R. F. W.; Srebrenik, S.; Nguyendang, T. T. *J. Chem. Phys.* **1978**, *68*, 3680.
- (88) Bader, R. F. W. *Phys. Rev. B* **1994**, *49*, 13348.
- (89) Bader, R. F. W. In *Quantum theory of atoms in molecules: recent progress in theory and application*; Matta, C. F., Boyd, R. J., Eds.; Wiley-VCH: Weinheim, 2007, p 37.

- (90) Nasertayoob, P.; Shahbazian, S. *Int. J. Quantum Chem.* **2010**, *110*, 1188.
- (91) Cioslowski, J.; Karwowski, J. In *Fundamentals of molecular similarity*; Kluwer: Dordrecht, 2001, p 101.
- (92) Cohen, L. *J.Math.Phys* **1966**, *7*, 781.
- (93) Cohen, L. *Philosophy of Science* **1966**, *33*, 317.
- (94) Weyl, H. *Zeitschrift fur Physik* **1927**, *46*, 1.
- (95) Margenau, H.; Hill, R. N. *Progress in Theoretical Physics* **1961**, *26*, 722.
- (96) Groblicki, R. *Journal of Mathematical Physics* **1983**, *24*, 841.
- (97) Cohen, L.; Zaporovanny, Y. I. *J.Math.Phys* **1980**, *21*, 794.
- (98) Cohen, L. *Journal of Mathematical Physics* **1984**, *25*, 2402.
- (99) Cohen, L. *J. Chem. Phys.* **1984**, *80*, 4277.
- (100) Cohen, L. *Phys. Lett. A* **1996**, *212*, 315.
- (101) More generally, the result is true whenever the second derivative of h_x at zero, vanishes.
- (102) I.e., one now allows coupling between the different Cartesian directions in the local kinetic energy formula.
- (103) Levy, M. *Phys. Rev. A* **1982**, *26*, 1200.
- (104) Levy, M.; Perdew, J. P. *Phys. Rev. A* **1985**, *32*, 2010.
- (105) Ghosh, S. K.; Parr, R. G. *J. Chem. Phys.* **1985**, *82*, 3307.
- (106) Kohn, W.; Sham, L. J. *Phys.Rev.* **1965**, *140*, A1133.
- (107) Ou-Yang, H.; Levy, M. *Phys. Rev. A* **1990**, *42*, 155.
- (108) Nagy, A. *Phys. Rev. A* **1993**, *47*, 2715.
- (109) Parr, R. G.; Liu, S. B.; Kugler, A. A.; Nagy, A. *Phys. Rev. A* **1995**, *52*, 969.
- (110) Nagy, A.; Liu, S. B.; Parr, R. G. *Phys. Rev. A* **1999**, *59*, 3349.
- (111) Gal, T.; Nagy, A. *Journal of Molecular Structure-Theochem* **2000**, *501*, 167.
- (112) Ayers, P. W.; Lucks, J. B.; Parr, R. G. *Acta.Chim.Phys.Debricina* **2002**, *34-35*, 223.
- (113) Levy, M.; Ayers, P. W. *Phys. Rev. A* **2009**, *79*, 064504.
- (114) Ayers, P. W.; Rodriguez, J. I. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **2009**, *87*, 1540.
- (115) Burke, K.; Cruz, F. G.; Lam, K. C. *J. Chem. Phys.* **1998**, *109*, 8161.
- (116) Sim, E.; Larkin, J.; Burke, K.; Bock, C. W. *J. Chem. Phys.* **2003**, *118*, 8140.
- (117) Hui, O. Y.; Levy, M. *Phys. Rev. Lett.* **1990**, *65*, 1036.
- (118) Rodriguez, J. I.; Ayers, P. W.; Gotz, A. W.; Castillo-Alvarado, F. L. *J. Chem. Phys.* **2009**, *131*, 021101.

(119) Gaiduk, A. P.; Chulkov, S. K.; Staroverov, V. N. *Journal of Chemical Theory and Computation* **2009**, *5*, 699.

(120) van Leeuwen, R.; Baerends, E. J. *Phys. Rev. A* **1995**, *51*, 170.

(121) It seems exceedingly unlikely that the local total kinetic energy would have the Laplacian form for the exact exchange-correlation energy functional. However, as this functional is unknown in any useful explicit form, it seems impossible to prove this.

(122) This postulate is based on the incredible diversity of local kinetic energy forms that are attainable if f is allowed to be a functional of the wavefunction. This leads the authors to believe that virtually any well-behaved local kinetic energy function can be reverse-engineered to generate a quasiprobability distribution function.

(123) Bader, R. F. W.; Heard, G. L. *J. Chem. Phys.* **1999**, *111*, 8789.

(124) Malcolm, N. O. J.; Popelier, P. L. A. *Faraday Discuss.* **2003**, *124*, 353.

(125) Gillespie, R. J.; Popelier, P. L. A. *Chemical Bonding and Molecular Geometry*; Oxford: New York 2001.

(126) Sagar, R. P.; Ku, A. C. T.; Smith, V. H.; Simas, A. M. *J. Chem. Phys.* **1988**, *88*, 4367.

(127) Bader, R. F. W.; Gillespie, R. J.; Macdougall, P. J. *J. Am. Chem. Soc.* **1988**, *110*, 7329.

(128) Gillespie, R. J.; Bytheway, I.; Dewitte, R. S.; Bader, R. F. W. *Inorg. Chem.* **1994**, *33*, 2115.

(129) Bader, R. F. W.; Macdougall, P. J.; Lau, C. D. H. *J. Am. Chem. Soc.* **1984**, *106*, 1594.

(130) Shi, Z.; Boyd, R. J. *J. Chem. Phys.* **1988**, *88*, 4375.

(131) Kiewisch, K.; Eickerling, G.; Reiher, M.; Neugebauer, J. *J. Chem. Phys.* **2008**, *128*, 044114.

(132) Eickerling, G.; Reiher, M. *Journal of Chemical Theory and Computation* **2008**, *4*, 286.

(133) Cassam-Chenai, P.; Jayatilaka, D. *Theor. Chem. Acc.* **2001**, *105*, 213.

(134) Cassam-Chenai, P.; Jayatilaka, D. *Theor. Chem. Acc.* **2002**, *107*, 383.

(135) Cassam-Chenai, P. *J. Math. Chem.* **2002**, *31*, 145.

(136) Bader, R. F. W. *Theor. Chem. Acc.* **2002**, *107*, 381.

(137) Delle Site, L. *Theor. Chem. Acc.* **2002**, *107*, 378.

(138) Kryachko, E. S. *Theor. Chem. Acc.* **2002**, *107*, 375.

(139) Mohallem, J. R. *Theor. Chem. Acc.* **2002**, *107*, 372.

(140) Nasertayoob, P.; Shahbazian, S. *Int. J. Quantum Chem.* **2008**, *108*, 1477.

- (141) Luana, V.; Costales, A.; Pendas, A. M. *Phys. Rev. B* **1997**, *55*, 4285.
- (142) Pendas, A. M.; Costales, A.; Luana, V. *Phys. Rev. B* **1997**, *55*, 4275.
- (143) F. Heidarzadeh; S. Shahbazian, *Int. J. Quantum. Chem.* DOI: 10.1002/qua.22629.
- (144) F. Heidarzadeh; S. Shahbazian, submitted (2010).

Chapter VIII

Breaking the Curse of Dimension for the Electronic Schrödinger Equation with Functional Analysis*

*The content of this chapter in preparation to be submitted: **J. S. M. Anderson**, P. W. Ayers
“Breaking the curse of dimension for the electronic Schrödinger equation with functional
analysis”; *J. Chem. Phys.*

VIII.I. Statement of the Problem

This chapter addresses the computational intractability of the full-configuration interaction (FCI) method. Specifically, this chapter reformulates FCI to break the curse of dimensionality using recent work in mathematics and computer science, specifically the theory of information-based complexity. By exploiting the differentiability of molecular wavefunctions, one can design methods that achieve the same accuracy as full-configuration interaction, but with polynomial cost. One practical algorithm for achieving this is derived here; it is a selected-CI algorithm that is motivated by my previous work on many-electron numerical integration grids. A detailed prescription for identifying the most relevant Slater determinants in a CI calculation is provided.

VIII.II. Limited Configuration Interaction Calculations

The primary goal of computational quantum chemistry is to solve the Schrödinger equation accurately enough. When moderate accuracy of a few kcal/mol is needed, established methods based on mean-field methods like density-functional theory or Hartree-Fock, and corrections to them, suffice. When higher accuracy is required, or when a single Slater determinant is a very bad model for the exact wavefunction, the computational complexity of existing methods explodes. The computational cost of benchmark methods like released-node quantum Monte Carlo^{1,2} and full configuration interaction (FCI) grows as the factorial of the number of electrons; such approaches are practical only for small molecules.

There has been sustained interest in restricting the space of Slater determinants, thereby breaking through the factorial scaling of FCI. Existing approaches can be broadly characterized according to whether Slater determinants are selected by the user based on chemical insight, selected automatically by an algorithm, or selected by some combination of these two strategies. For example, in restricted active-space methods (including MRCISD, RASSCF, and related methods),³⁻⁹ the set of Slater determinants is pruned in a problem-specific way based on chemical insight and practical experience about which Slater determinants are most important for the problem. In the hands of an expert, such methods give excellent accuracy with reasonable computational cost.

In algorithm-based selection methods, additional Slater determinants, including higher and higher excitations, are sequentially added according to some procedural rules. The iterative CI method is one approach of this type,^{10,11} where one includes more-and-more highly excited Slater determinants in an iterative fashion. Other examples include the “deadwood elimination” configuration interaction method of Ivancic and Ruedenberg,¹²⁻¹⁴ the full configuration interaction Monte Carlo (FCI-MC) method of Thom and Alavi,¹⁵⁻¹⁸ and the configuration interaction by perturbative iterative selection (CIPIS) approach and related perturbative-selection methods.¹⁹⁻²² Although they are not usually thought of in this way, computational approaches based on Shepard’s linear combination of graphically contracted functions (LCGCF),²³⁻²⁶ White’s density matrix renormalization group (DMRG),²⁷⁻³⁷ and Changlani *et al.*’s correlator product states (CPS) are also of this type. LCGCF, DMRG, and CPS are similar to iterative CI because they use a very compact representation of a wavefunction that lies, however, within the space of Slater determinants that is spanned in a FCI calculation.^{33,38} Methods like iterative CI, LCGCF, DMRG, and CPS are probably the best ways to formulate CI for general systems.

Even algorithm-based selection techniques require some chemical insight, however, because obtaining good results requires an intelligent choice for the input parameters and, in most cases, also for the orbital active space. While the cost of computing the energy with these methods does not grow as a factorial of the number of one-electron states, the number of iterations required to find the ground-state energy might not be polynomial. In the context of DMRG, this is known as the problem of “certifying” that the DMRG algorithm converges to the electronic ground state.³⁹ More generally, this problem afflicts all methods that optimize the orbital basis (e.g., RASSCF), because their associated energy hypersurface has many local minima.^{40,41} Moreover, the number of local minima may grow at a non-polynomial rate, so obtaining a good solution with polynomial computational scaling requires that the starting guess lies in the vicinity of a “good” local minimum. The “tricks” chemists use to ensure successful convergence in these cases are usually based on chemical insight, not mathematical rules.

VIII.III. Motivation: Balancing Errors in Configuration Interaction

Like most of the aforementioned methods, the method proposed here is algorithmic: we will prune the basis set for the N -electron wavefunction based on rules, not chemical information. Unlike the preceding approaches however, the

Slater determinants are selected based on the functional-analytic properties of the N -electron wavefunction; this means that the important Slater determinants can be selected at the very beginning of the algorithm, and so there is no iterative selection of determinants. The resulting method is a very particular flavor of limited-CI, where the configurations are selected based on functional-analytic principles. The basic mathematical tools that we are using have been previously employed, to good effect, in the context of density-functional theory,^{42-44,45} the few-electron Schrödinger equation,⁴⁶⁻⁴⁸ and the nuclear Schrödinger equation.⁴⁹

The motivation for this idea is the recognition that every FCI calculation has an error due to basis set incompleteness. I.e., there are Slater determinants that are not included in the FCI calculation because of basis set truncation (e.g., the determinant represented later in Fig. 1c) that make a larger contribution to the ground-state wavefunction than other determinants (e.g., the determinant represented later in the chapter by Fig. 1b) that are included in the FCI expansion of the wavefunction.

The ideal CI method would include only those Slater determinants that have a larger overlap with the wavefunction than the most highly overlapping determinant that would be included in the FCI were the orbital basis set extended. To make the ideal-CI concept more concrete, imagine extending the orbital basis set to the complete basis set limit. The only determinants included in the ideal-CI method are those that overlap more with the exact wavefunction than the first neglected determinant. Therefore, the decision on whether to include a Slater determinant from the limited-basis-set FCI calculation, $\Phi_i \in \text{FCI}$, is based on the criterion

$$\Phi_i \in \text{ideal-CI} \subset \text{FCI} \quad \leftrightarrow \quad \left| \langle \Phi_i | \Psi_{\text{exact}} \rangle \right| > \max_{\Phi'_i \notin \text{FCI}} \left| \langle \Phi'_i | \Psi_{\text{exact}} \rangle \right| \quad (1)$$

Because of the high differentiability of the electronic wavefunction, the number of Slater determinants that satisfy the criterion in Eq. (1) grows only as a polynomial, and not as a factorial, of the number of electrons. Notice that the ideal-CI method,

$$E_{\text{ideal-CI}} \equiv \min_{\mathbf{c}_i} \frac{\langle \Psi_{\text{ideal-CI}}(\mathbf{c}_i) | \hat{H} | \Psi_{\text{ideal-CI}}(\mathbf{c}_i) \rangle}{\langle \Psi_{\text{ideal-CI}}(\mathbf{c}_i) | \Psi_{\text{ideal-CI}}(\mathbf{c}_i) \rangle} \quad (2)$$
$$\Psi_{\text{ideal CI}}(\mathbf{c}_i) \equiv \sum_{i \in \text{ideal-CI}} c_i \Phi_i$$

is not manifestly size extensive. The size-extensivity error does not compromise the calculation, however, because it is less than or equal to the basis-set-truncation error. Ideal-CI can be formulated in a way that is *size consistent*, however,

because the energy of a set of dissociated fragments will be equal to the sum of fragment energies if the basis set is built from fragment-localized orbitals.

The ideal-CI method described in Eqs. (1) and (2) requires knowledge of the exact wavefunction; it also requires considering all possible determinants in the complete basis set limit. This is not practicable. Fortunately, by analyzing the functional-analytic properties of the wavefunction, we can propose a method that converges to ideal-CI in the high-accuracy limit. Before delving into the details, however, it is useful to define some notation and foreshadow the key results.

We will use the K spatial orbitals that are obtained by diagonalizing a mean-field Hamiltonian for the system as a basis set. List these orbitals in order of increasing orbital energy,

$$\phi_1(\mathbf{r}), \phi_2(\mathbf{r}), \dots, \phi_K(\mathbf{r}) \quad (3)$$

$$\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_K \quad (4)$$

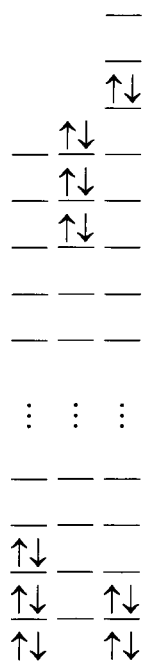
We denote an N -electron Slater determinant as

$$\Phi_i = \left| \phi_{i_{\alpha,1}} \alpha \quad \phi_{i_{\alpha,2}} \alpha \quad \dots \quad \phi_{i_{\alpha,N_\alpha}} \alpha \quad \phi_{i_{\beta,1}} \beta \quad \phi_{i_{\beta,2}} \beta \quad \dots \quad \phi_{i_{\beta,N_\beta}} \beta \right| \quad (5)$$

$$1 \leq i_{\sigma 1} < i_{\sigma 2} < \dots < i_{\sigma N_\sigma} \leq K \quad \sigma = \alpha, \beta \quad (6)$$

To make this notation clear, the indices for the Slater determinants depicted in Figure 1 are given as follows:

$$\begin{aligned} \mathbf{i}_{\text{Fig. 1a}} &= \left[1 \quad 2 \quad 3 \quad 1 \quad 2 \quad 3 \right]^T \\ \mathbf{i}_{\text{Fig. 1b}} &= \left[(K-2) \quad (K-1) \quad K \quad (K-2) \quad (K-1) \quad K \right]^T \\ \mathbf{i}_{\text{Fig. 1c}} &= \left[1 \quad 2 \quad (K+1) \quad 1 \quad 2 \quad (K+1) \right]^T \end{aligned} \quad (7)$$



(a) (b) (c)

Figure 1. Representation of Slater determinants used in a configuration interaction wavefunction. (a) The reference Slater determinant from, e.g., Hartree-Fock, in the truncated basis set. (b) A highly excited Slater determinant that is included in a full configuration interaction (FCI) calculation in the truncated basis set. (c) A Slater determinant included in FCI on an extended basis set that makes a larger contribution to the wavefunction than the Slater determinant in (b).

The FCI wavefunction can be written as:

$$\Psi_{\text{FCI}} = \sum_{\|\mathbf{i}\|_{\infty} \leq K} c_i \Phi_i \quad (8)$$

Recall that the ℓ^{∞} -norm of a vector is equal to the largest absolute value of a component in the vector,

$$\|\mathbf{i}\|_{\infty} = \max_{\substack{\sigma=\alpha,\beta \\ 1 \leq n \leq N_{\sigma}}} |i_{\sigma,n}| = \max_{\substack{\sigma=\alpha,\beta \\ 1 \leq n \leq N_{\sigma}}} i_{\sigma,n} \quad (9)$$

The proposed form for selected CI also depends on the ℓ^1 norm of a vector, which is

$$\|\mathbf{i}\|_1 = \sum_{\sigma=\alpha,\beta} \sum_{n=1}^N |i_{\sigma,n}| = i_{\alpha,1} + i_{\alpha,2} + \dots + i_{\alpha,N_{\alpha}} + i_{\beta,1} + i_{\beta,2} + \dots + i_{\beta,N_{\beta}}, \quad (10)$$

and has the general form,

$$\Psi_{\text{GKCI}} = \sum_{f_{\sigma\sigma'}(\mathbf{i}, T_{\sigma\sigma'}) < Q_{\sigma\sigma'}(K, T_{\sigma\sigma'})} c_i \Phi_i, \quad (11)$$

where

$$f_{\alpha\alpha}(\mathbf{i}, T) = \|\eta_{i_{\alpha}}\|_1 + T_{\alpha\alpha} \|\eta_{i_{\alpha}}\|_{\infty} \quad (12)$$

$$f_{\beta\beta}(\mathbf{i}, T) = \|\eta_{i_{\beta}}\|_1 + T_{\beta\beta} \|\eta_{i_{\beta}}\|_{\infty}$$

$$f_{\alpha\beta}(\mathbf{i}, T) = \|\eta_{i_{\alpha}}\|_1 + T_{\alpha\beta} \|\eta_{i_{\beta}}\|_1 + T_{\alpha\beta} \cdot \max(\|\eta_{i_{\alpha}}\|_{\infty}, \|\eta_{i_{\beta}}\|_{\infty}) \quad (13)$$

$$-1 < T_{\sigma\sigma'} < \infty$$

The components of the vector

$$\eta_{\mathbf{i}} = \left[\eta_{i_{\alpha,1}} \quad \eta_{i_{\alpha,2}} \quad \dots \quad \eta_{i_{\alpha,N_{\alpha}}} \quad \eta_{i_{\beta,1}} \quad \eta_{i_{\beta,2}} \quad \dots \quad \eta_{i_{\beta,N_{\beta}}} \right]^T \quad (14)$$

are the number of nodes in the respective orbitals. For example, if the thirteenth spatial orbital is an atomic $3p$ orbital, then $\eta_{13} = 2$. The number of nodes in the k^{th} spatial orbital can be approximated by solving the following equation for η_k

$$k = \frac{\Gamma(\eta_k + 3)}{6\Gamma(\eta_k + 1)}. \quad (15)$$

$\Gamma(x)$ is the Gamma Function.

This method reduces to FCI in the limit as $T \rightarrow +\infty$. As $T \rightarrow -1$, fewer and fewer configurations are used, and configurations in which the level of excitation is small, but the electrons are excited into high-lying orbitals, are preferred over configurations where the level of excitation is high, but the electrons are in low-lying orbitals. For $T = 0$, all Slater determinants in which the sum of the total number of orbital nodes is less than or equal to Q are included. As

pointed out by Avila and Carrington,⁴⁹ this choices of limited CI has been proven to be efficient for solving the vibrational problem;⁵⁰⁻⁵² in that context it is closely related to imposing an energy-cutoff on the product basis set.

We have used the general form in Eq. (12) for many-electron numerical quadrature grids, and our results in that context suggest that it is best to consider T a user-defined parameter and tune it to balance accuracy with computational cost.⁵³ On mathematical and practical grounds, it is optimal to choose $T < 0$.

Increasing Q increases the number of Slater determinants that are included in the method; small values of Q give very limited (and very fast) selected CI methods. Large values of Q give large and, for sufficiently large Q , exhaustive lists of Slater determinants. We recommend

$$\begin{aligned}
 Q_{\sigma\sigma}(K, T_{\sigma\sigma}) &= \sum_{i=1}^{N_{\sigma}} \eta_i - \underbrace{\max}_{1 \leq i \leq N_{\sigma}} \eta_i + (1 + T_{\sigma\sigma}) \eta_{K+1} \\
 Q_{\alpha\beta}(K) &= \sum_{\sigma=\alpha, \beta} \left(\sum_{i=1}^{N_{\sigma}} \eta_i - \underbrace{\max}_{1 \leq i \leq N_{\sigma}} \eta_i + \eta_{K+1} \right) \\
 &\quad + T_{\alpha\beta} \underbrace{\max}_{\sigma=\alpha, \beta} \left(\sum_{i=1}^{N_{\sigma}} \eta_i - \underbrace{\max}_{1 \leq i \leq N_{\sigma}} \eta_i + \eta_{K+1} \right)
 \end{aligned} \tag{16}$$

We believe that these are the largest values for Q that can be used before truncation of the orbital basis set introduces a larger error than the truncation of the CI expansion. When it is difficult to determine the number of nodes in the orbitals, one can approximate these formulae with

$$\begin{aligned}
 Q_{\sigma\sigma}(K, T_{\sigma\sigma}) &= \sum_{i=1}^{N_{\sigma}-1} \eta_i + (1 + T_{\sigma\sigma}) \eta_{K+1} \\
 Q_{\alpha\beta}(K) &= \sum_{\sigma=\alpha, \beta} \left(\sum_{i=1}^{N_{\sigma}-1} \eta_i + \eta_{K+1} \right) + T_{\alpha\beta} \underbrace{\max}_{\sigma=\alpha, \beta} \left(\sum_{i=1}^{N_{\sigma}-1} \eta_i + \eta_{K+1} \right)
 \end{aligned} \tag{17}$$

where η_k is approximated by solving Eq. (15).

We refer to this method as Griebel-Knapek CI by degrees (GK-CI), because it was Griebel and Knapek who first analyzed the cost function on which Eq. (12) is based^{54,55} and Eq. (12) is based on the idea that the number of nodes in an orbital can be used as an “effective polynomial degree.” The Griebel-Knapek cost function has been used by Griebel, Hamaekers, and Garcke in the context of the electronic Schrödinger equation, with promising results.^{46,56} The $T = 0$ case of this cost function was proposed by Smolyak,⁵⁷ and has proved useful for applications to density-functional theory⁴²⁻⁴⁴ and the nuclear Schrödinger equation.⁴⁹

In the next two sections, the essential mathematical background for understanding the GK-CI method will be presented. The motivation for this particular selected CI method will then be presented in detail in section VI, and analyzed in section VI.

VIII.IV. Background: Complexity Theory

A distinctive feature of the present approach is that it is based on the formalism of information-based computational complexity and tractability⁵⁸⁻⁶⁰ using the real-number model.⁶¹⁻⁶³ In this context, one writes the cost of the calculation as a function of the accuracy that is desired,

$$\text{cost} = f(\varepsilon). \quad (18)$$

Because one is using the real-number model, the “cost” is equal to the number of simple arithmetic operators, which we assume can be performed with infinite precision. In variational quantum Monte Carlo and grid-based methods, the cost is usually proportional to the number of integration/sampling points, M_{points} . Thus, the usual Monte Carlo expression for the accuracy,

$$\varepsilon = (M_{\text{points}})^{-1/2} \quad (19)$$

can be rewritten as

$$\text{cost}_{\text{MC}} \sim \varepsilon^{-2} \quad (20)$$

The notation in Eq. (20) means that the cost of a Monte Carlo calculation with accuracy ε is asymptotically proportional to $1/\varepsilon^2$. This is only an asymptotic result for the high-accuracy ($\varepsilon \rightarrow 0$) limit; computational prefactors and terms of higher order in $1/\varepsilon$ are not included. Equation (20) says that doubling the accuracy of a Monte-Carlo calculation quadruples the computational cost.

Equation (20) is the best possible result for general wavefunctions. For continuous wavefunctions, quasi-Monte Carlo sampling is theoretically more efficient,⁶⁴⁻⁶⁸ and can achieve

$$\text{cost}_{\text{quasi-MC}} \sim \varepsilon^{-1}. \quad (21)$$

For wavefunctions that are not only continuous but also have bounded s^{th} -order mixed derivatives, the best possible result is⁶⁹⁻⁷²

$$\text{cost} \sim \varepsilon^{-1/s}. \quad (22)$$

The results in Eqs. (21) and (22) are actually stronger than the Monte Carlo result because they pertain to the “worst case error,” while the error in Monte Carlo methods is the “average case error” (i.e., the expectation value of the error). Methods that are based on Eqs. (21) or (22) can certify the quality of result with mathematical rigor, not just provide probabilistic error bounds.

All of these results are purely mathematical, and they are only valid in the asymptotic limit $\varepsilon \rightarrow 0$. Additionally, computational prefactors have great practical implications, but they are not captured by these formulae. Thus, while quasi-Monte Carlo methods are superior to random Monte Carlo sampling in the limit of high accuracy (small ε limit), quasi-Monte Carlo methods might not be more efficient for the modest accuracy required in typical quantum chemistry applications. The goal of this paper is to chase the theoretical limits of computational efficiency by proposing a method that approaches optimal cost per unit accuracy, as expressed in in Eq. (22).

Established deterministic methods for solving the electronic Schrödinger equation are much less efficient than any of the results presented above. Consider, for example, the computational cost of straightforward numerical integration of the N -electron Schrödinger equation on a direct product grid with m points per dimension, $\text{cost} \propto M_{\text{points}} = m^{3N}$. If one uses the best possible one-dimensional integration formulae (many-dimensional Gaussian quadrature), the error is proportional to m^{-s} , presuming that the wavefunction has bounded derivatives of s^{th} order. One then has, for a direct product grid,

$$\text{cost}_{\text{direct prod. grid}} \propto \varepsilon^{-3N/s}. \quad (23)$$

or

$$\frac{\log(\text{cost}_{\text{direct prod. grid}})}{-\log \varepsilon} \propto \frac{N}{s}. \quad (24)$$

One may similarly argue that the cost of FCI in a K -orbital basis, with $K \gg N$, is

$$\frac{\log(\text{cost}_{\text{FCI}})}{-\log \varepsilon} \propto \frac{N}{s} \quad (25)$$

For such methods, the computational cost required to achieve a given accuracy, ε , grows exponentially with increasing dimension. Methods that suffer from this “curse of dimension” are said to be computationally intractable. By contrast, methods whose computational complexity does not grow with dimension (cf. Eqs. (20)-(22)) are said to be strongly tractable. Classical Monte Carlo methods are appealing because their cost/accuracy ratio does not depend on the dimensionality of the problem. The exploding computational cost of FCI with increasing electron number is captured in Eq. (25): the cost per unit accuracy ratio of FCI vanishes exponentially with increasing electron number.

VIII.V. Background: Convergence Rate of Basis Set Expansions

Consider a one-dimensional wavefunction, $\psi(x)$. By choosing a suitable set of w -orthogonal polynomials,

$$\delta_{ij} = \int_a^b P_i(x) P_j(x) w(x) dx \quad (26)$$

we can define an appropriate orthonormal basis set

$$\chi_i(x) = P_i(x) \sqrt{w(x)} \quad (27)$$

for expanding the wavefunction,

$$\psi(x) = \sum_{k=0}^{\infty} a_k \chi_k(x) \quad (28)$$

The k^{th} orthogonal polynomial, $P_k(x)$, has degree k and k nodes. Therefore, the basis function $\chi_k(x)$ also has k nodes, and it is reasonable to say that $\chi_k(x)$ has “effective polynomial degree” k .

If $\psi(x)$ is s times differentiable, that is, if

$$\left\| \frac{d^s \psi(x)}{dx^s} \right\|_{\infty} < \infty, \quad (29)$$

then the asymptotic rate of decay of the expansion coefficients in Eq. (28) is proportional to^{73,74}

$$|a_k| \sim \frac{1}{k^s} \quad (30)$$

Based on this, the error that is introduced by truncating the basis set is

$$\left\| \psi(x) - \sum_{k=0}^K a_k \chi_k(x) \right\| \sim \frac{1}{(K+1)^s} \sim \frac{1}{K^s} \quad (31)$$

Because this is an asymptotic result, it is only useful for sufficiently large K . The $\mathbb{L}^{\infty}(\mathbb{R})$ norm used in Eq. (29) is the essential supremum of the function’s magnitude,

$$\|f(x)\|_{\infty} = \text{ess sup} |f(x)| \quad (32)$$

For continuous functions, the essential supremum is just the least upper bound.

There are similar results for higher dimensions. One can expand the wavefunction in a basis set,

$$\Psi(x_1, x_2, \dots, x_d) = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_d=0}^{\infty} a_{k_1 k_2 \dots k_d} \chi_{k_1}(x_1) \chi_{k_2}(x_2) \cdots \chi_{k_d}(x_d). \quad (33)$$

If the wavefunction is antisymmetric with respect to interchange between its arguments, one can combine terms in Eq. (33) that differ only by permutation of the k_i 's and rewrite the expression as a linear combination of Slater determinants,

$$\Psi(x_1, x_2, \dots, x_d) = \sum_{0 \leq k_1 < k_2 < \dots < k_d} c_{\mathbf{k}} \begin{vmatrix} \chi_{k_1} & \chi_{k_2} & \dots & \chi_{k_d} \end{vmatrix} \quad (34)$$

In order to truncate the basis set expansion, one needs to study how the coefficients of the expansion, $c_{\mathbf{k}}$, decay. By analogy to Eq. (29), it is reasonable to analyze the case where the wavefunction has bounded s^{th} derivatives,

$$\left\| \frac{\partial^n \partial^{r_2} \dots \partial^{r_d} \Psi(\mathbf{x})}{\partial x_1^n \partial x_2^{r_2} \dots \partial x_d^{r_d}} \right\|_{\infty} < \infty \quad s \geq \|\mathbf{r}\|_{\infty} = r_1 + r_2 + \dots + r_d. \quad (35)$$

In this case,

$$|c_{\mathbf{k}}| \sim \left(\frac{1}{\|\mathbf{k}\|_{\infty}} \right)^s = \left(\frac{1}{\max_{1 \leq i \leq d} (k_i)} \right)^s \quad (36)$$

The guiding principle behind the ideal-CI method proposed in section III is that if we truncate the basis set at order K , then we should leave out all terms in the basis set expansion, Eq. (34), that contribute less to the expansion than the first neglected term. In the case described by Eqs. (35) and (36), this error-consistent truncation scheme reduces to FCI,

$$\Psi(x_1, x_2, \dots, x_d) = \sum_{\|\mathbf{k}\|_{\infty} \leq Q} c_{\mathbf{k}} \begin{vmatrix} \chi_{k_1} & \chi_{k_2} & \dots & \chi_{k_d} \end{vmatrix} \quad (37)$$

This is the ideal-CI expansion in the large K limit. If the wavefunction has only bounded s^{th} derivatives, and no additional smoothness, FCI is the asymptotically optimal basis-set expansion algorithm. In such cases, the only computationally tractable approach is Monte Carlo.

Fortunately, the molecular wavefunction has additional smoothness, called bounded mixed derivatives. A function has bounded mixed derivatives of s^{th} order if you can differentiate the function s times with respect to any coordinate,

$$\left\| \frac{\partial^n \partial^{r_2} \dots \partial^{r_d} \Psi(\mathbf{x})}{\partial x_1^n \partial x_2^{r_2} \dots \partial x_d^{r_d}} \right\|_{\infty} < \infty \quad s \geq \|\mathbf{r}\|_{\infty} = \max_{1 \leq i \leq d} r_i \quad (38)$$

This implies that $(s+1)^d$ derivatives exist; this exponential amount of differentiability can be exploited to break the curse of dimension. Notice that some of the derivatives in Eq. (38) are of very high order. For example, the following derivative of order s^d exists:

$$\left\| \frac{\partial^s \Psi(\mathbf{x})}{\partial x_1^s \partial x_2^s \cdots \partial x_d^s} \right\|_\infty < \infty \quad s \geq \|\mathbf{r}\|_\infty = \max_{1 \leq i \leq d} r_i \quad (39)$$

In this case, the asymptotic decay of the CI coefficients is

$$|c_{\mathbf{k}}| \sim \left(\frac{1}{\|\mathbf{k}\|_1} \right)^s = \left(\frac{1}{k_1 + k_2 + \cdots + k_d} \right)^s \quad (40)$$

and the ideal-CI expansion, in the large basis-set limit, has the form

$$\Psi(x_1, x_2, \dots, x_d) = \sum_{\|\mathbf{k}\|_1 \leq Q} c_{\mathbf{k}} \left| \chi_{k_1} \chi_{k_2} \cdots \chi_{k_d} \right| \quad (41)$$

This is identical to the $T=0$ case of Eq. (11). Compared to Eq. (37), there are many fewer terms in Eq. (41). For large K , the number of Slater determinants included in the FCI expansion (Eq. (37)) is proportional to the volume of a cube with side K , with a correction for antisymmetry. The number of Slater determinants included in Eq. (41) is proportional to the volume of the standard d -simplex with vertices at the origin, $(K, 0, 0, \dots, 0)$, $(0, K, 0, \dots, 0)$, and $(0, 0, \dots, 0, K)$, again with a correction for antisymmetry. There are thus $O(d!)$ fewer Slater determinants in Eq. (41) than in Eq. (37); the exploding computational cost of FCI is thus mitigated. In the context of the nuclear Schrödinger equation, the key result in Eq. (40) has been exploited by Avila and Carrington.⁴⁹

As a final example, we consider an even higher degree of differentiability called dominating mixed smoothness.^{54,55} This combines Eq. (35) and (39), so that

$$\left\| \frac{\partial^{r_1+q_1} \partial^{r_2+q_2} \cdots \partial^{r_d+q_d} \Psi(\mathbf{x})}{\partial x_1^{r_1+q_1} \partial x_2^{r_2+q_2} \cdots \partial x_d^{r_d+q_d}} \right\|_\infty < \infty \quad (42)$$

$$s \geq \|\mathbf{r}\|_\infty = \max_{1 \leq i \leq d} r_i$$

$$t \geq \|\mathbf{q}\|_\infty = q_1 + q_2 + \cdots + q_d$$

The corresponding asymptotic decay rate for the CI expansion coefficients looks like

$$|c_{\mathbf{k}}| \sim \left(\frac{1}{\|\mathbf{k}\|_1 + T \|\mathbf{k}\|_\infty} \right)^{s+t} \quad (43)$$

The optimal value of T , with $-1 < T < \infty$, is determined by the specific values of s and t in Eq. (42). The ideal-CI expansion, in the large basis set limit, has the form

$$\Psi(x_1, x_2, \dots, x_d) = \sum_{f(\mathbf{k}, T) \leq Q} c_{\mathbf{k}} \left| \chi_{k_1} \chi_{k_2} \cdots \chi_{k_d} \right| \quad (44)$$

where

$$f(\mathbf{k}, T) = \|\mathbf{k}\|_1 + T \|\mathbf{k}\|_\infty \quad (45)$$

$$-1 < T < \infty \quad (46)$$

Yserente showed that, for the electronic Schrödinger equation, one can choose $s = \frac{1}{2}$ and $t = 1$ in Eq. (42).⁷⁵ If all electrons have the same spin, then the electron-electron cusp is mitigated and an even stronger result holds, with $s = t = 1$.⁷⁵ The fact that s has different values for same-spin and opposite-spin electrons indicates that the optimal value of T will be different for same-spin and opposite-spin electrons. This is the motivation for the spin-resolved formulation in section III.

Because $s > 0$ and $t > 0$ in Eq. (42), we are justified in choosing $T < 0$, so the CI expansion in Eq. (44) has even fewer Slater determinants than the expansion in Eq. (41). This result has been exploited by Hamaekers and Griebel in the context of the electronic Schrödinger equation, but they based their approximation on the Galerkin approach, rather than the traditional CI expansion.^{46,47}

The results in this section are presented without mathematical justification and many nuances, especially related to Eqs. (42)-(46), have been glossed over. The interested reader is referred to the original literature for details. Equation (40) is based on the work of Smolyak,^{57,69,72,76} Eq. (45) is motivated by the work of Griebel and Knappek (who use a different sign convention for T).^{54,55} The link to expansions in orthogonal polynomials is most clearly explained by Novak and Ritter.⁷¹

The results in this section can be justified, however, on an intuitive level: remember that ℓ^1 and ℓ^∞ are dual to each other. If the order of bounded derivatives is given by the ℓ^1 norm (Eq. (35)), then the order of convergence in the series expansion is given by the ℓ^∞ norm (Eq. (36)), and vice versa (Eqs. (38) and (40)). Similarly, if the order of bounded derivatives is given by a combination of the ℓ^1 and ℓ^∞ norms (Eq. (42)), then so is the order of convergence of the series (Eq. (43)).

From the results in this section, the motivation for the algorithm proposed in section III should be clear.

VIII.VI. Algorithm Details: Griebel-Knappek CI by Degrees

The problem with directly translating the method in section V to CI calculations is the choice of basis. The key results in section V use a basis of one-dimensional orthogonal polynomials. Such a basis set can be used for atoms, but it

is not a popular approach in modern electronic structure theory.⁷⁷⁻⁸⁶ For many-electron molecules, using products of orthogonal polynomials as a basis set is not efficient. Fortunately, most of the key results that we cited for orthogonal polynomials can be extended to other basis sets of normalized, orthogonal functions.⁷⁴ The authors will not attempt to prove, mathematically, that the results presented in section V can still be applied when using the popular orbital basis sets in quantum chemistry (say, the Hartree-Fock orbitals expanded in some Gaussian basis set). However, plausibility can be established. Consider diagonalizing a one-electron Hamiltonian, $\hat{f}(\mathbf{r})$, using the Lanczos procedure. In the Lanczos procedure, one establishes a three-term recursion relation relative to the inner product,

$$\int \phi_i(\mathbf{r}) \left[\hat{f}(\mathbf{r}) - \varepsilon_0 + p \right] \phi_j(\mathbf{r}) d\mathbf{r}, \quad (47)$$

where the lowest eigenvalue of the one-electron Hamiltonian and $p > 0$ have been introduced to ensure that the bracketed operator is positive definite. This procedure is directly analogous to results in the theory of orthogonal polynomials.^{87,88} In addition, for one-electron Hamiltonians with a local potential (e.g., the Kohn-Sham Hamiltonian), Wintner's nodal theorem indicates that the nodes of the canonical orbitals are interlaced in the same pattern as the nodes of multidimensional orthogonal polynomials.⁸⁹ Specifically, Wintner showed that if $\phi_l(\mathbf{r})$ and $\phi_m(\mathbf{r})$ are eigenfunctions of $\hat{f}(\mathbf{r}) = -\frac{1}{2}\nabla^2(\mathbf{r}) + v_s(\mathbf{r})$ with $E_l < E_m$, then in any nodal region of $\phi_l(\mathbf{r})$, $\phi_m(\mathbf{r})$ must change sign. (A nodal region is a region of space bounded by a nodal surface, where $\phi_l(\mathbf{r}) = 0$.)⁹⁰ Since the nodal structure of the orbitals is the same as the nodal structure of multidimensional orthogonal polynomials, one expects that one can map the orbital basis set onto orthogonal polynomials by making a suitable transformation of coordinates. Special cases of this transformation are known, e.g., the relationship between the Laguerre and the Hermite orthogonal polynomials.

Thus motivated, we propose to use the number of nodes in the canonical Hartree-Fock or Kohn-Sham orbitals as a replacement for the degree of orthogonal polynomials. For atoms and linear molecules, the number of nodes in an orbital is easily established using Wintner's theorem and the angular symmetry of the orbitals (for atoms: s, p, d, \dots ; for linear molecules: $\sigma, \pi, \delta, \dots$). Given two orbitals with the same angular symmetry, the one with higher-energy has more nodes.

When there is no angular symmetry, deciphering the number of nodes in an orbital is more challenging. However, Wintner's theorem assures us that the

number of nodes generally increases with increasing energy. If we order the orbital basis set with increasing energy,

$$\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_K \quad (48)$$

then this closely (but not perfectly) corresponds to ordering the orbitals by increasing numbers of nodes.⁹¹ Dimensional arguments indicate that, barring special symmetries, the number of D -dimensional orbitals with η or fewer nodes is $\binom{\eta+D}{D}$; the number of orbitals with precisely η nodes is $\binom{\eta+D}{D} \left(1 - \frac{\eta}{\eta+D}\right)$.⁹²

Combining this result with Wintner's theorem allows us to estimate that the number of nodes in the k^{th} spatial-orbital by solving the following equation for η_k :

$$k = \frac{\Gamma(\eta_k + D + 1)}{D! \Gamma(\eta_k + 1)}. \quad (49)$$

This equation reduces to Eq. (15) in the common three-dimensional case, $D = 3$. An alternative to Eq. (49) is

$$k = \frac{(\eta_k + 1)(\eta_k + 2) \dots (\eta_k + D)}{D!} \quad (50)$$

For $D \leq 4$, this equation can be solved analytically.

Using the number of nodes in $\phi_k(\mathbf{r})$ as a proxy for the degree of the orthogonal polynomial, we can now exploit the results from section V to truncate the CI expansion. We use the most general, Griebel-Knapek, construction. First construct all possible Slater determinants containing only α -spin electrons that satisfy the criterion,

$$f_{\alpha\alpha}(\mathbf{i}_\alpha, T) = \|\eta_{\mathbf{i}_\alpha}\|_1 + T_{\alpha\alpha} \|\eta_{\mathbf{i}_\alpha}\|_\infty < Q_{\alpha\alpha}(K, T_{\alpha\alpha}). \quad (51)$$

Then form a similar set of Slater determinants containing only β -spin electrons. At this point we have a basis set of orthogonal functions for the α -spin and β -spin subspaces. The effective polynomial degree of a Slater determinant in the σ -spin subspace is,

$$\|\eta_{\mathbf{i}_\sigma}\|_1 = \eta_{i_{\sigma,1}} + \eta_{i_{\sigma,2}} + \dots + \eta_{i_{\sigma,N_\sigma}} \quad (52)$$

The final N -electron basis function is an (antisymmetric) product of the α -spin and β -spin Slater determinants,⁹³

$$\Psi = \sum_{f_{\alpha\beta}(\eta_{\mathbf{i}_\alpha}, \eta_{\mathbf{i}_\beta}) \leq Q_{\alpha\beta}(K, T_{\alpha\beta})} c_i \left| \phi_{i_{\alpha,1}} \alpha \ \phi_{i_{\alpha,2}} \alpha \ \dots \ \phi_{i_{\alpha,N_\alpha}} \alpha \right| \wedge \left| \phi_{i_{\beta,1}} \beta \ \phi_{i_{\beta,2}} \beta \ \dots \ \phi_{i_{\beta,N_\beta}} \beta \right| \quad (53)$$

The sum includes all of the spin-Slater determinants that satisfy the Griebel-Knapek condition,

$$f_{\alpha\beta}(\mathbf{i}, T_{\alpha\beta}) = \left(\|\boldsymbol{\eta}_{i_\alpha}\| + \|\boldsymbol{\eta}_{i_\beta}\| \right) + T_{\alpha\beta} \max_{\sigma=\alpha,\beta} \left(\|\boldsymbol{\eta}_{i_\alpha}\|, \|\boldsymbol{\eta}_{i_\beta}\| \right) < Q_{\alpha\beta}(K, T_{\alpha\beta}) \quad (54)$$

With these choices, the truncated CI takes the form given in Eqs. (11)-(14). Because we are defining an analogue of the degree of a polynomial by counting the nodes in the orbital basis functions, we call this method for building up the configuration space of Slater determinants “Griebel-Knapek CI by degrees” (GK-CI). The advantage of the spin-resolved formulation used here is that one can optionally exploit the greater degree of differentiability that arises when one considers only same-spin electrons. Based on this, it seems that most interesting and useful choices for the Griebel-Knapek T parameters will be characterized by

$$-1 < T_{\alpha\alpha} = T_{\beta\beta} \leq T_{\alpha\beta} \leq 0. \quad (55)$$

In the spirit of the ideal-CI method presented in section III, we want to truncate the CI expansion at the maximum value of

$$f_{\sigma\sigma'}(\boldsymbol{\eta}, T_{\sigma\sigma'}) = \|\boldsymbol{\eta}\| + T_{\sigma\sigma'} \|\boldsymbol{\eta}\|_{\infty} \quad (56)$$

that can be attained with our basis set. Equivalently, we should evaluate $f_{\sigma\sigma'}(\boldsymbol{\eta}, T_{\sigma\sigma'})$ for the most important Slater determinant that is not included in the basis set, and use the value of $f_{\sigma\sigma'}(\boldsymbol{\eta}_{\text{excluded}}, T_{\sigma\sigma'})$ to define a threshold for truncating the CI expansion. It is impossible to know precisely what is the most important Slater determinant, but a good candidate is the Slater determinant rendered in figure 1c, with electrons excited from the highest-occupied α -spin and β -spin molecular orbitals to the lowest-energy orbital that is not included in the basis set. Using the number of nodes as proxy for the degree, and assuming that the spatial orbitals are sorted in increasing order of nodes,

$$\eta_1 \leq \eta_2 \leq \dots \leq \eta_K \quad (57)$$

then the type of Slater determinant depicted in Figure 1c is indexed by

$$\mathbf{i}_{\text{excluded}} = \left[1 \quad \dots \quad (N_\alpha - 1) \quad (K+1) \quad 1 \quad \dots \quad (N_\beta - 1) \quad (K+1) \right]^T. \quad (58)$$

This gives

$$Q_{\sigma\sigma}(K, T_{\sigma\sigma}) = f_{\sigma\sigma}(\mathbf{i}_{\text{excluded}_\sigma}, T_{\sigma\sigma}) = \sum_{i=1}^{N_\sigma-1} \eta_i + (1 + T_{\sigma\sigma}) \eta_{K+1}$$

$$Q_{\alpha\beta}(K) = f_{\alpha\beta}(\mathbf{i}_{\text{excluded}}, T_{\alpha\beta}) = \sum_{\sigma=\alpha,\beta} \left(\sum_{i=1}^{N_\sigma-1} \eta_i + \eta_{K+1} \right) + T_{\alpha\beta} \max_{\sigma=\alpha,\beta} \left(\sum_{i=1}^{N_\sigma-1} \eta_i + \eta_{K+1} \right) \quad (59)$$

When the precise number of nodes are known, the formula in Eq. (16) allows one to compute $Q_{\sigma\sigma'}$ without sorting the orbitals into nodal order (as in Eq. (57)).

Ultimately one wants to know the full-CI energy in the limit of infinite basis set. We can extrapolate the energy obtained from a GK-CI calculation to the

$K \rightarrow \infty$ limit using the result in Eq. (43). Because the energy is stationary with respect to infinitesimal variations of the wavefunction, the error in the energy is square of the error in the wavefunction, so the asymptotic convergence of the energy can be fit to the form

$$E_{K \rightarrow \infty} - E_K = \left(\frac{a_K}{Q_{\alpha\beta}(K, T_{\alpha\beta})} \right)^{2u}, \quad (60)$$

where u characterizes the asymptotic decay rate of the CI coefficients. From Eq. (43), $u = (s + t)$. In practice, we would use u as a fit parameter. Depending on the value of T and the spin-multiplicity, one expects $1 \leq u \leq 2$.^{75,94,95,96} The value of u is limited by nondifferentiability at the electron-electron cusp, and if this feature is handled analytically using explicitly correlated basis functions (e.g., a Slater-Jastrow form), then the value of u , *ergo* the rate of convergence to the basis-set limit, will increase.

VIII.VII. Discussion

To clarify this approach, Table 1 lists the electron configurations that would be included in a doubly-occupied configuration interaction (DOCI) calculation on Beryllium and Neon, in a double- ζ basis set, with $T_{\alpha\alpha} = T_{\beta\beta} = 0$ and $T_{\alpha\beta} \rightarrow \infty$. Table 2 lists the electron configurations for DOCI in a triple- ζ basis set.⁹⁷ For $T_{\sigma\sigma} < 0$, fewer multiple excitations would be allowed. For $T_{\sigma\sigma} > 0$, higher-order excitations would be permitted.

Table 1 The electron configurations included in a doubly-occupied configuration interaction (DOCI) calculation on Beryllium and Neon, in a double- ζ basis set ($3s, 2p, 1d$), with $T_{\alpha\alpha} = T_{\beta\beta} = 0$; $T_{\alpha\beta} \rightarrow \infty$. The value of $f_{\sigma\sigma}$ is computed using Eq. (51). The final line of the table lists one of the most important electron configurations that is neglected because of basis-set truncation error.

Be	$f_{\sigma\sigma}$	Ne	$f_{\sigma\sigma}$
$1s^2 2s^2$	1	$1s^2 2s^2 2p^6$	4
$1s^2 2p^2$	1	$1s^2 2p^6 3s^2$	5
$1s^2 3s^2$	2	$1s^2 2p^6 3p^2$	5
$1s^2 3p^2$	2	$1s^2 2p^6 3d^2$	5
$1s^2 3d^2$	2	$1s^2 2s^2 2p^4 3s^2$	5
$2s^2 2p^2$	2	$1s^2 2s^2 2p^4 3p^2$	5
$2p^4$	2	$1s^2 2s^2 2p^4 3d^2$	5
$1s^2 4s^2$	3	$1s^2 2s^2 2p^4 4s^2$	6

Table 2 The electron configurations included in a doubly-occupied configuration interaction (DOCI) calculation on Beryllium and Neon, in a triple- ζ basis set ($4s,3p,2d,1f$), with $T_{\alpha\alpha} = T_{\beta\beta} = 0$; $T_{\alpha\beta} \rightarrow \infty$.

Be	$f_{\sigma\sigma}$	Ne	$f_{\sigma\sigma}$
$1s^2 2s^2$	1	$1s^2 2s^2 2p^6$	4
$1s^2 2p^2$	1	$1s^2 2p^6 3s^2$	5
$1s^2 3s^2$	2	$1s^2 2p^6 3p^2$	5
$1s^2 3p^2$	2	$1s^2 2p^6 3d^2$	5
$1s^2 3d^2$	2	$1s^2 2s^2 2p^4 3s^2$	5
$2s^2 2p^2$	2	$1s^2 2s^2 2p^4 3p^2$	5
$2p^4$	2	$1s^2 2s^2 2p^4 3d^2$	5
$1s^2 4s^2$	3	$1s^2 2p^6 4s^2$	6
$1s^2 4p^2$	3	$1s^2 2p^6 4p^2$	6
$1s^2 4d^2$	3	$1s^2 2p^6 4d^2$	6
$1s^2 4f^2$	3	$1s^2 2p^6 4f^2$	6
$2s^2 3s^2$	3	$1s^2 2s^2 2p^4 4s^2$	6
$2s^2 3p^2$	3	$1s^2 2s^2 2p^4 4p^2$	6
$2s^2 3d^2$	3	$1s^2 2s^2 2p^4 4d^2$	6
$2p^2 3s^2$	3	$1s^2 2s^2 2p^4 4f^2$	6
$2p^2 3p^2$	3	$2s^2 2p^6 3s^2$	6
$2p^2 3d^2$	3	$2s^2 2p^6 3p^2$	6
		$2s^2 2p^6 3d^2$	6
		$1s^2 2s^2 2p^2 3s^2 3p^2$	6
		$1s^2 2s^2 2p^2 3s^2 3d^2$	6
		$1s^2 2s^2 2p^2 3p^4$	6
		$1s^2 2s^2 2p^2 3d^4$	6
		$1s^2 2p^4 3s^2 3p^2$	6
		$1s^2 2p^4 3s^2 3d^2$	6
		$1s^2 2p^4 3p^4$	6
		$1s^2 2p^4 3d^4$	6
$1s^2 5s^2$	4	$1s^2 2s^2 2p^4 5s^2$	7

Several interesting features emerge. The near-degeneracy correlation of the Beryllium atom is accounted for automatically because the $2p$ and $2s$ orbitals have the same number of nodes. Therefore, whenever a Slater determinant with a $2s$ -orbital is included in the expansion, corresponding Slater determinants with $2p$ orbitals will also be included. Second, the expansion is severely truncated. This is most clear in the case of Neon in the double- ζ basis: because the $1s^2 2s^2 2p^4 4s^2$ Slater determinant is not available, only single, double, and triple excitations are included when $T_{\sigma\sigma} \leq 0$. Similarly, with a triple- ζ basis: because the $1s^2 2s^2 2p^4 5s^2$ determinant is not available, no excitations beyond fifth order are included for $T_{\sigma\sigma} \leq 0$.

We wrote a computer program to explicitly generate the allowed electronic configurations for GK-CI calculations. Figure 2 plots the number of Slater determinants for GK-CI calculation on an N -electron singlet state, for various sizes of the orbital basis and various choices for $T_{\sigma\sigma}$, with $T_{\alpha\beta} \rightarrow \infty$. For comparison, the scaling behavior of the analogous coupled-cluster (specifically, CCSD and CCSDT) and FCI calculations is shown.⁹⁸ Because the maximum number of nodes in the basis set grows very slowly as the basis set increases ($\eta_k \sim \sqrt[3]{k}$), the number of determinants is relatively insensitive to $T_{\sigma\sigma}$ in the range $-1 < T_{\sigma\sigma} \leq 0$. The benefit of using $T_{\sigma\sigma} < 0$ most apparent when the basis set is large; the benefits decrease as the level-of-filling increases. Figure 2 also demonstrates that, for $T_{\sigma\sigma} > 0$, the curse of dimension is not removed by the Griebel-Knapek construction, and the number of Slater determinants explodes exponentially.^{54,55}

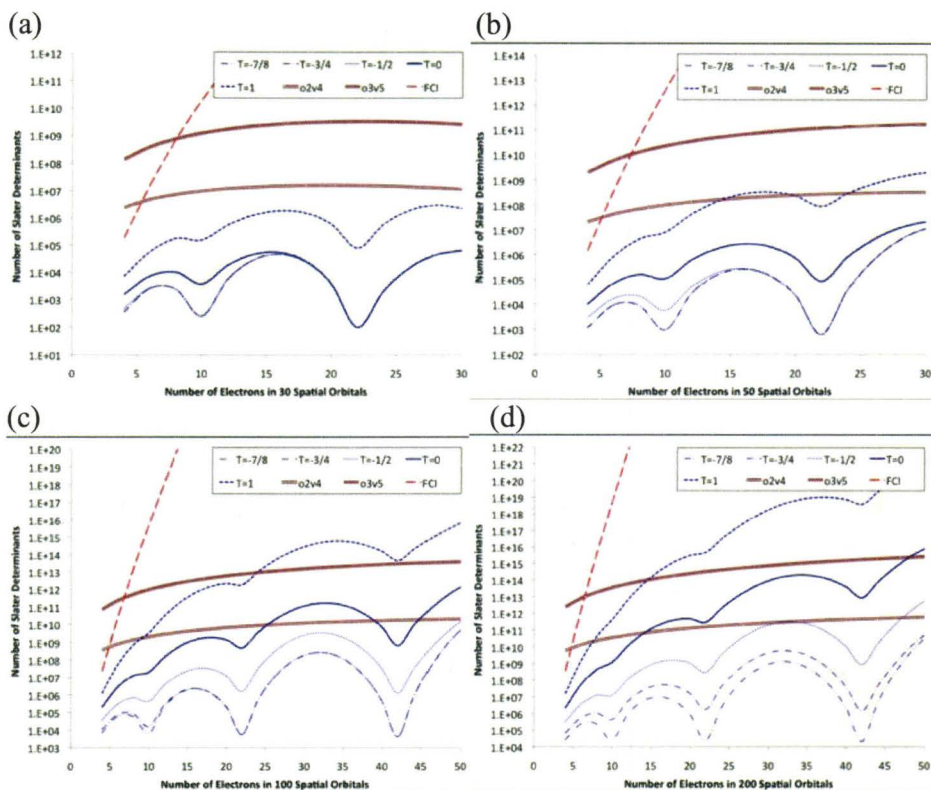


Figure 2 The number of Slater determinants in various types of Griebel-Knapek CI by degrees, for singlet states with various numbers of electrons, N , and sizes of the spatial-orbital basis, K . Different values of $T_{\alpha\alpha} = T_{\beta\beta} = T$ are used; $T_{\alpha\beta} \rightarrow \infty$. The values of o^2v^4 and o^3v^5 are also given; these numbers characterize the scaling of coupled-cluster singles and doubles (CCSD) and coupled-cluster singles, doubles, and triples (CCSDT), respectively. ($o = N_\alpha$ is the number of occupied states and $v = (K - N_\alpha)$ is the number of virtual states.) (a) 30 spatial orbitals; (b) 50 spatial orbitals; (c) 100 spatial orbitals; (d) 200 spatial orbitals.

The number of Slater determinants is even less sensitive to changes in $T_{\alpha\beta}$, at least as long as one is sensibly choosing $T_{\alpha\beta} > T_{\sigma\sigma}$. Indeed, as shown in the appendix, for any $T_{\alpha\alpha} = T_{\beta\beta} \leq 0$, all $T_{\alpha\beta} \geq T_{\sigma\sigma}$ give the same results. This is why we chose $T_{\alpha\beta} \rightarrow \infty$ in Figure 2. Only the $T_{\sigma\sigma} = 1$ curve, though, would have changed had we used $T_{\alpha\beta} = 0$ instead.

Notice that the number of Slater determinants included in a GK-CI calculation has a type of shell-structure. As discussed before in the context of the Beryllium atom, the determinants included in GK-CI include all possible ways of arranging electrons from orbitals with the maximum number of nodes in the reference determinant into orbitals with the same number of nodes. This “complete active space” family of determinants is largest when the last shell is half-filled, and smallest when the last shell is nearly empty. This gives a sizeable dip in the number of determinants when there are only two electrons outside a “closed shell” of orbitals with fewer nodes. As there are $(\eta+1)(\eta+2)(\eta+3)/3$ 3-dimensional orbitals with η or fewer nodes, “closed shells” occur at $K = 2, 8, 20, 40, 70, \dots$. The peaks and valleys in the number of GK-CI determinants reveal these shells.

Ultimately, no CI-type expansion based on *explicit* expansion in Slater determinants can compete with the very compact *implicit* representations for the wavefunction that are constructed using methods like the linear combination of graphically contracted functions (LCGCF)²⁶, the density matrix renormalization group (DMRG)^{27,28}, or correlator product states (CPS).⁹⁹ DMRG, in particular, has been applied to systems with massive dimensionality.^{34,100-102} However, while DMRG tends to behave as a polynomial-cost algorithm, it is still an exponential-cost algorithm from a mathematical standpoint.^{39,103,104} The analysis in this paper suggests that there might be ways, based on functional-analytic considerations and the Griebel-Knappek cost function, to reduce the dimensionality of the matrices that are multiplied together to form the matrices in DMRG or the correlators in CPS. Such a wavefunction might dramatically decrease the cost of DMRG and/or CPS and allow the inclusion of dynamical electron correlation, all without sacrificing accuracy. Approaches like this would be more efficient than explicit GK-CI.

VIII.VIII. **Summary**

While the mathematics is complicated, the key idea behind the Griebel-Knapek CI by degrees approach is very simple. The target is the ideal-CI method proposed in section III: in ideal-CI, the basis-set truncation error and the CI-expansion truncation error are matched. No effort is wasted on determinants that contribute less to the final wavefunction than the highest-contributing determinant that would be added were the basis set extended. The basic idea is clear from Figure 1: because the electron configuration in Figure 1b contributes much less to the wavefunction than the electron configuration in Figure 1c, the absence of Slater determinants like the one picture in Figure 1c will dominate the error in a FCI calculation. Ideal-CI can be formulated to be size-consistent. We do not believe that ideal-CI is size-extensive, but the deviation from size-extensivity is comparable to the error from basis-set truncation.

The ideal-CI method is merely formal. However, based on functional-analytic properties of the wavefunction, one can model the asymptotic decay of the CI-coefficients using the approach of sections V and VI. The key idea is to consider the number of nodes in the spatial orbitals to be a proxy for the polynomial degree, then use established results for the convergence rate of orthogonal-polynomial expansions. This defines the Griebel-Knapek CI by degrees method. In the basis-set limit, where these asymptotic formulae for the CI coefficients are exact, this method is equivalent to ideal-CI.

For an arbitrary N -particle Schrödinger equation, where the eigenfunctions are r -fold differentiable, ideal-CI and full CI are the same. This agrees with the fact that the general N -electron Schrödinger equation is computationally intractable.¹⁰⁵ However, the *molecular* electronic structure problem is special, and the eigenfunctions of the electronic Hamiltonian are not only differentiable, but they have bounded mixed derivatives (cf. Eq. (38)) and even dominating mixed smoothness (cf. Eq. (42)).⁷⁵ We do not need FCI in quantum chemistry: the fact that the electronic wavefunctions are mathematically “nice” allows us to dramatically truncate the CI expansion.¹⁰⁶ The parameter, $T_{\sigma\sigma'}$, can be used to adjust the assumptions about the differentiability of the wavefunction. With $T_{\sigma\sigma'} < 0$, the number of determinants in the GK-CI expansion is relatively small, and does not increase as a factorial of the number of electrons.

Acknowledgments. The authors acknowledge support from NSERC (graduate fellowship to JSMA, Discovery Grant to PWA), the Canada Research Chairs, and Sharcnet.

VIII.IX. Appendix

For any $T_{\sigma\sigma} = T_{\alpha\alpha} = T_{\beta\beta} \leq 0$, all values of $T_{\alpha\beta} \geq T_{\sigma\sigma}$ produce the same set of Slater determinants. To see this, recall that

$$Q_{\sigma\sigma} = \|\mathbf{n}_{\text{excluded},\sigma}\|_1 + T_{\sigma\sigma} \|\mathbf{n}_{\text{excluded},\sigma}\|_{\infty} \quad (61)$$

$$Q_{\alpha\beta} = \|\mathbf{n}_{\text{excluded},\alpha}\|_1 + \|\mathbf{n}_{\text{excluded},\beta}\|_1 + T_{\alpha\beta} \max\left(\|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty}, \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty}\right) \quad (62)$$

Suppose that $\{\mathbf{n}_{\alpha}, \mathbf{n}_{\beta}\}$ are the vectors containing the nodal numbers of two spin-Slater determinants. Then,

$$\begin{aligned} \|\mathbf{n}_{\alpha}\|_1 + T_{\sigma\sigma} \|\mathbf{n}_{\alpha}\|_{\infty} &< \|\mathbf{n}_{\text{excluded},\alpha}\|_1 + T_{\sigma\sigma} \|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty} \\ \|\mathbf{n}_{\beta}\|_1 + T_{\sigma\sigma} \|\mathbf{n}_{\beta}\|_{\infty} &< \|\mathbf{n}_{\text{excluded},\beta}\|_1 + T_{\sigma\sigma} \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty} \end{aligned} \quad (63)$$

Add together the two inequalities and rearrange them as,

$$-T_{\sigma\sigma} \left(\begin{array}{c} \|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty} - \|\mathbf{n}_{\alpha}\|_{\infty} \\ \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty} - \|\mathbf{n}_{\beta}\|_{\infty} \end{array} \right) < \left(\begin{array}{c} \|\mathbf{n}_{\text{excluded},\alpha}\|_1 + \|\mathbf{n}_{\text{excluded},\beta}\|_1 \\ -\|\mathbf{n}_{\alpha}\|_1 - \|\mathbf{n}_{\beta}\|_1 \end{array} \right) \quad (64)$$

The Slater determinant that is obtained when we combine the two spin-Slater determinants will be accepted if and only if

$$\left(\begin{array}{c} \|\mathbf{n}_{\alpha}\|_1 + \|\mathbf{n}_{\beta}\|_1 \\ +T_{\alpha\beta} \max\left(\|\mathbf{n}_{\alpha}\|_{\infty}, \|\mathbf{n}_{\beta}\|_{\infty}\right) \end{array} \right) < \left(\begin{array}{c} \|\mathbf{n}_{\text{excluded},\alpha}\|_1 + \|\mathbf{n}_{\text{excluded},\beta}\|_1 \\ +T_{\alpha\beta} \max\left(\|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty}, \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty}\right) \end{array} \right) \quad (65)$$

or, equivalently,

$$-T_{\alpha\beta} \left(\begin{array}{c} \max\left(\|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty}, \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty}\right) \\ -\max\left(\|\mathbf{n}_{\alpha}\|_{\infty}, \|\mathbf{n}_{\beta}\|_{\infty}\right) \end{array} \right) < \left(\begin{array}{c} \|\mathbf{n}_{\text{excluded},\alpha}\|_1 + \|\mathbf{n}_{\text{excluded},\beta}\|_1 \\ -\|\mathbf{n}_{\alpha}\|_1 - \|\mathbf{n}_{\beta}\|_1 \end{array} \right) \quad (66)$$

If the left-hand-side of inequality (64) is greater than or equal to the left-hand-side of equation (66), then the $\{\mathbf{n}_{\alpha}, \mathbf{n}_{\beta}\}$ pair will always be accepted as a Slater determinant. Thus, we need to analyze when:

$$-T_{\sigma\sigma} \left(\begin{array}{c} \|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty} - \|\mathbf{n}_{\alpha}\|_{\infty} \\ \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty} - \|\mathbf{n}_{\beta}\|_{\infty} \end{array} \right) \geq -T_{\alpha\beta} \left(\begin{array}{c} \max\left(\|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty}, \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty}\right) \\ -\max\left(\|\mathbf{n}_{\alpha}\|_{\infty}, \|\mathbf{n}_{\beta}\|_{\infty}\right) \end{array} \right) \quad (67)$$

The precise criterion is that

$$T_{\sigma\sigma} \left[\frac{\|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty} - \|\mathbf{n}_{\alpha}\|_{\infty} + \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty} - \|\mathbf{n}_{\beta}\|_{\infty}}{\max\left(\|\mathbf{n}_{\text{excluded},\alpha}\|_{\infty}, \|\mathbf{n}_{\text{excluded},\beta}\|_{\infty}\right) - \max\left(\|\mathbf{n}_{\alpha}\|_{\infty}, \|\mathbf{n}_{\beta}\|_{\infty}\right)} \right] \leq T_{\alpha\beta} \quad (68)$$

The term in brackets is greater than or equal to two. This allows us to state that all values of $T_{\alpha\beta}$ that satisfy Eq. (68) give the same set of Slater determinants and this set is simply the Cartesian product of the set of α -spin Slater determinants and the set of β -spin Slater determinants, which is the same result that one obtains in the $T_{\alpha\beta} \rightarrow \infty$ limit. In the special case where $T_{\sigma\sigma} \leq 0$, then we can furthermore say that all $T_{\alpha\beta} \geq 2T_{\sigma\sigma}$ give the Cartesian product of the spin-Slater determinant sets. Since sensible choices for the parameters are $T_{\alpha\beta} \geq T_{\sigma\sigma}$ and $T_{\sigma\sigma} \leq 0$, the Cartesian-product set is arguably the only sensible choice. This is why we used $T_{\alpha\beta} \rightarrow \infty$ in Figure 2.

VIII.X. References

- 1 D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).
- 2 J. B. Anderson, C. A. Traynor, and B. M. Boghosian, *J. Chem. Phys.* **95**, 7418 (1991).
- 3 J. Olsen, B. O. Roos, P. Jorgensen, and H. J. A. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).
- 4 J. Pitarch-Ruiz, J. Sanchez-Marin, and D. Maynau, *J. Comput. Chem.* **23**, 1157 (2002).
- 5 H. Nakatsuji, *J. Chem. Phys.* **83** (2), 713 (1985).
- 6 H. J. Werner and P. J. Knowles, *J. Chem. Phys.* **89** (9), 5803 (1988).
- 7 K. Tanaka, T. Sakai, and H. Terashima, *Theor. Chim. Act.* **76** (4), 213 (1989).
- 8 D. B. Knowles, J. R. Alvarezcollado, G. Hirsch, and R. J. Buenker, *J. Chem. Phys.* **92** (1), 585 (1990).
- 9 H. Lischka, R. Shepard, R. M. Pitzer, I. Shavitt, M. Dallos, T. Muller, P. G. Szalay, M. Seth, G. S. Kedziora, S. Yabushita, and Z. Y. Zhang, *PCCP* **3** (5), 664 (2001).
- 10 H. Nakatsuji and E. R. Davidson, *J. Chem. Phys.* **115** (5), 2000 (2001).
- 11 H. Nakatsuji, *J. Chem. Phys.* **113**, 2949 (2000).
- 12 L. Bytautas and K. Ruedenberg, *Chem. Phys.* **356**, 64 (2009).
- 13 J. Ivanic and K. Ruedenberg, *Theor. Chem. Acc.* **107** (4), 220 (2002).
- 14 J. Ivanic and K. Ruedenberg, *Theor. Chem. Acc.* **106**, 339 (2001).
- 15 G. H. Booth and A. Alavi, *J. Chem. Phys.* **132**, 174104 (2010).
- 16 D. Cleland, G. H. Booth, and A. Alavi, *J. Chem. Phys.* **132**, 041103 (2010).
- 17 G. H. Booth, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **131**, 054106 (2009).

- 18 A. J. W. Thom and A. Alavi, *J. Chem. Phys.* **123**, 204106 (2005).
- 19 B. Huron, P. Rancurel, and J. P. Malrieu, *J. Chem. Phys.* **75**, 5745 (1973).
- 20 O. Jitrik and C. F. Bunge, *Phys. Rev. A* **56** (4), 2614 (1997).
- 21 S. Evangelisti, J. P. Daudey, and J. P. Malrieu, *Chem. Phys.* **75**, 91 (1983).
- 22 F. Illas, J. Rubio, J. M. Ricart, and P. S. Bagus, *J. Chem. Phys.* **95**, 1877 (1991).
- 23 G. Gidofalvi and R. Shepard, *J. Comput. Chem.* **30**, 2414 (2009).
- 24 R. Shepard, M. Minkoff, and S. R. Brozell, *Int. J. Quantum Chem.* **107**, 3203 (2007).
- 25 R. Shepard and M. Minkoff, *Int. J. Quantum Chem.* **106**, 3190 (2006).
- 26 R. Shepard, *J. Phys. Chem. A* **109**, 11629 (2005).
- 27 S. R. White, *Phys. Rev. B* **48**, 10345 (1993).
- 28 S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- 29 S. Daul, I. Ciofini, C. Daul, and S. R. White, *Int. J. Quantum Chem.* **79** (6), 331 (2000).
- 30 S. R. White and R. L. Martin, *J. Chem. Phys.* **110** (9), 4127 (1999).
- 31 G. K. L. Chan, *J. Chem. Phys.* **120** (7), 3172 (2004).
- 32 G. K. L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).
- 33 G. Moritz and M. Reiher, *J. Chem. Phys.* **126**, 244109 (2007).
- 34 D. Ghosh, J. Hachmann, T. Yanai, and G. K. L. Chan, *J. Chem. Phys.* **128**, 144117 (2008).
- 35 D. Zgid and M. Nooijen, *J. Chem. Phys.* **128**, 114116 (2008).
- 36 K. H. Marti and M. Reiher, *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics* **224**, 583 (2010).
- 37 G. K. L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, and T. Yanai, in *Frontiers in Quantum Systems in Chemistry and Physics*, edited by S. Wilson, P. J. Grout, J. Maruani, G. DelgadoBarrio, and P. Piecuch (2008), Vol. 18, pp. 49.
- 38 F. Verstraete, D. Porras, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 227205 (2004).
- 39 N. Schuch, I. Cirac, and F. Verstraete, *Phys. Rev. Lett.* **100**, 250501 (2008).
- 40 M. W. Schmidt and M. S. Gordon, *Annu. Rev. Phys. Chem.* **49**, 233 (1998).
- 41 C. Angeli, C. J. Calzado, R. Cimiraglia, S. Evangelisti, and D. Maynau, *Mol. Phys.* **101**, 1937 (2003).
- 42 J. I. Rodriguez, D. C. Thompson, J. S. M. Anderson, J. W. Thomson, and P. W. Ayers, *Journal of Physics A* **41**, 365202 (2008).
- 43 J. I. Rodriguez, D. C. Thompson, P. W. Ayers, and A. M. Koster, *J. Chem. Phys.* **128**, 224103 (2008).

- 44 J. S. M. Anderson, J. I. Rodriguez, D. C. Thompson, and P. W. Ayers, in *Quantum Chemistry Research Trends* (Nova, Hauppauge, NY, 2007).
- 45 J. S. M. Anderson, R. Cuevas-Saavedra, D. C. Thompson, J. I. Rodriguez, and P. W. Ayers, in preparation.
- 46 M. Griebel and J. Hamaekers, *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics* **224**, 527 (2010).
- 47 M. Griebel and J. Hamaekers, *Esaim-Mathematical Modelling and Numerical Analysis-Modelisation Mathematique Et Analyse Numerique* **41** (2), 215 (2007).
- 48 J. Garcke and M. Griebel, *J. Comput. Phys.* **165** (2), 694 (2000).
- 49 G. Avila and T. Carrington, *J. Chem. Phys.* **131**, 174103 (2009).
- 50 R. J. Whitehead and N. C. Handy, *J. Mol. Spectrosc.* **59**, 459 (1976).
- 51 S. Carter and N. C. Handy, *Computer Physics Reports* **5**, 115 (1986).
- 52 X. G. Wang and T. Carrington, *J. Phys. Chem. A* **105**, 2575 (2001).
- 53 J. S. M. Anderson, R. Cuevas-Saavedra, and P. W. Ayers, unpublished.
- 54 M. Griebel and S. Knapek, *Constructive Approximation* **16**, 525 (2000).
- 55 M. Griebel and S. Knapek, *Mathematics of Computation* **78**, 2223 (2009).
- 56 J. Garcke and M. Griebel, *J. Comput. Phys.* **165**, 694 (2000).
- 57 S. A. Smolyak, *Dokl. Akad. Nauk* **4**, 240 (1963).
- 58 J. F. Traub and A. G. Werschulz, *Complexity and information*. (Cambridge UP, Cambridge, 1998).
- 59 J. F. Traub and H. Wozniakowski, *Bull.AMS* **26**, 29 (1992).
- 60 J. F. Traub, G. W. Wasilkowski, and H. Wozniakowski, *Information-Based Complexity*. (Academic Press, New York, 1988).
- 61 L. Blum, M. Shub, and S. Smale, *Bulletin of the American Mathematical Society* **21**, 1 (1989).
- 62 E. Novak, *Journal of Complexity* **11**, 57 (1995).
- 63 H. Wozniakowski, *Theoretical Computer Science* **219**, 451 (1999).
- 64 H. Niederreiter, *Bulletin of the American Mathematical Society* **84**, 957 (1978).
- 65 W. J. Morokoff and R. E. Caflisch, *J. Comput. Phys.* **122**, 218 (1995).
- 66 F. James, J. Hoogland, and R. Kleiss, *Comput. Phys. Commun.* **99**, 180 (1997).
- 67 H. Wozniakowski, *Bulletin of the American Mathematical Society* **24**, 185 (1991).
- 68 I. H. Sloan and H. Wozniakowski, *Journal of Complexity* **14**, 1 (1998).
- 69 G. W. Wasilkowski and H. Wozniakowski, *Journal of Complexity* **11**, 1 (1995).
- 70 R. Cools, E. Novak, and K. Ritter, *Computing* **62**, 147 (1999).
- 71 E. Novak and K. Ritter, *Constructive Approximation* **15**, 499 (1999).

- 72 E. Novak, *Nonlinear Analysis: Theory Methods & Applications* **30**, 1439 (1997).
- 73 N. I. Achieser, *Theory of Approximation*. (Dover, New York, 1992).
- 74 A. F. Timan, *Theory of Approximation of Functions of a Real Variable*. (Dover, New York, 1994).
- 75 H. Yserentant, *Numerische Mathematik* **98** (4), 731 (2004).
- 76 K. Petras, *Numerische Mathematik* **93**, 729 (2003).
- 77 H. Shull and P. O. Lowdin, *J. Chem. Phys.* **23**, 1362 (1955).
- 78 H. Shull and P. O. Lowdin, *J. Chem. Phys.* **23**, 1565 (1955).
- 79 P. O. Lowdin and H. Shull, *Phys. Rev.* **101**, 1730 (1956).
- 80 G. R. Taylor and R. G. Parr, *Proc. Natl. Acad. Sci.* **38**, 154 (1952).
- 81 Z. Xiong and N. Bacalis, *Chinese Physics* **15**, 992 (2006).
- 82 C. G. Bao, *Phys. Rev. A* **38**, 591 (1988).
- 83 J. Mitroy, M. W. J. Bromley, and K. Ratnavelu, *Int. J. Quantum Chem.* **107**, 907 (2007).
- 84 M. W. J. Bromley and J. Mitroy, *Int. J. Quantum Chem.* **107**, 1150 (2007).
- 85 H. Shull and P. O. Lowdin, *J. Chem. Phys.* **30**, 617 (1959).
- 86 E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).
- 87 W. Gautschi, *Orthogonal polynomials: Computation and approximation*. (Oxford UP, New York, 2004).
- 88 W. Gautschi, *Acta Numerica* **5**, 45 (1996).
- 89 A. Wintner, *J. Chem. Phys.* **16**, 405 (1948).
- 90 A guided derivation of Wintner's nodal theorem can be found in problem IX at www.chemistry.mcmaster.ca/courses/3bb3/z_HW/Set%202A.pdf
- 91 Wintner's theorem allows orbitals with more nodes to be lower than orbitals with fewer nodes; e.g., Wintner's theorem does not preclude the 4s orbital being lower in energy than the 3d orbital. But, in general, higher energy indicates additional nodes.
- 92 This is based on the nodal structure of orthogonal polynomials in Cartesian coordinates. Notice, however, that the same result holds even for linear molecules, albeit not for atoms.
- 93 Because electrons with different spin are distinguishable, one does not have to antisymmetrize this product. It is conventional because spin-eigenfunctions have this form.
- 94 H. Yserentant, *Numerische Mathematik* **105**, 659 (2007).
- 95 H. Yserentant, *Numerische Mathematik* **101** (2), 381 (2005).
- 96 P. W. Ayers, unpublished.
- 97 There is nothing special about DOCI. It is used as an illustration merely because with this restriction the number of Slater determinants is small enough to list.

- 98 T. Helgaker, P. Jørgensen, and J. Olsen, *Modern electronic structure theory*. (Wiley, Chichester, 2000).
- 99 H. J. Changlani, J. M. Kinder, C. J. Umrigar, and G. K. L. Chan, *Phys. Rev. B* **80**, 245116 (2009).
- 100 W. Mizukami, Y. Kurashige, and T. Yanai, *J. Chem. Phys.* **133**, 091101 (2010).
- 101 Y. Kurashige and T. Yanai, *J. Chem. Phys.* **130**, 234114 (2009).
- 102 T. Yanai, Y. Kurashige, D. Ghosh, and G. K. L. Chan, *Int. J. Quantum Chem.* **109**, 2178 (2009).
- 103 N. Schuch and J. I. Cirac, *Phys. Rev. A* **82**, 012314 (2010).
- 104 D. Aharonov, I. Arad, and S. Irani, *Phys. Rev. A* **82**, 012315 (2010).
- 105 N. Schuch and F. Verstraete, *Nature Physics* **5**, 732 (2009).
- 106 The key is that we are treating the true molecular Hamiltonian. The provocative ref. 96 treats ELECTRONS but uses a Hubbard-type Hamiltonian, not the actual molecular Hamiltonian of interest to chemists. It is perhaps surprising that solving the real-space Hamiltonian is EASIER than solving a discretized model Hamiltonian. However, there are many problems (e.g., in integer programming) that are computationally easy when the variables are allowed to be real numbers (like the electronic positions) but practically impossible when the variables are restricted to discrete states/sites. The conclusion of reference 96 does not apply for real molecular systems.

Chapter IX

Approaching the Theoretical Limits

of

Computational Efficiency with the

Local Schrödinger Equation^{*}

^{*}The content of this chapter is in preparation to be submitted: **J. S. M. Anderson**, H. Nakashima, P. W. Ayers, H. Nakatsuji “Approaching the theoretical limits of computational efficiency with the local Schrödinger equation”; *J. Chem. Phys.*

IX.I. Statement of the Problem

This chapter illustrates how to increase the efficiency of Nakatsuji's local Schrödinger equation (LSE) method. This key insight comes from an old theorem of Boys', which shows that using multidimensional quadrature grids not only increases the accuracy of the approach, but provides an explicit error estimate. The error estimate leads to a strategy for increasing the accuracy of the LSE. Boys' theorem suggests that using sparse grids in the LSE will produce a method that approaches the theoretical limits of computational efficiency for solving the molecular electronic structure problem. Combining sparse grids with this method might be the most effective method for exploiting the functional-analytic structure of the molecular wavefunction.

IX.II. Motivation

Describing how electrons “bind” atomic nuclei together to form molecules, predicting molecular properties and chemical reactivity, and quantifying the relative thermodynamic and kinetic stability of different molecules: these tasks, and many others at the heart of chemistry, are fundamentally problems in molecular electronic structure theory. Unfortunately, constructing accurate models for a molecule's electronic structure is difficult, with the biggest obstacle being the accurate description of electron correlation, which requires going beyond the usual orbital models (e.g. Hartree-Fock and approximate Kohn-Sham density- functional theory). Many approximate models for describing electron correlation exist, ranging from quantum Monte-Carlo to the configuration-interaction and coupled-cluster hierarchies.¹⁻³ Using these algorithms, the computational cost required to ensure that the error in the energy is less than ϵ typically grows exponentially with the number of electrons, N (cost $\sim \epsilon^{-pN}$, with $p \geq 1$), which is prohibitive.⁴ However, atomic and molecular systems are special, and possess certain simplifying features.⁵⁻⁷ This suggests that the most popular computational methods are less than optimal for the specific problem of determining molecules' electronic structure. In particular, conventional computational methods can be used for any electronic Hamiltonian, including model Hamiltonians (e.g., spin-Hamiltonians) that are known to be much more difficult to solve than the actual molecular Hamiltonian.⁸ Exploiting the special characteristics of molecular systems to develop more efficient methods for computational quantum chemistry is the chief goal of this research.

In order to design an algorithm for the electronic structure problem that has only polynomial scaling (cost $\sim \varepsilon^{-k}$, ideally $k \ll 1$), we must exploit the fact that the molecular wavefunction has a high degree of differentiability.⁵⁻⁷ Since differentiability is a real-space property, it is most natural to treat electrons in real space, as this will facilitate using the specific properties of the molecular Hamiltonian,

$$\hat{H}(\boldsymbol{\tau}) = \sum_{i=1}^N \left(-\frac{1}{2} \nabla_i^2 + v(\mathbf{r}_i) + \sum_{j=1}^{i-1} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \quad (1)$$

The approach we use in this work will be based on the Local Schrödinger equation,⁹⁻¹¹ which is essentially a collocation method for the Schrödinger equation. A similar approach has been used by Avila and Carrington for the nuclear Schrödinger equation.¹² Alternative methods for exploiting the differentiability of the wavefunction have been proposed by Garcke, Hamaeker, and Griebel,¹³⁻¹⁵ as well as the present authors.⁴

The Schrödinger equation can be written as

$$\left(\hat{H}(\boldsymbol{\tau}) - E_n \right) \Psi_n(\boldsymbol{\tau}) = 0 \quad (2)$$

This equation must hold at all points in the $3N$ -dimensional coordinate space of electrons. Let us pick some set of G points; these could be points from an integration grid. This gives a set of G simultaneous equations,

$$\left\{ \left(\hat{H} - E_n \right) \Psi_n(\boldsymbol{\tau}_g) = 0 \right\}_{g=1}^G \quad (4)$$

These equations are necessary conditions for the underlying Schrödinger equation, (2). Because $\hat{H}(\boldsymbol{\tau})$ is a linear operator, it can be approximated as,

$$\hat{H}\Psi(\boldsymbol{\tau}_f) \approx \sum_{g=1}^G H_{fg} \Psi(\boldsymbol{\tau}_g) \quad (5)$$

and Eq. (4) can be rewritten as

$$\left\{ \sum_{g=1}^G \left(H_{fg} - \delta_{fg} E_n \right) \Psi_n(\boldsymbol{\tau}_g) = 0 \right\}_{f=1}^G. \quad (6)$$

Computing the action of the Hamiltonian on a wavefunction implicitly requires numerical differentiation on the grid, and our previous studies suggests that the grids used for quantum chemical calculations are much more accurate for numerical integration than numerical differentiation.¹⁶ Direct implementation of Eq. (6) is likely to be problematic.

IX.III. Boys' Collocation for the Local Schrödinger Equation

We can avoid the need to resolve the Hamiltonian operator on the grid by expanding the wavefunctions in a basis set containing B properly antisymmetric basis functions,

$$\Psi_n(\boldsymbol{\tau}) = \sum_{b=1}^B c_{bn} \phi_b(\boldsymbol{\tau}). \quad (7)$$

Inserting Eq. (7) into Eq. (4) gives

$$\left\{ \sum_{b=1}^B (\hat{H} - E_n) \phi_b(\boldsymbol{\tau}_g) c_{bn} = 0 \right\}_{g=1}^G \quad (8)$$

Because the number of basis functions is usually much smaller than the number of grid points, this equation can only be solved in the least-squares sense. Symbolically, we can write Eq. (8) as

$$(\mathbf{A} - E_n \mathbf{F})^{(G \times B)} \mathbf{c}_n^{(B \times 1)} = \mathbf{0}^{(G \times 1)} \quad (9)$$

where we have explicitly indicated the dimensionality of the matrices and defined

$$\begin{aligned} \mathbf{A} &= \hat{H} \phi_b(\boldsymbol{\tau}_k) \\ \mathbf{F} &= \phi_b(\boldsymbol{\tau}_k) \end{aligned} \quad (10)$$

If the action of the Hamiltonian on the basis functions can be computed analytically, this formulation avoids the problem of numerical differentiation on the grid.

To solve Eq. (9), we need to multiply the equation on the left by a $B \times G$ matrix. The matrix must be chosen carefully to avoid introducing spurious solutions to the equations. The most obvious choice is to exploit the analogy to least squares, and define a weighted-least-squares system,

$$\left[(\mathbf{A} - E_n \mathbf{F})^\dagger \right]^{(B \times G)} \mathbf{W}^{(G \times G)} (\mathbf{A} - E_n \mathbf{F})^{(G \times B)} \mathbf{c}_n^{(B \times 1)} = \left[(\mathbf{A} - E_n \mathbf{F})^\dagger \right]^{(B \times G)} \mathbf{W}^{(G \times G)} \mathbf{0}^{(G \times 1)} \quad (11)$$

which simplifies to

$$\left[(\mathbf{A} - E_n \mathbf{F})^\dagger \right]^{(B \times G)} \mathbf{W}^{(G \times G)} (\mathbf{A} - E_n \mathbf{F})^{(G \times B)} \mathbf{c}_n^{(B \times 1)} = \mathbf{0}^{(B \times 1)} \quad (12)$$

\mathbf{W} can be any positive definite matrix.¹⁷ Suppose \mathbf{W} is diagonal, $\mathbf{W} = \delta_{jk} w_k$. Then Eq. (9) can be rewritten as

$$\left\{ \begin{array}{l} \sum_{b=1}^B \sum_{f=1}^G \sum_{g=1}^G \phi_a(\tau_f) (\hat{H} - E_n) w_g \delta_{fg} (\hat{H} - E_n) \phi_b(\tau_g) c_{bn} = \mathbf{0} \\ \sum_{b=1}^B \sum_{g=1}^G \phi_a(\tau_g) (\hat{H} - E_n) w_g (\hat{H} - E_n) \phi_b(\tau_g) c_{bn} = \mathbf{0} \end{array} \right\}_{\substack{a=1 \\ b=1}}^B \quad (13)$$

If the w_g are proportional to the quadrature weights for numerical integration on the grid, then this equation is the grid representation of the \hat{H} -squared condition for the solutions of the Schrödinger equation,

$$\sum_{b=1}^B \left\langle \phi_a \left| (\hat{H} - E_n)^2 \right| \phi_b \right\rangle c_{bn} = 0 = \left\langle \phi_a \left| (\hat{H} - E_n)^2 \right| \Psi_n \right\rangle. \quad (14)$$

Equation (13) is a standard “trick” in numerical analysis. The recent work of Boutry *et al.*¹⁸ provides a good algorithm for this problem, which was recently applied in the work of Manzhos *et al.* to the nuclear Schrödinger equation.^{19,20}

Recall that we can multiply Eq. (9) on the left by any $B \times G$ matrix. A different choice, which dates back to Boys,²¹ is to multiply on the left by $\mathbf{F}^\dagger = \phi_b^*(\tau_g)$ or, more generally, $\mathbf{F}^\dagger \mathbf{W}$, where \mathbf{W} is positive definite. With this choice, one has

$$\left[\mathbf{F}^\dagger \right]^{(B \times G)} \mathbf{W}^{(G \times G)} (\mathbf{A} - E_n \mathbf{F})^{(G \times B)} \mathbf{c}_n^{(B \times 1)} = \mathbf{0}^{(B \times 1)}. \quad (15)$$

If we assume the \mathbf{W} matrix is diagonal, then this equation has the simple and explicit form,

$$\left\{ \sum_{b=1}^B \sum_{g=1}^G \phi_a(\tau_g) w_g (\hat{H} - E_n) \phi_b(\tau_g) c_{bn} = \mathbf{0} \right\}_{a=1}^B \quad (16)$$

If the w_g are proportional to the quadrature weights for numerical integration on the grid, then this can be rewritten as

$$\sum_{b=1}^B \left\langle \phi_a \left| \hat{H} - E_n \right| \phi_b \right\rangle c_{bn} = 0 = \left\langle \phi_a \left| \hat{H} - E_n \right| \Psi \right\rangle \quad (17)$$

which is just what one obtains if one multiplies Eq. (2) by some “test functions,” $\{\phi_b(\tau)\}_{b=1}^B$ and integrates. Equation (16) is a type of collocation method for the Schrödinger equation.

The general idea of applying collocation to the Schrödinger equation was popularized by Yang and Peet (for the nuclear Schrödinger equation)²²⁻²⁶, but it dates back to the work of Boys, who noted the link to numerical integration. It was applied, in the context of density-functional theory, by McCormack *et al.*²⁷ More recently, this “local Schrödinger equation” approach has been used by

Nakatsuji *et al.* to obtain highly accurate solutions to the electronic Schrödinger equation for atoms and molecules.¹¹ It is this work that is most relevant to the present paper.

Any positive-definite \mathbf{W} matrix can be used. However, the error in the solutions to Eq. (16) will decrease if one makes a good choice for \mathbf{W} . A theorem of Boys is useful in this regard. He showed that the error in the energy eigenvalues from Eq. (16) can be written as

$$\varepsilon_n \sim (\mu_n + \mu_n^Q) \mu_n, \quad (18)$$

where ε_n is the error in the energy in Eq. (17), μ_n^Q is the largest fractional error in the numerical quadrature, and μ_n is the error in the approximate wavefunction,

$$\Psi_n(\boldsymbol{\tau}) \approx \sum_{b=1}^B c_{bn} \phi_b(\boldsymbol{\tau}), \quad (19)$$

due to basis-set truncation. To make the numerical integrations in Eq. (17) as accurate as possible, one should choose the collocation points from a numerical integration grid and the associated weights, w_g , should be proportional to the quadrature weights. Note that the method will converge to the right solution in the basis-set limit even if a very poor integration grid is used. However, convergence will be much faster if an accurate quadrature grid is used, ideally one with an error comparable to the error from basis-set truncation.

Boys' result is actually more general. He developed his collocation method in the context of the transcorrelated Hamiltonian, which is a non-Hermitian similarity-transformed Hamiltonian.²⁸ The transcorrelated Schrödinger equation can be rewritten as a special case of the scaled Schrödinger equation,²⁹

$$U(\boldsymbol{\tau}) (\hat{H}(\boldsymbol{\tau}) - E_n) \Psi_n(\boldsymbol{\tau}) = 0 \quad (20)$$

Though the transcorrelated Hamiltonian is still an active area of research,³⁰⁻³⁴ the primary goal of the transcorrelated approach—removing the Coulomb divergences from the Hamiltonian—may be achieved more simply by choosing $U(\boldsymbol{\tau}) = (V_{ne}(\boldsymbol{\tau}) + V_{ee}(\boldsymbol{\tau}))^{-1}$, $U(\boldsymbol{\tau}) = (V_{ne}(\boldsymbol{\tau}))^{-1} + (V_{ee}(\boldsymbol{\tau}))^{-1}$, or something similar. Here $V_{ne}(\boldsymbol{\tau})$ and $V_{ee}(\boldsymbol{\tau})$ are the electron-nuclear and electron-electron potentials, respectively.²⁹ With the scaled Schrödinger equation, Eq. (16) becomes

$$\left\{ \sum_{b=1}^B \sum_{g=1}^G \phi_a(\boldsymbol{\tau}_g) w_g \left[U(\hat{H} - E_n) \phi_b(\boldsymbol{\tau}_g) \right] c_{bn} = 0 \right\}_{a=1}^B \quad (21)$$

If the w_k represent integration weights, then this can be written as a local scaled-Schrödinger equation,

$$\left\{ \sum_{b=1}^B \langle \phi_a | U (\hat{H} - E_n) | \phi_b \rangle c_{bn} = 0 \right\}_{a=1}^B \quad (22)$$

Equation (22) may be cast in the form of a secular equation,

$$(\mathbf{H} - E_n \mathbf{S})^{(B \times B)} \mathbf{c}_n^{(B \times 1)} = \mathbf{0} \quad (23)$$

by defining,

$$\mathbf{H} = [h_{ab}]$$

$$h_{ab} = \sum_{g=1}^G w_g \phi_a(\tau_g) (U \hat{H} \phi_b(\tau_g)) \approx \langle \phi_a | U \hat{H} | \phi_b \rangle \quad (24)$$

$$\mathbf{S} = [s_{ab}]$$

$$s_{ab} = \sum_{g=1}^G w_g \phi_a(\tau_g) U(\tau_g) \phi_b(\tau_g) \approx \langle \phi_a | U | \phi_b \rangle \quad (25)$$

The error in the energy eigenvalues from the secular equation is

$$\varepsilon_n^U \sim (\mu_n^\dagger + \mu_n^Q) \mu_n \quad (26)$$

where μ_n and μ_n^\dagger are the errors in expanding the n^{th} right-hand-side and left-hand-side eigenvectors of the scaled Schrödinger equation. The conventional collocation approach to the Schrödinger equation is regained by setting $U(\tau) = 1$ in Eqs. (24) and (25). Notice that from the standpoint of Eq. (15), the scaled Schrödinger equation is just a very special choice of the weight matrix, with

$$\mathbf{W} = [w_{fg}] = w_g U(\tau_g) \delta_{fg}. \quad (27)$$

The disadvantage of the scaled Schrödinger equation is that, because $U \hat{H}$ is not Hermitian, the \mathbf{H} matrix is not Hermitian. Determining the \mathbf{H} matrix therefore requires B^2 numerical integrations, as opposed to only $B(B+1)/2$ when $U = 1$. So there are 50% more numerical integrations in the scaled Schrödinger equation case. This is compensated for by the fact that the integrands are now nonsingular, and so it will take many fewer points to achieve a given integration accuracy, μ^Q .

IX.IV. Algorithm: Sparse Numerical Integration Grids

Setting up the secular equation in Eq. (23) requires performing integrals over the basis functions,

$$h_{ab} = \int \phi_a^*(\tau) U(\tau) \hat{H}(\tau) \phi_b(\tau) d\tau \quad (28)$$

$$s_{ab} = \int \phi_a(\tau) U(\tau) \phi_b(\tau) d\tau \quad (29)$$

These are $3N$ dimensional integrals, where N is the number of electrons. Numerical integration in many dimensions is a challenging problem. One might try to extend the established numerical integration techniques from, for example, density-functional theory.³⁵⁻³⁹ These grids contain between 10^3 and 10^4 points per atom. Simply taking a direct product of these 3-dimensional grids to form a $3N$ -dimensional grid would then involve on the order of 10^{4N} points; this approach can be used for at most three electrons. This example illustrates that most methods for numerical integration in many dimensions and, in particular, all methods that use the direct products of lower-dimensional grids, suffer from the curse of dimension: the number of points required to obtain a given quadrature accuracy grows exponentially with increasing electron number,

$$G_{\text{direct product}} \sim (\mu^{\varrho})^{-N}. \quad (30)$$

The best-known way to overcome the curse of dimension is to use Monte Carlo integration, where the integration points, $\{\tau_g\}_{g=1}^G$ are weighted equally and randomly distributed according to some statistical distribution function. The resulting cost

$$G_{\text{Monte Carlo}} \sim (\mu^{\varrho})^{-2} \quad (31)$$

is optimal for the integration of general functions. Equation (31) indicates that doubling the accuracy of the numerical integrations requires quadrupling the number of integration points.

The only way to improve upon this scaling is to use the scaled Schrödinger equation; with the scaled Schrödinger equation one can choose $U(\tau)$ so that the integrands in Eq. (28) are continuous. For functions that are continuous but not differentiable, the optimal integration method is quasi-Monte Carlo, with

$$G_{\text{quasi-MC}} \sim (\mu^{\varrho})^{-1} \quad (32)$$

Doubling the integration accuracy now requires only doubling the grid size.

In quantum chemistry, however the integrands are usually not just continuous; they are actually differentiable. In particular, the integrands have bounded mixed derivatives.⁷ A function, $f(\tau)$ is said to have a bounded mixed derivative of l^{th} order if

$$\left\| \frac{\partial^{n_x} \partial^{n_y} \dots \partial^{n_N} f(x_1, y_1, z_1, \dots, x_N, y_N, z_N)}{\partial x_1^{n_x} \partial y_1^{n_y} \partial z_1^{n_z} \partial x_2^{n_{2x}} \partial y_2^{n_{2y}} \partial z_2^{n_{2z}} \dots \partial x_N^{n_{Nx}} \partial y_N^{n_{Ny}} \partial z_N^{n_{Nz}}} \right\|_\infty < \infty \quad (33)$$

for all derivatives for which

$$\|\mathbf{n}\|_\infty = \max_{\{1 \leq i \leq N\}} \left(\max(n_{ix}, n_{iy}, n_{iz}) \right) \quad (34)$$

Notice that the existence of a bounded mixed derivative implies that some derivatives of very high orders exist. For example, it implies that the $3Nt$ -th derivative is essentially bounded,

$$\left\| \frac{\partial^t \partial^t \dots \partial^t f(x_1, y_1, z_1, \dots, x_N, y_N, z_N)}{\partial x_1^t \partial y_1^t \partial z_1^t \dots \partial x_N^t \partial y_N^t \partial z_N^t} \right\|_\infty < \infty \quad (35)$$

Recall that the $\mathbb{L}^\infty(\mathbb{R}^{3N})$ norm of a function, $\|f(\boldsymbol{\tau})\|_\infty$ is the essential supremum of $|f(\boldsymbol{\tau})|$. For continuous functions, the essential supremum is the least-upper-bound on $|f(\boldsymbol{\tau})|$.

If the t^{th} bounded mixed derivatives exist, then there are $(t+1)^{3N}$ bounded derivatives in Eq. (33). Because the function has an exponential amount of differentiability, one can design formulas for integrating the function with exponentially fewer points than there are in a direct product grid. The optimal grid is the multidimensional Gaussian quadrature grid and the number of grid points that is needed is

$$G_{\text{Gaussian}} \sim (\mu^e)^{-1/t}. \quad (36)$$

For highly differentiable functions, Gaussian quadrature is immensely better than Monte-Carlo-type methods.

Unfortunately, multidimensional Gaussian quadrature grids do not always exist and, even when they do, the difficulty of finding them grows exponentially with the number of dimensions.⁴⁰ Fortunately there are methods that are “almost as good as” Gaussian quadrature methods, in the sense that they deviate from Eq. (36) by at most a logarithmic factor of N .⁴¹⁻⁴³ The first method of this type, and the prototype for all future methods, was designed by Smolyak.⁴⁴ The simple Smolyak-type formulas have been extended by the groups of Griebel, Petras, and others.⁴⁵⁻⁴⁸

Our approach to Smolyak-type integration has been described in detail elsewhere, and will only be briefly sketched here.^{16,49-51} The starting point is a one-dimensional integration grid for integration functions on $[0,1]$

$$\sum_{i=1}^{m(q)} w_i f(x_i) \approx \int_0^1 f(x) dx. \quad (37)$$

Here $\{x_i\}_{i=1}^{m(q)}$ are the integration points; $\{w_i\}_{i=1}^{m(q)}$ are the integration weights, $m(q)$ is the total number of points in the grid, and $q=1,2,\dots$ label the grids. The k^{th} grid must have polynomial degree of at least $2k-1$. It is possible, and indeed wise, to reuse the same grid; i.e., one can have $m(k) = m(k+1)$.

The one-dimensional formulas are then combined to form a 3-dimensional grid on the unit cube using a Smolyak-like formula. The key is to define the “difference” grids,

$$\Delta_x^{(q)} f = \sum_{i=1}^{m(q)} w_i f(x_i) - \sum_{i=1}^{m(q-1)} w_i f(x_i) = \sum_{i=1}^{m(q)} v_i f(x_i); \quad (38)$$

the $q = 0$ grid obviously has zero points, $m(0) = 0$. Eq. (38) corrects for the error in the integration grid of effort $q - 1$. The last equality is true only if one has nested grids, but we only use nested grids because non-nested grids are not nearly as efficient.

The three dimensional grids are then formed as

$$\sum_{I(q_x, q_y, q_z) \leq Q} \Delta_x^{(q_x)} \otimes \Delta_y^{(q_y)} \otimes \Delta_z^{(q_z)} f = \sum_{I=1}^{M(Q)} w_I f(x_I, y_I, z_I) \quad (39)$$

The specific index set, which controls which formulas appear in the integration grid, is chosen based on the smoothness of the function. The more differentiable the function is, the fewer integration points one needs. The simplest (and most expensive) case is the original Smolyak formula.⁴⁴ In that case, one uses

$$I(q_x, q_y, q_z) = q_x + q_y + q_z - 3 \quad (40)$$

or more generally, for d dimensions,

$$I(q_1, q_2, \dots, q_d) = \sum_{i=1}^d q_i - d. \quad (41)$$

From Eq. (39) we have a grid on the unit cube, but we would like to integrate over real space. We perform a coordinate transformation using the same importance sampling trick that is used to map random numbers on $[0,1]^d$ to those generated by a certain probability distribution in Monte Carlo theory,⁵²

$$\begin{aligned}
 \Theta_1(x) &= \frac{\int_{-\infty}^x \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(X, Y, Z) dZdYdX}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(X, Y, Z) dZdRYdX} \\
 \Theta_2(x, y) &= \frac{\int_{-\infty}^y \int_{-\infty}^{\infty} p(x, Y, Z) dZdY}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, Y, Z) dZdY} \\
 \Theta_3(x, y, z) &= \frac{\int_{-\infty}^z p(x, y, Z) dZ}{\int_{-\infty}^{\infty} p(x, y, Z) dZ}
 \end{aligned} \tag{42}$$

where $(\Theta_1, \Theta_2, \Theta_3) \in [0, 1]^3$ is a point in the unit cube and $p(x, y, z)$ is a probability distribution function that approximates the integrand. The quadrature formula becomes

$$\int f(\mathbf{r}) p(\mathbf{r}) d\mathbf{r} \approx \sum_{I=1}^{M(Q)} w_I f(\mathbf{r}_I); \tag{43}$$

This is most accurate when $f(\mathbf{r})$ is slowly varying.^{16,49,50} We use the promolecular electron density⁵³ or the valence electron density for $p(x, y, z)$.

We call Eq. (43) a one-electron grid; it is analogous to the one-electron grids in density-functional theory but, unlike those grids, it has a well-defined polynomial degree. Because the one-electron grids have a defined polynomial degree, we can apply the Smolyak formula in its general form (cf Eq. (41)) to form a $3N$ -dimensional grid,

$$\sum_{I(Q_1, Q_2, \dots, Q_N) \leq Q} \Delta_{\mathbf{r}_1}^{(Q_1)} \otimes \Delta_{\mathbf{r}_2}^{(Q_2)} \otimes \dots \otimes \Delta_{\mathbf{r}_N}^{(Q_N)} f = \sum_{I=1}^{M(Q)} w_I f(\mathbf{r}_{1I}, \mathbf{r}_{2I}, \dots, \mathbf{r}_{NI}) \tag{44}$$

In making this grid, one generates many points that are symmetry-equivalent. We generate only one permutation of the grid points by (a) forcing $Q_1 \leq Q_2 \leq \dots \leq Q_{N_\alpha}$ and $Q_1 \leq Q_2 \leq \dots \leq Q_{N_\beta}$ in Eq. (44), (b) using the indices for each one-electron difference grid, choosing only the points with $I_1 \leq I_2 \leq \dots \leq I_{N_\alpha}$ and $I_1 \leq I_2 \leq \dots \leq I_{N_\beta}$, and (c) respecting the symmetry with respect to exchange of α -spin and β -spin electrons in singlet states. Each symmetry-unique grid point is then reweighted by a factor that captures the number of possible permutations of the coordinates. Our approach to treating the antisymmetry of the electronic coordinates and treats all electrons equally. An unpleasant consequence of this approach is that points where electrons are at the same point in space appear in the grids. (I.e., because of point (b) above, the grids contain $\mathbf{r}_{I_k} = \mathbf{r}_{I_{k+1}}$. Because the

electron-electron potential diverges at these points, Smolyak-style grids can be applied only to the scaled Schrödinger equation.

IX.V. Algorithm: Free Iterative Complement Basis Functions

Boys' collocation method frees the calculation from the hegemony of Gaussian basis sets: the integrals do not have to be done analytically. One can envision using many different types of basis functions, e.g., Slater-Jastrow functions. Recall that the numerical integration error will be smallest when the integrands are highly differentiable (cf. Eq. (36)). This means that we should choose basis functions that accommodate the electron-electron and electron-nuclear cusps that arise from the Coulomb singularities in the Hamiltonian. The free iterative complement functions are one way to do this.¹¹

Imagine solving the scaled Schrödinger equation by a Krylov subspace method. One starts with an initial basis function, $\tilde{\Psi}^{(0)}(\tau)$, and applies the scaled-Schrödinger operator to it iteratively. This generates a Krylov subspace for the right-side eigenvectors,

$$\left\{ \tilde{\Psi}^{(0)}(\tau), U\hat{H}\tilde{\Psi}^{(0)}(\tau), (U\hat{H})^2\tilde{\Psi}^{(0)}(\tau), \dots \right\} \quad (45)$$

and for the left-side eigenvectors

$$\left\{ \tilde{\Psi}^{(0)}(\tau), \hat{H}U\tilde{\Psi}^{(0)}(\tau), (\hat{H}U)^2\tilde{\Psi}^{(0)}(\tau), \dots \right\} \quad (46)$$

In addition, there is a Krylov subspace for $U(\bar{\tau})$, which plays the role of a metric in this theory,

$$\left\{ \tilde{\Psi}^{(0)}(\tau), U\tilde{\Psi}^{(0)}(\tau), (U)^2\tilde{\Psi}^{(0)}(\tau), \dots \right\} \quad (47)$$

The iterative complement method is a method, with a short recursive length, for solving the Schrödinger equation in these subspaces. The free iterative complement method recognizes that because the wavefunction, $\tilde{\Psi}^{(0)}(\tau)$, is generally a sum of terms (e.g., a sum of several Slater determinants) and the Hamiltonian is also a sum of terms, $\hat{H}U\tilde{\Psi}^{(0)}$ is also a sum of terms. One way to expand the iterative subspace, then, is to treat every term with a new analytic structure as a basis function. Suppose the initial guess wavefunction was a sum of $J^{(0)}$ terms,

$$\tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(0)}} k_j^{(0)} \phi_j^{(0)}(\tau) \quad (48)$$

Applying $U\hat{H}$ to this function generates a new sum,

$$U\hat{H}\tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(1,\text{right})}} k_j^{(1,\text{right})} \phi_j^{(1,\text{right})}(\tau) \quad (49)$$

The number of terms in Eq. (49) is significantly greater than the number of terms in Eq. (48). Similarly, for the left-hand eigenspace, one has:

$$\hat{H}U\tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(1,\text{left})}} k_j^{(1,\text{left})} \phi_j^{(1,\text{left})}(\tau). \quad (50)$$

and for $\hat{U}(\bar{\tau})$ one has

$$U\tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(1,U)}} k_j^{(1,U)} \phi_j^{(1,U)}(\tau) \quad (51)$$

Iterating in this manner, one has

$$(U\hat{H})^n \tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(n,\text{right})}} k_j^{(n,\text{right})} \phi_j^{(n,\text{right})}(\tau) \quad (52)$$

$$(\hat{H}U)^n \tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(n,\text{left})}} k_j^{(n,\text{left})} \phi_j^{(n,\text{left})}(\tau). \quad (53)$$

$$(U)^n \tilde{\Psi}^{(0)}(\tau) = \sum_{j=1}^{j^{(n,U)}} k_j^{(n,U)} \phi_j^{(n,U)}(\tau) \quad (54)$$

At n^{th} order, an appropriate basis set of free iterative complement functions for Eq. (21) is

$$\begin{aligned} \mathcal{F}^{(n,\text{right})} &\equiv \left\{ \left\{ \phi_j^{(0)} \right\}_{j=1}^{j^{(0)}} \cup \left\{ \phi_j^{(1,\text{right})} \right\}_{j=1}^{j^{(1,\text{right})}} \cup \dots \cup \left\{ \phi_j^{(n,\text{right})} \right\}_{j=1}^{j^{(n,\text{right})}} \right\} \\ &\quad \cup \left\{ \left\{ \phi_j^{(0)} \right\}_{j=1}^{j^{(0)}} \cup \left\{ \phi_j^{(1,U)} \right\}_{j=1}^{j^{(1,U)}} \cup \dots \cup \left\{ \phi_j^{(n,U)} \right\}_{j=1}^{j^{(n,U)}} \right\} \\ &= \{ \phi_1^{\text{right}}, \phi_2^{\text{right}}, \dots, \phi_B^{\text{right}} \} \end{aligned} \quad (55)$$

$$\begin{aligned} \mathcal{F}^{(n,\text{left})} &\equiv \left\{ \left\{ \phi_j^{(0)} \right\}_{j=1}^{j^{(0)}} \cup \left\{ \phi_j^{(1,\text{left})} \right\}_{j=1}^{j^{(1,\text{left})}} \cup \dots \cup \left\{ \phi_j^{(n,\text{left})} \right\}_{j=1}^{j^{(n,\text{left})}} \right\} \\ &\quad \cup \left\{ \left\{ \phi_j^{(0)} \right\}_{j=1}^{j^{(0)}} \cup \left\{ \phi_j^{(1,U)} \right\}_{j=1}^{j^{(1,U)}} \cup \dots \cup \left\{ \phi_j^{(n,U)} \right\}_{j=1}^{j^{(n,U)}} \right\} \\ &= \{ \phi_1^{\text{left}}, \phi_2^{\text{left}}, \dots, \phi_B^{\text{left}} \} \end{aligned} \quad (56)$$

These basis sets are much larger than the union of the Krylov subspaces in Eq. (45). The number of basis functions grows geometrically with n , but the growth is

mitigated by the fact that many forms appear repeatedly. Because the free iterative complement basis set is defined as a union of analytic forms, any time a function reappears, it need not be used for future iterations. The free iterative complement basis is likely to be complete for the space of eigenfunctions that have the same symmetry as the initial seed, $\tilde{\Psi}^{(0)}$.

Equation (21) can now be written explicitly in terms of the free-complement basis,

$$\left\{ \sum_{b=1}^B \sum_{g=1}^G \phi_a^{\text{left}}(\tau_g) w_g \left[U(\hat{H} - E_n) \phi_b^{\text{right}}(\tau_g) \right] c_{bn} = 0 \right\}_{a=1}^B \quad (57)$$

Notice that one needs to use the appropriate (right- or left-side basis functions, respectively) in Eq. (57). However, because the error in the energy is more sensitive to the error in the right-hand-side basis set than it is to the error in the left-hand-side basis set (cf. Eq. (26)), it is possible to use either a smaller basis, or a basis set of lower quality, on the left. Because the left-hand-side basis functions may diverge when two electrons are at the same position, using the right-hand free iterative complement functions as left-side basis functions is suggested.⁹⁻¹¹ Alternatively, one can use only the subset of $\mathcal{F}^{(n,\text{left})}$ that does not diverge.

IX.VI. Summary

According to conventional wisdom, determining the wave functions and energies of molecular systems' stationary states is, even with the non-relativistic and Born-Oppenheimer approximations, intractable: the computational effort required to achieve a given level of accuracy grows exponentially with the number of electrons in the system. The present method challenges this paradigm. First of all, the free complement functions are an explicitly correlated superset of the Krylov subspace basis and, as such, provide a very compact representation for the wavefunction, without the factorial number of terms in ordinary wavefunction expansions. The free complement functions, however, are not analytically integrable. One way to overcome this problem is to use the local Schrödinger equation, or Boys' collocation method. Obtaining good results with Boys' collocation requires that one be able to accurately perform the integrals in Eqs. (28) and (29). This is problematic because the free-complement functions are not analytically integrable functions. This problem is circumvented using Smolyak-inspired methods for multidimensional numerical integration. In Smolyak integration, the number of grid points does not grow exponentially with increasing electron number. The combination of these two techniques: expansion of the wavefunction in a nearly-optimal basis set and numerical integration with a

nearly-optimal grid, approaches the theoretical limits of computational efficiency for solving the electronic Schrödinger equation.

IX.VII. References

- 1 K. Raghavachari and J. B. Anderson, *J. Phys. Chem.* **100**, 12960 (1996).
- 2 M. Head-Gordon, *J. Phys. Chem.* **100**, 13213 (1996).
- 3 T. Helgaker, P. Jørgensen, and J. Olsen, *Modern electronic structure theory*. (Wiley, Chichester, 2000).
- 4 J. S. M. Anderson and P. W. Ayers, in preparation.
- 5 H. Yserentant, *Numerische Mathematik* **105**, 659 (2007).
- 6 H. Yserentant, *Numerische Mathematik* **101** (2), 381 (2005).
- 7 H. Yserentant, *Numerische Mathematik* **98** (4), 731 (2004).
- 8 N. Schuch and F. Verstraete, *Nature Physics* **5**, 732 (2009).
- 9 H. Nakatsuji and H. Nakashima, *Int. J. Quantum Chem.* **109**, 2248 (2009).
- 10 H. Nakashima and H. Nakatsuji, *Phys. Rev. Lett.* **101** (240406) (2008).
- 11 H. Nakatsuji, H. Nakashima, Y. Kurokawa, and A. Ishikawa, *Phys. Rev. Lett.* **99**, 240402 (2007).
- 12 G. Avila and T. Carrington, *J. Chem. Phys.* **131**, 174103 (2009).
- 13 J. Garcke and M. Griebel, *J. Comput. Phys.* **165**, 694 (2000).
- 14 M. Griebel and J. Hamaekers, *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics* **224**, 527 (2010).
- 15 M. Griebel and J. Hamaekers, *Esaim-Mathematical Modelling and Numerical Analysis-Modelisation Mathematique Et Analyse Numerique* **41** (2), 215 (2007).
- 16 J. S. M. Anderson, J. I. Rodriguez, D. C. Thompson, and P. W. Ayers, in *Quantum Chemistry Research Trends* (Nova, Hauppauge, NY, 2007).
- 17 W can be even be positive semidefinite, as long as the number of positive eigenvalues is large enough. If W is indefinite then the equation could have additional solutions.
- 18 G. Boutry, M. Elad, G. H. Golub, and P. Milanfar, *Siam Journal on Matrix Analysis and Applications* **27**, 582 (2005).
- 19 S. Manzhos and T. Carrington, *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **87**, 864 (2009).
- 20 S. Manzhos, K. Yamashita, and T. Carrington, *Chem. Phys. Lett.* **474**, 217 (2009).
- 21 S. F. Boys, *Proc. R. Soc. London, Ser. A* **309**, 195 (1969).
- 22 W. Yang and A. C. Peet, *J. Chem. Phys.* **92** (1), 522 (1990).
- 23 A. C. Peet and W. Yang, *J. Chem. Phys.* **91** (11), 6598 (1989).

- 24 A. C. Peet and W. Yang, *J. Chem. Phys.* **90** (3), 1746 (1989).
25 W. Yang, A. C. Peet, and W. H. Miller, *J. Chem. Phys.* **91** (12), 7537 (1989).
26 W. Yang and A. C. Peet, *Chem. Phys. Lett.* **153** (1), 98 (1988).
27 D. A. McCormack, E. J. Baerends, E. van Lenthe, and N. C. Handy, *Theor. Chem. Acc.* **112**, 410 (2004).
28 S. F. Boys and N. C. Handy, *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences* **310**, 43 (1969).
29 H. Nakatsuji, *Phys. Rev. Lett.* **93** (3) (2004).
30 M. Kojo and K. Hirose, *Journal of Computational and Theoretical Nanoscience* **6**, 2567 (2009).
31 N. Umezawa and S. Tsuneyuki, *J. Chem. Phys.* **119**, 10015 (2003).
32 O. Hino, Y. Tanimura, and S. Ten-no, *Chem. Phys. Lett.* **353**, 317 (2002).
33 O. Hino, Y. Tanimura, and S. Ten-no, *J. Chem. Phys.* **115**, 7865 (2001).
34 S. Ten-no, *Chem. Phys. Lett.* **330**, 169 (2000).
35 J. M. Perez-Jorda, A. D. Becke, and E. San Fabian, *J. Chem. Phys.* **100**, 6520 (1994).
36 C. W. Murray, N. C. Handy, and G. J. Laming, *Mol. Phys.* **78**, 997 (1993).
37 A. D. Becke, *J. Chem. Phys.* **88** (4), 2547 (1988).
38 S. H. Chien and P. M. W. Gill, *J. Comput. Chem.* **27**, 730 (2006).
39 P. M. W. Gill, B. G. Johnson, and J. A. Pople, *Chem. Phys. Lett.* **209**, 506 (1993).
40 Y. Xu, *Common zeros of polynomials in several variables and higher dimensional quadrature*. (Wiley New York, 1994).
41 G. W. Wasilkowski and H. Wozniakowski, *Journal of Complexity* **11**, 1 (1995).
42 E. Novak and K. Ritter, *Numerische Mathematik* **75**, 79 (1996).
43 R. Cools, E. Novak, and K. Ritter, *Computing* **62**, 147 (1999).
44 S. A. Smolyak, *Dokl. Akad. Nauk* **4**, 240 (1963).
45 H. J. Bungartz and M. Griebel, *Journal of Complexity* **15**, 167 (1999).
46 M. Griebel and S. Knapek, *Constructive Approximation* **16**, 525 (2000).
47 K. Petras, *Numerische Mathematik* **93**, 729 (2003).
48 K. Petras, *Advances in Computational Mathematics* **12**, 71 (2000).
49 J. I. Rodriguez, D. C. Thompson, J. S. M. Anderson, J. W. Thomson, and P. W. Ayers, *Journal of Physics A* **41**, 365202 (2008).
50 J. I. Rodriguez, D. C. Thompson, P. W. Ayers, and A. M. Koster, *J. Chem. Phys.* **128**, 224103 (2008).
51 D. C. Thompson and P. W. Ayers, *Int. J. Quantum Chem.* **106**, 787 (2006).
52 L. Devroye, *Non-uniform Random Variate Generation*. (Springer-Verlag, New York, 1986).
53 F. L. Hirshfeld, *Theor. Chim. Act.* **44**, 129 (1977).

Chapter X

Conclusions

and

Prospects for Future Work

X.I. Conclusions

This dissertation presents new approaches to three of the fundamental problems facing quantum chemistry: predicting molecular reactivity, partitioning molecules into atoms, and computing accurate wavefunctions. For each of these topics, the mathematical groundwork is established. For the work on predicting molecular reactivity, illustrative applications were used to prove the utility of the approaches. The efficacy of these approaches is not in doubt, but each topic is a beginning, not an end: each topic developed here can, and should, be developed further.

Chapters III-V illustrate how to compute the relative reactivity of the atoms within a molecule using the general-purpose reactivity indicator. In Chapter III this indicator was derived from and by the fundamental equation of conceptual density-functional theory. In Chapter IV the indicator was shown to successfully predict the order of molecular reactivity for several pernicious examples: molecules where simple frontier orbital theory and the more general Fukui function both fail to predict reactivity correctly. Chapter V illustrates the most important feature of this indicator, the ability to distinguish between sites that react with hard reagents and those that react with soft reagents. This indicator is easily computed from any quantum chemistry software package that computes the electrostatic potential or atomic populations. I am currently preparing a paper that will explore some additional “pernicious examples;” with this paper I will also release my software for computing the general-purpose reactivity indicator.

The functional Taylor series that governs intermolecular interactions and chemical reactivity has an infinite number of terms. Up until the general-purpose reactivity indicator, these terms were treated only one-by-one. The resulting reactivity indicators are effective in most cases, but they fail in others. There is no reason to expect that there may not be several important terms in the Taylor series, which work together to control the reactivity. The general-purpose reactivity indicator is the first reactivity indicator to do this. Its method for achieving this—introducing a simplified “probe perturbation, then evaluating the full Taylor series for that perturbation—is innovative in its own right, and is now being used by many other research groups.

To truly understand the reactivity of molecules, one must understand the molecules’ structure. For example, we usually consider a coarse-grained atomic-level description of the general purpose reactivity indicator. Decomposing molecular properties into atomic and functional-group properties is the key to chemistry. But how should we define these atoms in a molecule?

Chapters VI and VII explore how different forms for the kinetic energy impact the partitioning of molecular properties into atomic contributions. Surprisingly and importantly, it is found that the atoms from the quantum theory of atoms in molecules (QTAIM) emerge for a plethora of wildly different forms of the kinetic energy. Chapter VI considers kinetic energy operators that include relativistic corrections. The atomic surfaces have the same “zero flux” form in this case. Chapter VII discusses how different forms of the kinetic energy operator can result in different values for the atomic kinetic energies if a molecule is partitioned arbitrarily. However, the partitioning from QTAIM generally does not suffer from this nonuniqueness: the same value of the atomic kinetic energy emerges from many forms. This is true for all of the conventional and intuitive forms of the kinetic energy. However, there are forms for the kinetic energy that are mathematically permissible by the postulates of quantum mechanics that give different atomic kinetic energies. This flexibility could be exploited to define new ways to partition molecules into open quantum subsystems, but these subsystems would not generally be acceptable representations of atoms in molecules.

All of the work in Chapters II-VII assumes that sufficiently accurate solutions to the molecular electronic structure problem can be obtained. This is not always the case. Chapters VIII and IX develop new approaches to solving the electronic Schrödinger equation for molecules using mathematical techniques from information-based complexity theory. These methods are designed to approach the theoretical limits of efficiency for solving the electronic Schrödinger equation. They do this by breaking the “curse of dimension” that plagues full-configuration interaction and released-node Monte Carlo approaches.

Chapter VIII presents a new selected configuration interaction method that has the same accuracy as full configuration interaction, but with only a polynomial (rather than a factorial) number of Slater determinants. The method in chapter VIII still follows Gaussian basis-set paradigm of modern quantum chemistry. The method in chapter IX breaks the hegemony of Gaussian basis sets by using the efficient integration methods I developed during my time at McMaster. (These contributions are not reproduced in the thesis, but are reviewed in the introductory chapter). These numerical integration grids can be applied to the local Schrödinger equation method of Nakatsuji and others. Though Nakatsuji’s initial implementation used randomly distributed, equally weighted points, it may be shown, via a theorem due to Boys, how the points should be located and how they should be weighted. It turns out that the points and weights from my many-electron quadrature grids are ideal for this purpose.

X.II. Prospects for Future Work on Reactivity

The general-purpose reactivity indicator is very simple because one of the design principles was that it should be easy to compute each atomic site's reactivity using only a very limited amount of information about its reaction partner. This led to an indicator that requires only the atomic charges of the molecule of interest and two parameters. The indicator is based on a model for the interaction energy between the molecule and its reaction partner, where the reaction partner is approximated. The model for the interaction energy, however, is very simple. In particular, it neglects polarizability. Including polarizability to define a more-general-purpose reactivity indicator would generally require two additional parameters. It might be interesting to approximate the polarizability with only one parameter or to express the polarizability contribution as a nonlinear function of the two parameters already in the model. Including polarizability would certainly increase the accuracy and applicability of the model. The interaction energy model also does not include steric hindrance. This was already a problem in the initial applications: reactivity indicators often predict sterically hindered sites as the most reactive. This means that the results of these indicators cannot be used without chemical intuition. Adding an approximate model for the steric hindrance to the model based, for example, on the Pauli repulsion energy, would give an even-more-general-purpose reactivity indicator. Such a reactivity indicator might come close to the "black box" tool for predicting molecular reactivity that chemists have been seeking for the last eighty years.

X.III. Prospects for Future Work on Kinetic Energy for Partitioning Atoms

In Chapter VII, I extended QTAIM by including scalar-relativistic corrections to the kinetic energy operator in the quantum mechanical Lagrangian/Hamiltonian. This had no effect on the zero-flux rule used to partition atoms. The next logical step is to include spin-orbit interactions, and to consider extending the method to an exact two component method such as X2C. One advantage of X2C is that there is no "picture change" error in this model. It will be interesting to see how this affects the definition of the atom in a molecule. I postulate that it will not change the definition much, if at all.

Chapter VIII considered exotic forms for the kinetic-energy operator in the non-relativistic context. Some of these forms, though they result in the correct value when evaluated over a molecule, may not be meaningful when evaluated

over an atom. For example, many kinetic energy densities take negative values. It seems absurd for the kinetic energy density to be negative. Cohen found a subset of the family I explored that always gives positive kinetic energy densities, and therefore always gives positive atomic kinetic energies. Investigating how this subset behaves within the quantum theory of atoms in molecules would be very interesting. The Cohen forms are nonlinear, wavefunction dependent, local kinetic energies; that form of local kinetic energy has never been considered previously.

X.IV. Prospects for Future Work on Grid Techniques

Chapter VIII can be considered a complete piece of work. There is much more potential for further development in Chapter IX. In particular, it is important to optimise the numerical integration grids *specifically* for the local Schrödinger equation. There are numerous parameters that can be changed, and it would be advantageous to explore and discover which combination of parameters produces the best grids for the local Schrödinger equation method. These parameters include (a) the kind of one-dimensional grid, (b) the Griebel-Knappek T value, and (c) the delay sequence (is Petras's ideal, or is something less aggressive better?). A detailed investigation of these parameters for three-dimensional integrals is almost finished. That will hopefully provide some insights that will be useful for the more daunting task of optimising many-electron grids.