Single-Determinant Theory of Electronic Excited States and Many-Electron Integrals for Explicitly Correlated Methods

Giuseppe Maria Junior Barca

June 2017

A thesis submitted for the degree of Doctor of Philosophy of the Australian National University

© Copyright by Giuseppe Maria Junior Barca 2017
All Rights Reserved
Declaration

The work described in this thesis, to the best of my knowledge, is original and does not contain material that has been submitted for a degree or diploma at any other University or College.

Giuseppe M. J. Barca
June, 2017
Acknowledgements

Now let me dispel a couple of rumors so they won’t fester into facts ... Yes! Before beginning this doctorate I had a passion for mathematics and quantum mechanics. And no, at the time when my supervisor, Prof. Peter Gill, first met me, I was not the man of science you see before you. He found me in a dying country, where only the lovely bones of an ancient intellectuality remain. “I was the intellectual equivalent of a 98-pound weakling. I’d go to the beach and people would kick copies of Byron in my face.”

This doctoral candidacy has represented the opportunity of my life and I am indebted with my deepest gratitude to my supervisor. Peter believed in me when my own people did not. During this three years and a half, not only he has been an one-off scientific mentor, but also and especially a friend and a father to me. His combination of scientific acumen and unshakeable patience provided a fertile ground where my creativity and analytic skills thrived reaching unprecedented levels.

My heartfelt thanks go also to Dr. Andrew Gilbert. His outstanding technical and theoretical aid and his extreme thoughtfulness were crucial factors for the successful fulfilment of my doctorate. I will always retain in my heart countless hours of scientific discussions that he endured with great lenience. I will never forget his precious friendship, but this is another story.

I am extremely grateful to my friend Dr. Pierre-Francois Loos (Titou). His very hard-working attitude and enthralling enthusiasm for the field engaged me in many exciting and productive ‘theoretical’ adventures. Without Titou’s support and collaboration a good part of the research work included in this thesis would have not been possible at all.

Thanks to the past and present members of the Gill group whom I have had the pleasure to know personally and have provided me with support and friendship — in
particular Dr. Caleb Ball, who has been a very enjoyable and benevolent companion for more than three years now, and Simon McKenzie, who, inter alia, together with Dr. Ball, had the bravery of proofreading this thesis.

Special thanks to Prof. Jacob Katriel for providing me with many very useful comments on this thesis.

I am also grateful for the generous financial support I received for the duration of my doctoral studies in Canberra, and, in particular, the University Research Scholarship from the Australian National University.

I wish also to thank Dr. Jia Deng for his unselfish friendship and his extraordinary private support.

Furthermore, I wish to thank the entire Gill family for welcoming me as one of them.

I am greatly thankful to my wife for morally sustaining me in this and many other hard choices, and for tolerating me during these years. Thank you for your courage and your love.

I thank my parents who supported with great patience my pursuit of manifold interests and encouraged me to endure when all seemed unsurmountable.

Last but not least, I thank my grandfather. He inspired me throughout my life, instilling the seeds of perseverance, courage and dignity. I owe to him who I am today.

Giuseppe M. J. Barca

Canberra, June, 2017
“Thus the problem is not so much to see what nobody has yet seen, as to think what nobody has yet thought concerning that which everybody sees.”

Arthur Schopenhauer — Parerga und Paralipomena, essay # 76
Abstract

The aim of this thesis is twofold.

Its first part, Part A, is concerned with the development and assessment of a single-determinant theory for electronic excited states. The theory is based on two simple algorithms for finding excited-state solutions to self-consistent field (SCF) equations, the Maximum Overlap Method (MOM) and the Initial Maximum Overlap Method (IMOM). The extent to which these higher SCF solutions are useful approximations to excited states is examined in diverse case studies, including challenging instances such as double excitations, conical intersections and charge-transfer states. Results indicate that single-determinant models yield, in most cases, accurate approximations to electronic excited states, even for difficult excitations where other low-cost excited-state methods either perform poorly or fail completely.

In Part B, we present efficient methods for the accurate evaluation of many-electron integrals arising in the explicitly correlated electronic structure theory. In our computational schemes efficient screening techniques, which adopt newly developed upper bounds, are used to sift out the tiny fraction of integrals which are significant. Then, non-negligible integrals are evaluated via recurrence relations that represent the generalization to three and four-electron integrals of two-electron integrals contraction-efficient schemes such as the Head-Gordon-Pople and PRISM algorithms. In this way, we developed general computational schemes for integrals arising from the use of a wide class of multiplicative correlation factors of the form $f_{12} = f(|r_1 - r_2|)$ and more specific methods for many electron integrals involving Gaussian Geminals. Our results support the evidence that our Gaussian-Geminal-based schemes yield a dramatic reduction of the computational complexity of these integrals.
List of Publications

The following manuscripts have been published as a direct consequence of the work undertaken for this thesis:

Chapter 3 : G. M. J. Barca, A. T. B. Gilbert, and P. M. W. Gill
“Hartree-Fock description of excited states of H₂”,

Chapter 4 : G. M. J. Barca, A. T. B. Gilbert, and P. M. W. Gill
“Simple models for difficult electronic excitations”,

Chapter 5 : G. M. J. Barca, A. T. B. Gilbert, and P. M. W. Gill
“The excitation number: A tale of misassigned multiply excited states”,

Chapter 7 : G. M. J. Barca, P.-F. Loos, and P. M. W. Gill
“Many-electron integrals over Gaussian basis functions. II. Upper bounds”,

Chapter 8 : G. M. J. Barca, P.-F. Loos, and P. M. W. Gill
“Many-electron integrals over Gaussian basis functions. I. Recurrence relations for three-electron integrals”,

Chapter 9 : G. M. J. Barca, and P.-F. Loos
“Recurrence relations for four-electron integrals over Gaussian basis functions”,
Chapter 10: G. M. J. Barca, and P. M. W. Gill

“Two-electron integrals over Gaussian geminals”,


Chapter 11: G. M. J. Barca, and P.-F. Loos

“Three- and four-electron integrals involving Gaussian Geminals: fundamental integrals, upper bounds and recurrence relations”,

# Contents

Acknowledgements v  
Abstract ix  
List of Publications xi  

1 Fundamentals of Quantum Chemistry 1  

1.1 From Schrödinger to Quantum Chemistry 1  
1.2 Preamble on notation and units 2  
1.3 Canonical Approximations 6  
1.3.1 The Born-Oppenheimer Approximation 7  
1.3.2 The Variation Method 8  
1.3.3 Orbitals 11  
1.3.4 Slater Determinants 11  
1.3.5 Linear Combination of Atomic Orbitals, Basis Functions and Basis Sets 13  

1.4 Wave-Function and Density Formalisms 16  
1.5 Wave-Function Methods 18  
1.5.1 Hartree-Fock Theory 18  
Restricted Hartree-Fock: The Roothan-Hall Equations 20  
Unrestricted Hartree-Fock: The Pople-Nesbet Equations 22  
The Self-Consistent Field Procedure 24  
Forming trial density matrices and the aufbau protcol 26
1.5.2 Electron Correlation ........................................ 27
    Unrestricted Hartree-Fock, correlation and spin contamination 29
1.5.3 Configuration Interaction ................................. 30
1.5.4 Coupled Cluster ........................................... 33
1.5.5 Perturbation Theory ....................................... 37
1.6 Density-Functional Methods ................................ 39
    1.6.1 Early approaches ....................................... 39
    1.6.2 The Kohn-Sham Method ................................ 41
    1.6.3 Exchange-Correlation Functionals ..................... 42

PART A — SINGLE-DETERMINANT THEORY OF ELECTRONIC EXCITED STATES 43

    INTRODUCTION TO PART A ................................... 45

2 A Short Introduction to Excited-State Methods ................. 47
    2.1 Introduction ............................................. 47
        Multi-Reference and Multi-Configurational Methods .... 49
    2.2 Single-Reference Methods ................................ 51
        2.2.1 CIS .................................................. 51
        2.2.2 TD-HF .............................................. 52
        2.2.3 TD-DFT .............................................. 55
            Fundamentals of TD-DFT ................................. 55
            Time-Dependent Kohn-Sham Method .................... 56
            TD-DFT for Excited States ............................ 57
        2.2.4 Coupled-Cluster Methods ............................. 58
            Equation-Of-Motion Coupled Cluster .................. 58
            Linear-Response Coupled Cluster (CCn) ............... 61
            Performance of EOM-CC and CCn Methods ............. 61
    2.3 Multi-Reference Methods .................................. 63
        2.3.1 MRSCF .............................................. 63
        2.3.2 MRCI and MRPT .................................... 65

3 Single-Determinant Approximation for Excited States .......... 67
    3.1 Introduction ............................................. 67
CONTENTS

3.2 The Maximum Overlap Method ........................................... 71
3.3 Hartree-Fock Description of Excited States of H\(_2\) .................. 72
  3.3.1 Method and Results ............................................... 72
  3.3.2 Discussion .......................................................... 74
3.4 Concluding Remarks ..................................................... 78

4 Single-Determinant Models for Difficult Excitations — The Initial
  Maximum Overlap Method .................................................. 81
  4.1 Introduction ............................................................ 81
  4.2 The Initial Maximum Overlap Method ................................ 82
  4.3 Double Excitations ..................................................... 85
    4.3.1 The H\(_2\) molecule ............................................. 85
    4.3.2 Polycyclic hydrocarbons ..................................... 86
  4.4 Charge-Transfer States ............................................... 88
    4.4.1 Ethylene + Tetrafluoroethylene ................................ 89
    4.4.2 Bacteriochlorin + Zn-Bacteriochlorin ....................... 90
  4.5 Conical Intersections ................................................ 92
    4.5.1 The H\(_3\) molecule ............................................. 92
    4.5.2 Retinal ............................................................ 93
  4.6 Concluding Remarks .................................................... 96

5 The excitation number ..................................................... 97
  5.1 Introduction ............................................................ 97
  5.2 The Excitation Number \(\eta\) .......................................... 99
  5.3 Hole and Particle Densities .......................................... 101
  5.4 Numerical Results ..................................................... 104
  5.5 Discussion and Misassignments ...................................... 106
    5.5.1 The \(2^{1}A_{g}\) state of trans-(1,3)-butadiene (C\(_4\)H\(_6\)) ... 107
    5.5.2 The \(1^{1}E_{2g}\) valence state of benzene (C\(_6\)H\(_6\)) .......... 108
    5.5.3 A note on the low-lying double excitation of MnO\(_4\)^{−} ....... 108
  5.6 Concluding Remarks .................................................... 109
PART B — MANY-ELECTRON INTEGRALS FOR EXPLICITLY CORRELATED METHODS

INTRODUCTION TO PART B .................................................. 112

6 An Overview of Explicitly Correlated Methods ........................................ 115
  6.1 Introduction ........................................................................ 115
  6.2 Hylleraas-Type Wave Functions: the Emergence of Three- and Four-Electron Integrals .......................................................... 117
  6.3 Slater Geminal ................................................................ 119
  6.4 Gaussian Geminals .............................................................. 119
  6.5 Exponentially Correlated Gaussians Method ...................... 120
  6.6 Gaussian-Geminals MP2 ...................................................... 120
  6.7 MP2-R12 ........................................................................ 122
  6.8 MP2-GG with Pre-Optimized GGs (GG(n) method) ............. 126
  6.9 MP2-F12 ........................................................................ 128
     6.9.1 SP Ansatz with the Rational Generator ......................... 129
     6.9.2 MP2-R12 versus MP2-F12 ........................................... 129
  6.10 F12-Coupled-Cluster ........................................................ 131
  6.11 Strategies for the Approximated Evaluation of Many-Electron Integrals 132
     6.11.1 Resolution of the Identity ........................................... 132
     6.11.2 Density Fitting .......................................................... 133
     6.11.3 Numerical Quadrature ............................................... 135

7 General Notation and Upper Bounds .................................................... 137
  7.1 Introduction ...................................................................... 137
  7.2 Gaussians, Shells and Integrals .......................................... 139
  7.3 Operators ........................................................................ 141
  7.4 Bounding Gaussians ........................................................... 142
     7.4.1 Shell-bounding Gaussians .......................................... 142
     7.4.2 Shell-pair-bounding Gaussians .................................... 143
  7.5 Types of Upper Bound ........................................................ 144
     7.5.1 Integral bounds .......................................................... 145
CONTENTS

7.5.2 Class bounds ................................................. 145
7.5.3 Shell-m-tuplet bounds ....................................... 146
7.6 Intermediates .................................................... 146
7.7 Upper Bounds ..................................................... 148
  7.7.1 Theory ........................................................ 149
  7.7.2 Performance ................................................. 152
    One-electron integrals ......................................... 152
    Two-electron integrals ......................................... 153
    Three-electron integrals ....................................... 156
7.8 The Contraction Problem ....................................... 156
7.9 Algorithms ....................................................... 159
7.10 Concluding Remarks ........................................... 161
7.11 Appendix A: Bounding Slater Geminal Integrals .......... 164
  7.11.1 Upper bound for the product of two Slater functions .... 164
  7.11.2 Primitive factors ......................................... 164
  7.11.3 Contracted factors ........................................ 165
7.12 Appendix B: Computation of $\langle 1 | f^9 \rangle$ and $\langle \bar{1} | f^9 \rangle$ .... 165

8 Recurrence Relations for Three-Electron Integrals 167

8.1 Introduction ..................................................... 167
8.2 Three-Electron Integrals ....................................... 168
  8.2.1 Three-electron operators ................................ 168
  8.2.2 Permutational symmetry .................................. 169
8.3 Algorithm ......................................................... 169
  8.3.1 Construct shell-pairs, -quartets and -sextets .......... 170
  8.3.2 Step O. Form fundamental integrals .................... 171
  8.3.3 Step T₁. Build momentum on center $A_3$ .............. 173
  8.3.4 Step T₂. Build momentum on center $A_2$ .............. 174
  8.3.5 Step T₃. Build momentum on center $A_1$ .............. 175
  8.3.6 Step C. Contraction ....................................... 175
  8.3.7 Steps T₄ to T₆. Shift momentum to ket centers .......... 176
8.4 Concluding Remarks ............................................ 176
8.5 Appendix: Derivation of VRRs ........................................ 176

9 Recurrence Relations for Four-Electron Integrals ............ 179
  9.1 Introduction ............................................................. 179
  9.2 Four-Electron Integrals .............................................. 180
    9.2.1 Four-electron operators ......................................... 180
  9.3 Fundamental Integrals ............................................... 180
  9.4 Recurrence Relations ................................................. 183
    9.4.1 Vertical recurrence relations ................................... 184
    9.4.2 Transfer recurrence relations .................................. 186
    9.4.3 Horizontal recurrence relations ................................. 187
  9.5 Algorithm ............................................................... 188
  9.6 Concluding Remarks ................................................... 190

10 Two-Electron Integrals Over Gaussian Geminals ............ 193
  10.1 Introduction ........................................................... 193
  10.2 Additional Notation and Rys Integrals ........................... 196
  10.3 Algorithm ............................................................... 197
    10.3.1 Significant shell-pairs ......................................... 197
    10.3.2 Significant shell-quartets ...................................... 199
    10.3.3 Construct $[00]$ .................................................. 199
    10.3.4 Construct $\langle c_1c_2|00\rangle$ by late contraction .......... 200
    10.3.5 Construct $\langle c_1c_2|00\rangle$ by early contraction .......... 201
    10.3.6 Construct $\langle a_1a_2|b_1b_2\rangle$ ............................... 203
  10.4 Examples: Forming a $\langle pp|pp\rangle$ Class .................. 203
    10.4.1 The late path ................................................... 204
    10.4.2 The early path .................................................. 204
  10.5 Computational Costs .................................................. 207
  10.6 Concluding Remarks .................................................. 209
  10.7 Appendix: Forming $\langle dd|dd\rangle$ & $\langle ff|ff\rangle$ on the Late Path 210

11 Three- and Four-Electron Integrals Involving Gaussian Geminals 213
  11.1 Introduction ........................................................... 213
CONTENTS

11.2 Three- and Four-Electron Integrals ................................. 215
  11.2.1 Three- and four-electron operators ............................ 215
11.3 Fundamental Integrals ............................................ 216
11.4 Upper Bounds ..................................................... 218
  11.4.1 Primitive bounds .............................................. 219
  11.4.2 Contracted bounds ............................................ 221
11.5 Recurrence Relations .............................................. 223
  11.5.1 Vertical recurrence relations .................................. 223
  11.5.2 Transfer recurrence relations ................................ 224
  11.5.3 Horizontal recurrence relations ................................ 225
11.6 Algorithm .......................................................... 225
11.7 Concluding Remarks ............................................... 229

12 Conclusions ....................................................... 231
Chapter 1

Fundamentals of Quantum Chemistry
1.1 From Schrödinger to Quantum Chemistry

In principle, the electronic structure and the properties of any molecule in whichever of its quantized states can be determined by solving the Schrödinger equation. The equation reads

$$i\hbar \frac{\partial \Psi}{\partial t} = \mathcal{H}\Psi$$

(1.1)

where $\mathcal{H}$ is the Hamiltonian for the system, $\hbar$ is Planck’s constant divided by $2\pi$ and $i$ is the imaginary unit. The elusive entity denoted by $\Psi$ is the so-called wave function, a function of the spatial and spin coordinates of the particles (and of time), and it encloses all the information about the system. The latter fact is also known as the first postulate of quantum mechanics.

It is certainly one of the eccentricities of quantum mechanics that even if the state of a system is completely specified by $\Psi$, the wave function itself does not have any physical interpretation. However, the accepted view, which goes back to Born, is that the square of the absolute value of wave function $|\Psi|^2$ represents a probability density distribution.

At first glance Eq. (1.1) might seem innocuous to the reader, however much complexity is hidden in the Hamiltonian term. For isolated molecules, ignoring relativity and electromagnetism, the Hamiltonian takes the form

$$\mathcal{H} \equiv \mathcal{H}(\mathbf{r}, \mathbf{R})$$

$$= -\sum_i^N \frac{\hbar^2}{2m_e} \nabla^2_i - \sum_A^M \frac{\hbar^2}{2M_A} \nabla^2_A - \sum_i^N \sum_A^M \frac{Z_A e^2}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i<j}^N \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{A<B}^M \frac{Z_A Z_B e^2}{|\mathbf{R}_A - \mathbf{R}_B|}$$

(1.2)

where $\nabla^2_i$ is the Laplacian operator,

$$\nabla^2_i = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2},$$

(1.3)

$M$ is the number of atomic nuclei, $N$ is the number of electrons in the system, $\mathbf{R}_A$ is the coordinate of the $A$-th nucleus and $\mathbf{r}_i$ the coordinate of the $i$-th electron. Furthermore
\(m_e\) and \(e\) are the electronic mass and charge as well as \(M_A\) and \(Z_A\) are the nuclear mass and charge of the \(A\)-th nucleus.

The notation \(\mathcal{H}(\mathbf{r}, \mathbf{R})\) in Eq. (1.2) is to explicitly highlight the dependence of the Hamiltonian operator on the entire set of electronic and nuclear coordinates \(\mathbf{r} \equiv \{r_i\}\) and \(\mathbf{R} \equiv \{R_A\}\). It also suggests a significant fact: the Hamiltonian is time-independent. This implies that Eq. (1.1) is separable in time and space-spin parts, i.e. its solution can be written as the product

\[
\Psi(\mathbf{x}, \mathbf{R}, t) = \Psi(\mathbf{x}, \mathbf{R})\Theta(t)
\]

of a purely spatial-spin part \(\Psi(\mathbf{x}, \mathbf{R})\) and a purely temporal term \(\Theta(t)\), where \(\mathbf{x} \equiv \{x_i\}\) is a shorthand notation to compactly represent the dependence on all the variables \(x_i \equiv \{r_i, s_i\}\), with \(s_i\) being the spin of the \(i\)-th electron.

Using the ansatz of Eq. (1.4) for the wave function and solving the differential equation for \(\Theta(t)\) yields

\[
\Theta(t) = \exp \left( -\frac{iEt}{\hbar} \right)
\]

which leaves us with the time-independent wave equation\[1, 2\]

\[
\mathcal{H}(\mathbf{r}, \mathbf{R}) \Psi(\mathbf{x}, \mathbf{R}) = E \Psi(\mathbf{x}, \mathbf{R})
\]

Equation (1.6) is an eigenvalue equation. For any physically meaningful system the \(\mathcal{H}(\mathbf{r}, \mathbf{R})\) operator is Hermitian, and in order for Eq. (1.6) to yield physically acceptable solutions, the wave function \(\Psi(\mathbf{x}, \mathbf{R})\) must satisfy the following normalization condition

\[
\int |\Psi(\mathbf{x}, \mathbf{R})|^2 d\mathbf{x} d\mathbf{R} = 1
\]

which ultimately stems from the probabilistic interpretation. Since the integral must be finite, the wave function is also required to comply with a square integrability boundary condition. Solving the equation is then guaranteed to yield a spectrum of real eigenvalues \(E_k\) and of corresponding eigenfunctions \(\Psi_k\), the latter forming a complete orthogonal set. Each eigenfunction \(\Psi_k\) provides a complete description of the \(k\)-th quantized stationary state of the system, while the corresponding real eigenvalue \(E_k\) is its energy.
There is no doubt that the time-independent Schrödinger equation (TISE) (1.6) is the key to the entire molecular structure problem. On the one hand its solutions would yield a complete description of the system, on the other solving the equation exactly is impossible, barring a lilliputian number of exceptions.

This situation was known since the early days of quantum mechanics, in fact Dirac himself in 1929 stated[3]

“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. 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.”

Here is the foundation of an entire discipline: quantum chemistry. This is a branch of applied mathematics entailing the development and the implementation of electronic structure methods which offer an approximate and computationally efficient solution of the Schrödinger equation.
CHAPTER 1. FUNDAMENTALS OF QUANTUM CHEMISTRY

Figure 1.2 is the two-dimensional chart of quantum chemistry introduced by Pople in 1965[4] to present “a projection of the whole subject in which each calculation of a molecular wave function is represented by a point with two coordinates”. The horizontal coordinate corresponds to the size of the molecule measured by the number of electrons. The vertical coordinate measures the level of sophistication of the quantum-mechanical method used, which is also a measure of its accuracy. The computational complexity of an electronic structure approximation increases considerably with both the size of the system and the level of accuracy required. In particular, the curve correlating the level of sophistication with the number of electrons of a quantum chemical calculation is a hyperbola. In fact, due to their characteristic computational cost dependence, highly accurate calculations (in extremis solving exactly the Schrödinger equation) are possible only for small system scales, and vice-versa large molecular sizes are treatable only with lower accuracy methods.

This is the Gordian knot of quantum chemistry, a discipline which is torn between two extreme forces: accuracy and scalability.

As computer architecture and calculative capabilities evolve, the role of the quantum chemist is to devise new methods which offer a computationally increasingly beneficial
compromise between accuracy and scalability.

This is the ultimate goal of this thesis. In this pursuit, we first briefly present a number of well-established and ubiquitous approximations and methods that constitute the fundamentals of quantum chemistry, laying the condicio sine qua non on which the new theory is based.

1.2 Preamble on notation and units

At this stage, it is beneficial to clarify our notation for integrals. Unfortunately, in quantum chemistry two notations are commonly adopted for one- and two-electron integrals: physicists’ notation and chemists’ notation.

In physicists’ notation one- and two-electron integrals are defined by

$$\langle i | j \rangle = \langle \chi_i | \chi_j \rangle = \int \chi_i(x_1) \dagger \chi_j(x_1) dx_1$$

$$\langle ij | kl \rangle = \langle \chi_i \chi_j | f_{12} | \chi_k \chi_l \rangle = \iint \chi_i(x_1) \dagger \chi_j(x_2) \dagger f_{12} \chi_k(x_1) \chi_l(x_2) dx_1 dx_2$$

where $\chi_i(x_1)$ is a function of the electronic coordinates $x_1$ of electron “1” and $f_{12} = f(x_1, x_2)$ is a function of the electronic coordinates of both electron “1” and “2”. Also antisymmetrized two-electron integrals of the kind

$$\langle ij | kl \rangle = \langle ij | kl \rangle - \langle ij | lk \rangle$$

are common. In chemists’ notation these integrals read

$$\langle i | j \rangle = \langle \chi_i \chi_j \rangle = \int \chi_i(x_1) \dagger \chi_j(x_1) dx_1$$

$$\langle ij | kl \rangle = \langle \chi_i \chi_j | f_{12} | \chi_k \chi_l \rangle = \iint \chi_i(x_1) \dagger \chi_j(x_2) \dagger f_{12} \chi_k(x_1) \chi_l(x_2) dx_1 dx_2$$

In order to avoid any inconsistency we eschew chemists’ notation altogether. All integrals will be in physicists’ notation.

Atomic units are used throughout.
1.3 Canonical Approximations

1.3.1 The Born-Oppenheimer Approximation

The “clamped nuclei” approximation proposed by Born and Oppenheimer in 1927\[5\] is applied in the vast majority of quantum-chemical methods. In order to elucidate it, we rewrite the Hamiltonian in Eq. (1.2) as

\[
\mathcal{H}(\mathbf{r}, \mathbf{R}) = \mathcal{T}_e(\mathbf{r}) + \mathcal{T}_N(\mathbf{R}) + V_{eN}(\mathbf{r}, \mathbf{R}) + V_{ee}(\mathbf{r}) + V_{NN}(\mathbf{R}) \tag{3.13}
\]

where \(\mathbf{r}\) and \(\mathbf{R}\) denote the sets of electronic and nuclear coordinates, respectively. The terms \(\mathcal{T}_e\) and \(\mathcal{T}_N\) are the electronic and the nuclear kinetic energy operators, while \(V_{eN}\), \(V_{ee}\) and \(V_{NN}\) are the electron-nuclei attraction potential and the electron-electron and nuclei-nuclei repulsion potentials, respectively.

The solution of the TISE (1.6) is not separable in electronic and nuclear parts because of the coupling term \(V_{eN}(\mathbf{r}, \mathbf{R})\). The Born-Oppenheimer approximation assumes that this nuclear and electronic separation is approximately correct, that is

\[
\Psi(\mathbf{r}, \mathbf{R}) \approx \Psi_e(\mathbf{r}; \mathbf{R})\Psi_N(\mathbf{R}) \tag{3.14}
\]

where \(\Psi_e\) and \(\Psi_N\) are the electronic and the nuclear part of the wave function and the notation \((\mathbf{r}; \mathbf{R})\) indicates that \(\Psi_e\) depends only parametrically on the location of the nuclei. The approximation relies on the observation that nuclei are roughly 2000 times heavier than electrons. This means that in practice the electrons will respond instantaneously to any movement of the nuclei, and therefore we can fix the nuclear configuration at some set of values \(\mathbf{R}^*\) and solve the “clamped nuclei” TISE for the electronic wave function

\[
\mathcal{H}_e(\mathbf{r}; \mathbf{R}^*)\Psi_e(\mathbf{r}; \mathbf{R}^*) = E_e(\mathbf{R}^*)\Psi_e(\mathbf{r}; \mathbf{R}^*) \tag{3.15}
\]

where the electronic Hamiltonian

\[
\mathcal{H}_e(\mathbf{r}; \mathbf{R}^*) = \mathcal{T}_e(\mathbf{r}) + V_{eN}(\mathbf{r}; \mathbf{R}^*) + V_{ee}(\mathbf{r}) + V_{NN}(\mathbf{R}^*) \tag{3.16}
\]
is obtained from $\mathcal{H}(r, R)$ by neglecting the nuclear kinetic energy operator $T_N(R)$, as the nuclei are assumed to be fixed. Note that the total energy $E_e$ includes the electronic energy and also the nuclear repulsion energy $V_{NN}$.

Once the electronic problem is solved, the nuclear problem can be cast as

$$\mathcal{H}_N(R)\Psi_N(R) = E\Psi_N(R)$$  \hspace{1cm} (3.17)

where the nuclear Hamiltonian is

$$\mathcal{H}_N(R) = T_N(R) + E_e(R).$$  \hspace{1cm} (3.18)

The total energy $E_e(R)$ provides a potential for the nuclear motion which is usually referred to as the potential energy surface (PES) of the system. Solutions to the nuclear TISE (3.17) describe the vibrational, rotational and translational modes of a molecule, and $E$, which is the Born-Oppenheimer approximation to the total energy, includes electronic, vibrational, rotational and translational energy.

In this thesis we are concerned solely with the electronic problem (Eq. (3.15)). Therefore in the remainder of this work the electronic wave function will be reported simply as $\Psi$, that is omitting the subscript “e”.

### 1.3.2 The Variation Method

The variation method is a powerful tool for obtaining approximate solutions to eigenvalue equations which is based on a theorem known as the variational principle.[6] The theorem states that, for any trial wave function $|\tilde{\Psi}\rangle$ that satisfies the appropriate boundary conditions of the problem, the expectation value of the Hamiltonian (i.e. the energy) is an upper bound to the exact ground-state energy, $E_{\text{exact}}$.

$$E_{\text{exact}} \leq E[\tilde{\Psi}] = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}$$  \hspace{1cm} (3.19)

where the notation $E[\tilde{\Psi}]$ emphasises that the energy is a functional of the trial wave function. Thus the minimization of $E[\tilde{\Psi}]$ with respect to all allowed trial wave functions yields the exact ground-state wave function and energy. Since the Schrödinger equation
cannot be solved exactly except in a very small number of special cases, such a principle is of paramount importance as it provides a criterion to assess the quality of a trial wave function and allows for a systematic improvement towards the exact solution.

The Linear Variational Problem

The typical application of the variation method is to the particular case in which only linear variations of the trial wave function are allowed, that is

$$|\tilde{\Psi}\rangle = \sum_{i=1}^{N} c_{i} |\Phi_{i}\rangle$$  \hspace{1cm} (3.20)

where \{\Phi_{i}\} is a fixed set of \(N\) orthonormal \(n\)-electron basis functions.

Using the Lagrangian multiplier formalism

$$\mathcal{L}(c_{1}, \ldots, c_{N}, E) = \langle \tilde{\Psi} | \mathcal{H} | \tilde{\Psi} \rangle - E\langle \tilde{\Psi} | \tilde{\Psi} \rangle - 1$$  \hspace{1cm} (3.21)

where the multiplier \(E\) imposes the normalization condition

$$\langle \tilde{\Psi} | \tilde{\Psi} \rangle = \sum_{i} c_{i}^{2} = 1$$  \hspace{1cm} (3.22)

then the variational problem takes the form

$$\frac{\partial \mathcal{L}}{\partial c_{i}} = 0 \quad i = 1, \ldots, N$$  \hspace{1cm} (3.23)

The goal is now to find the optimum set of coefficients \(c_{i}\). If we define the matrix representation of the Hamiltonian in the \{\Phi_{i}\} basis as

$$(H)_{ij} = H_{ij} = \langle \Phi_{i} | \mathcal{H} | \Phi_{j} \rangle$$  \hspace{1cm} (3.24)

the problem can be recast in the standard eigenvalue problem for the \(H\) matrix

$$H c = Ec$$  \hspace{1cm} (3.25)

where \(c\) is a column vector with elements \(c_{i}\).
Since $H$ is a Hermitian matrix, Eq. (3.25) can be solved to yield $N$ orthonormal eigenvectors $c_\alpha$ and corresponding eigenvalues $E_\alpha$, which for convenience are arranged in ascending order, that is

$$Hc_\alpha = E_\alpha c_\alpha \quad \alpha = 0, 1, \ldots, N - 1$$

(3.26)

with

$$E_0 \leq E_1 \leq \ldots \leq E_{N-1}. \quad (3.27)$$

**Hylleraas-Undheim-MacDonald theorem**

Remarkably, the variation method can be extended to all the eigenstates of a system,* if they are modeled via a linear vector space, by virtue of a special case of Cauchy’s interlace theorem,[7] known as the Hylleraas-Undheim-MacDonald theorem.[8, 9] In order to illustrate the theorem, consider two linear variational spaces which constitute two orthonormal sets of $n$-electron basis functions $S' = \{\Phi'_i\}$ and $S'' = \{\Phi''_i\}$, where the first space is a subset of the second, that is $S' \subset S''$. The matrix eigenvalue equations in the two basis set are

$$H' c'_\alpha = E'_\alpha c'_\alpha \quad \alpha = 0, 1, \ldots, N - 1$$

(3.28)

$$H'' c''_\alpha = E''_\alpha c''_\alpha \quad \alpha = 0, 1, \ldots, N - 1, N$$

(3.29)

where we have again assumed that the eigenvalues are sorted in ascending order and, without loss of generality, that $S''$ contains one more basis function than $S'$. Then the Hylleraas-Undheim-MacDonald theorem states that

$$E_0'' \leq E'_0 \leq E'_1 \leq \cdots \leq E''_{N-1} \leq E'_{N-1} \leq E''_N$$

(3.30)

which relates the eigenvalues of two variational spaces $S' \subset S''$ in such a way that those of the larger space are always lower bounds to their homologous in the smaller basis set.

---

*This statement is restricted to those eigenstates which have a finite number of eigenstates below them. Thus, states which are embedded in the continuum, such as resonances, are excluded.
1.3.3 Orbitals

A common form for the trial wave function $\tilde{\Psi}$ is a linear combination of orthonormal $n$-electron basis functions $\Phi_i$, however we have not discussed yet the mathematical structure of the latter. We will soon see that quantum chemists mostly use the Slater determinant to this purpose, and before considering many-electron wave functions, it is beneficial to discuss how wave functions for a single electron are modelled.

In order to completely describe an electron it is necessary to specify its spatial distribution and its spin.

A spatial orbital $\psi_i(r)$ is a function of the position vector $r$ and describes the spatial distribution of an electron. The physical implication of this definition is that $|\psi_i(r)|^2 dr$ represents the probability of finding the electron in the infinitesimal volume element $dr$. Since we are concerned with the molecular electronic structure, from now on we will refer to the function $\psi_i(r)$ as a molecular orbital (MO).

Once the spatial distribution of the electron is determined we need to specify its spin. A complete set for describing the spin of an electron consists of the two orthonormal functions $\alpha(s)$ and $\beta(s)$, one for spin up and the other for spin down.

The wave function which describes a single electron in its entirety is a spin orbital $\chi_i(x)$, where $x$, as previously reported, represents both space and spin coordinates. Thus, from a set of $N$ MOs $\{\psi_i\}$ one can form a set of $2N$ spin orbitals by multiplication with either the $\alpha(s)$ or $\beta(s)$ spin functions, that is

$$\chi_i(x) = \psi_i(r) \alpha(s) \quad (3.31)$$
$$\chi_{i+1}(x) = \psi_i(r) \beta(s) \quad (3.32)$$

Assuming that the set of MOs $\{\psi_i\}$ is orthonormal so is the homologous set of spin orbitals $\{\chi_k\}$.

1.3.4 Slater Determinants

Suppose that we want to reasonably approximate the wave function of an $n$-electron system by using combinations of spin orbitals. Then adopting an appropriate set of $n$ spin orbitals, each hosting a single electron, it might seem reasonable to use just their
product as an \( n \)-electron wave function

\[
\Phi^H(x_1, \ldots, x_n) = \prod_{i=1}^{n} \chi_i(x_i) \quad (3.33)
\]

The wave function in Eq. (3.33) is known as a “Hartree product” as it was first proposed by Hartree in 1927.\(^{[10]}\) The physical picture originating from it definitely has its allure as each electron simply occupies a single spin orbital.

Unfortunately nature is not as simple and intuitive in this case. Electrons are fermions and as such these particles i) are indistinguishable, ii) follow Fermi-Dirac statistics.\(^{[11], [12]}\) Hartree products do not satisfy either of these two physical aspects of electrons.

Since the Hamiltonian is spin free, the fermionic behaviour must be built directly in the wave function. A satisfactory theory can be obtained if we make the additional requirement that the many-electron wave function be antisymmetric with respect to the interchange of any two electronic coordinates, that is

\[
\Phi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = -\Phi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) \quad (3.34)
\]

This requirement is the so-called \textit{antisymmetry principle}, which is essentially a mathematical consequence of the well known Pauli exclusion principle.\(^{[13]}\)

In 1930 Fock\(^{[14]}\) and Slater\(^{[15]}\) independently proposed a \( n \)-electron wave function that takes the form of a so-called \textit{Slater determinant}

\[
\Phi^S(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix}
\chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_n(x_1) \\
\chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_n(x_2) \\
\vdots & \vdots & \ddots & \vdots \\
\chi_1(x_n) & \chi_2(x_n) & \cdots & \chi_n(x_n)
\end{vmatrix}
\]

This determinant easily satisfies the antisymmetry requirement because if two rows are exchanged, which is equivalent to exchanging two electrons, then the sign of the determinant, and thus the wave function, is changed.

Since then Slater determinants have become the building blocks of the vast majority of quantum chemical methods. In essence, in many of the most accurate approximations
for the trial wave function $\tilde{\Psi}$ the fixed set of $N$ orthonormal $n$-electron basis functions in Eq. (3.20) is made of $N$ orthonormal Slater determinants, namely

$$\Phi_i = \Phi_i^S$$

(3.35)

thus, usually $\tilde{\Psi}$ is just a linear combination of Slater determinants

$$|\tilde{\Psi}| = \sum_{i=1}^{N} c_i |\Phi_i^S\rangle$$

(3.36)

For this reason in the remainder of this thesis the symbol $\Phi$ indicates a Slater determinant, where we omit the superscript “S” for brevity.

1.3.5 Linear Combination of Atomic Orbitals, Basis Functions and Basis Sets

We have just seen that ordinarily the trial many-electron wave function $\tilde{\Psi}$ is a linear combination of Slater determinants $\Phi_i$. These are normalized determinants of spin orbitals $\chi_i(x_i) = \psi_i(r_i)\sigma(s_i)$, where the MO $\psi_i(r_i)$ is a spatial distribution and $\sigma(s_i)$ is a spin function (either $\alpha(s_i)$ or $\beta(s_i)$).

In general, it is convenient to introduce a further approximation in which MOs are expanded in a basis $\{\phi_{\mu}(r)\}$ of $N$ atomic orbitals (AOs)

$$\psi_i(r) = \sum_{\mu=1}^{N} C_{\mu i} \phi_{\mu}(r)$$

(3.37)

This expansion is known as a Linear Combination of Atomic Orbitals (LCAO).

The atomic basis set is usually referred to just as the ‘basis set’. In principle, there exists a wide array of candidates that one can use to construct such a basis set. Ideally, the basis functions should be chosen to allow for a systematic and rapid improvement towards completeness, easy algebraic manipulation and efficient computer implementation. However, in practice, compromises have to be made in order to strike a balance between accuracy and computational cost.

In 1930 Slater[15] and Zener[16] proposed the so-called Slater (or exponential)
functions (SFs) as atomic basis functions

\[ |a|^{SF} \equiv \zeta_a^A(r) = n_a^{SF} \left[ \prod_{t=x,y,z} (t - A_t)^{a_t} \right] e^{-\alpha |r-A|} \]  

(3.38)

with exponent \( \alpha \), center \( A = (A_x, A_y, A_z) \), normalization constant \( n_a^{SF} \), and, with an abuse of notation, angular momentum \( a = (a_x, a_y, a_z) \) and total angular momentum \( a = (a_x + a_y + a_z) \). SFs are the natural choice for atomic systems as for \( a = 0, 1 \), and with the correct choice of \( \alpha \) they are exact solutions to the hydrogenic Schrödinger equation. Another attractive feature of these basis functions is that they can reproduce two characteristic behaviours of any exact wave function i) the cusp at the electron-nucleus coalescence point,[17] ii) the exponential decay.[18]

Despite their useful physical features, after an initial successful use in atomic calculations, SFs could not be easily extended to molecular systems because of difficulties in the evaluation of multi-centre two-electron integrals, which are of crucial importance to modern quantum chemistry.

A breakthrough came in 1950 with Boys’ proposal of approximating Slater functions with contracted Gaussian functions (CGFs)[19]

\[ |a| \equiv \phi_a^A(r) = \sum_{i=1}^{K_a} |a|_i \]  

(3.39)

which are linear combinations of \( K_a \) (known as the “degree of contraction” of the CGF) primitive Gaussian functions (PGFs)

\[ |a|_i \equiv \varphi_a^A(r) = D_a^n n_i^a \left[ \prod_{t=x,y,z} (t - A_t)^{a_t} \right] e^{-\alpha_i |r-A|^2} \]  

(3.40)

where \( D_a^n \) is a contraction coefficient and \( n_i^a \) a normalization factor. We will usually suppress the primitive index \( i \).

PGFs, unlike SFs, offer neither the correct nuclear cusp nor decay behaviour. Nonetheless, it has been established both theoretically and empirically that a ‘brute force’ saturation of function space, including Gaussians with suitably large and small exponents, can reduce these errors to any desired level.

14
Formally this means that many more integrals must be computed than for SFs. However, GFs have other very desirable and unique properties that still make the evaluation of integrals over CGFs far superior to those over SFs. In Part B we will explore these properties in detail.

Thus GF basis sets are the most used basis sets in quantum chemical calculations. Among them, the most popular are those developed by Pople and co-workers, the correlation-consistent families of Dunning and co-workers, and the polarization-consistent families of Jensen.

The Pople basis sets are arguably the most widely used for routine applications because of the good compromise between computational expense and the resulting accuracy. The Pople basis sets are generally of a split-valence nature. Symbols like 6-31G or 6-311G represent double and triple splits of the valence shell. The acronym is interpreted as 6 being the number of GFs for the inner shells and 31 or 311 being the number of GFs in the contractions of the valence shells. The Pople basis sets can be augmented with diffuse (+) and polarization functions for an improved description of electronic structure.

The Dunning correlation-consistent basis sets, denoted cc-pVXZ (X = D, T, Q), are developed specifically for recovering the correlation energy. These basis sets are constructed by adding shells of s, p, d, ... functions to the atomic Hartree-Fock orbitals in a systematic manner. Since each function within each shell contributes nearly an equal amount to the correlation energy, the correlation-consistent basis sets are well-defined with respect to increases in size and accuracy.

Finally, there are the Jensen polarization consistent basis sets, denoted pc-x (x=0, 1, 2, 3, 4), whose design philosophy is similar to that of cc-pVXZ, but with a main emphasis on Hartree-Fock and Density Functional Theory calculations.

In general the correlation-consistent and polarization-consistent basis sets can be expected to converge smoothly to the complete basis set (CBS) limit. This allows extrapolation techniques to be used for calculating molecular energies at the CBS limits.
1.4 Wave-Function and Density Formalisms

As discussed in Section 1.3.2, the total energy for a molecular system is a functional of the wave function, that is \( E = E[\tilde{\Psi}] \). Furthermore, every eigenstate with energy \( E_k \) and wave function \( \Psi_k \) is an extremum of the functional \( E[\tilde{\Psi}] \). In other words, satisfying the Schrödinger equation (1.6) is equivalent to solving the variational problem

\[
\delta E[\tilde{\Psi}] = 0 \tag{4.41}
\]

with the constraint of \( \tilde{\Psi} \) being normalized. Most of the contemporary calculations on electronic structures use this wave-function formalism and for this reason they are known as wave-function methods.

In 1964 Hohenberg and Kohn\[29\] showed that there exists an alternative but rigorous formalism that allows to replace the complicated \( n \)-electron wave function \( \Psi(x_1, x_2, \ldots, x_n) \) and the associated variational procedure with the much simpler one-electron density

\[
\rho(r_1) = n \int \ldots \int \Psi(x_1, x_2, \ldots, x_n)^\dagger \Psi(x_1, x_2, \ldots, x_n) ds_1 dx_2 \ldots dx_n \tag{4.42}
\]

and its associated calculation scheme. Despite the huge simplification made in going from the wave function, which is a \( 4n \)-dimensional mathematical construct, to the density, which is just three-dimensional, no intrinsic information is lost.\[29\] In fact the electronic Hamiltonian in Eq. (3.16) can be rewritten as

\[
\mathcal{H} = T_e + \sum_{i=1}^{n} V_{ext}(r_i) + V_{ex} + V_{NN} \tag{4.43}
\]

where it is completely determined by the external potential \( V_{ext} \), which usually can be identified with the nuclear potential.

In particular, Hohenberg-Kohn\[29\] proved that for an \( n \)-electron system with a given external potential \( V_{ext} \) the variational theorem establishes a procedure to uniquely determine the ground-state wave function \( \Psi_0 \), where the ground state energy \( E_0 \) is a functional of \( n \) and \( V_{ext} \)

\[
E_0 = E[n, V_{ext}] \tag{4.44}
\]
Their first theorem states

“The external potential $V_{\text{ext}}$ is determined within a trivial additive constant by the electron density.”

and, in turn, the ground state energy is determined by $V_{\text{ext}}$, then the ground state energy itself must be a functional of the density

$$E_0 = E[\rho] \quad (4.45)$$

This is more formally stated by the second Hohenberg-Kohn theorem which defines the variational principle for the density

“For a trial density $\tilde{\rho}(\mathbf{r})$ such that $\tilde{\rho}(\mathbf{r}) \geq 0$ and $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = n$,

$$E[\tilde{\rho}] \geq E_0 \quad (4.46)$$

Over the last few decades, quantum chemical methods that use this density formalism, known as density-functional methods, have not only become widespread but arguably the most widely used for routine applications because of the good compromise between the computational cost and the resultant accuracy.

The aim of the remaining sections of this chapter is to introduce the reader to the most common and established wave-function and density-functional methods.
1.5 Wave-Function Methods

1.5.1 Hartree-Fock Theory

In the Hartree-Fock (HF) approximation the trial wave function takes its simplest form, a single Slater determinant

$$\tilde{\Psi}^{HF} = |\chi_1\chi_2 \ldots \chi_a\chi_b \ldots \chi_n\rangle$$

(5.47)

where $|\chi_1\chi_2 \ldots \chi_a\chi_b \ldots \chi_n\rangle$ is a shorthand notation for a Slater determinant of $n$ spin orbitals $\{\chi_a\}$.

According to the variational principle, we are interested in minimizing the energy expectation value calculated with $\Phi$ as a functional of the occupied spin orbitals $\chi_a$, subject to the constraint that the set of orbitals remains orthonormal during the variation.

Assuming an $n$-electron closed-shell system, this is achieved by minimizing with respect to the spin orbitals $\{\chi_a\}$ the Lagrangian

$$\mathcal{L}[\{\chi_a\}] = E[\{\chi_a\}] - \sum_{a=1}^{n} \sum_{b=1}^{n} \epsilon_{ab}(\langle \chi_a | \chi_b \rangle - \delta_{ab})$$

(5.48)

where $\epsilon_{ab}$ constitute a set of unknown Langrange multipliers and

$$E[\{\chi_a\}] = \sum_{a=1}^{n} \langle \chi_a | \chi_a \rangle + \frac{1}{2} \sum_{a=1}^{n} \sum_{b=1}^{n} \langle \chi_a \chi_b | \chi_a \chi_b \rangle - \langle \chi_a \chi_b | \chi_b \chi_a \rangle$$

(5.49)

is the expectation value of the Hamiltonian over the single determinant $\Phi$.

Performing this variation (that is solving $\delta\mathcal{L}[\{\chi_a\}] = 0$) one arrives at the following set of integro-differential equations, known as the Hartree-Fock equations,

$$f(x_1)\chi_a(x_1) = \sum_{b=1}^{n} \epsilon_{ba} \chi_b(x_1) \quad a = 1, 2, \ldots, n$$

(5.50)

where the Fock operator

$$f(x_1) = [h(x_1) + \sum_{b=1}^{n} J_b(x_1) - K_b(x_1)]$$

(5.51)
is the combination of other three operators. The core-Hamiltonian operator

\[ h(x_1) = -\frac{1}{2} \nabla^2 + \sum_A \frac{Z_A}{|r_1 - R_A|} \]  

(5.52)

which is constituted by the kinetic energy operator and the nuclear potential of a single electron chosen to be electron “1”. The Coulomb operator

\[ J_b(x_1) = \int \chi_b^\dagger(x_2)\chi_b(x_2) \frac{d\mathbf{x}_2}{|r_1 - r_2|} \]  

(5.53)

which represents the average local potential at \( x_1 \) arising from an electron in \( \chi_b \), and the exchange operator

\[ K_b(x_1)\chi_a(x_1) = \left[ \int \chi_b^\dagger(x_2)\chi_a(x_2) \frac{d\mathbf{x}_2}{|r_1 - r_2|} \right] \chi_b(x_1) \]  

(5.54)

which has no classical interpretation as it arises from the antisymmetry of the wave function.

From the form of the Fock operator in Eq. (5.51), it can be shown that the essence of the HF approximation is to replace the many-electron problem by a one-electron problem in which the electron repulsion is treated in an averaged way. In fact the quantity

\[ V^{HF}(x_i) = \sum_{b \neq i} J_b(x_i) - K_b(x_i) \]  

(5.55)

known also as the HF potential, is the average field experienced by electron “\( i \)” due to the presence of the remaining electrons. For this reason the HF method is often defined as a mean-field approximation.

The wave function is a determinant and as such it is invariant to unitary transformations of the spin orbitals. It follows that the Fock operator is also invariant to such transformations and consequently the optimal spin orbitals are not uniquely defined. It is possible therefore to obtain an equivalent set of Lagrange multipliers \( \{ \epsilon_{ab}^c \} \) and of spin orbitals \( \{ \chi_a^c \} \), which are referred to as the canonical spin orbitals, from the the old
sets \{\epsilon_{ab}\} and \{\chi_a\}, by means of a unitary transformation

\[
\chi_a^c = \sum_b \chi_b U_{ba} \tag{5.56}
\]

\[
\epsilon^c = U^\dagger \epsilon U \quad \text{with} \quad U^\dagger = U^{-1} \tag{5.57}
\]
such that the matrix \(\epsilon^c\) which contains the new multipliers \(\{\epsilon^c_{ab}\}\) is diagonal. In this way the Hartree-Fock equations reduce to their simpler, canonical form

\[
f^c(\mathbf{x}_1)\chi_a^c(\mathbf{x}_1) = \epsilon^c_a \chi_a^c(\mathbf{x}_1) \quad a = 1, 2, \ldots, n \tag{5.58}
\]

Henceforth, we will drop the superscript “c” for conciseness.

**Restricted Hartree-Fock: The Roothan-Hall Equations**

Although Hartree-Fock theory was formulated in the 1930s, for a long time numerical solutions were only possible for atoms. The breakthrough in molecular calculations was due to Hall and Roothaan in 1951,[30, 31] who transformed the Hartree-Fock system of integro-differential equations into a much more computationally effective matrix eigenvalue equation.

Let us assume that we are dealing with a closed-shell system with \(2n\) electrons, \(n\) with \(\alpha\) spin and \(n\) with \(\beta\) spin. We can choose two types of spin orbitals for this problem: restricted spin orbitals, which are constrained to have the same MOs (spatial functions) for \(\alpha\) and \(\beta\) electrons; and unrestricted spin orbitals, which have different MOs for \(\alpha\) and \(\beta\) spins.

The Roothan-Hall method adopts restricted spin orbitals, which have the form

\[
\chi_i(\mathbf{x}) = \psi_j(\mathbf{r})\alpha(s) \tag{5.59}
\]

\[
\chi_{i+1}(\mathbf{x}) = \bar{\psi}_j(\mathbf{r})\beta(s) \tag{5.60}
\]

Thus, the closed-shell determinant can be written as

\[
\Phi = |\chi_1 \chi_2 \cdots \chi_{2n-1} \chi_{2n} \rangle = |\psi_1 \bar{\psi}_1 \cdots \psi_n \bar{\psi}_n \rangle \tag{5.61}
\]
By using this ansatz for the wave function, it is possible to integrate over all spin coordinates to obtain the closed-shell spatial Hartree-Fock or simply restricted Hartree-Fock (RHF) equations

\[ f(r_1)\psi_j(r_1) = \epsilon_j\psi_j(r_1) \quad j = 1, 2, \ldots, n \]  

(5.62)

where the closed-shell Fock operator has the form

\[ f(r_1) = h(r_1) + \sum_a n \int J_a(r_1) - K_a(r_1) \]  

(5.63)

and the closed-shell Coulomb and exchange operators are

\[ J_a(r_1) = \int \frac{\psi_a(r_2)\psi_a^*(r_2)}{|r_1 - r_2|} dr_2 \]  

(5.64)

\[ K_a(r_1)\psi_i(r_1) = \left[ \int \frac{\psi_a(r_2)\psi_i(r_2)}{|r_1 - r_2|} dr_2 \right] \psi_a(r_1) \]  

(5.65)

Even if spin-free, (5.62) still remains a system of integro-differential equations with no practical procedures for obtaining numerical solutions for molecules.

Roothaan’s contribution was to show that by expanding the MOs in a set of \( N \) known spatial basis functions, the differential equations would become a set of algebraic equations which is solvable by standard matrix techniques. In particular, if we apply the LCAO approximation

\[ \psi_i = \sum_{\mu=1}^{N} C_{i\mu} \phi_\mu \quad i = 1, 2, \ldots, N \]  

(5.66)

then Eq. (5.62) takes the form

\[ \sum_{\nu=1}^{N} F_{\mu\nu}C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu}C_{\nu i} \quad i = 1, 2, \ldots, N \]  

(5.67)

where the elements of the overlap matrix are

\[ S_{\mu\nu} = \int \phi_\mu(r_1)\phi_\nu(r_1) dr_1 \]  

(5.68)
and those of the Fock matrix are

\[ F_{\mu\nu} = \int \phi_{\mu}(r_1) \phi_{\nu}^*(r_1) f(r_1) \, dr_1 \]  

(5.69)

Equations (5.67) are known as the Roothaan-Hall equations, and they can be written more compactly in the famous matrix form

\[ FC = SC\epsilon \]  

(5.70)

where \( C \) is the MO coefficients matrix containing all the orbital expansion coefficients \( \{C_{\mu i}\} \).

**Unrestricted Hartree-Fock: The Pople-Nesbet Equations**

Obviously, not all molecules, nor all states are can be described by a closed-shell electronic configuration. In order to model open-shell systems, we need to use an unrestricted set of spin orbitals

\[ \chi_i(x) = \psi_\alpha^i(x) \equiv \psi_\alpha^i(r) \alpha(s) \]  

(5.71)

\[ \chi_{i+1}(x) = \bar{\psi}_\beta^i(x) \equiv \bar{\psi}_\beta^i(r) \beta(s) \]  

(5.72)

That is the \( \alpha \) electrons are described by a set of spatial orbitals \( \{\psi_\alpha^i\} \), and the \( \beta \) electrons are described by a different set of spatial orbitals \( \{\psi_\beta^i\} \).

In this case, the wave function is the unrestricted determinant

\[ \Phi = |\chi_1 \chi_2 \cdots \chi_i \chi_{i+1} \cdots \rangle = |\bar{\psi}_1^\alpha \bar{\psi}_1^\beta \cdots \psi_1^\alpha \bar{\psi}_1^\beta \cdots \rangle \]  

(5.73)

The adoption of these unrestricted spin orbitals in Eq. (5.58) yields, after integration over spin variables, the following *unrestricted Hartree-Fock* (UHF) equations

\[ f^\alpha(r_1) \psi_\alpha^i(r_1) = \epsilon_\alpha^i \psi_\alpha^i(r_1) \]  

(5.74)

\[ f^\beta(r_1) \psi_\beta^i(r_1) = \epsilon_\beta^i \psi_\beta^i(r_1) \]  

(5.75)
where the $\alpha$ and $\beta$ spatial Fock operators are

$$f^\alpha(r_1) = h(r_1) + \sum_{\alpha} J^\alpha_{\alpha}(r_1) - K^\alpha_{\alpha}(r_1) + \sum_{\alpha} J^\beta_{\alpha}(r_1)$$  \hspace{1cm} (5.76)

$$f^\beta(r_1) = h(r_1) + \sum_{\alpha} J^\beta_{\alpha}(r_1) - K^\beta_{\alpha}(r_1) + \sum_{\alpha} J^\beta_{\alpha}(r_1)$$  \hspace{1cm} (5.77)

with $n_\alpha$ and $n_\beta$ being the number of $\alpha$ and $\beta$ electrons, respectively. The new spin-specific Coulomb and exchange operators are defined as

$$J^\sigma_{\alpha}(r_1) = \int \frac{\psi^\sigma_{\alpha}(r_2)}{|r_1 - r_2|} \psi^\sigma_{\alpha}(r_2) dr_2$$  \hspace{1cm} (5.78)

$$K^\sigma_{\alpha}(r_1) \psi^\sigma_{\alpha}(r_1) = \left[ \int \frac{\psi^\sigma_{\alpha}(r_2)}{|r_1 - r_2|} \psi^\sigma_{\alpha}(r_2) dr_2 \right] \psi^\sigma_{\alpha}(r_1)$$  \hspace{1cm} (5.79)

where $\sigma \in \{\alpha, \beta\}$.

From the definitions for $f^\alpha$ and $f^\beta$, we notice that the two sets of equations (5.74) and (5.75) are coupled and cannot be solved independently.

In order to solve the UHF equations, we proceed as in the RHF case by introducing an atomic orbital basis for the expansion of the MOs

$$\psi_i^\alpha = \sum_{\mu=1}^{N} C^\alpha_{\mu i} \phi_\mu \quad i = 1, 2, \ldots, N$$  \hspace{1cm} (5.80)

$$\psi_i^\beta = \sum_{\mu=1}^{N} C^\beta_{\mu i} \phi_\mu \quad i = 1, 2, \ldots, N$$  \hspace{1cm} (5.81)

This leads to the formulation of the UHF equation in the matrix form

$$F^\alpha C^\alpha = SC^\alpha \epsilon^\alpha$$  \hspace{1cm} (5.82)

$$F^\beta C^\beta = SC^\beta \epsilon^\beta$$  \hspace{1cm} (5.83)

which are known as the Pople-Nesbet equations.

The Pople-Nesbet equations represent a generalization of the Roothan-Hall equations, in fact the former collapse into the latter when $\{\psi_i^\alpha\} = \{\psi_i^\beta\}$, that is when the $\alpha$ and the $\beta$ MOs are identical. For this reason we will now discuss only the procedure to solve the Pople-Nesbet equations.
The Self-Consistent Field Procedure

As just discussed, the problem of determining the unrestricted Hartree-Fock molecular orbitals \( \{ \psi_i^\alpha \} \) and \( \{ \psi_i^\beta \} \) and their orbital energies \( \{ \epsilon_i^\alpha \} \) and \( \{ \epsilon_i^\beta \} \), involves solving the matrix equations (5.82) and (5.83).

However, to proceed with the solution methodology we first need the explicit expressions for the ground state energy \( E_0 \) and for the Fock matrices. These are best expressed in terms of the \( \alpha \) and \( \beta \) density matrices \( P^\alpha \) and \( P^\beta \), whose elements are defined by the equations

\[
P^\alpha_{\mu\nu} = \sum_{a=1}^{n_\alpha} C^\alpha_{\mu a} (C^\alpha_{va})^\dagger \tag{5.84}
\]

\[
P^\beta_{\mu\nu} = \sum_{a=1}^{n_\beta} C^\beta_{\mu a} (C^\beta_{va})^\dagger \tag{5.85}
\]

The definition of a total density matrix, \( P^T \), as

\[
P^T = P^\alpha + P^\beta \tag{5.86}
\]

is beneficial to our task.

Using these definitions the ground state energy can be written as

\[
E_0 = \frac{1}{2} \sum_{\mu=1}^N \sum_{\nu=1}^N \left[ P^T_{\mu\nu} H^\text{core}_{\mu\nu} + P^\alpha_{\mu\nu} F^\alpha_{\mu\nu} + P^\beta_{\mu\nu} F^\beta_{\mu\nu} \right] \tag{5.87}
\]

with the \( \alpha \) and \( \beta \) Fock matrix elements being

\[
F^\alpha_{\mu\nu} = H^\text{core}_{\mu\nu} + \sum_{\lambda=1}^N \sum_{\sigma=1}^N P^T_{\lambda\sigma} \langle \mu\sigma | \nu\lambda \rangle - P^\alpha_{\lambda\sigma} \langle \mu\sigma | \lambda\nu \rangle \tag{5.88}
\]

\[
F^\beta_{\mu\nu} = H^\text{core}_{\mu\nu} + \sum_{\lambda=1}^N \sum_{\sigma=1}^N P^T_{\lambda\sigma} \langle \mu\sigma | \nu\lambda \rangle - P^\beta_{\lambda\sigma} \langle \mu\sigma | \lambda\nu \rangle \tag{5.89}
\]

where the elements of the core-Hamiltonian matrix are trivially

\[
H^\text{core}_{\mu\nu} = \int \phi_\mu (r_1) \dagger h(r_1) \phi_\nu (r_1) dr_1 \tag{5.90}
\]
The UHF problem is solved when the optimal spin orbitals are found. Once the atomic basis is introduced, the problem translates in finding the optimal MO coefficients \( \{C_{\mu}^\alpha\} \) and \( \{C_{\mu}^\beta\} \). In fact, from these one can form the density matrices \( P^\alpha \), \( P^\beta \) and \( P^T \), which are needed for the evaluation of the Fock matrices \( F^\alpha \) and \( F^\beta \), and therefore of the energy \( E_0 \).

However, since the Fock matrices depend on the density matrices or equivalently on the MO coefficients, which are also the solution to the problem, that is \( F^\alpha = F^\alpha(C^\alpha, C^\beta) \) and \( F^\beta = F^\beta(C^\alpha, C^\beta) \), the Pople-Nesbet equations are nonlinear

\[
F^\alpha(C^\alpha, C^\beta)C^\alpha = SC^\alpha \epsilon^\alpha
\]

\[\text{(5.91)}\]

\[
F^\alpha(C^\alpha, C^\beta)C^\beta = SC^\beta \epsilon^\beta
\]

\[\text{(5.92)}\]

and must be solved iteratively.

The iterative scheme adopted to solve the Pople-Nesbet equations in known as the self-consistent-field (SCF) procedure. For a given molecular system and atomic basis set \( \{\phi_\mu\} \), in broad terms, it consists of the following steps:

1. Calculate all required integrals, \( S_{\mu\nu} \), \( H^\text{core}_{\mu\nu} \) and \( \langle \mu\sigma | \nu\lambda \rangle \);
2. Obtain an initial guess for the MO coefficients and therefore for the density matrices \( P^\alpha \) and \( P^\beta \);
3. Calculate the Fock matrices \( F^\alpha \) and \( F^\beta \);
4. Diagonalize the Fock matrices to obtain \( C^\alpha \), \( C^\beta \), \( \epsilon^\alpha \) and \( \epsilon^\beta \);
5. Form two new density matrices \( P^\alpha \) and \( P^\beta \) using \( C^\alpha \) and \( C^\beta \);
6. Determine whether the procedure has converged, where the most common convergence criterion relies on the fact that at convergence the commutator relation \( [S P F, F P S] = 0 \) becomes true;

(a) If \( [S P F, F P S] = 0 \), then the procedure has converged and the resultant solution represented by \( C \), \( P \), \( F \) can be used to compute the energy \( E_0 \) in Eq. (5.87) and other properties of interest;
(b) If $[SPF, FPS] \neq 0$, then the procedure has not converged, therefore it returns to step (4).

**Forming trial density matrices and the aufbau protocol.** Before concluding our treatment of Hartree-Fock theory, it is beneficial to analyse step (5) of the SCF more in depth. Suppose we are dealing with a molecular system containing $n_\alpha$ $\alpha$-spin electrons and $n_\beta$ $\beta$-spin electrons. Every time the Fock matrices are diagonalized in step (4) of the SCF, $N$ $\alpha$-spin and $N$ $\beta$-spin MOs are obtained, where $N$ is the dimension of the atomic basis set. In particular, the MOs are specified by the column of the $C^\alpha$ and $C^\beta$ matrices resulting from the diagonalization. Now, in order to ensure a decent accuracy $N > n_\alpha, n_\beta$. This implies that there are more MOs than electrons.

Since the density matrices use only the occupied spatial orbitals (see Eqs. (5.84) and (5.84)), in step (5) one needs to choose which MOs to occupy in order to form the new density matrices.

In order to converge to the ground state wave function one follows the aufbau protocol which dictates that one simply occupies, at each cycle in step (5), the $n_\alpha$ and $n_\beta$ orbitals with the associated lowest orbital energies $\epsilon_j^\alpha$ and $\epsilon_j^\beta$.

![RHF molecular orbital diagram](image)

Figure 1.3: RHF molecular orbital diagram.
The resulting orbital picture is sketched in Fig. 1.3, where the MOs of an RHF calculation for a closed-shell $2n$-electron system are ordered according to their orbital energies. Following the aufbau protocol the $2n$ lowest energy MOs are used to form the new density matrices at step (5).

Once the calculation is converged, the orbitals which are occupied are called (guess what!) the occupied orbitals, while the remaining, unoccupied MOs are the virtual orbitals (or simply virtuals). The highest energy occupied MO is known as the Highest Occupied Molecular Orbital (HOMO) and the lowest energy virtual is the Lowest Unoccupied Molecular Orbital (LUMO). We will see soon that virtual MOs play an important role in correlated methods.

1.5.2 Electron Correlation

The HF method treats the electron-electron interaction using a mean-field approximation, that is the interaction is accounted only in an average fashion. In a large basis set HF accounts for 99% of the total energy of a system, however it turns out that the remaining 1% is extremely important for describing its chemistry.

The difference between the exact non-relativistic energy of a system, $E_{\text{exact}}$, and the restricted Hartree-Fock energy $E_{\text{RHF}}$ obtained in the limit that the basis set approaches completeness was defined by Löwdin in 1959 [32] as the correlation energy

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{RHF}}$$

This energy is due to a specific kind of correlation (Coulomb correlation) of the electrons, which is not taken into account by the HF potential $V^{HF}$ in Eq. (5.55). In fact, quoting Löwdin,[33] between two electrons “$i$ and $j$ there is in reality a potential $\mathcal{H}_{ij}$ which, particularly for small distances $r_{ij} \approx 0$, may be tremendously large. If this potential is repulsive, like the Coulomb potential $\mathcal{H}_{ij} = e^2/r_{ij}$, it tries naturally to keep the particles apart, and, since this correlation is entirely neglected in forming the Slater determinant, the corresponding energy is affected by an error which is usually called the ‘correlation energy’.”

Electron correlation has a dual origin:
1. **Coulomb correlation**: as just highlighted by Löwdin’s quote, there is a Coulomb interaction between the electrons that generally decreases the probability of finding two electrons (of any spin) close to each other. Thus, for each couple of electrons “i” and “j” the correct wave function must be such that its $ij$-pair probability density

$$
\rho_{ij}(\mathbf{r}_i, \mathbf{r}_j) = \int \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_n)^\dagger \Psi(\mathbf{x}_1, \ldots, \mathbf{x}_n) d\mathbf{x}_1 \ldots d\mathbf{x}_{i-1} ds_i \ldots d\mathbf{x}_{j-1} ds_j \ldots d\mathbf{x}_n \quad (5.94)
$$

is zero at their coalescence point, that is

$$
\lim_{\mathbf{r}_{ij} \to 0} \rho_{ij}(\mathbf{r}_i, \mathbf{r}_j) = 0 \quad (5.95)
$$

This depression in the pair probability density is called the *Coulomb hole* and it is at the origin of the famous and paramount electron-electron coalescence conditions derived by Kato in 1957.[17]

2. **Fermi correlation**: electrons are indistinguishable and obey Fermi-Dirac statistics, which imposes the antisymmetry requirement to the wave function. This implies that two same-spin electrons cannot be found simultaneously at the same point in space and consequently that the probability of finding one in the immediate vicinity of the other is close to zero. In this sense the correct wave function is such that its pair probability density exhibits a “hole”, the *Fermi hole*.

Fermi correlation is taken into account “automatically” by using Slater determinants. Thus, the correlation energy in Eq. (5.93) refers only to Coulomb correlation.

A minor distinction is worth making between *same-spin* correlation and *opposite-spin* correlation. Same-spin correlation is the Coulomb correlation between electrons having the same spin (not to be mistaken for Fermi correlation). Instead, opposite-spin correlation is between electrons having opposite spin. Since the mathematics of a Slater determinant is such that same-spin electrons are already “kept apart”, the contribution to the correlation energy $E_{\text{corr}}$ due to same-spin correlation is usually minor and its major part is due to opposite-spin correlation.
Another distinction is between dynamic and static correlation. Dynamic correlation is associated with an instantaneous repulsion between the electrons, such as those occupying either the same MOs or nearly spatially equal MOs. The static correlation is usually associated with electrons avoiding each other on a permanent basis, such as those occupying two energetically degenerate orbitals with different spatial distributions. A typical example of where static correlation is very significant is bond breaking. Eventually, the distinction between static and dynamic correlation remains nebulous and mathematically not well-defined.

The coming sections dedicated to wave-function methods will be dealing with “correlated methods”, that is methods which take into account either partially or fully electronic correlation.

**Unrestricted Hartree-Fock, correlation and spin contamination**

Before proceeding with the established correlated methods, it is worth making a note on the correlation problem in the context of unrestricted Hartree-Fock theory. A well-known observation is that RHF cannot properly describe the bond breakage as it usually yields wrong dissociation limits, that is the RHF energy of the system, close to dissociation, is larger than the sum of energies of the isolated component molecules. On the other hand, UHF is known to correctly model dissociation.

This success of UHF is due to the fact that the $\alpha$ and $\beta$ MOs involved into the bond are not constrained to be spatially degenerate. As the bond is stretched the MOs break their spatial degeneracy allowing for the localisation of the electrons on the molecular fragments. Since the electronic localisation is energetically favourable during the bond breakage, this event is associated with a lowering of the UHF energy below the RHF one. The point where the RHF and UHF descriptions start to differ is often referred to as a RHF/UHF instability point.

Since, in this situation, the UHF energy is lower than the RHF one, according to definition (5.93) the UHF wave function has partly introduced electron correlation. However, this correlation comes with a price: spin contamination. Once the MOs spatial symmetry is broken, the wave function is not a pure spin state anymore, but it is contaminated by higher spin multiplicities. The amount of spin contamination can be
1.5. WAVE-FUNCTION METHODS

inferred from the expectation value of the $S^2$ operator, usually denoted by the symbol $\langle S^2 \rangle$. Its correct expectation value for a pure spin state is $\langle S^2 \rangle = S_z(S_z + 1)$, that is 0 for a singlet, 0.75 for a doublet, 2 for a triplet, etc. An UHF singlet, for example, might contain some amount of triplet, quintet and so on.

The $\langle S^2 \rangle$ value of the UHF wave function can be calculated using the formula

$$\langle S^2 \rangle = S_z(S_z + 1) + n_\beta \sum_{ij} |\langle \psi_i^\alpha | \psi_j^\beta \rangle|^2$$

from which it can be easily verified that the RHF wave function will always yield, for example, pure singlet spin states.

1.5.3 Configuration Interaction

The HF method provides the energetically best single-determinant many-electron trial wave function within a one-electron basis of atomic orbitals. As the size of the basis set is increased up to its completeness, the lowest possible single-determinant energy, known as the Hartree-Fock limit, is reached. No further amelioration is possible by variation of the one-electron basis.

However, as previously discussed, a converged $n$-electron closed-shell RHF calculation within an $N$-dimensional basis set yields $N$ MOs, of which $n$ are doubly occupied and the remaining $(N - n)$ virtuals are left unoccupied. As sketched in Fig. 1.4, from these $N$ converged MOs, a set of additional $\binom{N}{n} - 1$ orthonormal Slater determinants can be constructed by promotion of a number of electrons from the occupied to the virtual MOs. Determinants which are obtained by promoting one, two, three, . . . , $n$ electrons in this way are named singly, doubly, triply, . . . , $n$-tuply substituted (or excited) determinants.
In the Configuration Interaction (CI) method, which was devised by Condon in 1930,[34] the many-electron basis for the trial wave function is systematically improved by linearly combining the reference wave function $\Phi_0$ (usually the HF ground state determinant) with a certain number of substituted determinants

$$
\Psi_{CI}^{\text{CL}} = c_0 \Phi_0 + \sum_{i} c_i^r \Phi_i^r + \sum_{i<j}^r c_{ij}^{rs} \Phi_{ij}^{rs} + \cdots
$$

(5.97)

where the $c_0$, $c_i^r$, $c_{ij}^{rs}$, ..., coefficients are the so-called CI amplitudes, $\Phi_i^r$ are singly substituted determinants in which an electron has been promoted from the reference state occupied MO $\psi_i$ to the virtual MO $\psi_r$, $\Phi_{ij}^{rs}$ are doubly substituted determinants and so forth.

Since the CI wavefunction is just a linear combination of Slater determinants, the TISE can be solved in this basis by following the “linear variational method” described in Section 1.3.2, that is by diagonalization of the Hamiltonian matrix $H$ as defined in Eq. (3.24).

If expansion (5.97) includes only all the singly substituted determinants, the method
is referred to as CI-Singles (CIS), if it includes all the singly and doubly substituted determinant CI-Singles-Doubles (CISD) and so forth, up to when all possible substituted determinants are incorporated to form the so-called Full CI (FCI) wave function.

The FCI solution yields the best possible solution within a given one-electron basis. To be more specific it recovers the maximum amount of correlation energy within the basis set. Furthermore, as the size $N$ of the basis set increases the resulting many-electron basis of substituted determinants approaches completeness. It can be shown that, in the limit $N \to \infty$, that is for a sufficiently large one-electron basis, the FCI wave function converges to the exact solution of the TISE.

![Figure 1.5: CI convergence to the exact solution.](image)

In this sense, the major advantage of CI over other correlated methods is that it offers a systematic manner to improve the accuracy of the trial wave function towards exactitude. However, accuracy in CI comes at a high price. In fact, it can be shown that a CI which includes up to $m$-tuply substituted determinant scales as

$$
\mathcal{O}(n^m V^{2+m}) = \begin{cases} 
\mathcal{O}(N^{2m+2}) & \text{if } n \approx V >> m \\
\mathcal{O}(N^{m+2}) & \text{if } V >> n, m
\end{cases}
$$

(5.98)

where $n$ is the number of electrons, $N$ is the size of the one-electron basis and $V$ is the number of virtual orbitals. Equation (5.98) shows that, for example, CIS, CISD and
CISDTQ scale as $O(N^4)$, $O(N^6)$ and $O(N^{10})$, respectively, for typical cases.

When the level of electron promotion $m$ becomes comparable with the size of the basis $N$, that is for FCI or nearly FCI, the computational scaling becomes $O(N_{\text{det}}n^2N^2)$, where $N_{\text{det}}$ is the product of the number of determinants for the alpha and beta electrons separately. Thus, for a system with $k$ alpha and $k$ beta electrons, where $k$ is a large number, the following number of determinants can be obtained via the Stirling approximation for factorials

$$N_{\text{det}}(N = 2k) \approx \frac{16^k}{k\pi} \quad k \text{ large}$$  \hspace{1cm} (5.99)

This exponential scaling makes FCI unrealistic except for quite small molecules. Furthermore, even if the truncated versions (CISD, CISDT, etc.) formally are much lower scaling, these methods violate both size-consistency,[35] which establishes that the energy of a system should equal the energy sum of the separate subsystems at the bond dissociation limit, and the more general size-extensivity requirement,[36] for which the energy should scale linearly with the number of particles in the system.

It is in this necessity of size-extensivity that the next wave-function method finds its origin.

### 1.5.4 Coupled Cluster

The Coupled-Cluster (CC) set of methods stems from a motivation to find an improved theory to truncated CI theory which is inherently size-extensive. It invokes a non-linear expansion of the wave function, which defines a set of non-linear partial differential equations that need to be solved using iterative techniques. For this reason, CC theory is much more involved than the linear expansion analogue of configuration interaction. Since its introduction, first in nuclear physics[37] and then into quantum chemistry in 1966 by Cizek,[38] CC has established its status as perhaps the most reliable yet computationally affordable approximation to the full CI expansion.

The CC wave function is best expressed in terms of excitation operators $\hat{X}$ that
operating on a reference determinant $\Phi_0$ generate substituted determinants, that is

$$\hat{X}_i^r \Phi_0 = c_i^r \Phi_i$$  \hspace{1cm} (5.100)$$

$$\hat{X}_{rs}^{ij} \Phi_0 = c_{ij}^{rs} \Phi_{ij}$$  \hspace{1cm} (5.101)$$

$$\vdots$$

$$\hat{X}_{rs...}^{ij...} \Phi_0 = c_{ij...}^{rs...} \Phi_{ij...}$$  \hspace{1cm} (5.102)$$

For example, using this notation the CI wave function in Eq. (5.97) can be rewritten as

$$\tilde{\Psi}_{\text{CI}} = \left[ 1 + \sum_{ir} \hat{X}_i^r + \sum_{i<j}^{r<s} \hat{X}_{rs}^{ij} + \cdots \right] \Phi_0$$  \hspace{1cm} (5.103)$$

In CC the correlated state is generated from the reference $\Phi_0$ (usually the HF ground state) by applying all possible operators $(1 + \hat{X})$:

$$\tilde{\Psi}_{\text{CC}} = \prod_{ir} (1 + \hat{X}_i^r) \prod_{i<j}^{r<s} (1 + \hat{X}_{rs}^{ij}) \cdots \Phi_0$$  \hspace{1cm} (5.104)$$

$$= \phi_0 + \sum_{ir} \hat{X}_i^r \Phi_0 + \sum_{i<j}^{r<s} (\hat{X}_{rs}^{ij} + \hat{X}_i^r \hat{X}_j^s) \Phi_0 + \cdots$$  \hspace{1cm} (5.105)$$

From Eq. (5.104) a hierarchy of truncated CC wave functions is established as in CI: CC with all single and double excitation operators (CCSD),[39] CC with all single, double and triple excitations (CCSDT)[40] and so forth. However, the analogy stops here. In fact, while the truncated CI model is not size-extensive, truncated CC is. The reason for this substantial difference is brought out by comparing Eq. (5.103) with Eq. (5.105): a CISD wave function includes only connected double excitations arising from the $\hat{X}_{rs}^{ij}$ operators, while the CCSD wave function encloses both connected and disconnected (that is independent) double excitations, which arise from the combination of single excitation operators $\hat{X}_i^r \hat{X}_j^s$. Furthermore, while in CISD excitation higher than doubles are forbidden, CCSD includes higher order disconnected excitations, such as the triples and the quadruples arising from $\hat{X}_i^r \hat{X}_j^t \hat{X}_k^u$ and $\hat{X}_{rs}^{ij} \hat{X}_{kl}^{tu}$, respectively.

As the size of molecular systems is increased or during chemical dissociation, the
disconnected excitations become more and more important in terms of correlation effects.

Thus, it is for the absence of disconnected excitations that truncated CI is not size extensive, in particular the largest error is typically associated with the lack of the disconnected quadruples arising from the product of double excitation operators, that is \( \hat{X}_{ij}^{rs} \hat{X}_{kl}^{tu} \) terms.

If the CC wave function was evaluated in the form of Eq. (5.105), immediately the exponential cost problem would arise. In fact, the number of disconnected cluster coefficients grows factorially with the excitation level.[41]

The solution to this problem is found by considering the excitation operators labeled \( T \), where

\[
T_1 = \sum_{ir} \hat{X}_{ir}^r \\
T_2 = \sum_{i<j} \sum_{r<s} \hat{X}_{rs}^{rs} \\
\vdots
\]

It is now possible to write the CC wave function not as a linear expansion, but by using the following \textit{exponential ansatz}

\[
\hat{\Psi}^{\text{CC}} = \exp(T)\Phi_0
\]

where the \textit{cluster operator} \( T = T_1 + T_2 + \ldots \) and

\[
\exp(T) = 1 + T_1 + T_2 + \ldots + \frac{1}{2!} T_1^2 + T_1 T_2 + \frac{1}{3!} T_1^3 + \ldots
\]

Thus the exponential ansatz automatically generates the disconnected clusters.

Even if the energy cannot be evaluated variationally for technical reasons, one approach[42] is to perform a similarity transformation on the Hamiltonian multiplying it on the left by \( \exp(-T) \), that is

\[
\exp(-T)\mathcal{H}\exp(T)\Phi_0 = E_{\text{CC}}\Phi_0
\]
where \( E_{CC} \) is the coupled-cluster energy. Using the Baker-Campbell-Hausdorff (BCH) expansion, it can be shown that the similarity-transformed Hamiltonian is given \textit{exactly} by the following sum of five terms

\[
\exp(-T)\mathcal{H}\exp(T) = \mathcal{H} + [\mathcal{H}, T] + \frac{1}{2!} [[\mathcal{H}, T], T] + \frac{1}{3!} [[[\mathcal{H}, T], T], T] + \frac{1}{4!} [[[[[\mathcal{H}, T], T], T], T], T]
\]  

(5.111)

Inserting this result into Eq. (5.110) and then projecting on the left either on the reference state \(|\Phi_0\rangle\) or on substituted-determinants \(|\Phi_k\rangle\) yields

\[
\langle \Phi_0 | \mathcal{H} + [\mathcal{H}, T] + \frac{1}{2!} [[\mathcal{H}, T], T] + \frac{1}{3!} [[[\mathcal{H}, T], T], T] | \Phi_0 \rangle = E_{CC}
\]  

(5.112)

\[
\langle \Phi_k | \mathcal{H} + [\mathcal{H}, T] + \frac{1}{2!} [[\mathcal{H}, T], T] + \frac{1}{3!} [[[\mathcal{H}, T], T], T] | \Phi_0 \rangle = 0
\]  

(5.113)

From these equations the CC energy and amplitudes are determined, respectively.

In practice, the cluster operator \( T \) is usually truncated at some level to reduce the computational cost. Truncation after the single and double excitations gives CCSD, which scales as \( \mathcal{O}(N^6) \) with respect to system size. As discussed previously, the CCSD model neglects the connected triple excitations that have been shown to be quite important in many chemical applications. Hence, for very accurate calculations, it is necessary to go beyond singles and doubles and to include triples, performing CCSDT. However, the explicit treatment of triple excitations has an \( \mathcal{O}(N^8) \) complexity and is too computationally expensive for routine applications. In 1989 Raghavachari \textit{et al.} \cite{43} developed a CCSD with a perturbative treatment of the triples which is known as CCSD(T). With a computational complexity of \( \mathcal{O}(N^6) \) in iterative steps and one non-iterative step of complexity \( \mathcal{O}(N^7) \), CCSD(T) is acknowledged as the "\textit{golden standard of quantum chemistry}" , offering an outstanding compromise between very high accuracy and computational cost.
1.5.5 Perturbation Theory

The basic idea of perturbation theory (PT) is to obtain an approximate solution of a complicated quantum mechanical system by breaking its Hamiltonian into a sum consisting of a weak disturbance to a simpler system for which the solution is known. If the disturbance, or perturbation, is not too large, the various physical quantities associated with the complicated system can be studied based on knowledge of the simpler one.

Rayleigh-Schrödinger perturbation theory is among the most popular perturbation theories used in quantum mechanics.\[44, 45\] Mathematically, this technique expresses the exact Hamiltonian $H$ as an unperturbed part $H_0$ and a small perturbation $V$ multiplied by a dimensionless parameter $\lambda$.

$$ H_0 + V $ = \sum_{n=0}^{\infty} \lambda^n \Psi^{(n)} $ = E \Psi $$

The eigenfunctions $\Psi^{(0)}$ and the eigenvalues $E^{(0)}$ of $H_0$ are known,

$$ H_0 \Psi^{(0)} = E^{(0)} \Psi^{(0)} $$

The unknown wavefunction and the energy are then expanded in the power series,

$$ \Psi = \sum_{n=0}^{\infty} \lambda^n \Psi^{(n)} $$

$$ E = \sum_{n=0}^{\infty} \lambda^n E^{(n)} $$

By inserting the power series into the Schrödinger equation, and equating coefficients of
1.5. WAVE-FUNCTION METHODS

each power of $\lambda$, various orders of correction for the energy can be determined as

\begin{align*}
E^{(0)} &= \langle \Psi^{(0)} | H_0 | \Psi^{(0)} \rangle \\
E^{(1)} &= \langle \Psi^{(0)} | V | \Psi^{(0)} \rangle \\
E^{(2)} &= \langle \Psi^{(0)} | V | \Psi^{(1)} \rangle \\
E^{(3)} &= \langle \Psi^{(0)} | V | \Psi^{(2)} \rangle \\
& \vdots \\
E^{(n)} &= \langle \Psi^{(0)} | V | \Psi^{(n-1)} \rangle
\end{align*}

where the $n$-th order energy correction is determined by the $(n-1)$-th order correction to the wave function.

In 1934, Møller and Plesset described in a short note of just five pages how Rayleigh-Schrödinger perturbation theory can be applied to the calculation of electron correlation.\[46\] These authors suggested using the HF wave function as the starting point of the perturbation expansion, that is $\Psi^{(0)} = \Phi$. Therefore the zeroth-order problem becomes

$$H_0 \Phi = E_0^{(0)} \Phi$$

where the zeroth-order, unperturbed Hamiltonian $H^{(0)}$ is defined as the sum of Fock operators

$$H_0 = \sum_i f(r_i)$$

and the perturbation is given by

$$V = H - H^{(0)}$$

One can then evaluate the energy corrections by using Eqs. (5.118)--(5.122). According to these equations, to the zeroth-order (MP0), the energy is the sum of the HF eigenvalues, and the first order of perturbation (MP1) recovers the HF energy. The correction at the second-order (MP2) is more interesting. One in fact needs to determine the first order correction $\Psi^{(1)}$ in order to evaluate $E^{(2)}$. Since the reference wave function is the HF one, singly substituted determinants give no contribution to
ψ(1) due to Brillouin theorem. In addition, triple and higher excitations do not mix with Φ because of the two-particle nature of the perturbation. Therefore, the first-order correction can be exactly expanded in terms of only doubly substituted determinants |Φrs\rangle, yielding the famous MP2 energy correction

\[ E^{(2)} = \frac{1}{4} \sum_{ij} \sum_{rs} |\langle \psi_i \psi_j | \psi_r \psi_s \rangle|^2 \]

The theory can be applied up to any level of perturbation, but at the second-order it already recovers a large fraction of the correlation energy, while incurring a relatively inexpensive O(N^5) cost with respect to the size of basis set. Moreover, it has been shown that the convergence of the perturbation series is poor,[47, 48] and so higher-order MPn theory is not commonly used.

The two past decades have seen intensive development focused on MP2, where the goals of accomplishing low-order scaling and increasing the reliability of the results were pursued. These developments have been extremely successful and, as a result, have established MP2 as a valuable tool for treating large molecules with hundreds of atoms.

1.6 Density-Functional Methods

Density functional theory (DFT)[29, 49] has essentially become the de facto method of choice among computational chemists in recent years. This is because the present-day hybrid functionals (such as B3LYP) offer a good balance between accuracy and computational cost, thereby allowing much larger systems to be treated compared to traditional ab initio methods, while retaining much of their accuracy. The significance of DFT has been recognised through the award of a shared 1998 Nobel Prize in chemistry to its founding father Walter Kohn.

1.6.1 Early approaches

As discussed in Section 1.4, the key idea of DFT is to express the electronic energy as a functional E[ρ] of the electronic density ρ. In particular, supported by the theoretical
1.6. DENSITY-FUNCTIONAL METHODS

justification of the first Hohenberg-Kohn theorem one may write

$$E[\rho] = V_{ne}[\rho] + T[\rho] + V_{ee}[\rho] = \int \rho(r) V_{ext}(r) \, dr + F[\rho] \quad (6.127)$$

where $T[\rho]$ is the kinetic energy, $V_{ee}[\rho]$ is the electron-electron interaction energy and $F[\rho]$ is a universal functional of $\rho$

$$F[\rho] = T[\rho] + V_{ee}[\rho] = \langle \Psi | T_{e} + V_{ee} | \Psi \rangle \quad (6.128)$$

The electron-electron interaction can be decomposed further as

$$V_{ee}[\rho] = J[\rho] + E_{X}[\rho] + E_{C}[\rho] \quad (6.129)$$

where $E_{X}[\rho]$ and $E_{C}[\rho]$ are the exchange and the correlation energy functionals, while $J[\rho]$ is the classical electron-electron repulsion energy

$$J[\rho] = \frac{1}{2} \int \rho(r_1) \rho(r_2) \, \frac{d \rho_1}{\rho_1 - \rho_2} \, dr_1 dr_2 \quad (6.130)$$

Since $V_{ne}[\rho]$ and $J[\rho]$ are explicit functionals of the density, the unknown functionals $T[\rho]$, $E_{X}[\rho]$ and $E_{C}[\rho]$ constitute the Gordian knot of DFT.

The earliest DFT approaches of the 1920s (before Hohenberg and Kohn) by Thomas, Fermi and Dirac\[50, 51, 52\] attempted to find accurate forms for the kinetic and exchange energy functionals

$$T[\rho] = C_{T} \int \rho^{5/3}(r) \, dr \quad (6.131)$$

$$E_{X}[\rho] = C_{X} \int \rho^{4/3}(r) \, dr \quad (6.132)$$

These approaches were of little quantitative use, especially because of unacceptably large errors in the kinetic energy $T[\rho]$. In fact, one might recall from the virial theorem that the kinetic energy is of the order of the total energy; even a small error in such a large quantity renders the theory inaccurate. Another major defect is that no binding is predicted by this formalism.\[53\] It will be only in 1965 that these errors will be reduced to acceptable levels thanks to the revolutionary Kohn-Sham method.
1.6.2 The Kohn-Sham Method

In 1965, realising that the failure of the Thomas-Fermi model was associated with the determination of the kinetic energy, Kohn and Sham circumvented this problem by devising a methodology which is the foundation of the use of DFT in computational chemistry.[49]

The key idea was to replace the true kinetic energy by the kinetic energy of a model system having the same density but non-interacting electrons. Thus, Kohn and Sham defined

$$F[\rho] = T_s[\rho] + J[\rho] + E_{XC}[\rho]$$

(6.133)

where $T_s[\rho]$ is the kinetic energy of the non-interacting system. Since $F[\rho]$ is given by Eq. (6.128) it follows that

$$E_{XC}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]$$

(6.134)

This exchange-correlation energy $E_{XC}[\rho]$, which is the key quantity in modern DFT, constitutes just a residual correction as $T[\rho] - T_s[\rho] \ll T[\rho]$. Thus, as opposed to Thomas-Fermi approach, a small error in $E_{XC}[\rho]$ would yield still acceptably accurate electronic energies.

Expressing the electronic structure problem for an $n$-electron non-interacting system is trivial

$$\mathcal{H}_s = -\frac{1}{2} \sum_{i=1}^{n} \nabla_i^2 + \sum_{i=1}^{n} \mathcal{V}_s(r_i)$$

(6.135)

where $\mathcal{V}_s$ in the external potential in which the non-interacting electrons are moving. It can be shown that such system has the same electronic density as the real system if it complies with the following condition on its external potential

$$\mathcal{V}_s(r) = \mathcal{V}_{ext} + \frac{\delta J[\rho]}{\delta \rho(r)} + \frac{\delta E_{XC}[\rho]}{\delta \rho(r)}$$

(6.136)

Since Hamiltonian (6.135) is separable, its exact wave function is simply a single determinant constructed from orbitals that are solutions to the eigensystem

$$\left[ -\frac{1}{2} \nabla_i^2 + \mathcal{V}_s(r) \right] \psi_i(r) = \epsilon_i \psi_i(r)$$

(6.137)
defining the density of the non-interacting system – which is identical to the density of the real system – as

\[ \rho(r) = \sum_{i}^{n} \psi_i(r) \dagger \psi_i(r) \]  

(6.138)

and finally its kinetic energy

\[ T[\rho] = \sum_{i}^{n} \langle \psi_i | - \frac{1}{2} \nabla_i^2 | \psi_i \rangle \]  

(6.139)

This formalism recasts the DFT problem similar to the HF one. In fact the optimal Kohn-Sham MOs \( \psi_i(r) \), and therefore the energy, have to be determined via the SCF procedure.

### 1.6.3 Exchange-Correlation Functionals

The performance of DFT relies mainly on the sophistication of the approximate exchange-correlation functional \( E_{XC}[\rho] \), whose development has been a significant aspect of research in DFT. The current types of exchange-correlation functionals can be classified into five relatively well-established classes: i) functionals including the local spin density approximation (LSDA), which relies on the electron density; ii) those based on the generalised gradient approximation (GGA), which involve not only the density but also its gradient; iii) hybrid functionals that linearly combine GGA and exact HF exchange; iv) the meta-GGAs that incorporate the kinetic energy density; v) double-hybrid functionals that combine DFT with correlation energy evaluated using the MP2 method. The relative proportion of each component in the functional is normally obtained by fitting to experimental data.

Hybrid functionals are known to significantly improve the accuracy of DFT results. Unfortunately, the inclusion of exact exchange introduces non-locality into the exchange-correlation (XC) potential and precludes the use of fast evaluation techniques such as the Continuous Fast Multipole Method (CFMM).\[54\] This means that calculations using hybrid functionals are significantly slower than those using LSDA or GGA functionals, a weakness that has restricted their application to relatively smaller systems.
PART A

SINGLE-DETERMINANT THEORY OF ELECTRONIC EXCITED STATES
Introduction to Part A

Understanding electronically excited states is important in many fields such as photovoltaics, optics, synthetic chemistry, and biology. Quantum chemistry has played a key role in improving this understanding, but while ground states are usually studied using a small set of well-established methods such as Density Functional Theory (DFT)\cite{29, 49}, Møller-Plesset Perturbation Theory and Coupled-Cluster Theory, the arsenal of approaches to excited states is extremely large.

The abundance of excited state methods results partly from the belief that, whereas ground states are often described well by single-determinant methods, excited states are usually multi-reference in character, especially if the state is doubly-excited or in the vicinity of a conical intersection.\cite{55, 56} However, specification of a multi-reference wavefunction is more difficult and often requires a delicate combination of chemical intuition, experience, trial and error, and incurs a substantially higher computational cost.

The single-reference CIS and TD-DFT methods address some of these concerns but, notwithstanding their success, have some important weaknesses. Both CIS and TD-DFT (within the adiabatic local density approximation\cite{57}) are incapable of describing doubly-excited states\cite{58, 59, 60, 61} and often fail near conical intersections.\cite{61} Moreover, CIS can overestimate excitation energies by as much as 2 eV\cite{57} and TD-DFT performs poorly for Rydberg\cite{62, 63, 64, 65} and charge-transfer states.\cite{66, 67, 68, 69}

These weaknesses arise because neither CIS nor TD-DFT allow the molecular orbitals (MOs) to relax in the excited state and, consequently, they struggle in cases where that relaxation would be significant. In fact, these methods treat excited and ground states on a different footing, whereas properties of the former are obtained via a linear response approach in which the ground-state determinant is adopted as reference for all the excitations.

One may reasonably ask, however, whether it is either necessary or desirable to treat ground and excited states as if they were fundamentally different. Such a methodological discontinuity is difficult to justify on either physical or mathematical grounds and, indeed, by singling out the ground state for special treatment, we introduce a bias that can be hard to remove later. Is it not more natural, one may wonder, to obtain excited states,
at least in a first-order approximation, by an SCF procedure?

In this Part A, we present a single-determinant theory for electronic excited states. The theory is based on simple algorithms which replace the *aufbau* protocol, the Maximum Overlap Method (MOM) and the Initial Maximum Overlap Method (IMOM), for finding excited-state solutions to self-consistent field (SCF) equations. In this way, the resulting excited-state models are obtained and can be treated in the same way as ground-state solutions.

In Chapter 2 we present a short introduction to the state-of-the-art excited-state methods, with particular emphasis on those approaches that will be encountered in the remainder of this Part A. Chapter 3 introduces our single-determinant theory of excited states, therefore the MOM algorithm, and examines the extent to which higher single-determinant Hartree-Fock solutions are reasonable first-order approximations to excited states of the H₂ molecule. In Chapter 4, we ask how well the single-determinant approximation can handle challenging cases, such as double excitations, conical intersections and charge-transfer states, where other low-cost excited-state methods either perform poorly or fail completely. Finally, in Chapter 5 we exploit the structural simplicity of the single-determinant scheme to rigorously define multiply excited states.
Chapter 2

A Short Introduction to
Excited-State Methods

“One, No One, and One Hundred Thousand”

Luigi Pirandello

2.1 Introduction

Ground states are usually studied using a small set of well-established methods such as Density Functional Theory (DFT),[29, 49] Møller-Plesset Perturbation Theory and Coupled-Cluster Theory. On the other hand, the arsenal of approaches to excited states is very large, including Multi-Reference Configuration Interaction (MRCI),[70, 71, 72] Complete Active Space Configuration Interaction (CASCI),[73] Complete Active Space Self-Consistent Field (CASSCF),[74, 75] Restricted Active Space Self-Consistent Field (RASSCF),[76] CASPT2,[77, 78, 79] Multi-Reference Møller-Plesset Perturbation theory (MRMP),[80] Symmetry-Adapted Cluster-Configuration Interaction (SAC-CI),[81, 82, 83] Equation-Of-Motion Coupled Cluster (EOM-CC),[84, 85, 86] Linear Response Coupled Cluster (LR-CC),[87, 88] Configuration Interaction Singles (CIS),[89, 90, 91] CIS with perturbative treatment of doubles (CIS(D)),[90] Time-Dependent DFT (TDDFT),[92, 93, 94, 95] Constrained DFT,[96] many-body Green’s functions methods
2.1. INTRODUCTION

This plethora of excited state methods ultimately results from the observation or the belief that electronic excited states are intrinsically much more complex than ground states. Quoting Serrano-Andrés and Serrano-Pérez:

"Computing electronic excited states with quantum chemical methods is much more complex than doing it for ground states, because it implies not only coping with higher solutions of the electronic Hamiltonian with diverse character and therefore requiring more complex methods, but also to solve a plethora of new situations, such as hypersurface crossings and coupling between states, that usually requires to abandon convenient approximations like Born-Oppenheimer, for instance."

This Chapter 2 provides a compact description of some of the most relevant and employed quantum chemical methods for electronic excited states, with particular emphasis on those approaches that will be encountered in the remainder of this thesis. Many of these methods rely on a reference wave function, which is used as a baseline for the calculation. To this aim, it is beneficial to introduce a classification for the methods (sketched in Fig. 2.1) based on the number of electronic configurations used to build the reference wave function.

- **Single-Reference Methods.** Within these methods the reference is a single-determinant wave function (that is either the HF or a DFT one), which determines an optimal set of MOs. Typically the reference wave function is optimized for the ground state. If the wave function is the HF determinant, the excited state calculation is performed via post-HF methods such as CI, CC or PT using the MOs of the reference. Similarly, if the wave function is a Kohn-Sham determinant, the excited state calculation is a post-SCF procedure typically based on a linear response approach.

- **Multi-Reference Methods.** The reference wave function is a linear combination of a number of selected determinants. Both the expansion amplitudes and the MOs are optimized for the entire linear combination of determinants via a Multi-
Reference Self-Consistent Field (MRSCF) procedure. The reference wave function can be optimized for the ground state or for the target excited state, or the optimization can be “averaged” over a given set of states. Such a wave function is able to inherently model the static correlation effects, which are believed to be paramount in the excited state chemistry. Post-MRSCF methods are frequently applied in order to correct the poor dynamic correlation treatment of the MRSCF wave function. The latter methods are usually either CI-based (MRCI) or perturbative approaches (MRPT).

<table>
<thead>
<tr>
<th>Excited-state methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single-Reference</strong></td>
</tr>
<tr>
<td>Variational</td>
</tr>
<tr>
<td>CI</td>
</tr>
<tr>
<td>(CIS, CISD, ...)</td>
</tr>
<tr>
<td>Non-Variational</td>
</tr>
<tr>
<td>CIS(D)</td>
</tr>
<tr>
<td>TD-DFT</td>
</tr>
<tr>
<td>EOM-CC</td>
</tr>
<tr>
<td>CCn</td>
</tr>
<tr>
<td>SAC-CI</td>
</tr>
<tr>
<td>Single-Determinant</td>
</tr>
<tr>
<td>(HF, DFT)</td>
</tr>
<tr>
<td><strong>Multi-Reference</strong></td>
</tr>
<tr>
<td>Variational</td>
</tr>
<tr>
<td>MRSCF</td>
</tr>
<tr>
<td>(CASSCF, RASSCF)</td>
</tr>
<tr>
<td>MRCI</td>
</tr>
<tr>
<td>(CASCI, RASCI)</td>
</tr>
<tr>
<td>Non-Variational</td>
</tr>
<tr>
<td>MRPT</td>
</tr>
<tr>
<td>(CASPT2)</td>
</tr>
</tbody>
</table>

Figure 2.1: Classification of the most common excited-state methods. The single-determinant approach introduced in Chapter 3 is an intrinsically single-reference method.

**Multi-Reference and Multi-Configurational Methods**

“Multi-Reference Methods” are often referred to as “Multi-Configurational Methods”. Actually these terms are pretty much used interchangeably. The reason is that usually Multi-Reference wave functions are made of linear combinations of Configuration State Functions (CSFs).

CGFs are defined in terms of the *occupation-number* (ON) operators $N_p$, with each
operator counting the number of electrons in a spin orbital $\chi_p$, that is

$$N_p|\Phi_k\rangle = \delta_{kp}|\Phi_k\rangle = k_p|\Phi_k\rangle \quad (1.1)$$

where $|\Phi_k\rangle$ is a Slater determinant, $\delta_{kp}$ is a Kronecker delta, and $k_p$ is equal to one if $\chi_p$ is occupied and it is equal to zero if $\chi_p$ is unoccupied. Now, the following orbital ON operators

$$N^\alpha_p = N^\alpha_{p\alpha} + N^\alpha_{p\beta} \quad (1.2)$$

with $\alpha$ and $\beta$ denoting the orbital spin functions, commute with the total spin ($S^2$) and the projected spin ($S_z$) operators

$$[N^\alpha_p, S^2] = 0 \quad (1.3)$$
$$[N^\alpha_p, S_z] = 0 \quad (1.4)$$

Since the spin-symmetry of the exact wave function requires it to be an eigenfunction of both $S^2$ and $S_z$, and given Eqs. (1.3) and (1.4), one can set up a many-electron basis of functions that are simultaneously eigenfunctions of the orbital ON operators as well as of the projected and total spin operators. Such spin-adapted functions are the CGFs.[70]

Clearly, different Slater determinants may have the same orbital occupation number $N^\alpha_p$, but different spin-orbital occupation numbers $N^\alpha_{p\alpha}$ and $N^\alpha_{p\beta}$. The set of all determinants with the same orbital ONs but different spin-orbital ONs constitutes an orbital configuration. Thus, a configuration is made of all the CSFs which are degenerate with respect to the orbital ON operators.

A Multi-Configurational wave function is a linear combination of CSFs belonging to different configurations. A Multi-Reference wave function is just a linear combination of Slater determinants. The first is a special case of the latter.
2.2 Single-Reference Methods

2.2.1 CIS

Configuration Interaction Singles (CIS) is the conceptually simplest wave-function based method for the calculation of electronic excitation energies and excited-state properties. The CIS wave function is a linear combination of singly-substituted determinants

\[ \Psi^{\text{CIS}} = \sum_{ir} c_{ir} \Phi_{ir}^{r} \]  

(2.5)

This ansatz for the wave function is substituted into the exact TISE, then the projection onto the space of singly-substituted determinants \( \langle \Phi_{j}^{s} | \) yields

\[ \sum_{ir} c_{ir}^{r} \langle \Phi_{j}^{s} | \mathcal{H} | \Phi_{ir}^{r} \rangle = E_{\omega} \sum_{ir} c_{ir}^{r} \delta_{ij} \delta_{rs} \]  

(2.6)

with

\[ \langle \Phi_{j}^{s} | \mathcal{H} | \Phi_{ir}^{r} \rangle = (E_{0} + \epsilon_{r} - \epsilon_{i}) \delta_{ij} \delta_{rs} + \langle \psi_{r} \psi_{i} | \psi_{j} \psi_{s} \rangle \]  

(2.7)

Substituting back Eq. (2.7) into Eq. (2.6), the CIS problem can be recast into the following matrix eigenvalue problem

\[ AX = \omega X \]  

(2.8)

where \( A \) is the Hamiltonian matrix in the space of singles minus a unit matrix multiplied by the constant \( E_{0} \) (that is the ground state energy)

\[ A_{ir,js} = \langle \Phi_{i}^{r} | \mathcal{H} | \Phi_{s}^{j} \rangle - E_{0} \delta_{ij} \delta_{rs} \]  

(2.9)

\[ = \langle \psi_{i} \psi_{j} | \psi_{r} \psi_{s} \rangle + \delta_{ij} \delta_{rs} (\epsilon_{r} - \epsilon_{i}) \]  

(2.10)

The matrix \( X \) contains the CIS expansion coefficients and \( \omega \) is a diagonal matrix of excitation energies \( E_{\omega} - E_{0} \). Thus, the excitation energies are obtained simply solving the secular equation

\[ (A - \omega)X = 0 \]  

(2.11)

Here are some good properties of the CIS methodology. CIS is a relatively cheap
method, formally scaling as HF, that is as $O(N^4)$, though the scaling prefactor can be much larger. Since CIS solutions are obtained variationally, the method complies with the Hylleraas-Undheim-MacDonald theorem for which each solution represents an upper bound to the exact energy of the homologous excited state. It is also remarkable that, due to Brillouin’s theorem, CIS is size-consistent. Furthermore, the method can be set to yield pure spin-states for closed-shell systems.

Most of the disadvantages of CIS are related to the lack of orbital relaxation, that is the ground-state HF orbitals are not a particularly good basis for the expansion of the correlated excited-state wave function. This leads CIS to usually overestimate excitation energies by about $0.5 - 2$ eV compared to the experimental values. Additionally, given that the singly-substituted determinants bring little or no dynamic correlation, the method fails whenever such correlation does not cancel appropriately between the excited and the ground state. In principle, this problem can be solved by performing CIS(D), where correlation is added perturbatively to the CIS wave function at the MP2 cost. Finally, CIS is structurally incapable of modelling double excitations (for obvious reasons) and, except for very few exceptions, fails to reproduce excited states in proximity of conical intersections.

2.2.2 TD-HF

Time-Dependent Hartree-Fock (TD-HF) is not an extremely common method for electronic excited states. However, we have decided to introduce it here as it represents the theoretical connection between CIS and TD-DFT.

The time-dependent Hartree-Fock equations, that is the set of equations obtained by using a time-dependent single-determinant ansatz for the TDSE, were first derived in 1930 by Dirac. However, what today is commonly meant by TD-HF is not Dirac’s approach, but equations that are obtained in first-order time-dependent perturbation theory (also known as linear response theory) using Dirac’s wave function as a reference.

As anticipated, the starting point is the TDSE

$$i \frac{\partial \Psi(r,t)}{\partial t} = \mathcal{H}(r,t)\Psi(r,t)$$

(2.12)
where the time-dependent electronic Hamiltonian

\[ \mathcal{H}(\mathbf{r}, t) = \mathcal{H}(\mathbf{r}) + \mathcal{V}(\mathbf{r}, t) \]  

(2.13)

includes an arbitrary one-particle time-dependent operator, for example a time-dependent electric field

\[ \mathcal{V}(\mathbf{r}, t) = \sum_i v_i(\mathbf{r}_i, t) \]  

(2.14)

Using the time-dependent HF ansatz for the wave function

\[ \tilde{\Psi}^{\text{TDHF}}(\mathbf{r}, t) = \phi(\mathbf{r}_1, t) \ldots \phi(\mathbf{r}_n, t) \]  

(2.15)

and minimizing the energy yields the time-dependent variant of the HF equations

\[ i \frac{\partial \Phi(\mathbf{r}, t)}{\partial t} = \mathcal{F}(\mathbf{r}, t)\Phi(\mathbf{r}, t) \]  

(2.16)

where the Fock operator \( \mathcal{F}(\mathbf{r}, t) \) contains the mean-field electronic potential, as usual, but also the \( \mathcal{V}(\mathbf{r}, t) \) potential. For a closed-shell system, adopting the time-dependent LCAO approximation

\[ \psi_i(\mathbf{r}, t) = \sum_{\mu} C_{i\mu}(t) \phi_{\mu}(\mathbf{r}) \]  

(2.17)

yields the following time-dependent HF equations

\[ i \frac{\partial P_{\mu i}}{\partial t} = \sum_{\nu=1}^N [F_{\mu\nu} P_{\nu i} - P_{\mu\nu} F_{\nu i}] \quad i = 1, 2, \ldots, N \]  

(2.18)

Assuming that at \( t = 0 \) the molecular system is in a stationary state given by the ground state HF determinant \( \Phi_0(\mathbf{r}) \), if a small time-dependent perturbation is applied the Slater determinant will respond to this perturbation but by changing only slightly, since the perturbation is weak. Thus, if the time-dependent density and the Fock matrices

\[ F_{\mu\nu} = F_{\mu\nu}^{(0)} + F_{\mu\nu}^{(1)} \]  

(2.19)

\[ P_{\mu\nu} = P_{\mu\nu}^{(0)} + P_{\mu\nu}^{(1)} \]  

(2.20)
are the sum of unperturbed \((F_{\mu\nu}^{(0)}, P_{\mu\nu}^{(0)})\) and perturbed time-dependent terms \((F_{\mu\nu}^{(1)}, P_{\mu\nu}^{(1)})\),
then for \(t = 0\) Eq. (2.18) simplifies to

\[
\sum_{\nu=1}^{N} \left[ F_{\mu\nu}^{(0)} P_{\nu i}^{(0)} - P_{\mu\nu}^{(0)} F_{\nu i}^{(0)} \right] = 0 \quad i = 1, 2, \ldots, N
\]  

(2.21)

while for \(t > 0\)

\[
\frac{d P_{\mu i}^{(1)}}{dt} = \sum_{\nu=1}^{N} \left[ F_{\mu\nu}^{(0)} P_{\nu i}^{(1)} - P_{\mu\nu}^{(0)} F_{\nu i}^{(0)} + F_{\mu\nu}^{(1)} P_{\nu i}^{(0)} - P_{\mu\nu}^{(0)} F_{\nu i}^{(1)} \right] \quad i = 1, 2, \ldots, N
\]  

(2.22)

At this point a time-dependent potential is specified, typically as a single-component
time-dependent electric field. Since its contribution is in the first order terms then \(F_{\mu\nu}^{(1)}\)
and \(P_{\mu\nu}^{(1)}\) are completely specified and Eqs. (2.21) and (2.22) can be recast into the
following non-Hermitian matrix eigenvalue problem

\[
\begin{bmatrix}
A & B \\
B^\dagger & A^\dagger
\end{bmatrix}
\begin{bmatrix}
X \\
Y
\end{bmatrix}
= \omega
\begin{bmatrix}
I & 0 \\
0 & -I
\end{bmatrix}
\begin{bmatrix}
X \\
Y
\end{bmatrix}
\]  

(2.23)

where

\[
A_{ir,js} = \langle \Phi_r^* | \mathcal{H} | \Phi_j^s \rangle - E_0 \delta_{ij} \delta_{rs}
\]  

(2.24)

\[
= \langle \psi_r \psi_j | \psi_r \psi_s \rangle + \delta_{ij} \delta_{rs} (\epsilon_r - \epsilon_i)
\]  

(2.25)

\[
B_{ir,js} = - \langle \Phi_0 | \mathcal{H} | \Phi_{ir}^s \rangle
\]  

(2.26)

\[
= \langle \psi_r \psi_j | \psi_s \psi_r \rangle
\]  

(2.27)

and \(\omega\) is the diagonal matrix of excitation energies \(E_\omega - E_0\).

The leading term of \(A\) is just the difference of the energies of orbitals \(\psi_r\) and \(\psi_i\),
from which and to which the electron is excited, respectively. The second term of \(A\)
and the elements of \(B\) stem from the linear response of the Coulomb and the exchange
operators to the first-order changes in the spin orbitals.

While matrix \(A\) provides the coupling between singly-substituted determinants,
the elements of \(B\) involve the interaction of the reference state with doubly excited
configurations and, together with the magnitude of the \(Y\) amplitudes, are a measure of
the missing correlation into the ground state. However, the TD-HF approximation is in
general reasonable only if such correlation effects in the ground state are small.\[57\]

Last but not least, if the $B$ matrix is set to zero, the TD-HF scheme in Eq. (2.23)
reduces to the CIS one in Eq. (2.8). This represents the well-known Tamm-Dancoff
approximation (TDA),\[99\] which is often used in TD-DFT.

TD-HF is roughly twice as computationally expensive as CIS, while still maintaining
a formal scaling of $O(N^4)$. Although in some sense an extension of CIS, TD-HF yields
excitation energies which are usually only slightly lower than CIS and therefore still
severely overestimated. In fact, the effects of the $B$ matrix are small and are supposed
to be small in order for TD-HF to be reasonably accurate. TD-HF can show very
poorly predicted triplet spectra, especially when the ground state wave function is UHF
unstable.\[100\]

2.2.3 TD-DFT

The current use of Time-Dependent DFT (TD-DFT) for excited-state calculations stems
from the “actual” time-dependent theory, however, in a deceptive manner analogous to
TD-HF, it is a linear response approach which is somewhat improperly defined TD-DFT.
For this reason this Section is organised in three parts. The first two parts lay the
fundamentals of the time-dependent theory which are the reference on which the third
part treating the linear response theory, used for excited states, is based.

Fundamentals of TD-DFT

The Hohenberg-Kohn theorems rely on a stationarity assumption, as such the original
Kohn-Sham DFT theory is not applicable to time-dependent systems. Instead, the
formal foundation of time-dependent DFT are the Runge-Gross theorems,\[92\] which
may be summarised as follows

“For a fixed initial state and given an analytic time dependent potential, the mapping
to the time dependent probability density is injective. That is, for the same initial
state, two different external potentials can not give the same probability density function
$\rho(r, t)$.”
In particular Runge and Gross proved that the exact time-dependent electron density, \( \rho(r, t) \), determines the time-dependent potential, \( V_{\text{ext}}(r, t) \), up to an additive constant \( C(t) \), and thus the time-dependent wave function, \( \Psi(r, t) \), up to a time-dependent phase factor \( e^{-i\alpha(t)} \). This establishes the one-to-one mapping between time-dependent densities and time-dependent external potentials, consequently the potential and the wave function themselves are functionals of the density

\[
\rho(r, t) \leftrightarrow V_{\text{ext}}[\rho(r, t)] + C(t) \leftrightarrow \Psi[\rho(r, t)]e^{-i\alpha(t)}
\] (2.28)

Using this result, it can be shown that the exact density can be found by minimization of the action integral

\[
A[\rho(r, t)] = \int_{t_0}^{t_1} \langle \Psi[\rho(r, t)] | i \frac{\partial}{\partial t} - \mathcal{H}(r, t) | \Psi[\rho(r, t)] \rangle \, dt
\] (2.29)

that is by solving the Euler equation

\[
\frac{\delta A[\rho(r, t)]}{\delta \rho(r, t)} = 0
\] (2.30)

Similarly to time-independent DFT, the action integral can be conveniently reformulated as the difference

\[
A[\rho(r, t)] = B[\rho(r, t)] - \int_{t_0}^{t_1} \int \rho(r, t) V_{\text{ext}}(r, t) \, dr dt
\] (2.31)

where

\[
B[\rho(r, t)] = \int_{t_0}^{t_1} \langle \Psi[\rho(r, t)] | i \frac{\partial}{\partial t} - \mathcal{T}_c(r) - V_{ee}(r) | \Psi[\rho(r, t)] \rangle \, dt
\] (2.32)

is a universal functional, with \( V_{\text{ext}}(r, t) = V_{\text{ne}}(r) + V(r, t) \), where \( V(r, t) \) is the time-dependent perturbation.

**Time-Dependent Kohn-Sham Method**

At this point, the Kohn-Sham approach (see Section 1.6.2) is applied to the TDSE. Therefore, the time-dependent electronic structure problem is recast in terms of a non-interacting \( n \)-electron system, whose density is determined by the orbitals \( \{ \psi_i(r, t) \} \),
obtained by solving the time-dependent Kohn-Sham (TD-KS) equations

\[
\frac{i}{\hbar} \frac{\partial \psi_i(r, t)}{\partial t} = \left[ -\frac{1}{2} \nabla_i^2 + V_s(r, t) \right] \psi_i(r, t) \quad (2.33)
\]

In order that the two systems, the interacting and the non-interacting one, have the same density, the \( V_s(r, t) \) potential must comply with the equation

\[
V_s(r, t) = V_{\text{ext}}(r, t) + \int \frac{\rho(r', t)}{|r - r'|} \, dr' + \frac{\delta A_{\text{XC}}[\rho(r, t)]}{\delta \rho(r, t)} \quad (2.34)
\]

where, in analogy to the traditional time-independent Kohn-Sham scheme, all the exchange-correlation effects are collected in the term \( [\delta A_{\text{XC}}[\rho(r, t)]]/[\delta \rho(r, t)] \).

The exact time-dependent exchange-correlation action functional \( A_{\text{XC}}[\rho(r, t)] \) is unknown, thus approximations to it are introduced. The common approximation, which is also one of the most limiting approximations of the current form of TD-DFT, is the so-called adiabatic local density approximation (ALDA). Within this approximation, the time-dependent action functional \( A_{\text{XC}}[\rho(r, t)] \) is replaced with a time-independent local one, based on the assumption that the density varies slowly with time. In practice, the standard ground-state exchange-correlation functionals are adopted.

**TD-DFT for Excited States**

The universally adopted TD-DFT method for excited states is a linear response approach based on the TD-KS equations. The working equations are derived using the same procedure as in Section 2.2.2 for TD-HF. This yields a non-Hermitian matrix eigenvalue problem exactly of the form reported in Eq. (2.23). The difference between TD-DFT and TD-HF resides in the matrix elements of the \( A \) and \( B \) matrices

\[
A_{ir,js} = \langle \psi_i \psi_j | \psi_r \psi_s \rangle + \langle \psi_i \psi_j | V_{\text{XC}} | \psi_s \psi_r \rangle + \delta_{ij} \delta_{rs} (\epsilon_r - \epsilon_i) \quad (2.35)
\]

\[
B_{ir,js} = \langle \psi_i \psi_j | \psi_s \psi_r \rangle + \langle \psi_i \psi_j | V_{\text{XC}} | \psi_r \psi_s \rangle \quad (2.36)
\]

While in TD-HF the second terms correspond to the response of the non-local HF exchange potential, which yields a Coulomb-like term, in TD-DFT they correspond to the response of the chosen exchange-correlation potential \( V_{\text{XC}} \). The vast majority of
2.2. SINGLE-REFERENCE METHODS

the implementations adopt the ALDA approximation, where the exchange-correlation potential is obtained as the second functional derivative of the exchange-correlation functional

$$
\varphi_{XC} = \frac{\delta E_{XC}}{\delta \rho(r) \delta \rho(r')}
$$

Analogously to TD-HF, setting to zero the $B$ matrix yields the TDA/TD-DFT approximation, which is usually a very good approximation to TD-DFT. This is due to the fact that correlation is already included in the DFT ground state and therefore the $Y$ amplitudes tend to be small.[101, 102]

TD-DFT is one of the cheapest excited-state methods with a formal scaling of $O(N^4)$. The method, within the ubiquitous ALDA approximation, yields accurate results for valence excitations whose excitation energies lie well below the ionization potential. For such states, the typical error of TDDFT is in the range $0.1 - 0.5$ eV.[57]

Unfortunately, TD-DFT (within ALDA) is incapable of describing doubly-excited states[58, 59, 60, 61] and often fails near conical intersections.[61] Moreover, it performs poorly for Rydberg,[62, 63, 64, 65] charge-transfer states,[66, 67, 68, 69] and for valence states of molecules exhibiting extended $\pi$ systems.[103, 104]

2.2.4 Coupled-Cluster Methods

The most accurate single-reference family of methods for excited states is that of Coupled-Cluster (CC) approaches. Within this family the most successful are the Symmetry-Adapted Cluster-Configuration Interaction (SAC-CI),[81, 82, 83] the Equation-Of-Motion Coupled Cluster (EOM-CC),[84, 85, 86] and the Linear Response Coupled Cluster (LR-CC or CCn).[87, 88] SAC-CI is comparable to the EOM-CCSD approach, which includes up to doubly-excited cluster operators.[105] For this reason, in this Section we will discuss only the EOM-CC and the CCn methods.

Equation-Of-Motion Coupled Cluster

Referring to the original work by Stanton[105] and to the excellent review by Krylov[106], one can formulate the EOM problem as the diagonalization of the CC similarity-
transformed Hamiltonian $\mathcal{H}_T \equiv \exp(-T)\mathcal{H}\exp(T)$ introduced in Eq. (5.110)

$$\mathcal{H}_T R^f |\Phi_0\rangle = E_f R^f |\Phi_0\rangle \quad (2.38)$$

$$\langle \Phi_0 | L^g \mathcal{H}_T = \langle \Phi_0 | L^g E_g \quad (2.39)$$

$$\langle \Phi_0 | L^g R^f |\Phi_0\rangle = \delta_{gf} \quad (2.40)$$

where $R^f$ and $L^g$ are general excitation and de-excitation operators with respect to the reference $|\Phi_0\rangle$, which in single-reference EOM is the ground-state Slater determinant.

Since the similarity-transformed $\mathcal{H}_T$ in no longer an Hermitian Hamiltonian, its left and right eigenstates, $\langle \Phi_0 | L^g$ and $R^f |\Phi_0\rangle$, are not Hermitian conjugates, but according to Eq. (2.40) they are chosen to form a biorthogonal set.

If $R^f$ contain all possible excitations and $L^g$ all the de-excitations, then $\mathcal{H}_T$ is diagonalized in the complete basis set yielding the exact energies and states. In a more practical case the operators are truncated, for example

$$R^f = R_0^f + R_1^f + R_2^f + \ldots + R_n^f \quad (2.41)$$

$$= \sum_k r_k^f e_k \quad (2.42)$$

$$L^f = L_0^f + L_1^f + L_2^f + \ldots + L_n^f \quad (2.43)$$

$$= \sum_k l_k^f d_k \quad (2.44)$$

where $r_k^f$ and $l_k^f$ are expansion coefficients, $e_k$ are excitation operators ($\hat{X}_i^r$, $\hat{X}_{ij}^{rs}$, etc.) and $d_k$ are de-excitation operators ($\hat{X}_i^s$, $\hat{X}_{ij}^{sr}$, etc.).[106]

Now, the defined $R^f$ operator has the remarkable commutator property

$$[\mathcal{H}_T, R^f] |\Phi_0\rangle = \omega_f R^f |\Phi_0\rangle \quad (2.45)$$

where $\omega_f = E_f - E_0$, is the transition energy of the electronic excitation. By introducing the de-excitation operator $L^f$, we can write the transition energy as a general expectation value of the non-Hermitian operator $\mathcal{H}_T$

$$\omega_f = \frac{\langle \Phi_0 | L^f [\mathcal{H}_T, R^f] |\Phi_0\rangle}{\langle \Phi_0 | L^f R^f |\Phi_0\rangle} \quad (2.46)$$
By considering the first variation of the functional in Eq. (2.46) with respect to the coefficients of the right $R^f$ and the left $L^f$ vectors, and assuming that the variations $\delta R^f$ and $\delta L^f$ are independent, we arrive at the non-Hermitian secular problem for the coefficients $\{r^f_k\}$ and $\{l^f_k\}$

\[
(\tilde{H} - E_0)R = R\omega 
\]
\[
L(\tilde{H} - E_0) = \omega L 
\]

where $\tilde{H}$ is the matrix representation of the similarity-transformed Hamiltonian in the basis of the configurations generated by $\{e_k, d_k\}$ from the reference $|\Phi_0\rangle$. $R$ and $L$ are constructed from the expansion coefficients $\{r^f_k\}$ and $\{l^f_k\}$, and $\omega$ is the diagonal matrix of the transition energies.

The different specification of the $R^f$ ($L^f$) operator gives rise to different families of EOM-CC methods. In the EOM for Electronically Excited States (EOM-EE) the EOM operators are simply made of linear combinations excitation operators as in CI, that is

\[
R^f_{EE} = \sum_{ir} \hat{X}^r_i + \sum_{i<j \atop r<s} \hat{X}^{rs}_{ij} + \cdots
\]

When target states are multi-configurational owing to orbital degeneracies, the EOM operator includes a spin-flip (SF) operator which allows to use a high-spin reference for the wave function. This gives rise to the EOM-SF family[107, 108] where

\[
R^f_{SF} = \sum_{ir} \hat{X}^{rs}_{ir} + \cdots
\]

with $\hat{X}^{rs}_{ir}$ being a SF single excitation operator, which, with respect to the reference determinant, annihilates the electron in the $i$-th $alpha$-spin occupied MO and creates an electron into the $r$-th $beta$-spin virtual MO. The EOM-SF family is usually used for describing bond-breaking or di-radicals and tri-radicals, where the homologous wave functions are formally single excitations with respect to the high-spin references.
Linear-Response Coupled Cluster (CCn)

The Linear-Response Coupled Cluster method arises from the study of the response of the CC wave function when exposed to a time-dependent electromagnetic perturbation. In particular, if we consider a time-dependent perturbation in a Fourier expanded form

$$V(t) = \sum_y \exp(-i\omega_y t) f_y(\omega_y) Y$$

(2.51)

where $Y$ is an Hermitian operator, $\omega_y$ is the oscillator frequency, and $f_y(\omega_y)$ is a strength parameter, then the time-dependent expectation value of an observable, $X$, can be expanded in orders of perturbation as

$$\langle \Psi_0(t)|X|\Psi_0(t) \rangle = \langle \Psi_0|X|\Psi_0 \rangle + \sum_y f_y(\omega_y) \langle\langle X,Y \rangle\rangle_{\omega_y} \exp(-i\omega_y t) + \cdots$$

(2.52)

where $\langle\langle X,Y \rangle\rangle_{\omega_y}$ is the so-called linear response function, and $\Psi_0 = \Psi_0(t = 0)$ is the (stationary) ground state wave function.

The linear response function can be written in terms of the unperturbed eigenstates $\{\Psi_k\}$ as

$$\langle\langle X,Y \rangle\rangle_{\omega_y} = \mathcal{P}_{XY} \sum_k \frac{\langle X|\Psi_k \rangle \langle \Psi_k|Y|\Psi_0 \rangle}{\omega_y - \omega_k}$$

(2.53)

where $\mathcal{P}_{XY}(g_{xy}) = g_{xy} + g_{yx}$ and $\omega_x + \omega_y = 0$. Thus, the linear response function has poles when the external frequency, $\omega_y$, is equal to the excitation energies $\omega_k = E_k - E_0$ of the system.

The derivation of the CC linear response functions is rather complex and the reader may study Ref.[109] and references therein for further details. To our aim suffice it to say that the derivation leads to the excitation energies being determined by a non-Hermitian eigenvalue problem of the form in Eqs. (2.47) and (2.48).

Performance of EOM-CC and CCn Methods

According to Krylov[106], for any EOM model the quality of the target-state wave functions depends on i) how well the reference state is described by a single-reference CC method — it is important that a single Slater determinant provides a qualitatively correct zero-order wave function; ii) the leading configurations in the target wave
functions should appear at the same excitation level. These conditions are usually met by singly excited states of closed-shell molecules around their ground-state equilibrium geometries.

If the reference state acquires considerable multi-configurational character (this is usually associated with bonds breakage, molecules with di-radical character, transition metals and other cases with small HOMO-LUMO gaps), the quality of the CC reference state deteriorates, leading to less accurate excitation energies. Nonetheless, even when the reference state is heavily multi-configurational, some EOM states can still be well described, provided that single excitations from the reference produce all the leading configurations for the target states.

Furthermore, even though doubly-substituted determinants are formally present in EOM-CCSD models, the EOM treatment of doubly excited states is disastrous because the important configurations appear at different excitation level and, therefore, they are not treated in a balanced fashion.

It is also worth mentioning that the EE variant of EOM tends to perform much better with RHF references, except in cases of bond-breaking or di- and tri-radicals, where EOM-SF should be adopted.

For the pure CC models such as CCSD and CCSDT, EOM-CC and LR-CC excitation energies are identical and their performance can be analysed at once.[110] As reported by Sneskov et al.,[110] extensive energy benchmark calculations have been performed which place CC with doubles (CCSD and CC2) at a mean error in the range of 0.2 – 0.4 eV, with tails in the distribution presenting absolute errors larger than 1 eV. The addition of triples has a significant effect shifting the mean error to 0.1 eV and the maximum absolute error roughly to 0.4 eV. When the benchmark is limited to singly excited states,* these methods tend to perform much better, with mean errors within 0 – 0.2 eV and maximum absolute errors between 0.1 – 0.6 eV for CC2 and CCSD, as well as 0.1 eV for CCSD with perturbative triples correction.[111, 112, 113]

The picture that emerges from these benchmarks is that CC methods are probably the most accurate family of methods when the reference is single-configurational and the target excited state is mainly composed by single excitations.

---

*over 90 % of the wave function is made of single excitations
The excited-state CC approaches have an applicability that is restricted by their computational complexity analogous to their ground-state cousins. In fact, CC2 formally scales as $O(N^5)$, CCSD as $O(N^6)$, CC3 as $O(N^7)$ and CCSDT as $O(N^8)$. This limits their practical use to small and medium size systems.

2.3 Multi-Reference Methods

2.3.1 MRSCF

As discussed previously, the reference state or the target excited state, or both, may acquire multi-reference character, for example during bond-breaking, in di-radicals, in transition metals, and when there are orbital degeneracies (small HOMO-LUMO gaps). In these cases the use of a single reference or even of a single configuration may be inadequate for a quantitative description of the electronic state.

The reference wave function can be improved by including more than one determinant or electronic configuration. In a Multi-Reference Self-Consistent Field (MRSCF) calculation the wave function can be viewed as a truncated CI expansion

$$\tilde{\Psi}^{\text{MRSCF}} = \sum_i c_i \Phi_i$$

in which both the amplitudes $c_i$ and the collective set of orbitals used to construct all the $\Phi_i$ are optimized simultaneously.[73, 114, 115, 116, 56] The optimization procedure is iterative and since the number of MRSCF iterations required for achieving convergence tends to increase with the number of determinants (configurations) included, the size of MRSCF wave functions that can be practically treated is somewhat smaller than for CI methods.[116]

Once the target state(s) is established, the major problem of a MRSCF calculation becomes the selection of the determinants which are necessary for including the properties of interest. Due to the colossal number of possible combinations the MRSCF method is not a black-box approach. Hence, typically the selection is made by “sifting” out the orbitals which are believed to contribute most to the chemical properties of the target state.

The most successful MRSCF method is the Complete Active Space Self-Consistent
2.3. MULTI-REFERENCE METHODS

Field (CASSCF) developed by Roos and coworkers in 1980.[74, 75, 117] Here the selection of determinants is done by partitioning the MOs into active and inactive spaces. Within the active space a FCI is performed including all the appropriate CSFs in the MRSCF optimization. A common notation is \([n, m]\)-CASSCF, which indicates that \(n\) electrons are distributed in all possible ways in \(m\) active orbitals.[116] Still, the structure of the active space must be decided manually and it is certainly not a well-defined task. Quoting González et al.,[55] “Obtaining an appropriate active space requires chemical intuition, experience, as well as some trial and error”.

As CASSCF is FCI within the active space, it becomes unmanageably expensive even for quite small active spaces. This has led to the development of variations such as Restricted-Active Space Self-Consistent Field (RASSCF),[118] Occupation Restricted Multiple Active Space (ORMAS) CI,[119] SplitCAS,[120] where the active space is divided into several sections, each having restrictions on the excitations allowed.

Typically, in order to assure the orthogonality of the desired MRSCF solutions, the orbitals are optimized for a functional of the energy, which is defined as an average of a certain number of states each associated with a certain weight. This is the working principle of the state-averaged (SA) CASSCF and RASSCF. In this way it is also possible to overcome the “root flipping” problem. This is the interchange of the order of the roots, which may happen if state-specific optimizations are performed independently.

The main advantage of MRSCF approaches is that potentially they can model very accurately static correlation effects. Conversely, since the expansion in Eq. (3.54) is in general not large, dynamic correlation is poorly described by MRSCF methods, which therefore do not usually reach chemical or spectroscopic accuracy.[55] Furthermore, the MRSCF iterative procedure is much harder to converge than traditional SCF, and much more prone to converge to solutions that are not minima. For this reason MRSCF wave function optimizations are normally carried out by expanding the energy to the second order in the orbital and the determinant coefficients and using more expensive Newton-Raphson-based methods to force the convergence to a minimum.[116] Finally, as any form of truncated CI, MRSCF methods are in general not size extensive.
2.3.2 MRCI and MRPT

Two common routes to account for both multi-reference character (i.e. static correlation) and dynamic correlation effects are the Multi-Reference Configuration Interaction (MRCI) and the Multi-Reference Perturbation Theory (MRPT).†

In the MRCI method[70, 71, 72, 73, 121] a MRSCF wave function is used as reference for a CI calculation. Thus, the substituted determinants are generated by considering excitations with respect to all the determinants that enter the MCSCF reference. Therefore, in MRCI the number of configurations is increased with respect to single-reference CI by a factor roughly equal to the number of determinants included in the MRSCF. MRCI is highly accurate, but also incredibly computationally expensive limiting its applicability to fairly small systems. Furthermore, typically MRCI methods truncate the CI expansion and therefore are not size extensive (see [122] and references therein).

Due to their combination of reduced cost compared to MRCI and of size extensivity, MRPT methods are arguably the most successful and systematic procedures to account for the correlation energy of MR states. Analogously to the single-reference case in Section 1.5.5, the MRSCF wave function is used as a reference for a perturbative treatment which yields corrections up to $n$ order following a Taylor series expansion.

Amongst the MRPT approaches, the CASPT2 by Andersson and coworkers[77, 78, 79] and the Multi-Reference Møller-Plesset Perturbation theory (MRMP) by Hirao[80] stand out for their theoretical achievements. The more recently introduced RASPT2[123] has shown results similar to CASPT2 for reduced computational costs.[124]

MRPT approaches can be remarkably accurate depending of course on the reliability of the reference wave function,[124] although outliers in the vertical excitation energies can exceed 1.0 eV absolute errors. Other shortcomings of these methods are the “intruder state” problem and the non-orthogonality among state-specific MRMP solutions. The first is a general problem of perturbative methods, but in MRPT is much more common due to a smaller separation between the energies of the occupied (active in MRSCF) and the external molecular orbitals (inactive in MRSCF). A possibility is to use level-shifting techniques as in the LS-CASPT2 method.[125] The effects of the second problem are

†To this aim there are also other approaches, the most important being certainly Multi-Reference Coupled Cluster (MRCC), but we will not encounter these methods in this work.
usually negligible, except for Rydberg state solutions which can cause larger overlaps. In these cases the MS-CASPT2\[126\] (and MS-RASPT2) seems to provide a viable solution to the problem.\[55\]
Chapter 3

Single-Determinant Approximation for Excited States

“Less is more”

Robert Browning — Men and Woman, Andrea del Sarto

3.1 Introduction

As previously mentioned, it is common to study the electronic ground states of molecules using a small set of well-established methods such as Density Functional Theory (DFT), Møller-Plesset Perturbation Theory and Coupled-Cluster Theory. Since all of these methods are single-reference in nature, the “standard ground-state approach” may be simply summarised as follows:

1. Find a single-determinant (SCF) wave function and energy. This can be a reasonable first-order approximation if the Hartree-Fock (HF) method is used or a sufficiently accurate approximation if the Kohn-Sham method (DFT) is used;

2. If the HF approximation has been adopted, refine that energy and wave function using post-Hartree-Fock methods such as MP2 and Coupled Cluster.
Although, as discussed in Chapter 2, there is no “standard excited-state approach” we may identify two main excited-state approaches:

1. Single-reference procedure
   
   (a) Find the ground-state single-determinant wave function;
   
   (b) Compute the properties or the wave functions for an entire set of excited states using the ground-state determinant as reference.

2. Multi-reference procedure
   
   (a) Find a multi-determinant wave function which is a reasonable first-order approximation;
   
   (b) Refine that energy and wave function using post-MRSCF methods such as MRPT and MRCI;

Both the excited-state protocols are radically different from the ground-state one.

The single-reference procedure treats excited and ground states on a different footing by adopting the ground-state determinant as a reference for all the excitations, instead of using different excited-state determinants for each excited state. This disparity has often significant negative consequences such as lowering the performance and increasing the computational complexity of single-reference methods for excited states compared to ground-state ones. Both CIS and TD-DFT (within the adiabatic local density approximation[57]) are incapable of describing doubly-excited states[58, 59, 60, 61] and often fail near conical intersections.[61] Moreover, CIS can overestimate excitation energies by as much as 2 eV[57] and TD-DFT performs poorly for Rydberg[62, 63, 64, 65] and charge-transfer states.[66, 67, 68, 69] At the same time, both CIS and TD-DFT require large additional programming efforts and are computationally more expensive than traditional SCF method. Even highly correlated methods, such as those based on Coupled Cluster, are affected by the ground/excited state imbalance. For example, the EOM-CC treatment (for truncated expansions of the cluster operator) of doubly excited states is disastrous.[106] These weaknesses arise because the molecular orbitals (MOs) are not allowed to properly relax in the excited state and, consequently, such methods struggle in cases where that relaxation would be significant.
The multi-reference procedure treats excited and ground states as if they were physically different. This intrinsic disparity results from the belief that, whereas ground states are often described well by single-determinant method, excited states are usually multi-reference in character.\[55, 56\] However, as discussed in Chapter 2, the specification of a multi-reference wave function is much more difficult and often requires a delicate combination of chemical intuition, experience, trial and error that incurs a substantially higher computational cost.

Therefore, we think it is appropriate to ask whether it is either necessary or desirable to treat ground and excited states as if they were fundamentally different. Indeed, such a methodological discontinuity is difficult to justify on either physical or mathematical grounds. Is it not more natural to obtain excited states, at least in a first-order approximation, by an SCF procedure?

After a ground-state SCF has converged, crude approximations to excited states can be obtained by promoting electrons from occupied to virtual orbitals to yield singly- or multiply-excited states. This approach has well-known deficiencies, the most obvious of which is that the orbitals are optimal only for the ground state and are not able to relax after the electronic excitation. In separate studies, Hunt and Goddard\[127\] and Huzinaga and Arnau\[128, 129\] suggested methods for optimizing these virtual orbitals for excited states. They allow rotations within the virtual manifold but do not permit the occupied orbitals to relax. Morokuma and Iwata\[130\] later removed this constraint and allowed both the occupied and virtual orbitals to relax within their respective subspaces. In both cases, mixing between occupied and virtual subspaces is forbidden and this ensures that the resulting excited states are rigorously orthogonal to the ground state.

In 2008, Gilbert \textit{et al}. contended that the importance of orthogonality had been overstated.\[131\] In fact, quantum mechanics requires exact wave functions to be orthogonal, that is \(\langle \Psi_i | \Psi_j \rangle = 0, i \neq j\), but it makes no such demand on SCF wave functions and, in fact, one should expect \(\langle \Phi_i^{\text{SCF}} | \Phi_j^{\text{SFC}} \rangle \neq 0\). Keeping this in mind, Gilbert \textit{et al}. abandoned orthogonality and, instead, decided to approximate excited states with genuine higher solutions of the SCF equation

\[
\mathcal{H}^{\text{SCF}} \Phi_k^{\text{SCF}} = E_k^{\text{SCF}} \Phi_k^{\text{SCF}}
\]
3.1. INTRODUCTION

where $\mathcal{H}^{\text{SCF}}$ is either the Hartree-Fock or the Kohn-Sham Hamiltonian and

$$\Psi_k^{\text{SCF}} = \det[\chi_i(r_j, s_j)]$$ (1.2)

are just Slater determinants in a spin-unrestricted formalism.

There are two objections to the proposal that excited states be modeled in this way. First, because it can be difficult to find higher solutions of (1.1) and, second, because conventional wisdom holds that most excited states are intrinsically multi-configurational in nature.

The first objection has been addressed by the recent introduction of a number of interesting numerical techniques.\cite{131, 132, 133, 134, 135} The second objection requires further study and, in this Chapter, we examine the extent to which single-determinant HF solutions are reasonable approximations to the ground state and ten singly excited states of the $\text{H}_2$ molecule. In doing so, we will adopt the Maximum Overlap Method (MOM) developed by Gilbert et al.\cite{131} which will be described in the next Section.

We should emphasise from the outset that, within a single-determinant model, the open-shell singlets are necessarily contaminated by the associated triplet. Attempts to remedy this situation have been made,\cite{136} but these approaches move beyond the single-determinant HF framework. Notwithstanding this spin contamination, we will refer to these contaminated solutions as "singlets" to differentiate them from the triplets.

Because the electrons no longer share an orbital in the excited states of $\text{H}_2$, dynamic correlation is less important than in the ground state and it is plausible that the excited states may, in fact, be modeled more accurately than the ground state. By a similar argument based on the Fermi hole which separates same-spin electrons in the HF approximation, it is possible that triplet states will be more accurately modeled than singlets.\cite{116} However, in states where static correlation plays a prominent role, it is possible that neither of these expectations may be met.

In order to describe homolytic bond-breaking, we have used spin-unrestricted Hartree-Fock (UHF) theory\cite{137} throughout. Most of the states that we will consider dissociate into dissimilar fragments, e.g. $\text{H}(1s) + \text{H}(2p_x)$, and we therefore face the “symmetry dilemma” discussed by Löwdin\cite{138} many years ago. However, because our primary goal is to model the energetic characteristics of the various states of $\text{H}_2$, whenever a solution
splits into symmetric and broken-symmetry solutions, we choose the broken-symmetry branch, and whenever the solution splits into restricted and unrestricted solutions, we choose the unrestricted branch.

### 3.2 The Maximum Overlap Method

As described in Section 1.5.1, on each iteration of the SCF procedure, the current MO coefficient matrix \( C^{\text{old}} \) is used to build a Fock (or Kohn-Sham) matrix \( F \) and the generalized eigenvalue problem

\[
FC^{\text{new}} = SC^{\text{new}} \epsilon
\]

is then solved to obtain a new MO coefficient matrix \( C^{\text{new}} \) and orbital energies matrix \( \epsilon \). At that point there are \( \binom{N}{n} \) ways to choose which of the new orbitals to occupy but this large space is rarely, if ever, explored exhaustively. Instead, one usually follows the aufbau protocol, which dictates that one simply occupies the \( n \) orbitals with the lowest orbital energies \( \epsilon_j \).

An alternative protocol, which Gilbert et al.\cite{131} defined the Maximum Overlap Method (MOM), states that the new occupied orbitals should be those that overlap most with the span of the old occupied orbitals. If we define the orbital overlap matrix

\[
O = (C^{\text{old}})^\dagger S C^{\text{new}}
\]

then \( O_{ij} \) is the overlap between the \( i \)-th old orbital and the \( j \)-th new orbital, and the projection of the \( j \)-th new orbital onto the old occupied space is

\[
p_j = \sum_{i} |O_{ij}| = \sum_{i} \left| \sum_{\mu \nu} C_{i\mu}^{\text{old}} S_{\mu \nu} C_{j\nu}^{\text{new}} \right|
\]

In this way, the full set of \( p_j \) values can be found by three matrix-vector multiplications, at \( \mathcal{O}(N^2) \) cost, and this adds negligibly to the cost of each SCF cycle. One then occupies the \( n \) orbitals with the largest projections \( p_j \).

To use the MOM, the SCF calculation must begin with orbitals that lie within the basin of attraction of the target excited solution. Often, it is sufficient to perform a
3.3 Hartree-Fock Description of Excited States of H\textsubscript{2}

In this Section we use the MOM to locate 11 low-energy solutions of the HF equations for the H\textsubscript{2} molecule and we find that, with only one exception, these yield surprisingly accurate models for the low-lying excited states of this molecule.

3.3.1 Method and Results

First, we constructed a large basis set by adding two diffuse \(p\)-shells with exponents \(\alpha_1 = 0.0513930\) and \(\alpha_2 = 0.0185657\) to the aug-pc4 basis.[139] These exponents were obtained by fitting the existing seven \(p\)-shell exponents to an exponential function, and extrapolating to obtain exponents for the eighth and ninth \(p\)-shells. Addition of a tenth \(p\)-shell had no effect on the results. The quality of the resulting basis set is indicated by the following HF energies (in atomic units): \(H(1s) = -0.4999996\), \(H(2s) = -0.1249831\),
H(2p) = -0.1249875, and He(1s^2) = -2.8615949. The complete-basis limit for the last of these, which is the united-atom limit, is -2.8616799.[140]

Using a modified version of the Q-Chem 4.0 package,[141] we then used the MOM to compute the UHF energy curves of eleven states of the H_2 molecule. The success of the MOM depends critically on the initial guesses for the occupied MOs and we found that the appropriate ground-state orbitals of the H_2^+ cation were satisfactory in all cases. Solutions for both short and long bond lengths were obtained in this way, and these were propagated for other bond lengths by moving in steps of 0.1 a.u., reading in the orbitals from the previous step, and using the MOM to maintain the desired state. By propagating in both directions, we were able to obtain both symmetric and broken-symmetry solutions for several of the states. The Allard–Kielkopf correlation diagram[142] was used to confirm that our assigned states dissociated correctly.

Figure 3.1: MO diagram for the H_2 molecule.

Figure 3.1 shows the MO diagram for H_2 and each of the excited states considered has a single electron in the 1σ_g orbital and another in an higher orbital, as indicated in Table 3.1. All of these states have been well studied, both experimentally and theoretically.[143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155] Exact
3.3. HARTREE-FOCK DESCRIPTION OF EXCITED STATES OF H$_2$

equilibrium bond lengths, $r_e$, dissociation energies, $D_e$ and harmonic frequencies, $\omega_e$, from Huber and Herzberg[143] are listed in Table 3.1, along with HF errors, $\Delta = \text{HF} - \text{Exact}$. The potential energy curve for each state is shown in Fig. 3.2.

Table 3.1: Exact values[143] and HF deviations, $\Delta$, for equilibrium bond lengths, $r_e$, dissociation energies, $D_e$, and harmonic vibrational frequencies, $\omega_e$, in the H$_2$ molecule.

<table>
<thead>
<tr>
<th>HOMO</th>
<th>State</th>
<th>$r_e$ (Å)</th>
<th>$D_e$ (eV)</th>
<th>$\omega_e$ (cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Exact</td>
<td>$\Delta$</td>
<td>Exact</td>
</tr>
<tr>
<td>$1\sigma_g$</td>
<td>$X^1\Sigma_g^+$</td>
<td>0.741</td>
<td>-0.008</td>
<td>4.748</td>
</tr>
<tr>
<td>$1\sigma_u$</td>
<td>$b^3\Sigma_u^+$</td>
<td>1.293</td>
<td>0.529</td>
<td>3.582</td>
</tr>
<tr>
<td>$2\sigma_g$</td>
<td>$a^3\Sigma_g^+$</td>
<td>0.989</td>
<td>-0.004</td>
<td>3.057</td>
</tr>
<tr>
<td></td>
<td>$E^1\Sigma_g^+$</td>
<td>1.011</td>
<td>-0.015</td>
<td>2.543</td>
</tr>
<tr>
<td>$1\sigma_u$</td>
<td>$EF^1\Sigma_u^+$</td>
<td>2.315</td>
<td>-0.043</td>
<td>2.440</td>
</tr>
<tr>
<td>$2\sigma_u$</td>
<td>$e^3\Sigma_u^+$</td>
<td>1.107</td>
<td>-0.011</td>
<td>1.589</td>
</tr>
<tr>
<td></td>
<td>$B^1\Sigma_u^+$</td>
<td>1.119</td>
<td>0.000</td>
<td>1.109</td>
</tr>
<tr>
<td>$1\pi_u$</td>
<td>$c^3\Pi_u$</td>
<td>1.037</td>
<td>-0.006</td>
<td>3.069</td>
</tr>
<tr>
<td></td>
<td>$C^1\Pi_u$</td>
<td>1.033</td>
<td>-0.006</td>
<td>2.542</td>
</tr>
<tr>
<td>$1\pi_g$</td>
<td>$i^3\Pi_g$</td>
<td>1.069</td>
<td>-0.009</td>
<td>0.924</td>
</tr>
</tbody>
</table>

3.3.2 Discussion

The strengths and weaknesses of HF theory when describing the $X^1\Sigma_g^+$ ground state of H$_2$ are well-known. HF underestimates the equilibrium bond length by 0.008 Å, underestimates the dissociation energy by 1.111 eV and overestimates the harmonic frequency by 90 cm$^{-1}$. The large error in $D_e$ reflects the large correlation energy ($E_c = -1.11$ eV [147, 156]) at the equilibrium geometry. (The correlation energy of the dissociated atoms is exactly zero.) These results provide a benchmark for determining how well HF models the excited states.

The lowest-energy triplet state, $b^3\Sigma_u^+$, dissociates correctly and agrees qualitatively with the exact curve. The exact curve has a very shallow minimum at $r_e \approx 7.8$ a.u.[144] due to dispersion, but HF is incapable of modelling this and the HF solution is therefore purely repulsive.

The B $^1\Sigma_u^+$ state is poorly modeled by HF. All of the open-shell singlets are heavily spin-contaminated but, in this case, the contaminating state ($b^3\Sigma_u^+$) is repulsive and grossly distorts the energy curve of the B state. Like the spin-restricted HF solution for
the ground state, it has half an electron of each spin on each center and is qualitatively incorrect in the dissociation limit. The correct dissociation products for this state are \( H(1s) + H(2p) \) which have a total energy of \(-0.625\) a.u.

The deviations in the \( r_e \) and \( \omega_e \) value for the \( ^3\Sigma_g^+ \) state are about half those for the ground state and the deviation in \( D_e \) is much smaller. Furthermore, its dissociation into \( H(1s) + H(2p) \) is qualitatively correct.

The singly-excited \( E \, ^1\Sigma_g^+ \) and doubly-excited \( F \, ^1\Sigma_g^+ \) states have the electron configurations \( 1\sigma_g^22\sigma_g^1 \) and \( 1\sigma_u^2 \), respectively. These interact strongly to form an EF \( ^1\Sigma_g^+ \) state\[155, 148\] with a double-minimum potential energy curve. The HF solution for the \( E \, 1\sigma_g^22\sigma_g^1 \) configuration yields \( r_e, D_e \) and \( \omega_e \) values which agree well with the exact values at the first EF minimum. The HF solution for the \( F \, 1\sigma_u^2 \) configuration is in poorer agreement with the exact curve at the second EF minimum (for example, \( D_e \) is underestimated by 1 eV) but this is not surprising because, as in the \( X \, 1\sigma_g^2 \) state, the electrons share an orbital and are therefore strongly correlated. In HF theory, these two solutions cannot mix. However, it is known from elementary quantum mechanics that two same-symmetry solutions form an avoided crossing. Thus, these two HF solutions combine to form upper- and lower-energy curves and we choose the lower one to model the EF state. The HF energy curve for the EF state has a cusp at \( R = 3.52 \) a.u. (the maximum on the true EF curve lies at \( R = 3.13 \) a.u.) An analogous cusp arises in the HF energy curve for the torsional rotation in ethene.[157]

The \( r_e, D_e \) and \( \omega_e \) parameters for the \( e \, ^3\Sigma_u^+ \) state are well modeled by HF and the correct dissociation products, \( H(1s) + H(2s) \), are also obtained.

For the \( B' \, ^1\Sigma_u^+ \) state, HF predicts \( r_e \) accurately but overestimates \( D_e \) by about 15\% (0.006 a.u.). This \( D_e \) error is among the largest in our study, but is still much smaller than that for the ground state. Although the dissociation limit for this state, \( H(1s) + H(2s) \), can be accurately modeled using a broken-symmetry UHF wave function, we were unable to locate an HF solution connecting these fragments to the equilibrium structure. Thus, the HF curve for this state shown in Fig. 3.2 dissociates to an unphysical solution with half an alpha electron and half a beta electron on each centre.

The HF descriptions of the \( c \, ^3\Pi_u \) and \( C \, ^1\Pi_u \) states are similar. The deviation in both bond lengths is 0.006 Å, which is similar to, but slightly better than, the ground
3.3. HARTREE-FOCK DESCRIPTION OF EXCITED STATES OF H\textsubscript{2}

Figure 3.2: UHF potential energy curves for low-lying states of the H\textsubscript{2} molecule.

state deviation of 0.008 Å. The $D_e$ and $\omega_e$ errors are also smaller than those of the ground state, with the triplet state being reproduced slightly better. The dissociation products for both states are H(1s) + H(2p\subscript{\pi}) and both HF solutions dissociate correctly.

The I\textsuperscript{1}\Pi\subscript{g} and i\textsuperscript{3}\Pi\subscript{g} states have almost identical $r_e$, $D_e$ and $\omega_e$ values. This similarity is not unexpected because the electron in the diffuse 1\pi\subscript{g} orbital rarely encounters the electron in the compact 1\sigma\subscript{g} orbital and their spins are therefore largely irrelevant. HF reproduces this, giving almost the same errors for the two states. However, neither solution dissociates correctly to the H(1s) + H(2p\subscript{\pi}) fragments. These states have small well depths and the errors of roughly 0.07 eV, although comparable to those for other states, are relatively larger.

It is interesting to note that HF theory overestimates the dissociation energies of four of the singlet states that we studied, viz. B\textsuperscript{1}\Sigma\subscript{u}+, EF\textsuperscript{1}\Sigma\subscript{g}+, B'\textsuperscript{1}\Sigma\subscript{u}+ and C\textsuperscript{1}\Pi\subscript{u}. This implies that each of these states has a positive correlation energy and, therefore, that the higher solutions of the HF equation are non-variational.
We also note that, when a HF solution breaks symmetry as the bond is stretched, this always occurs beyond the equilibrium bond length of the target solution, thereby ensuring that the data in Table 3.1 are unique. To understand the symmetry-breaking phenomenon in greater detail, we have analyzed the C $^{1}\Pi_{u}$ and I $^{1}\Pi_{g}$ states. Both states dissociate to H(1s) + H(2$p_{\pi}$) but only the former breaks symmetry and can dissociate to the correct energy. The C $^{1}\Pi_{u}$ solution bifurcates at 4.3 a.u. into a higher-energy symmetric and a lower-energy broken-symmetry solution. No such splitting occurs for the I $^{1}\Pi_{g}$ solution.

Figure 3.3: Lowest eigenvalues $\lambda$ of the HF MO rotation Hessian for the symmetric C $^{1}\Pi_{u}$ solution in the H$_{2}$ molecule.

Figure 3.3 shows the lowest eigenvalues of the MO rotation Hessian[158, 159] for the symmetric C $^{1}\Pi_{u}$ as a function of bond length. The bifurcation arises at the bond length where one of the eigenvalues becomes negative, indicating the existence of a new, lower-energy solution.[160, 161] Closer inspection of the corresponding eigenvector reveals that the rotations involved include mixing of the occupied $\pi_{u}$ orbital with the virtual $\pi_{g}$ orbitals, thus breaking the symmetry of the solution. The eigenvalues for the symmetric I $^{1}\Pi_{g}$ solution (Fig. 3.4) do not exhibit the same behaviour, which explains the absence of a broken symmetry solution for this state. In each case, the zero eigenvalue indicates that each of the II solutions is degenerate.
3.4 Concluding Remarks

The single-determinant approach treats ground and excited states on an equal footing, enabling to extend the well-established quantum chemistry ground-state methods and technologies to excited-state calculations.

We discussed two possible objection to the proposal that excited states be modeled in this way. The first, related to the difficulty of finding higher solutions of the SCF equations, can be addressed by replacing, within the SCF procedure, the aufbau protocol with the MOM.

The second objection concerns the quality of these higher SCF solutions as first-order approximations to electronically excited states, as they are often reported to be intrinsically multi-reference. In order to study the accuracy of the single-determinant models as first-order approximations, we used the MOM to locate 11 low-energy solutions of the HF equations for the $\text{H}_2$ molecule. Hartree-Fock theory (spin-unrestricted and allowing broken symmetry) is surprisingly effective for modelling the low-lying excited states of the $\text{H}_2$ molecule. We find that it provides useful first-order descriptions of ten of the eleven states considered and, in fact, with the single exception of the $\text{B} \ ^1\Sigma_u$ state, it is more accurate for the excited states (near their equilibrium bond lengths) than it
is for the ground state. We also find, as anticipated, that the triplet states are usually more accurately modeled than the singlets.

The accuracy of the results in Table 3.1 strongly suggests that there is a correspondence between solutions of the HF equation and solutions of the Schrödinger equation. It appears likely that the higher HF solutions may often (though not always) be a satisfactory first-order approximation for excited states, which can subsequently be improved, if desired, using standard correlation methods. We investigated this, for both singly- and multiply-excited states, and will present our findings in the next Chapter.
Chapter 4

Single-Determinant Models for Difficult Excitations — The Initial Maximum Overlap Method

“Once the anchor of reason has been cut, one’s craft may go anywhere. One may become a St Francis or equally a Hitler.”

Brand Blanshard — Reason and Analysis, essay # 21

4.1 Introduction

In Chapter 2, we presented evidence supporting the fact that single-determinant models may be satisfactory first-order approximations to electronic excited states.

In this Chapter, we ask how well the single-determinant approximation can handle challenging cases where other low-cost excited-state methods either perform poorly or fail completely. In fact, we recall that both CIS and TD-DFT (within the adiabatic local density approximation[57]) are incapable of describing doubly-excited states[58, 59, 60,
4.2. THE INITIAL MAXIMUM OVERLAP METHOD

61] and often fail near conical intersections.[61] Moreover, CIS can overestimate excitation energies by as much as 2 eV[57] and TD-DFT performs poorly for Rydberg[62, 63, 64, 65] and charge-transfer states.[66, 67, 68, 69] Indeed, also much more sophisticated methods, such as EOM-CC, fail the treatment of doubly excited states.[106]

We present a single-determinant approach to three of the most challenging topics in the chemistry of excited states: double excitations, charge-transfer states and conical intersections. The results are obtained by using the Initial Maximum Overlap Method (IMOM) which is a modified version of the Maximum Overlap Method (MOM). The new algorithm converges much better than the original, enabling to obtain solutions which were either much more SCF-expensive before or out of reach of the original MOM, especially for these difficult problems.

All of the excitation energies reported here are vertical, i.e. the energy of the excited state is calculated at the structure of the ground state, and no attempt has been made to correct for zero-point vibration energy.

If we use our modified MOM to find an excited-state solution of the self-consistent field (SCF) equations, we will refer to the resulting energy using the unadorned name of the functional (e.g. BLYP). On the other hand, if we use the conventional time-dependent approach to estimate the excitation energy from the ground state, we will refer to the energy by prefixing “TD” to the name of the functional (e.g. TD-BLYP). As we will show, these two approaches often yield strikingly different models of excited-state energetics.

4.2 The Initial Maximum Overlap Method

The MOM provides an alternative to the aufbau principle for determining which MOs to occupy on each cycle of an SCF calculation. Rather than choose the lowest energy MOs, we choose those with the largest projection into the span of the occupied MOs of the previous SCF cycle. Given the MO overlap matrix

\[ O = (C^{\text{previous}})^\dagger S C^{\text{new}} \]  

(2.1)
and projections $p_j = (\sum_i |O_{ij}|)$, we simply choose to occupy the MOs with the largest $p_j$ values.

The \textit{aufbau} criterion drives the SCF towards the lowest-energy solution of the SCF equations. If it is replaced by the MOM criterion, the SCF can discover higher-energy solutions of those equations and we have demonstrated that these correspond to the excited states of the system.\cite{131, 162, 163} This approach is not limited to the lowest-energy solutions of each symmetry type.

Notwithstanding the success of the MOM, cases sometimes arise wherein the SCF converges either slowly or to an undesired state, and this is especially common in systems with near-degeneracies. Such behavior led us to develop a modified MOM protocol in which we choose to occupy the MOs with the largest projection into the space spanned by the occupied MOs of the \textit{initial guess}. Thus, the MO overlap matrix becomes

$$ O = (C^{\text{initial}})^\dagger SC^{\text{new}} \quad (2.2) $$

This strategy – the Initial Maximum Overlap Method (IMOM) – encourages the SCF to find a solution of the SCF equations in the neighbourhood of the initial guess. We implemented it in the Q-Chem 4.4.1 package\cite{164} and used it to generate all the results reported in this Chapter.

The success of the MOM and IMOM depend critically on the quality of the initial guess. However, even with a good guess, it is possible for the MOs in a MOM calculation gradually to drift away from the initial guess. The IMOM prevents this by anchoring the SCF to that guess and this simple algorithmic modification greatly enhances its reliability, allowing it to locate troublesome states which the MOM struggles to find. Figure 4.1 shows the energy behaviour of the SCF iterations, using IMOM and MOM, for the same initial guess orbitals. During the first 40 iterations, the MOM algorithm progressively drifts away from the basin of attraction of the target solution, then the SCF spends roughly 400 iterations in the neighbourhood of a different solution. Successively, it drifts away also from this solution to finally converge, after another couple of hundred cycles, to the ground state. IMOM converges within 32 cycles to the target solution, which is the EF $1\Sigma_g^+$ state\cite{155, 148} of $\text{H}_2$.

All calculations in the remainder of this Chapter use the spin-unrestricted formalism.\cite{137, 83}
Figure 4.1: Energy behaviour of the SCF iterations yielded by IMOM and MOM using an initial guess for the EF $^1\Sigma_g^+$ state\cite{155,148} of H$_2$. The bond length is 1.8 Å. Both calculations were performed at the HF/aug-pc4 level of theory. Red dots (*) indicate SCF convergence within a Direct Iterative Inversion Subspace (DIIS) error of $10^{-6}$. 

4.2. THE INITIAL MAXIMUM OVERLAP METHOD
4.3 Double Excitations

The term “doubly excited” is commonly used to describe states whose configuration interaction (CI) expansions include doubly-substituted configurations with large amplitudes. However, being dependent on the chosen reference configuration, this definition is ambiguous.

If the ground-state wave function $\Psi$ and excited-state wave function $\hat{\Psi}$ for an $N$-electron system are single determinants of the MOs $\psi_i$ and $\hat{\psi}_j$, respectively, the excitation number[166]

$$\eta = N - \sum_i^n \sum_j^n |\langle \psi_i | \hat{\psi}_j \rangle|^2$$

measures the number of electrons in the excited state which occupy the space spanned by the virtual orbitals of the ground state. Thus, for example, doubly-substituted versions of $\Psi$ would have $\eta = 2$. Post-excitation orbital relaxation leads to small deviations from this ideal value but $\eta$ allows us easily to decide whether or not a state is doubly excited. We will discuss the excitation number in detail in Chapter 5.

4.3.1 The H$_2$ molecule

The $1s^2$ ground state of the H$_2$ molecule has been a benchmark for quantum chemical methods since the dawn of quantum mechanics.[167] Accurate energies of the lowest doubly-excited $1s^2$ state, an auto-ionizing resonance, were first obtained by Bottcher and Docken[168] and later by others[169, 170, 171, 172].

We performed SCF calculations on both states with $R = 1.4$ bohr using a modified aug-mcc-pV8Z basis[173] to which additional diffuse $s$, $p$ and $d$ shells were added and from which the $g$ and higher shells were removed.

The HF and CI excitation energies (Table 4.1) are similar because excitation preserves the electron pair and so the correlation energies of the states are comparable. Because the self-interaction errors[174]

$$E_{\text{SIE}} = \frac{1}{2}E_J + E_X$$

85
4.3. DOUBLE EXCITATIONS

<table>
<thead>
<tr>
<th>%Fock</th>
<th>BLYP</th>
<th>B3LYP</th>
<th>HF-LYP</th>
<th>HF</th>
<th>Full CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(1\sigma_u^2)$</td>
<td>-0.16250</td>
<td>-0.16016</td>
<td>-0.09950</td>
<td>-0.08076</td>
<td>-0.11755</td>
</tr>
<tr>
<td>$E(1\sigma_g^2)$</td>
<td>-1.17031</td>
<td>-1.18071</td>
<td>-1.17198</td>
<td>-1.13363</td>
<td>-1.17428</td>
</tr>
<tr>
<td>$\Delta E$</td>
<td>27.42</td>
<td>27.77</td>
<td>29.18</td>
<td>28.65</td>
<td>28.75</td>
</tr>
<tr>
<td>$E_{\text{SIE}}(1\sigma_u^2)$</td>
<td>-1.781</td>
<td>-1.219</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$E_{\text{SIE}}(1\sigma_g^2)$</td>
<td>0.007</td>
<td>0.217</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\eta$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 4.1: Total energies ($E$, in hartree), excitation energies ($\Delta E$, in eV), self-interaction energies ($E_{\text{SIE}}$, in eV) and $\eta$ values for the $1\sigma_g^2 \rightarrow 1\sigma_u^2$ excitation in H$_2$.

(where $E_J$ and $E_X$ are Coulomb and exchange energies) are much larger in the excited state than in the ground state, DFT results improve as the percentage of Fock exchange increases. This highlights the need for exchange functionals that are accurate for excited states.

Because the electrons are excited from a gerade to an ungerade MO, the overlap integrals in (3.3) vanish and $\eta$ is predicted by all levels of theory to be exactly 2.

4.3.2 Polycyclic hydrocarbons

It is difficult to model $\pi \rightarrow \pi^*$ excited states of benzene and polyacenes accurately. While semi-empirical methods, such as the Pariser-Parr-Pople (PPP) method,[175, 176, 177] can give good results, ab initio methods often struggle to obtain comparable performance.

Early CI studies gave excitation energy errors exceeding 1 eV for some valence states of benzene.[178, 179, 180] Multireference configuration interaction (MRCI)[181] and symmetry-adapted cluster configuration interaction (SAC-CI)[182] reduce the errors to 0.5 eV but only multireference perturbative treatments (e.g. MRMP and CASPT2) were able to achieve errors of 0.1 – 0.29 eV for benzene,[183, 184] 0.27 – 0.54 eV for naphthalene,[183] 0.15 eV for anthracene and 0.25 eV for naphthacene.[185]

For consistency with the other molecules, we classify the states of benzene using $D_{2h}$ symmetry, and its first singlet excited state, which is $1^1E_{2g}$ in $D_{6h}$, becomes the totally symmetric $2^1A_g$ in $D_{2h}$.

The lowest totally symmetric, $\pi \rightarrow \pi^*$, singly and doubly excited singlet states are interesting. In benzene, the single excitation is much lower in energy than the double, but the ordering reverses in anthracene[183, 185] and larger molecules.[186]
To discover whether standard DFT methods reproduce this reordering, we used the IMOM to compute the BLYP/6-311G* energies of these states in benzene, naphthalene, anthracene and pleiadene (Fig. 4.2) at ground-state BLYP/6-311G structures. Initial guesses were obtained from the ground-state MOs by promoting two electrons from the HOMO to the lowest $\pi^*$ virtual orbital.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>State</th>
<th>Present work $\eta$</th>
<th>BLYP</th>
<th>Previous work</th>
<th>Experimental work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>$2 , ^1A_g$</td>
<td>1.0058</td>
<td>7.70</td>
<td>8.01[183]</td>
<td>7.73[183]</td>
</tr>
<tr>
<td></td>
<td>$5 , ^1A_g$</td>
<td>2.0034</td>
<td>10.21</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Naphthalene</td>
<td>$2 , ^1A_g$</td>
<td>1.0191</td>
<td>5.66</td>
<td>5.86[183]</td>
<td>5.65[183]</td>
</tr>
<tr>
<td></td>
<td>$4 , ^1A_g$</td>
<td>2.0019</td>
<td>6.77</td>
<td>6.75[189]</td>
<td>6.76[189]</td>
</tr>
<tr>
<td>Anthracene</td>
<td>$2 , ^1A_g$</td>
<td>2.0013</td>
<td>4.62</td>
<td>5.42[183]</td>
<td>5.03[183]</td>
</tr>
<tr>
<td></td>
<td>$3 , ^1A_g$</td>
<td>1.0141</td>
<td>4.92</td>
<td>6.57[185]</td>
<td>5.28[185]</td>
</tr>
<tr>
<td>Pleiadene</td>
<td>$2 , ^1A_1$</td>
<td>2.0118</td>
<td>2.46</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>$3 , ^1A_1$</td>
<td>1.0818</td>
<td>3.43</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 4.2: Excitation energies of the lowest totally symmetric singly- and doubly-excited states of aromatic systems.

Table 4.2 compares the resulting excitation energies with CASSCF, MRMP and experimental values, and Fig. 4.3 reveals the energy reordering as the system size increases. In all cases, the BLYP calculations give the correct ordering and pleasingly
accurate excitation energies, with a mean absolute deviation from the experimental values of only 0.15 eV. The worst result, an error of 0.41 eV for the singly-excited $3^1A_g$ state of anthracene, can be compared with CASSCF and MRMP errors of 1.24 eV and 0.32 eV, respectively.

![Figure 4.3: The lowest valence totally symmetric singly and doubly excited states for C$_6$H$_6$, C$_{10}$H$_8$, C$_{14}$H$_{10}$ and C$_{18}$H$_{12}$ using BLYP/6-311G*.

The $1^1E_{2g}$ state in benzene has been reported to be a doubly-excited state arising from the (HOMO)$^2 \rightarrow$ (LUMO)$^2$ transition. This assignment is based on large amplitudes of doubly-substituted configurations that appear in the CASSCF and MRMP wave functions.[183] However, this state can be accurately modelled with a single determinant, and its $\eta$ value (1.0058) strongly suggests that, in fact, it is only singly excited. We infer from this that the important doubles in the CASSCF wave function serve largely to describe correlation and relaxation, and should not be interpreted as indicating that the state is doubly excited.

### 4.4 Charge-Transfer States

A typical electronic excitation creates an electron-hole pair as the electron moves from one MO to another. If the electron and hole are separated by a significant distance $R$, the result is termed a charge-transfer (CT) state. Because of their charges, the electron
and hole attract, causing the energy of the system to rise as $1/R$ as the donor and acceptor separate.

CIS is able to reproduce the $1/R$ dependence of the excitation energy for CT states, despite giving large (0.5 - 2 eV) excitation energy errors.\cite{57} In contrast, TD-DFT calculations fail to capture the $1/R$ behavior, not because of flaws in TD-DFT itself but because of the adiabatic approximation universally adopted in its implementations.\cite{68, 69} Because excited-state DFT energies obtained using the IMOM do not make this approximation, one anticipates that they will avoid the undesirable features of TD-DFT. To test this, we have compared the performance of DFT with CIS and TD-DFT for the lowest CT states of two supramolecular systems: ethylene + tetrafluoroethylene and bacteriochlorin + zinc-bacteriochlorin.

4.4.1 Ethylene + Tetrafluoroethylene

![Figure 4.4: The ethylene + tetrafluoroethylene complex.](image)

CT states of the $\text{C}_2\text{H}_4 + \text{C}_2\text{F}_4$ complex (Figure 4.4) have been studied previously by Dreuw \textit{et al.} using TD-DFT and CIS.\cite{68} To capture the correct $1/R$ behaviour, they proposed a hybrid approach which combines TD-DFT and CIS and which yielded reasonable estimates for the CT excitation energies. However, because of its reliance on CIS, their approach is not very accurate.

Figure 4.5 compares the excitation energies of the first CT state predicted by EOM-CCSD, CIS, M08-HX, B3LYP and TD-B3LYP as the distance $R$ is varied. The 6-31G* basis set was used for all calculations. As anticipated, TD-B3LYP is qualitatively wrong, while EOM-CCSD, CIS, M08-HX and B3LYP all reproduce the correct $1/R$ behaviour. The failure of TD-B3LYP can be traced to an incompletely modeled interaction between the electron and the hole.

Although CIS captures the correct decay behaviour, it predicts an excitation energy at $R = 4.6$ which is 0.65 eV greater than the EOM-CCSD reference (Table 4.3).
4.4. CHARGE-TRANSFER STATES

Figure 4.5: Variation of the excitation energy $\Delta E$ of the first CT state of C$_2$H$_4$ + C$_2$F$_4$ with the distance $R$.

<table>
<thead>
<tr>
<th>Complex</th>
<th>$R$/Å</th>
<th>M08-HX</th>
<th>B3LYP</th>
<th>CIS</th>
<th>TD-B3LYP</th>
<th>EOM-CCSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$_2$H$_4$ + C$_2$F$_4$</td>
<td>4.60</td>
<td>10.84</td>
<td>10.48</td>
<td>11.61</td>
<td>6.79</td>
<td>10.96</td>
</tr>
<tr>
<td>BC + ZnBC</td>
<td>11.20</td>
<td>3.28</td>
<td>3.33</td>
<td>3.69</td>
<td>1.82</td>
<td>—</td>
</tr>
<tr>
<td>BC + ZnBC</td>
<td>12.04</td>
<td>3.41</td>
<td>3.49</td>
<td>3.78</td>
<td>1.84</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 4.3: Excitation energies (in eV) of the first CT transfer states of the C$_2$H$_4$ + C$_2$F$_4$ and BC + ZnBC complexes.

B3LYP error at this point is also large (~0.48 eV), but the M08-HX functional[192] reduces this to 0.12 eV.

4.4.2 Bacteriochlorin + Zn-Bacteriochlorin

Bacteriochlorins (7,8,17,18-tetrahydroporphyrins) are the chromophoric moiety of bacteriochlorophylls (BChl) which are found in purple bacteria, green bacteria and heliobacteria.[193] Their photochemistry has aroused broad scientific interest, from the development of artificial light-harvesting antennae for photoactive devices [194, 195, 196, 197] to photodynamic therapy for cancer destruction [198]. The key step in the process involves the absorption of light and a transfer of the singlet excitation energy, via protein-BChl...
complexes, to the photosynthetic reaction centre[195]. It is obvious, therefore, that the theoretical study of such excitations requires a detailed understanding of charge-transfer (CT) states.

In nature, the zinc-bacteriochlorin (Zn-BC) is linked to the bacteriochlorin (BC) through a phenylene bridge. However, the phenylene group has only a minor influence on the CT states and, in our study, we follow the approach of Dreuw et al.,[69] adopting the model shown in Fig. 4.6. This allows the distance between the chromophores to be varied from that determined by the bridge, which is 12.04 Å. (Note that our definition of $R$ differs from theirs.)

![Figure 4.6: Bacteriochlorin + zinc-bacteriochlorin complex](image)

![Figure 4.7: Variation of the excitation energy $\Delta E$ of the first CT state of BC + Zn-BC with the distance $R$. The dashed line shows the natural separation $R = 12.04$ Å.](image)
B3LYP/6-31G* structures of the complex were optimized for several $R$ between 11.2 and 16.2 Å. Excitation energies of the first CT state were calculated using B3LYP, M08-HX, CIS and TD-B3LYP with the 6-31G* basis, and these are shown in Figure 4.7. As we saw for the C$_2$H$_4$+C$_2$F$_4$ complex, the DFT and CIS methods predict the correct 1/$R$ dependence but TD-B3LYP fails.

At $R = 12.04$ Å, the CIS excitation energy of 3.78 eV (Table 4.3) is close to the value (3.79 eV) obtained by Dreuw et al. using their hybrid approach.[69] Higher levels of theory are prohibitive for the BC/Zn-BC system, and no experimental results are available. However, our results for the C$_2$H$_4$ + C$_2$F$_4$ complex suggest that CIS probably overestimates the excitation energy and that the M08-HX results are probably the most accurate.

### 4.5 Conical Intersections

A conical intersection (ConInt) is a subset of the nuclear coordinate space where the adiabatic potential energy surfaces (PESs) of two electronic states of a molecule are degenerate. ConInts frequently play a key role in the reactions, spectroscopy and dynamics of molecules, especially those of biochemical interest.[199, 200]

Due to the degeneracy of the PESs, excited-state calculations involving ConInts are challenging. In particular, linear response-based methods, e.g. CIS and TD-DFT, fail when the HOMO-LUMO gap is small or zero. We test the ability of single-determinant methods to model PESs in the vicinity of ConInts by considering the H$_3$ and retinal molecules.

#### 4.5.1 The H$_3$ molecule

The study of ConInts in H$_3$ has a long history, both theoretical and experimental, and we encourage the interested reader to study the survey[201] by Halász et al. for further details. There are four ConInts involving the three lowest-energy electronic states but we will focus on the ConInt with $D_{3h}$ symmetry as it is characterised by a single interatomic distance $R$. For most values of $R$, the doublet ground state $D_0$ has $^2E'$ symmetry. However, for very small $R$, the $^2A'_1$ state is lower in energy, thus creating the ConInt. Mielke et al. reported[202] accurate PESs for the H+H$_2$ reaction.
at a highly correlated level with the aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ basis sets. They found the minimal energy ConInt (MECI) at $R = 0.495$ Å and we use this value to assess the performance of various DFT, TD-DFT and CIS methods.

The energies of the two states were computed using the conventional CIS method, four IMOM-based methods (HF, HF-LYP, BLYP and B3LYP) and three TD-DFT methods (TD-HF-LYP, TD-BLYP and TD-B3LYP). In all cases, the aug-cc-pVDZ basis set[203] was used. The energy differences are plotted in Fig. 4.8 and, on such plots, any ConInts appear as cusps on the horizontal axis.

Using the IMOM, we were able to find the ConInt at all four single-determinant levels considered. CIS is also able to model the ConInt and, for the narrow domain of $R$ shown in Fig. 4.8, the CIS and HF energies are very similar. This is no coincidence for, if the CIS state intersected the ground state at a different value of $R$, the CIS solution would have a lower energy than the ground state, leading to a contradiction.

None of the TD-DFT models yields a ConInt. Both TD-BLYP and TD-B3LYP show discontinuities at the ConInt because the ground-state reference changes from $^2A'_1$ to $^2E'$ at this point. These different references give different excitation energies for the $D_1$ state, leading to the discontinuity. The TD-HF-LYP PES is continuous and has a cusp that coincides with that obtained using HF-LYP. However, it does not correspond to a ConInt as the solutions are not degenerate at this point.

HF-LYP, the most accurate of the methods considered, predicts an MECI at $R = 0.483$ Å which is only 0.012 Å below the Mielke value ($R = 0.495$ Å) and HF and CIS predict an MECI at 0.016 Å above Mielke’s. Both BLYP and B3LYP predict an MECI at a bond length that is almost 0.1 Å shorter than Mielke’s, again showing that standard exchange functionals struggle when applied to excited-state electron densities.

4.5.2 Retinal

The photoisomerization of the 11-cis retinal chromophore to its all-trans form in the rhodopsin protein is the primary process involved in vision.[204] Many attempts have been made to explain this process from an electronic structure point of view, including pioneering ab initio calculations by Du and Davidson[205] in 1990 on the excited states of the protonated Schiff base of retinal (PSBR). Early theoretical studies of
photoisomerization of protonated Schiff base cations were reported by Bonačić-Koutecký et al. [206]. They showed that the isomerization of the formaldiminium cation (CH$_2$NH$_2^+$) occurs through a ConInt between the $S_1$ and $S_0$ states at an N–C bond twist-angle of 90°. Since then, many theoretical studies [207, 208, 209, 210] have provided evidence that the $S_1 - S_0$ ConInt in retinal is responsible for the ultrafast photoisomerization of the molecule and this was later corroborated by experiments by Polli et al. [211]. Unfortunately, Levine et al. [61] report that ConInts involving a closed-shell singlet ground state cannot be found by either TD-DFT or CIS because “matrix elements connecting the initial state and the response states are excluded from the formulation”.

Figure 4.8: Energy difference $\Delta E$ between the $D_0$ and $D_1$ states of equilateral H$_3$ as the interatomic distance $R$ varies. Green triangles show the MECI predicted by Mielke et al. [202].
The reaction path for the photoisomerization of retinal undoubtedly involves complicated motions of all of the nuclei. However, it is dominated by the torsion rotation about the C_{15}–C_{16} double bond[208, 209, 212] and the MECI is expected to lie near $\phi = 90^\circ$ (Fig. 4.9).

The equilibrium geometry was found at B3LYP/6-31G* and frozen-geometry scans for $0 \leq \phi \leq 180^\circ$ were then performed. Using the 6-31G* basis, the ground ($S_0$) and excited ($S_1$) energies were computed using the IMOM at the BLYP and B3LYP levels and the resulting PESs are shown in Fig. 4.10. Both BLYP and B3LYP predict a ConInt near $\phi = 90^\circ$, lying approximately 2 eV above the equilibrium structure. This is slightly lower than CASSCF-based estimates of around 2.3 eV in the work of Molnar et al.[208] and of Andruniów et al.[212]
4.6 Concluding Remarks

In this Chapter, we have examined single-determinant approximations for excited states involving double excitations, charge-transfer and conical intersections. These determinants correspond to higher-energy solutions of the SCF equations, and are widely believed to be difficult to obtain. However, the new IMOM protocol provides a straightforward and reliable method for obtaining these solutions, and we have shown that they may be preferable to other low-cost excited-state methods.

For double excitations, which cannot be described by CIS or TD-DFT, IMOM-based HF or DFT calculations are among the few low-cost options available. Moreover, we find that the single-determinant energies obtained in this way are remarkably accurate and can rival far more expensive methods such as CASSCF and MRMP. It is especially surprising to discover how accurately the $1\sigma_2^u$ resonance state in $H_2$ is modelled by HF theory.

Charge-transfer states are also modelled well by single determinants and the correct $1/R$ behaviour is predicted even for functionals whose potentials are asymptotically incorrect.

Finally, conical intersections, which are particularly challenging for both CIS and TD-DFT, are satisfactorily treated by IMOM-based DFT.
Chapter 5

The excitation number

“That’s right!” shouted Vroomfondel, “we demand rigidly defined areas of doubt and uncertainty!”

Douglas Adams — The Hitchhiker’s Guide to the Galaxy

5.1 Introduction

In Chapters 3 and 4, we have shown that single-determinant approximations are often reliable models for electronic excited states, even for difficult cases such as double excitations, charge-transfer states and conical intersections.

In this Chapter, we exploit the structural simplicity of the orbital picture emerging from the single-determinant scheme to rigorously respond to a fundamental, unanswered question in chemistry: how many electrons are excited during an electronic transition?

The answer to this question is not straightforward as the absorption of even a single photon can excite more than one electron.[213, 214] Moreover, a two-photon experiment can result in either a single electron excited multiple times, or several electrons excited individually.[215, 216]

Multiply excited states play a key role in influential fields such as optoelectronics.[217] but it is rare that the number of excited electrons can be measured experimentally, and no reliable theoretical methods exist that can compute an “excitation number”.

97
5.1. INTRODUCTION

In part, the reason for this theoretical shortcoming is that the most successful approaches for excited states are based on configuration interaction (CI) expansions of the wave functions \( \Psi_k \)

\[
\Psi_k \approx c_0 \Phi_0 + \sum_{i} \sum_{r} c_i^r \Phi_i^r + \sum_{i<j} \sum_{r<s} c_{ij}^{rs} \Phi_{ij}^{rs} + \cdots
\]  

(1.1)

where \( c_0, c_i^r, c_{ij}^{rs}, \ldots \) are the CI amplitudes and \( \Phi_0 \) is the reference configuration, typically the Hartree-Fock (HF) ground-state determinant. We recall from Chapter 1, that the \( \Phi_i^r \) are singly-substituted determinants in which an electron has been promoted from the occupied molecular orbital (MO) \( |\psi_i\rangle \) to the virtual MO \( |\psi_r\rangle \), \( \Phi_{ij}^{rs} \) are doubly-substituted determinants and so forth.

The addition of excited determinants in this CI expansion performs two major functions: i) they introduce electronic correlation; ii) they model relaxation effects, i.e. they account for the fact that the MO basis is optimised for the reference state and is non-optimal for the target excited state;[115]

The form of the CI-based approaches leads one to naturally define an \( m \)-tuply excited state as one whose CI expansion includes significant amplitudes for the \( m \)-tuply-substituted determinants. Unfortunately, this definition is deceptive for two reasons: i) the nature of the excitation depends on the reference configuration, so that two different reference configurations may yield two inconsistent classifications of the excitation; ii) the value at which amplitudes become “significant” is somewhat arbitrary.

A better approach was devised in 1995 by Head-Gordon et al.[218] These authors suggested the number of electrons promoted during an electronic transition is given by the promotion number \( p_n = \text{Tr} A = \text{Tr} D \), where \( A \) and \( D \) are the attachment and detachment densities, respectively. Unfortunately the promotion number yields values that deviate significantly from integers and this complicates the assignment of the excitation number.

In Section 5.2 of this Chapter, we introduce a new quantity, the excitation number \( \eta \), as a metric for the characterisation of multiply excited states. In Section 5.3, based on the mathematical features of the excitation number, we define the hole and particle densities \( \hat{\rho}_{AB}(r) \) and \( \hat{\rho}_{AB}(r) \), which enable the visualization of the part of space left and
occupied by the excited electrons, respectively. Sections 5.4 and 5.5 present results which show that the new metric is rigorous and remarkably quantitative, whereas previous definitions were either inconsistent or just qualitative, leading to cases of misassignment. Throughout the latter two Sections, we provide evidence of how the excitation number can be generally used in chemistry to markedly improve the characterisation of excited states.

5.2 The Excitation Number \( \eta \)

We have previously shown,[131, 163, 219] that an alternative to CI-based approaches is to model excited states by approximating the wave function with a single determinant

\[
\Psi_k \approx \Phi_k = \det[\chi_i(r_i, s_i)],
\]

of spin-orbitals \( \chi_i(r, s) = \psi_i(r)\sigma(s) \), with MOs \( \psi_i(r) \) and spin functions \( \sigma(s) \). This is achieved by replacing the aufbau protocol with the Maximum Overlap Method (MOM)[131] or, more reliably, with the Initial Maximum Overlap Method (IMOM).[219] This approximation can be remarkably accurate,[163] even for very challenging excitations.[219] As outlined in Figure 5.1, the reason for the accuracy of IMOM is that it allows a full relaxation of the MOs, which are optimized specifically for the target excited state.

Thus, by adopting single-determinant approximations for excited states the relaxation problem is solved. There is yet another important consequence of this approach: the excitation number can now be rigorously defined.

In fact, if \( \Phi_A \) and \( \Phi_B \) are the determinants for the two states \( A \) and \( B \), and \( |a_i\rangle \) and \( |b_i\rangle \) are their occupied spin orbitals, respectively, then the projection of \( |b_i\rangle \) into the occupied space of state \( A \) is

\[
|b_i^A\rangle = \sum_j |a_j\rangle \langle a_j|b_i\rangle
\]

(2.3)

Therefore, the part of the density of state \( B \) that lies in the occupied space of \( A \) is

\[
\rho^A_B(r, s) = \sum_i \sum_j \sum_k \langle a_j|b_i\rangle \langle b_i|a_k\rangle a_j(r, s) a_k(r, s)
\]

(2.4)

Consequently, the number of electrons which have been excited in the transition from \( A \)
5.2. THE EXCITATION NUMBER \( \eta \)

Figure 5.1: Molecular orbital diagram for single-determinant approximations to ground and doubly-excited states. The doubly-substituted determinant of ground state orbitals is allowed to relax its orbitals through the iterative IMOM protocol.

The excitation number to B can be defined as the difference

\[
\eta_{AB} = n - n^A_B
\]  
(2.5)

where \( n \) is the total number of electrons in the system and

\[
\eta^A_B = \sum_i \sum_j \langle a_j | b_i \rangle \langle b_i | a_j \rangle
\]  
(2.6)

is the number of electrons in state B that lie in the occupied space of A, obtained by integrating the \( \rho^A_B(r,s) \) density.

Some useful properties of the excitation number \( \eta_{AB} \) are: i) it is invariant to an arbitrary unitary transformation of either the occupied or the virtual MOs of both the states; ii) it is symmetric in A and B, such that \( \eta_{AB} = \eta_{BA} \); iii) it is exactly an integer.
if A and B have the same sets of MOs, that is if there is no relaxation effect; iv) it is extremely simple to evaluate.

5.3 Hole and Particle Densities

Figure 5.2: Hole and particle densities for a “double charge-transfer state” of Tris(bipyridine)ruthenium(II) ([Ru(bipy)_3]^{2+}) calculated at the BLYP/3-21G level of theory. The double excitation (\( \eta = 2.072 \)) involves a metal to ligand (spin up) and a ligand to ligand (spin down) charge transfer. The computed vertical excitation energy is of 6.43 eV.

The “projected” density defined in Eq. (2.4) can be used as an additional source of information. Specifically, considering, without loss of generality, only spin-up electrons, the density of the (spin-up) electronic hole left in the A \( \rightarrow \) B transition and of the paired excited particle(s) are respectively

\[
\tilde{\rho}_{AB}(r) = \rho_A(r) - \rho^A_B(r) \tag{3.7}
\]

\[
\rho_{AB}(r) = \rho_B(r) - \rho^A_B(r) \tag{3.8}
\]
5.3. HOLE AND PARTICLE DENSITIES

where \( \rho_A(r) \) and \( \rho_B(r) \) are just the (spin-up) densities of state A and B. As shown in Fig. 5.2, the hole and particle densities can be easily visualised, offering valuable information about the spatial distribution of the electrons involved in the excitation.

If we assume the following LCAOs for the MOs

\[
a_i(r) = \sum_{\mu=1}^{N} C_{i\mu}^A \phi_\mu(r) \\
b_j(r) = \sum_{\nu=1}^{N} C_{j\nu}^B \phi_\nu(r)
\]

where \( \{ \phi_i \} \) defines the atomic basis set, Eq. (2.4) becomes

\[
\rho_B^A(r) = \sum_i \sum_j \sum_k p_{ji}^A p_{ki}^A \sum_{\mu\nu} C_{j\mu}^A C_{k\nu}^B \phi_\mu(r) \phi_\nu(r)
\]

where \( p_{ji}^A = (a_j | b_i) \). Thus the matrix elements of the representation of \( \rho_B^A(r) \) within the AO basis are

\[
(D_B^A)_{\mu\nu} = \sum_i \sum_j \sum_k p_{ji}^A p_{ki}^A C_{j\mu}^A C_{k\nu}^B
\]

By defining the matrix \( P^A = p^A(p^A)^\dagger \), the \( D_B^A \) matrix is given by the expression

\[
D_B^A = (C^A)^\dagger P^A C^A
\]

Using Eqs. (3.7) and (3.13), the following expression for the hole density matrix \( \tilde{D}_{AB} \) is obtained

\[
\tilde{D}_{AB} = D_A - D_B^A \\
= (C^A)^\dagger(I - P^A) C^A
\]

where \( D_A \) is the density matrix for state “A” and \( I \) is the identity matrix.

Analogously, it can be shown that the particle density matrix \( \tilde{D}_{AB} \) is

\[
\tilde{D}_{AB} = D_B - D_B^A \\
= (C^B)^\dagger(I - P^B) C^B
\]
where $P^B = p^B(p^B)^\dagger$ and $p^B_{ji} = \langle b_j | a_i \rangle = p^A_{ij}$. 
5.4 Numerical Results

Table 5.1 reports \( \eta \) values and vertical excitation energies for different singlet and triplet excited states of several molecules obtained at the BLYP/6-31+G* level of theory. The mean absolute deviations from the reference values for the excitation energies is less than 0.3 eV. Errors for other DFT functionals are less than 0.5 eV, while larger deviations were seen for HF. It is both remarkable and pleasing to observe how close the \( \eta \) values are to integers. This is not specific to the BLYP functional as the values change little when other DFT functionals are used.

From property (iii) of \( \eta \), the deviations from integers can be certainly ascribed to orbital relaxation effects. However, small or zero deviations from integers do not necessarily imply negligible or no orbital relaxation, as the relaxation of each MO may yield either positive or negative values, thus opening the possibility for cancellation.

Table 5.1 highlights that \( \eta \) values tend to be slightly larger than the nearest integer, the only exception (out of 39 cases) being the doubly excited state of LiH with configuration \( (1\sigma^+)^2(3\sigma^+)^2 \). This anomaly is due to a combination of there being few electrons in the system together with very little relaxation associated with the lowest energy \( 1\sigma^+ \) orbital, and very large relaxation of the \( 3\sigma^+ \) orbital. This large relaxation is achieved mainly by mixing with the \( 2\sigma^+ \) orbital, which is part of the ground state occupied space. These conditions are uncommon and, therefore, it is expected that \( \eta \) values smaller than the nearest integer are rare.

We use the data in Table 5.1 to justify our definition of \( m \)-tuply excited states: they are states with \( \eta \approx m \).

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Transition</th>
<th>State</th>
<th>( \eta )</th>
<th>( p_n )</th>
<th>( \Delta E (eV) )</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH(_2)O</td>
<td>( 2, b_1 \rightarrow 2, b_2 )</td>
<td>( ^1A_2 )</td>
<td>1.024</td>
<td>1.372(1.211)</td>
<td>3.49</td>
<td>4.07[220]</td>
</tr>
<tr>
<td>CH(_2)O</td>
<td>( 2, b_1 \rightarrow 2, b_2 )</td>
<td>( ^3A_2 )</td>
<td>1.018</td>
<td>1.413(1.231)</td>
<td>3.36</td>
<td>3.50[220]</td>
</tr>
<tr>
<td>CH(_2)O</td>
<td>( (2, b_1)^2 \rightarrow (2, b_2)^2 )</td>
<td>( ^1A_1 )</td>
<td>2.020</td>
<td>2.431</td>
<td>9.94</td>
<td></td>
</tr>
<tr>
<td>C(_4)H(_6)</td>
<td>( 1, b_{1g} \rightarrow 2, a_u )</td>
<td>( ^1B_{1u} )</td>
<td>1.002</td>
<td>1.146(1.347)</td>
<td>5.54</td>
<td>5.92[221]</td>
</tr>
<tr>
<td>C(_4)H(_6)</td>
<td>( 1, b_{1g} \rightarrow 2, a_u )</td>
<td>( ^3B_{1u} )</td>
<td>1.006</td>
<td>1.281(1.097)</td>
<td>3.21</td>
<td>3.22[221]</td>
</tr>
<tr>
<td>C(_4)H(_6)</td>
<td>( 1, a_u \rightarrow 2, a_u )</td>
<td>( ^1A_g )</td>
<td>1.022</td>
<td>1.384</td>
<td>6.11</td>
<td>6.25[222]</td>
</tr>
<tr>
<td>C(_4)H(_6)</td>
<td>( 1, a_u \rightarrow 2, a_u )</td>
<td>( ^3A_g )</td>
<td>1.012</td>
<td>1.385</td>
<td>5.05</td>
<td>4.91[221]</td>
</tr>
</tbody>
</table>
### CHAPTER 5. THE EXCITATION NUMBER

$$(1\, b_{1g})^2 \rightarrow (2\, a_u)^2$$  
\[
\begin{array}{cccc}
1\, e_{1g} \rightarrow 1\, e_{2u} & 1\, e_{1g} \rightarrow 1\, e_{2u} & 1\, a_{2u} \rightarrow 1\, e_{2u} & 1\, a_{2u} \rightarrow 1\, e_{2u} \\
1\, b_{2g} \rightarrow 2\, b_{1u} & 1\, b_{2g} \rightarrow 2\, b_{1u} & 2\, b_{2g} \rightarrow 2\, b_{1u} & (2\, b_{2g})^2 \rightarrow (2\, b_{1u})^2 \\
4\, a_2 \rightarrow 6\, b_2 & 4\, a_2 \rightarrow 6\, b_2 & 3\, a_2 \rightarrow 6\, b_2 & (4\, a_2)^2 \rightarrow (6\, b_2)^2 \\
4\, a_2 \rightarrow 5\, a_2 & (4\, a_2)^2 \rightarrow (6\, b_2)^2 & 3\, a_2 \rightarrow 5\, a_2 & (3\, a_2)^2 \rightarrow (6\, b_2)^2 \\
3\, e_{2g} \rightarrow 4\, e_{1g} & 3\, e_{2g} \rightarrow 4\, e_{1g} & (3\, e_{2g})^2 \rightarrow (4\, e_{1g})^2 & 1\, t_1 \rightarrow 2\, e \\
& & & (1\, t_1)^2 \rightarrow (2\, e)^2 \\
\end{array}
\]

\[
\begin{array}{cccc}
1\, B_{2u} & 1\, B_{2u} & 1\, E_{2g} & 1\, B_{3u} \\
1.001 & 1.013 & 1.006 & 1.008 \\
1.165(1.082) & 1.426(1.049) & 1.284 & 1.547(1.239) \\
5.00 & 4.33 & 7.66 & 3.32 \\
& & 7.80 & 3.30 \[185]
\end{array}
\]

\[
\begin{array}{cccc}
1\, A_{1g} & 1\, A_{1g} & 2\, a_u & 2\, b_{1u} \\
2.004 & 2.001 & 2.294 & 2.242 \\
2.30 & 4.62 & 10.19 & 4.71 \[227]
\end{array}
\]

\[
\begin{array}{cccc}
1\, B_1 & 1\, B_1 & 2\, B_1 & 1\, B_1 \\
1.014 & 1.017 & 1.014 & 1.014 \\
1.510(1.497) & 1.647(1.339) & 1.629 & 1.642 \\
1.08 & 0.85 & 2.94 & 2.79 \\
& & 3.29 \[191]
\end{array}
\]

\[
\begin{array}{cccc}
1\, A_1 & 2\, A_1 & 3\, A_1 & 3\, A_1 \\
2.012 & 2.102 & 3.006 & 3.006 \\
2.677 & 2.677 & 1.916 & 1.916 \\
2.47 & 2.47 & 6.25 & 6.25 \\
& & 2.46 \[191]
\end{array}
\]

\[
\begin{array}{cccc}
1\, A_1 & 4.005 & 1\, A_1 & 1\, A_1 \\
4.005 & 4.724 & 4.005 & 4.005 \\
4.724 & 10.72 & 4.724 & 10.72 \\
& & 10.72 & 10.72 \\
& & 10.72 & 10.72 \\
\end{array}
\]

\[
\begin{array}{cccc}
1\, E_{1g} & 1\, E_{1g} & 1\, A_{1g} & 1\, A_{1g} \\
1.023 & 1.030 & 2.044 & 2.044 \\
1.573(1.246) & 1.638(1.258) & 2.862 & 2.862 \\
2.62 & 2.11 & 7.16 & 7.16 \\
& & 2.70 \[228]
\end{array}
\]

\[
\begin{array}{cccc}
1\, T_2 & 1\, T_2 & 1\, T_2 & 1\, T_2 \\
1.055 & 1.037 & 1.055 & 1.055 \\
1.657(1.471) & 1.706(1.433) & 1.657(1.471) & 1.657(1.471) \\
2.23 & 1.96 & 2.23 & 2.23 \\
& & 2.23 & 2.23 \\
\end{array}
\]

\[
\begin{array}{cccc}
1\, A_1 & 2\, A_1 & 1\, A_1 & 1\, A_1 \\
2.115 & 3.171 & 2.115 & 2.115 \\
3.171 & 6.54 & 3.171 & 6.54 \\
6.54 & 6.54 & 6.54 & 6.54 \\
\end{array}
\]

*Not assigned
5.5 Discussion and Misassignements

The excitation number can be used to significantly enhance the theoretical characterisation of excited states. For example, without \( \eta \) it would be very difficult to assign the \( 2 \, ^1A_g \) state of anthracene (\( C_{14}H_{10} \)) and the \( 2 \, ^1A_1 \) state of pleiadene (\( C_{18}H_{12} \)) as low-lying doubly-excited states, facts which are corroborated by strong agreement with the experimental results.[227, 191]

The excitation number also allows us to determine if states that have been assigned as multiply excited states based on CI amplitudes have, in fact, been misassigned in the literature. This appears to be the case for the controversial \( 2 \, ^1A_g \) state of \emph{trans}-butadiene.
CHAPTER 5. THE EXCITATION NUMBER

and the $1^1E_{2g}$ state of benzene.

5.5.1 The $2^1A_g$ state of trans-(1, 3)-butadiene ($C_4H_6$)

The $2^1A_g$ state of trans-butadiene ($C_4H_6$ in Table 5.1) has been extensively studied both experimentally and theoretically (for a survey of its theoretical and experimental history the reader is referred to Saha et al.[233] and references therein). The most sophisticated and accurate CI-based studies all agree that the state shows considerable contributions from doubly-substituted determinants. For example, 42% of the accurate wave function obtained by Serrano-Andrés et al.[221] consists of doubly-substituted configurations, and yields a vertical excitation energy of 6.27 eV, to be compared with the experimental value of 6.25 eV.[222] Later, Hsu et al.[102] reported remarkably accurate excitation energies for this state obtained using TD-DFT and a range of functionals (0.18 eV MAD from 6.25 eV). This was surprising since it is well known that TD-DFT is structurally incapable of describing double excitations within the commonly adopted Adiabatic Local Density Approximation.[58, 60, 59, 61] Since then, others[60, 61] have highlighted the double character of $2^1A_g$, culminating in 2006 with the study by Starcke et al. where they concluded that “for short polyenes (in particular for butadiene) the lowest excited $2^1A_g$ state can clearly be classified as doubly excited”, while the ability of TD-DFT to describe the state is ascribed to a “fortuitous cancellation of errors in the ground and excited state wave functions”.[234]

What does $\eta$ tell us about the $2^1A_g$ state of trans-butadiene? Table 5.1 shows that a simple BLYP/6-31+G* single-determinant calculation yields an excitation energy in very good agreement with experiment and an $\eta$ value of 1.022, leading us to conclude $2^1A_g$ is actually a singly excited state!

The reader may be wondering why, if this state is only singly excited, accurate CI-based wave functions all show significant double excitation contributions. The reason is that in all these approaches the MOs are not optimized for the excited state and, as a result, many excited determinants (including the doubly-substituted ones) come into play in order to take care of the relaxation of the non-optimal MOs. This still does not explain why TD-DFT can actually describe the $2^1A_g$ excitation. In order to do so, we must discern between two effects of TD-DFT: i) a partial relaxation of the MOs which
is obtained exclusively via single excitations; ii) the effect of the exchange-correlation functional. Now, one may argue, based on the fact that accurate CI-based wave functions all show a prevalent doubles contribution, that the partial relaxation of TD-DFT is certainly not suited to model the 2\(^1\)A\(_g\) state. Still, such a comparison between TD-DFT and CI is physically meaningless because of the exchange-correlation functional effects. In particular, it is very likely that, because of the latter, in TDDFT doubly excited determinants are not significant at all for describing 2\(^1\)A\(_g\). This is clearly evidenced by the outstanding results of Hsu et al.\[102\]

5.5.2 The 1\(^1\)E\(_{2g}\) valence state of benzene (C\(_6\)H\(_6\))

A similar story to the 2\(^1\)A\(_g\) state of butadiene exists for the 1\(^1\)E\(_{2g}\) valence state of benzene. Benzene is one of the most theoretically studied, and at the same time of the most challenging, molecules in computational chemistry (for a survey of its theoretical history the reader is referred to Matos et al.[235] and references therein). Early single-reference CI studies on benzene\[71, 179, 180\] gave errors larger than 1.0 eV for some of the valence states. These results were improved by multi-reference CI (MRCI)[181] and symmetry adapted cluster CI (SAC-CI)[182] but still gave errors of 0.5 eV. It was only with the use of perturbation theory applied to a MRCI wave function that accuracies of 0.1-0.29 eV were achieved.[183, 184] The most accurate vertical excitation energies for the 1\(^1\)E\(_{2g}\) valence state have been obtained in this way by Lorentson et al. and by Hashimoto et al. both with \(\Delta E = 7.73\) eV,[183, 184] where the experimental vertical excitation energy is 7.80 eV.[187] Since in both studies the wave function has a strong doubly excited character (30% and 33%, respectively), 1\(^1\)E\(_{2g}\) is described as the double excitation arising from the (HOMO)\(^2\) \(\rightarrow\) (LUMO)\(^2\) transition.[183]

Again, Table 5.1 shows that, with an \(\eta\) value of 1.006, the 1\(^1\)E\(_{2g}\) state of benzene is only a singly excited state, and the significant amplitudes of doubly-substituted determinants have to be attributed to relaxation and correlation effects.

5.5.3 A note on the low-lying double excitation of MnO\(_4^-\)

In closing this Chapter, we draw attention to the doubly excited (\(\eta = 2.115\)) \(^1\)A\(_1\) state of MnO\(_4^-\) arising from the \((1t_1)^2 \rightarrow (2\ e)^2\) transition (highlighted in bold in Table
5.1). The experimental spectrum shows a considerable absorption peak at 6.56 eV,[230] however, the transition associated with this absorption has not been definitively assigned. Given that the experimental value strongly agrees with our theoretical result of 6.54 eV, we believe the state can be identified as the doubly excited $^1A_1$. This is corroborated by the strong computed oscillator strength of 0.141.

5.6 Concluding Remarks

In this Chapter, we have introduced the excitation number, $\eta$, which is a measure of the number of electrons that are involved in an electronic transition. The physical motivation behind $\eta$ also leads to natural definitions for particle and hole densities, which allow the parts of space occupied and vacated by the excited electrons to be visualised. Furthermore, we have shown that $\eta$ values are pleasingly close to integers, which allows for unambiguous determination of how many electrons have been excited. We have used this to correctly describe several multiply excited states that have been misassigned in the past.
PART B

Many-Electron Integrals for Explicitly Correlated Methods
### Introduction to Part B

In 1985, starting from the Hylleraas functional\[236, 237, 238\] and using the interelectronic distance \( r_{12} = |r_1 - r_2| \) as a correlation factor, Kutzelnigg derived the seminal form of the MP2-R12 equations.\[239\] This explicitly correlated method (more formally stated together with Klopper in 1987\[240\]) was later extended to higher levels of theory and more accurate correlation factors \( f_{12} = f(r_{12}) \),\[241, 242, 243, 244, 245, 246, 247, 248, 249, 250\] such as linear combinations of Gaussian Geminals

\[
f_{12} = \sum_{g=1}^{K_G} D_g \exp(-\lambda_g r_{12}^2),
\]

or a single Slater Geminal

\[
f_{12} = \exp(-\lambda r_{12}).
\]

The resulting ‘F12 methods’ achieve chemical accuracy for small organic molecules with relatively small Gaussian basis sets\[251, 252, 253, 254, 255, 256\] and are quickly becoming the first-choice method for high accuracy.\[254, 255\]

The inclusion of the correlation factor \( f_{12} \) dredges up an old problem: in addition to two-electron integrals (traditional ones and new ones), three-electron integrals over \( f_{13}f_{23}, r_{12}^{-1}f_{13} \) and \( r_{12}^{-1}f_{13}f_{23} \), as well as four-electron integrals over \( r_{12}^{-1}f_{14}f_{23}, r_{12}^{-1}f_{13}f_{34} \) and \( r_{12}^{-1}f_{13}f_{14} \) arise. These integrals are not known analytically except when Gaussian Geminals are used,\[257, 258\] and, at that time, the only way to evaluate them would have been via expensive Gauss-Legendre quadratures.\[259, 260\] Additionally, citing Kutzelnigg and Klopper: \[261\] “even if fast procedures for the evaluation of these integrals were available, one would have to face the problem of the large number of these integrals; while that of two-electron integrals is \( \sim N^4 \), there are \( \sim N^6 \) three-electron and \( \sim N^8 \) four-electron integrals. The storing and manipulating of these integrals could be handled only for extremely small basis sets.”

For this reason the success of the R12 method was determined by the decision of avoiding three- and four-electron integrals entirely through the insertion of the *resolution*
of the identity \textit{(RI)}[261, 252, 262]

\[
\hat{I} \approx \sum_{\mu}^{N_{RI}} |\chi_{\mu}| \langle \chi_{\mu} | .
\]  

(6.18)

In this way three- and four-electron integrals are approximated as linear combinations of products of more conventional two-electron integrals. Of course, the accuracy of the RI approximation (7.19) relies entirely on the assumption that the auxiliary basis set is sufficiently large \((N_{RI} \gg N)\). Therefore, in the general context of explicitly correlated methods, it is reasonable to ask: what is the most suitable method to evaluate three- and four-electron integrals?

The use of RI as the method of choice does not seem definitive to us. In fact, eschewing the RI approximation would offer at least two advantages: i) smaller one-electron basis as the larger auxiliary basis set would not be required anymore; ii) the three- and four-electron integrals could be computed exactly. Moreover, one could avoid the commutator rearrangements involved in the computation of integrals over the kinetic energy operator.[263]

In the late 80’s two-electron integrals technology was still in development.[264, 265, 266, 267, 268, 269, 270, 20] Nowadays, though still challenging, we believe these integrals can be computed effectively via recursive schemes.

Another important remark is that the actual number of significant (i.e. greater than a given threshold) three- and four-electron integrals in a large system is, at worst, \(O(N^{3})\) or \(O(N^{4})\). Furthermore, if the correlation factor is a short-range operator (as in modern F12 methods[245, 246, 247, 248, 271, 249, 250]) it can be shown that the number drops to only \(O(N^{2})\). These kinds of scaling are achievable, for example, partly by exploiting robust density fitting[272] or, as we will show later, fully by rigorous upper-bound-based screening methods. If we can develop good algorithms for identifying and computing the tiny fraction of many-electron integrals that are significant, large-scale calculations using F12 methods will become feasible, without the need for the RI approximation.

In Part B we present new, effective schemes for the evaluation of the many-electron integrals over gaussian basis functions required for explicitly correlated methods. Such schemes exploit a combination of efficient screening methods and recurrence relations,
and represent a generalisation of contraction-efficient algorithms for the evaluation of two-electron integrals, such as the Head-Gordon-Pople and the PRISM schemes,[270, 20] to many-electron integrals.

Chapter 6 offers a brief overview of explicitly correlated methods, with a particular emphasis on the many-electron integrals’ problem. In Chapter 7, after introducing our notation, we construct and test a variety of novel two- and three-electron integral upper bounds and discuss their implementation in effective screening algorithms, which enable one to compute only the significant fraction of integrals. Chapters 8 and 9 present efficient algorithms to evaluate the tiny fraction of significant integrals over a general class of multiplicative three- and four-electron operators using recurrence relations. In Chapter 10 we show that two-electron integrals over Gaussian Geminals, \(\exp(-\lambda r_{12}^2)\), due to their unique factorization properties, can be computed via a combination of new screening and recursion techniques, much more cheaply than in any other previous scheme. Finally, in Chapter 11, we construct screening and recurrence schemes for three- and four-electron integrals arising when the correlation factor \(f_{12}\) is a Gaussian Geminal. We present evidence that the Gaussian Geminal choice yields a dramatic reduction of the computational complexity of these integrals.

We believe our approach represents a major step towards an accurate and efficient computational scheme for three- and four-electron integrals.
Chapter 6

An Overview of Explicitly Correlated Methods

“To create greater convergence, we need more integration.”

Emmanuel Macron

6.1 Introduction

Quantum chemistry is torn between two extreme forces: accuracy and scalability. It is well known that accurate wave functions are required to encompass electron correlation effects, which are often of central importance in the energetics of the system. On the other hand, scalability demands from a quantum chemical method low computational complexity with respect to system size. High parallelizability is also very desirable, given the modern computer architecture.

For correlated wave functions that are expanded in terms of products of one-electron basis functions, accuracy and scalability are definitely irreconcilable. In particular, all standard post-Hartree-Fock calculations (e.g. many-body perturbation theory, configuration interaction or coupled-cluster theory) suffer from a slow convergence of electron correlation with respect to the size of the one-electronic basis set adopted.
6.1. INTRODUCTION

The slow convergence problem is linked to the presence of the Coulomb hole in the exact wave function $\Psi$. As discussed in Section 1.5.2, the Coulomb potential $1/r_{ij}$, between two electrons “$i$” and “$j$”, is divergent at the electron-coalescence point $r_{ij} = 0$. Nonetheless, the local energy $E = \mathcal{H}\Psi/\Psi$ is constant everywhere, and thus the divergence in the Coulomb energy much be exactly cancelled by the kinetic energy operator.[251] This is achieved only if the wave function is linear in the inter-electronic distance as the electrons coalesce, because $-(1/2)\nabla^2 r = -1/r$. A more rigorous exploration of such properties of the exact many-particle wave function led Kato to the formulation of the famous cusp condition[17]

$$\lim_{r_{ij} \to 0} \left( \frac{\partial \Psi}{\partial r_{ij}} \right) = \frac{1}{2} \Psi(r_{ij} = 0) \quad (1.1)$$

where $\Psi$ denotes spherical averaging over the hypersphere $r_{ij} = \text{const.}$, and, later, Pack and Brown to the formulation of more general conditions.[273]

It is the electronic cusp required by Eq. (1.1) that products of one-electron basis functions struggle to model, due to their smoothness. This yields an extremely slow convergence as $O \left[ (l + 1/2)^{-4} \right]$, with the angular momentum $l$, in the increment of the correlation energy, or as $O \left[ (L + 1)^{-3} \right]$ in the truncation error at some maximum value of the angular momentum $L$.[239] Furthermore, for each individual $l$ the saturation of the $l$-subspace is required from the basis. Interestingly the slow convergence of CI expansions is not only due to the $O \left[ (l + 1/2)^{-4} \right]$ dependence, but also to the fact that the saturation of each $l$-subspace requires a combinatorial number of configurations.[239]

In this Chapter we present a brief overview of explicitly correlated electronic structure theories, which circumvent the slow convergence problem by explicitly including $r_{12}$ dependencies into the wave function. Theory is presented following a chronological order, which we believe beneficial for understanding the evolution of the matter.

For more detailed treatises on the subject, the interested reader may study Refs.[251, 262, 252, 253, 254].

116
6.2 Hylleraas-Type Wave Functions: the Emergence of Three- and Four-Electron Integrals

The story of explicitly correlated methods begins in the 1930s with Hylleraas\textsuperscript{236, 237} and James and Coolidge,\textsuperscript{167} who introduced explicitly the inter-electronic coordinate $r_{12}$ as an additional two-electron function to expand the wave functions of He and H$_2$. As Kutzelnigg later showed, this not only leads to a prominent improvement of the energy convergence with respect to the angular momentum, from $\mathcal{O} \left( (l + 1/2)^{-4} \right)$ to $\mathcal{O} \left( (l + 1/2)^{-8} \right)$, but it also avoids the second source of the slow convergence of CI.\textsuperscript{239}

In particular, Hylleraas’ calculations were carried out on the $1^1S$ state of He using the coordinates

$$s = r_1 + r_2; \quad t = r_1 - r_2; \quad u = r_{12}$$

(2.2)

where $r_1$ and $r_2$ are the distances of electron ‘1’ and ‘2’ from the nucleus and $r_{12}$ is the distance between the two electrons. The spatial part of the He ground-state wave function was written as

$$\tilde{\Psi}_{\text{Hyl}} \equiv \exp(-\zeta s) \sum_{k=1}^{N} c_k s^{l_k} t^{2m_k} u^{n_k}$$

(2.3)

as only even powers of $t$ contribute to singlet states.

With only 3 terms Hylleraas obtained $E = -2.90243 E_h$ after variationally optimizing both the linear parameters $c_k$ and the nonlinear parameter $\zeta$ (by hand!). In a systematic approach the He ground-state energy can be computed from an $N$-term Hylleraas expansion with all terms that satisfy $l_k + 2m_k + n_k \leq L$. With $L = 6$ the Hylleraas expansion has 50 terms and obtains microhartree accuracy in the energy; with $L = 13$, the expansion has 308 terms and yields nanohartree accuracy.\textsuperscript{252}

Since the 1930s much progress has been achieved in Hylleraas-type calculations for He and He-like ions,\textsuperscript{274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287} leading to results with up to 40-significant-digits accuracy.\textsuperscript{288} Consequently, the more interesting question becomes how Hylleraas-type trial wave functions may be designed for many-electron atoms and polyatomic molecules.

Hylleraas’ ansatz can be straightforwardly extended to many-electron atoms by
including in the wave function all the possible inter-electronic coordinates of the system, that is $r_{12}$, $r_{13}$, $r_{23}$ and so forth. This is the essence of the Hylleraas-configuration-interaction (Hyl-CI) method, whose trial wave function for an $n$-electron atom is defined as follows

$$\tilde{\Psi}^{\text{Hyl-CI}} = \sum_{m=1}^{K} c_m \hat{O}(L^2) A \left( \sigma_j r_{12}^{v_m} \prod_{i=1}^{n} \phi_{i_m} \right)$$  \hspace{1cm} (2.4)$$

where $\hat{O}(L^2)$ is an idempotent orbital angular momentum projection operator, $A$ is the $n$-electron antisymmetrization operator, $\sigma_j$ is an $n$-electron spin function and $\phi_{i_m}$ is a one-electron spatial orbital (typically a Slater-type orbital).

Note that in Hyl-CI only configurations with $v_m = 0, 1$ are included. In this case, not only are traditional one- and two-electron integrals needed for the calculation, but new, additional three- and four-electron integrals arise. The complexity of the integrals will not increase when the number of electrons becomes larger than four. In particular two- and three-electron integrals over the following operators (where $k$ is 1 or 2)

$$1/r_{12}; \quad r_{12}^{k}; \quad r_{12}r_{13}; \quad r_{12}^{k}/r_{13}; \quad r_{12}r_{13}/r_{23}$$  \hspace{1cm} (2.5)$$

and four-electron integrals over the operators

$$r_{12}r_{13}/r_{14}; \quad r_{12}r_{13}/r_{34}; \quad r_{12}r_{34}/r_{23}$$  \hspace{1cm} (2.6)$$

are required.

Thus, the major price to pay for the enhanced convergence rate brought by the insertion of the linear $r_{12}$ factor in the trial wave function is that now numerous, complicated three- and four-electron integrals must be evaluated. Much progress in the evaluation of such integrals has been achieved during the last 60 years.[289, 290, 291, 292, 293, 294, 259, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311] However, despite the progress, Hyl-CI calculations have been thus far restricted to either small atomic systems or very small molecules, most certainly due to the difficulty of evaluation and the numerosity of the integrals.[252, 253]

The efficient evaluation of integrals over the operators in Eqs. (2.5) and (2.6), and of more general three- and four-electron integrals arising in explicitly correlated
methods will be the main object of this Part B of the thesis. For the reminder of this Chapter though, we will continue our overview of explicitly correlated electronic structure theories.

### 6.3 Slater Geminal

At almost the same time as Hylleraas, James and Coolidge, Slater showed, in a study on the Rydberg series of the He atom, that “the correct wave function for $S$ states should approach $\exp(-2(r_1 + r_2) + r_{12}/2)$ at small $r$’s”. [312] Slater suggested multiplying the wave functions of atoms by a prototypical form of what is known today as a Slater Geminal (SG) function

$$S_{12} = \exp(-\lambda_{12} r_{12}).$$

(3.7)

Indeed many years later SGs will be proven much more effective in enhancing the convergence behaviour than the linear $r_{12}$. [248, 247, 250] We will return to this point in our discussion of MP2-F12 theory.

### 6.4 Gaussian Geminals

The increase in mathematical complexity brought by the introduction of such correlation factors, either $r_{12}$ or $S_{12}$, to perform Hylleraas-type calculations is computationally overwhelming. This has been clear since the dawn of explicitly correlated methods. Quoting Boys, from 1960, [257] “the mathematical treatment was of such nature as to offer no hope of generalization to more than two electrons, except perhaps with extreme difficulties for three electrons.”

In the wake of Hylleraas and Slater, Boys [257] and Singer [258], in 1960, independently proposed the inclusion of the Gaussian Geminal (GG) function,

$$G_{12} = \exp(-\lambda_{12} r_{12}^2),$$

(4.8)

instead of the bare $r_{12}$ or $S_{12}$ terms. The use of GGs significantly reduces the amount of calculation required for explicitly correlated wave functions as “there are explicit formulas for all of the necessary many-dimensional integrals”. [257] Interestingly, in the
same article a visionary Boys argued that even if GGs do not fulfil the cusp conditions exactly, they could be used fit the conditions based on functions of the $S_{12}$ type. Remarkably, he also noted that the most fruitful application of GGs might not be for Hylleraas-type calculations, but rather “for corrections of the perturbation type to the wave functions calculated by other methods.”

### 6.5 Exponentially Correlated Gaussians Method

During the years following the initial proposal by Boys[257] and Singer[258], variational calculations adopting GGs flourished, giving birth to many of the most accurate results for small-scale systems. In these calculations the wave function was expanded in terms of antisymmetrized linear combinations of Exponentially Correlated Gaussians (ECG)[313]

$$
\phi_k = \prod_{i=1}^{N} \exp\left(-\alpha_{ki}|r_i - A_{ki}|^2\right) \prod_{i>j}^{N} \exp\left(-\beta_{kij}|r_i - r_j|^2\right).
$$

(5.9)

Since then, extremely accurate ECG applications treated many atomic and some diatomic systems.[314, 315, 316] Modern variational calculations on systems with more than four electrons usually restrict each $n$-electron basis function to at most one $r_{ij}$,[317] in complete analogy to Hyl-CI. Thus, the resulting matrix elements include up to four-electron integrals only.

Even if three- and four-electron integrals can be obtained in closed-form, the main difficulty with ECG wave functions is the need to optimize not only the linear expansion coefficients but also the non-linear parameters of each $\phi_k$, that is $\alpha_{ki}$, $A_{ki}$ and $\beta_{kij}$. There is a total of $N(N+7)/2$ such parameters for a molecule without special symmetry, which make the non-linear optimization already extremely expensive for systems with few electrons. For this reason these calculations still linger on small molecular scales.[252, 253]

### 6.6 Gaussian-Geminals MP2

For MP2 calculations it is conventional to take advantage of the decomposition of the second-order energy into a sum of pair contributions.[238] This decomposition requires
the first-order pair functions $|u_{ij}\rangle$, which constitute the first-order correction to the wave function, to be strongly orthogonal to the occupied space. In orbital-based methods this is trivially achieved by expanding the $|u_{ij}\rangle$ in products of virtual orbitals. In general though, in order to ensure the orthogonality, it is necessary to use an explicitly projected form $Q_{12}|u_{ij}\rangle$, where $Q_{12} = (1 - O_1)(1 - O_2)$ is a projector onto the complete virtual space, and

$$
O_k = \sum_{i=1}^{\text{occ.}} \chi_i(x_k) \int dx_3 \chi_i(x_3) \mathcal{P}_{sk}
$$

(6.10)

are projectors onto the occupied spin-orbital space, with $\mathcal{P}_{sk}$ replacing the coordinates of electron $k$ by 3. Then the functional known as the Hylleraas functional or strong orthogonality functional (SOF), originally introduced by Sinagoğlu,[318, 238]

$$
H[u_{ij}] = \langle u_{ij}|Q_{12}(\mathcal{F}_1 + \mathcal{F}_2 - \epsilon_i - \epsilon_j)Q_{12}|u_{ij}\rangle + 2\langle u_{ij}|Q_{12}^{-1}r_{12}^{-1}|ij\rangle \geq \epsilon_{ij}^{(2)}
$$

(6.11)

can be minimized for each pair to provide upper bounds for each MP2-pair energy $\epsilon_{ij}^{(2)}$. In Eq. (6.11), according to the formalism of Klopper et al.,[251] $i$ and $j$ are occupied spin-orbitals, the kets (bras) $|u_{ij}\rangle$ and $|ij\rangle$ are understood to be properly antisymmetrized, $\mathcal{F}$ is the Fock operator and $\epsilon_i$ is an orbital energy.

In the early 1970s, based on Sinagoğlu’s earlier work,[238] Pan and King proposed the use of GGs for the expansion of the pair functions $|u_{ij}\rangle$ of the first-order correction to the wave function of Moller-Plesset perturbation theory.[319, 320] A modified, slightly cheaper version of this approach was developed by Adamowicz and Sadlej which allowed the extension of Pan and King’s methodology to the $\text{H}_2$, LiH and BH molecules.[321, 322, 323]

Even if these methods represented a substantial step forward in terms of computational burden (compared to Hyl-CI), they still suffered from the major limitation of requiring an extensive optimization of the non-linear parameters in the correlated GGs. In particular, the bottleneck of these calculations was determined by the large number of four-electron integrals arising from the combined presence of the Fock operator, the GGs and the $Q_{12}$ projector in the strong orthogonality functional. These needed to be evaluated at each optimization step expending $\mathcal{O}(N^6N_{GG}^2)$ time, where $N_{GG}$ was the number of GGs used.[324]
A major breakthrough was brought in 1982 by Szalewicz et al.\cite{325,326} with the introduction of a \textit{weak orthogonality functional} (WOF) which reduced the time of a single calculation by a factor of $O(N^2)$, as only the evaluation of two- and three-electron integrals is required. The WOF, which retains a strict upper bound to the second-order energy, has the form

\begin{equation}
W[u_{ij}] = \langle u_{ij} | \mathcal{F}_1 + \mathcal{F}_2 - \epsilon_i - \epsilon_j | u_{ij} \rangle + 2 \langle u_{ij} | Q_{12} r_{12}^{-1} | ij \rangle + \eta_{ij} \langle u_{ij} | O_1 + O_2 | u_{ij} \rangle \geq \epsilon_{ij}^{(2)} \tag{6.12}
\end{equation}

where $\eta_{ij}$ is a positive parameter satisfying the inequality $\eta_{ij} \geq (\epsilon_i + \epsilon_j)/2 - \epsilon_1$, with $\epsilon_1$ being the orbital eigenvalue of the lowest energy occupied orbital. The ansatz used for the pair function in WOF calculations is

\begin{equation}
|u_{ij}\rangle = \mathcal{A} \left[ \sigma_1 \sigma_2 \sum_k c_k \exp \left( -\alpha_k |r_1 - A_k|^2 - \beta_k |r_2 - B_k|^2 - \gamma_k |r_1 - r_2|^2 \right) \right] \tag{6.13}
\end{equation}

where $\sigma_1$ and $\sigma_2$ are spin functions. Analogously to the ECG method, both the linear coefficients $c_k$ and the non-linear parameters $\alpha_k, \beta_k, \gamma_k, A_k$ and $B_k$ are fully optimized for each pair function $|u_{ij}\rangle$. In this way Szalewicz et al. managed to obtain highly converged MP2, MP3 and Coupled-Cluster results for a variety of small molecules.\cite{326,327,328}

There are two main flaws in GGs-MP2-WOF calculations; i) the optimization of non-linear parameters is still extremely expensive, precluding the possibility of applying this method to medium- or large-sized systems and restricting such calculations to systems not larger than 10 electrons;\cite{252,253} ii) the penalty function $\eta_{ij} \langle u_{ij} | O_1 + O_2 | u_{ij} \rangle$ in the WOF may lead to poor results if small basis sets are used.\cite{251,252}

6.7 MP2-R12

In 1985, independently, Kutzelnigg derived a first form of the MP2-R12 method starting from the Hylleraas’ functional.\cite{239} Kutzelnigg’s idea, which was more formally stated together with Klopper in 1987,\cite{240} rested upon the direct usage of the linear $r_{12}$ correlation factor, given that it exactly satisfies the cusp condition unlike $G_{12}$. The
original ansatz used for the pair functions in Ref. [240] was

\[ |u_{ij} \rangle = \frac{1}{2} c_{ij} \mathcal{Q}_{12} r_{12} |ij \rangle + |w_{ij} \rangle \] (7.14)

with

\[ |w_{ij} \rangle = \sum_{a<b} d_{ij}^{ab} |ab \rangle \] (7.15)

where labels \( i \) and \( j \) refer to occupied spin-orbitals, labels \( a \) and \( b \) to virtuals, and \( c_{ij} \) and \( d_{ij}^{ab} \) are linear variational parameters. Later, in 1990,[329] it was suggested to orthogonalize the two-electron basis functions \( \mathcal{Q}_{12} r_{12} |ij \rangle \) against all orbital products constructible in the given one-particle basis, which is achieved by using the projection operator

\[ \mathcal{Q}_{12} = (1 - \mathcal{O}_1)(1 - \mathcal{O}_2) - \mathcal{V}_1 \mathcal{V}_2 \] (7.16)

where

\[ \mathcal{V}_k = \sum_{a=1}^{\text{virt.}} \chi_a(x_k) \int dx_3 \chi_a(x_3) \mathcal{P}_{3a} \] (7.17)

are projectors onto the virtual spin-orbital space. Thus, the SOF is recast in the form

\[ 2 \Re \langle ij | r_{12}^{-1} | u_{ij} \rangle + \langle u_{ij} | \mathcal{F}_1 + \mathcal{F}_2 - \epsilon_i - \epsilon_j | u_{ij} \rangle \] (7.18)

and should be minimized with respect to the variational parameters.

However, the inclusion of \( r_{12} \) dredged up an old problem: in addition to two-electron integrals (traditional ones and new ones), three-electron integrals over the \( r_{12} r_{13}^{-1}, r_{12} r_{13}^{-1} r_{23} \) operators and four-electron integrals over \( r_{12} r_{23}^{-1} r_{34} \) arise. Unlike their Gaussian cousins \( (G_{12} r_{13}^{-1}, G_{12} r_{13}^{-1} r_{23} \text{ and } G_{12} r_{23}^{-1} G_{34} ) \), and even within a Gaussian one-electron basis, these integrals do not have analytical expressions and, at that time, the only way to evaluate them would have been via either one- or two-dimensional Gaussian-Legendre quadratures.[259, 296, 297] Additionally, quoting Kutzelnigg and Klopper,[261]

“even if fast procedures for the evaluation of these integrals were available […] there are \( \sim N^6 \) three-electron and \( \sim N^8 \) four-electron integrals. The storing and manipulating of these integrals could be handled only for extremely small basis sets.”
Undoubtedly the two-electron integral technology was still in development in late 1980s: McMurchie and Davidson derived Hermite-Gaussian-based recurrence relations (RRs) in 1978,[264] Pople and Hehre in the same year,[265] Dupuis, Rys and King developed Rys quadrature between 1976 and 1983,[266, 267, 268] Obara and Saika published their first paper on RRs in 1986;[269] the Head-Gordon-Pople (HGP) algorithm,[270] which began to face more rigorously the contraction problem by using Obara and Saika’s RRs in tandem with newly introduced horizontal RRs, was born in 1988. Contraction-efficient schemes, such as PRISM, made their first appearance on the scene only in the early 1990s (see Ref. [20] and therein). Thus, it should not be surprising that Klopper and Kutzelnigg’s three- and four-electron integrals could be evaluated only via an expensive quadrature for any angular momentum of the basis functions. Nowadays, even if still challenging, these integrals could be computed much more effectively via judicious recursive schemes, relegating the quadrature only to the fundamental integrals.[330] Another considerable remark is that even if the task of computing $O(N^6)$ or $O(N^8)$ integrals in a large system is still overwhelming, the number of significant three- and four-electron integrals in such a system is, at worst, $O(N^3)$ or $O(N^4)$. These kind of scalings are achievable, for example, by exploiting robust density fitting[272] or upper bounds-based screening methods. We will get back to these issues in the coming Chapters.

Nevertheless, the MP2-R12 method was successful due to some ingenious approximations:

1. The occupied spin-orbitals $|i\rangle$ are assumed to be eigenfunctions of the exact Fock operator, that is $\mathcal{F}|i\rangle = \epsilon_i|i\rangle$. This, which is often referred to in the literature as the Generalized Brillouin Condition (GBC),[261] allows the reformulation $(\mathcal{F}_1 + \mathcal{F}_2 - \epsilon_1 - \epsilon_2)r_{12}|ij\rangle = [\mathcal{F}_1 + \mathcal{F}_2, r_{12}]|ij\rangle$ and further simplifications by deleting commutators of the form $[\mathcal{F}_1 + \mathcal{F}_2, (1 - \phi_1)(1 - \phi_2)]$. [250]

2. The virtual spin-orbitals $|a\rangle$ are assumed to be eigenfunctions of the exact Fock operator, that is $\mathcal{F}|a\rangle = \epsilon_a|a\rangle$, which is known as the Extended Brillouin Condition (EBC)[261] and allows simplifications by deletion of commutators of the form $[\mathcal{F}_1 + \mathcal{F}_2, \eta_1 \eta_2]$. 

124
3. The commutator relation \([K_1 + K_2, r_{12}] = 0\) is assumed to be true, yielding the approximation \([\mathcal{F}_1 + \mathcal{F}_2, r_{12}] \approx [t_1 + t_2, r_{12}]\).

4. Lastly, though arguably the most important approximation, three- and four-electron integrals are entirely avoided through the insertion of the resolution of the identity (RI)

\[
\hat{I} \approx \sum_{\mu}^{N_{RI}} |\chi_\mu\rangle \langle \chi_\mu| 
\]  

(7.19)

into the three- or four-electron operators.[261, 252, 262] In this way, three- and four-electron integrals are approximated as linear combinations of products of more conventional two-electron integrals.

Further approximations and improvements to the original MP2-R12 approach have been derived in the last decades.[331, 252, 262, 253] Within the orbital invariant ansatz, where all the excitation amplitudes are optimized, the MP2-R12 method scales as \(O(N^6)\), where \(N\) is a measure for the molecular size. More specifically, this cost is associated with the construction and the iterative inversion of the so-called “B matrix”* necessary for the optimization of the amplitudes. These procedures scale respectively as \(O(N_{\text{occ}}^4N_{RI}^2)\) and \(O(N_{\text{occ}}^4N_{\text{virt}}^2)\), where \(N_{\text{occ}}\) and \(N_{\text{virt}}\) are the number of occupied and virtual molecular orbitals.[331, 262]

Concerning the performance, Klopper, Manby, Ten-no and Valeev in their study of 2006[251] reported that for a set of calculations using the aug-cc-pVQZ on 28 small molecules, the various flavours of MP2-R12 all yield more than 98% of the complete basis set (CBS) valence-shell MP2 energies. This percentage rises to over 99% if aug-cc-pV5Z is used.[332]

Despite its success MP2-R12 has some important flaws. First, Eq. (7.19) is a good approximation only if the auxiliary basis set \(\{\chi_\mu\}_{1 \leq \mu \leq N_{RI}}\) is almost complete and therefore the auxiliary basis has to be very large to achieve chemically acceptable accuracy. Although more efficient RI formulations based on auxiliary basis set (ABS)[332] and its orthogonal component, complementary ABS (CABS),[333] have been developed, in practice one finds that it is difficult to achieve millihartree accuracy in total energies using the RI approach.[252] In fact, in the best-case scenario,[334] both ABS or CABS

---

*It is actually a third order tensor.
become exact, for atoms, when the basis set is saturated to $2L_{\text{occ}}$, where $L_{\text{occ}}$ is the maximum angular momentum of the occupied orbitals. Although this stringent statement cannot be rigorously extended to more complicated electronic charge distributions such as molecules due to the lack of spherical symmetry, in practice the same guidelines for the RI basis set apply.

Second, in a conclusive study of 2005, May, Valeev, Polly and Manby\cite{May2005} reported that MP2-R12 theory fails to deliver MP2 correlation energies in AO basis sets that are practical for large molecules. They also showed that the main factor responsible for such failure is the use of $r_{12}$ as explicitly correlated factor in the ansatz. Quoting Valeev,\cite{Valeev2005}

“Although in principle even a double-zeta basis set is sufficient to compute correlation energies in error by less than 1%, practice showed that MP2-R12 energies computed with an aug-cc-pVXZ basis were only as precise as the standard MP2/aug-cc-pV(X+1)Z energy.”

6.8 MP2-GG with Pre-Optimized GGs (GG(n) method)

As previously discussed, the accuracy of the RI approximation (7.19) relies entirely on the assumption that the auxiliary basis set is sufficiently large ($N_{RI} \gg N$). Based on this fact and on the analytic integrability of three- and four-electron integrals over GGs, in 1996 Persson and Taylor found a way to kill two birds with one stone: to use pre-optimized GGs expansions to eschew RI and the non-linear optimization at the same time.\cite{Persson1996} In fact, the main motivation for the introduction of RI in MP2-R12 was the complexity of the integrals introduced by the linear correlation factor $r_{12}$. Persson and Taylor suggested to fit $r_{12}$ using some more tractable functions, such as GGs

$$r_{12} \approx \sum_{\nu} b_{\nu} \left[ 1 - \exp(-\lambda_{\nu} r_{12}^2) \right]$$

and then to compute the arising three- and four-electron integrals in closed form. In this way they were able to show that even a six-term fit introduces errors of only 0.5 mE_h, while ten terms reduce the error to 20 $\mu$E_h.\cite{Persson1996}

One can rationalise the rapid basis set convergence with GGs by noting that the cusp at $r_{12} = 0$ should be of minor importance for the calculation of the electronic energy
because the volume element $4\pi r_{12}^2$ suppresses contributions for very short $r_{12}$.[252]

Unfortunately, further attempts to improve the approximations used by Klopper and Kutzelnigg so that smaller basis sets could be used were not very fruitful. In fact, the use of more elaborate fits, over a larger range of $r_{12}$ in the attempt of recovering high percentages of the total correlation energy actually resulted in a slight degradation of the results.[245] We know now, that this deterioration is not ascribable to the GG functions per se, but rather to the failure of $r_{12}$ as correlation factor.[249, 250, 335] It was realised, therefore, that a way to further improve the results is to optimize the linear coefficients of the GGs, for example, by means of the variational principle, rather than constraining them by fit.[245, 246]

The main advantage of this approach is that all variational parameters are linear, circumventing the most time-consuming steps of all the precedent GG-based approaches.

As in the work of Szalewicz et al.[325, 326] the WOF was used to avoid four-electron integrals. In Ref. [252], Hättig, Klopper, Köhn and Tew state

“The intention of Persson-Taylor ansatz was to develop a method that would provide results similar to the MP2-R12 method but without the need to insert resolution of the identity approximations that violate the upper-bound property of the Hylleraas functional. Even though evaluation of the necessary three-electron integrals is tedious and time-consuming, it was argued that in an integral-direct manner this evaluation would perform favourably on massive parallel computer architectures and that due to the short range of the Gaussians $\exp(-\gamma_n r_{12}^2)$ efficient integral screening, in particular, in the framework of local-correlation methods, would eventually lead to a method that scales linearly with the size of the system.”

Persson-Taylor approach, which is usually referred to as the GG($n$) method, was implemented by Dahle et al.[336, 337, 338] for different levels of sophistication and accuracy indicated by increasing $n$ values, that is GG(0), GG(1) and so forth.

Unexpectedly, the results for Persson-Taylor method given by Dahle et al.[336, 337, 338] were particularly poor. Such failure has been later attributed to the fact that GGs expansions are not well suited for use with the WOF as they are unable to form strong
orthogonal pairs.\[252\] Hence, the poor performance of these GG calculations was not due to the Gaussian Geminals \textit{per se}, but rather to their use in conjunction with the WOF.

### 6.9 MP2-F12

In 2004 Ten-no showed that the Slater Geminal $S_{12}$ is much more effective in recovering the MP2 correlation energy than the linear $r_{12}$ and that a 10-term GGs fit to the SG yields very similar results, suggesting that the cusp behaviour can be accurately represented by GGs.\[248, 339\] This was the beginning of modern MP2-F12 theory, where the ansatz for the pair functions takes the form

\[ |u_{ij} \rangle = \frac{1}{2} \sum_{kl} c_{ij}^{kl} Q_{12} f_{12} |kl\rangle + \frac{1}{2} \sum_{ab} d_{ij}^{ab} |ab\rangle \]  

in which $|kl\rangle$ and $|ab\rangle$ are appropriately antisymmetrized products of occupied and virtual orbitals, respectively, and $f_{12} = f(r_{12})$ is a spherically symmetric Geminal correlation factor which models the correlation holes appropriately. The terminology F12 was first used by May and Manby to distinguish the use of non-linear correlation factors.\[247\]

A variety of correlation factors have been examined during the last two decades,\[249, 340\] however modern MP2-F12 methods almost exclusively employ SGs or contracted GGs as correlation factors, that is respectively

\[ f_{12} = -\lambda^{-1} \exp(-\lambda r_{12}) \]  

\[ f_{12} = \sum_g K_g D_g \exp(-\lambda g r_{12}^2) \]

The main advantages of these correlation factors are related to the fact that they cover a wider domain of the correlation hole better than other factors such as linear $r_{12}$ functions. In fact, the long-range behaviour of the linear $r_{12}$ has been shown to be unphysical.\[248, 255\]

In practical evaluations of the necessary integrals, it is common practice to represent SG by a linear combination of GGs.
All the approximations presented for MP2-R12 are applied also for MP2-F12 methods (including RI).

### 6.9.1 SP Ansatz with the Rational Generator

In 2004, Ten-no brought another major breakthrough by proving that the excitation amplitudes multiplying the Geminal integrals could be determined by the cusp conditions.[248, 271] This is achieved by using the so-called SP ansatz for the pair functions

\[
|u_{ij}\rangle = \frac{1}{2} \sum_{ab} d_{ij}^{ab}|ab\rangle + \mathcal{R}_{12}|ij\rangle
\]  

(9.24)

where \(\mathcal{R}_{12}\) is the rational generator for the cusp conditions.[271]

\[
\mathcal{R}_{12} = Q_{12}f_{12}\left(\frac{3}{8} + \frac{1}{8}p_{12}\right)
\]  

(9.25)

and \(p_{12}\) is the permutation operator over the position vectors \(r_1\) and \(r_2\), that is

\[
p_{12}\Phi^{(s)} = (-1)^s\Phi^{(s)}
\]  

(9.26)

where \(s\) takes 0 and 1 values for singlet and triplet pairs, respectively.[255]

Using the commutability between the Fock and strong orthogonality projection operators, it is ensured that such an ansatz eliminates the Coulomb singularity of the perturbation in the first-order equation. Comparison between (9.21) and (9.24) leads to the condition for the amplitudes

\[
c_{ij}^{kl} = \frac{3}{8} \delta_{ik}\delta_{jl} + \frac{1}{8}\langle kl|p_{12}|ij\rangle
\]  

(9.27)

Ten-no’s fixed-amplitude method brings the scaling of MP2-F12 down to \(O(N^5)\) (equivalent to traditional MP2), because the construction and inversion of the \(B\) matrix is no longer required for obtaining the \(c_{ij}^{kl}\) amplitudes.[252, 248, 262]

### 6.9.2 MP2-R12 versus MP2-F12

In 2006 Valeev published a clear benchmark study (Ref. [335]) comparing the performance of MP2-R12 with MP2-F12, using both SGs and a set of (contracted) GGs with
fixed exponents, whose coefficients are optimized linearly. Figure 6.1 is taken from this study. Here the basis set incompleteness is analysed in terms of the relative error

$$\epsilon = \frac{E_{CBS} - E}{E_{CBS}}$$  (9.28)

where $E$ and $E_{CBS}$ are the correlation energies computed with the given and the complete basis set, respectively.

From Fig. 6.1 it is very clear that $r_{12}$ is by far the worst among the correlation factors: it takes an aug-cc-pVTZ basis to reduce the basis set error below the MP2/aug-cc-pVQZ level. Referring to Fig. 6.1, three GGs with an aug-cc-pVDZ one-electron basis are sufficient to best the precision of conventional MP2/aug-cc-pVQZ energy. The method with a single SG with unitary exponent can approach the precision of standard MP2/aug-cc-pV6Z, but not quite match it. The precision of the best standard MP2 calculation is matched or exceeded when five or seven GGs are used with only an aug-cc-pVDZ basis set.

Figure 6.1: From Ref. [335]. Frozen-core MP2 energies for several molecules computed with conventional and explicitly correlated wave functions. The notation aXZ stands for the aug-cc-pVXZ basis set.
6.10 F12-Coupled-Cluster

The technology of F12 methods in the previous Section can be applied to extend Coupled-Cluster methods. Noga et al.\cite{341, 241} introduced the F12-CC ansatz by extending the standard CC cluster operator with an additional F12 operator that takes care of the electron-electron cusp. For example, the F12-CCSD wave function is parametrized as

$$|\tilde{\Psi}\rangle^{\text{F12-CCSD}} \equiv \exp(T_1 + T_2 + \mathcal{J}_{12})|\Phi_0\rangle$$

(10.29)

where $|\Phi_0\rangle$ is the reference wave function, the operators $T_1$ and $T_2$ are the standard single- and double-excitation CC operators (see Eqs. (5.106) and (5.106)), and the Geminal cluster operator is

$$\mathcal{J}_{12} = \left(\frac{1}{2!}\right)^2 \sum_{ij} \sum_{kl} c_{ij}^{kl} \hat{Y}_{ij}^{kl}$$

(10.30)

whereas the effect of the $\hat{Y}_{ij}^{kl}$ operator is

$$\hat{Y}_{ij}^{kl} |ij\rangle = Q_{12} f_{12} |kl\rangle$$

(10.31)

Thus, the procedure for obtaining the non-variational energy and the amplitudes is very similar to the standard CC one. This leads to additional amplitude equations, whose implementation and computational complexities are vastly greater than that of standard CCSD and MP2-F12. In particular, the most expensive steps of F12-CCSD are associated with the evaluation of so-called intermediate $Z$ arising from the additional amplitude equations.\cite{342} This requires $O(N^8)$ operations and $O(N^6)$ storage, which is a dramatic increase with respect to standard CCSD.

Besides the special intermediates necessary for F12-CCSD, the sheer number of terms that appear in the amplitude equations has been defined as “staggering”.\cite{253} This is why most F12-CCSD implementations are based on approximations which lead to simplifications and cost reductions. The treatment of such approximate F12-CC methods certainly escapes the aim of this Chapter, however the interested reader may study Refs. \cite{342, 343}.
6.11 Strategies for the Approximated Evaluation of Many-Electron Integrals

As we have discussed throughout this Chapter, the slow convergence of dynamic correlation energies using conventional expansions of one-electron basis functions is effectively avoided in explicitly correlated electronic structure theory. The penalty is the introduction of many-electron integrals. For example, explicitly correlated MP2 methods involve the evaluation of three-electron integrals of the form

\[
\langle ij| r^{-1}_{12} O_2 f_{12} | kl \rangle = \sum_m \langle ijm| r^{-1}_{13} f_{13} | kml \rangle
\]  
(11.32)

\[
\langle ij| [T_1 + T_2, f_{12}] O_2 f_{12} | kl \rangle = \sum_m \langle ijm|[T_1 + T_2, f_{12}] f_{13} | kml \rangle
\]  
(11.33)

in addition to new two-electron integrals over \( f_{12} r^{-1}_{12} \), \([T_1 + T_2, f_{12}]\), and \([[[T_1 + T_2, f_{12}], f_{12}]\), where \( T_1 \) and \( T_2 \) are the kinetic energy operators for electrons “1” and “2”, respectively.

One can see that in Eqs. (11.32) and (11.33) and from now on, bra and kets are no longer antisymmetrized. Hence, \( \langle ijm| = \psi_i(r_1)^\dagger \psi_j(r_2)^\dagger \psi_m(r_3)^\dagger \) and \( |kml\rangle = \psi_k(r_1)\psi_m(r_2)\psi_l(r_3) \).

Different treatments of the commutator between the exchange operator and the correlation factor require additional integrals over the three-electron operator \( f_{13} r^{-1}_{12} f_{23} \) and the four-electron operator \( f_{13} r^{-1}_{12} f_{23} \).

The entire success of F12 methods relies on effective strategies for the treatment of such integrals. In the remainder of this Section, we discuss current approximated strategies for the evaluation of many-electron integrals. Exact or nearly-exact evaluation will be treated in the coming Chapters.

6.11.1 Resolution of the Identity

As anticipated before, three- and four-electron integrals that appear in matrix elements are typically simplified via the insertion of the RI. For example, the following three-electron integral

\[
\langle ijm|f_{12} r^{-1}_{13}|kml\rangle \approx \sum_p^{N_{RI}} \langle ij|f_{12}|pm\rangle \langle pm|r^{-1}_{12}|kl\rangle
\]  
(11.34)
is approximated by the sum of products of two-electron integrals.

If the set of orthonormal orbitals \( \{ \psi_p \} \) was complete the RI would be exact. Kutzen-nigg realised that the partial wave analysis of the RI error in atoms establishes basic requirements on the RI basis \( \{ \psi_p \} \) for atoms and possibly also for molecules.\[239\]

For example, the three-electron integral in Eq. (11.34) has a partial wave expansion that truncates at angular momentum \( 3L_{\text{occ}} \) for atoms, where \( L_{\text{occ}} \) is the maximum angular momentum associated with the set of occupied spin orbitals. Hence, the RI approximation for such an integral becomes exact if the RI basis is saturated up to \( 3L_{\text{occ}} \).

Three-electron integrals over “cyclic” operators, such as \( f_{12}r_{13}^{-1}f_{23} \), are even more computationally demanding, as they are approximated via a triple RI, that is

\[
\langle ijm | f_{12}r_{13}^{-1} f_{23} | mkl \rangle \approx \sum_{pq} \langle ij | f_{12} | pq \rangle \langle pm | r_{12}^{-1} | qr \rangle \langle r_{12} | kl \rangle
\] (11.35)

Furthermore, the partial wave expansion of these three-electron integrals does not truncate, as in the previous case, but it is reported to converge quickly, that is the error in truncating the basis at some \( L \) goes as \( (L + 1)^{-7} \).\[261\]

Ten-no and Manby\[334\] have proposed that the \( 3L_{\text{occ}} \) angular momentum requirement could be relaxed somewhat to \( 2L_{\text{occ}} \) by a clever rearrangement of the three-electron integrals prior to the RI application. Additionally, more efficient RI formulations based on auxiliary basis set (ABS)\[332\] and its orthogonal component, complementary ABS (CABS)\[333\] have been developed. However, one finds that in practice, given the basis set requirements, it is difficult to achieve millihartree accuracy in total energies using the RI approach, especially for large systems.\[252\]

### 6.11.2 Density Fitting

Explicitly correlated calculations based on the RI approximation are rather expensive in comparison with the usual orbital-based expansion because of the size of the RI basis and the required new types of two-electron integrals. Manby has improved this situation by introducing the density fitting (DF) technique into explicitly correlated theory.\[247, 344\] In particular, in the DF-MP2-F12 theory, robust density fitting (RDF)
has been employed, where products of orbitals hosting the same electron are fit using an auxiliary basis set

$$\psi_p(r)\psi_q(r) \equiv |pq\rangle = D_{A}^{pq}|A\rangle$$ (11.36)

and the coefficients $D_{A}^{pq}$ are obtained by minimizing a sensible metric. For example, if for simplicity we consider the following RDF approximation of two-electron repulsion integrals

$$(pq|r_{12}^{-1}|rs) \approx D_{A}^{pq}(A|r_{12}^{-1}|B)D_{B}^{rs}$$ (11.37)

then the coefficients are obtained by minimizing the Coulomb energy of the orbital product fitting residuals

$$\Delta_{pq} = (pq - \tilde{p}q|r_{12}^{-1}|rs - \tilde{r}s)$$ (11.38)

giving

$$D_{A}^{pq} = [J^{-1}]_{AB}J_{pq}^{B}$$ (11.39)

where $J_{AB} = (A|r_{12}^{-1}|B)$ and $J_{pq}^{B} = (B|r_{12}^{-1}|pq)$. Thus the Coulomb integral is approximated as

$$(\tilde{p}q|r_{12}^{-1}|\tilde{r}s) = D_{A}^{pq}J_{AB}D_{B}^{rs} = J_{pq}^{A}[J^{-1}]_{AB}J_{pq}^{B}$$ (11.40)

and the error in the fitted integral

$$(pq|r_{12}^{-1}|rs) - (\tilde{p}q|r_{12}^{-1}|\tilde{r}s) = (pq - \tilde{p}q|r_{12}^{-1}|rs - \tilde{r}s)$$ (11.41)

is quadratic in the error of the fitted densities, which is why the fit is called “robust”, characterized by avoiding errors linear in the error of the fitted densities.

In DF-MP2-F12 it is necessary to use formulae for the fitted integrals that are explicitly robust,[344] because the same density fitting coefficients are used for multiple target integral types. For example, a robust fit for a general three-electron integral,

$$\langle prt|qsu \rangle \equiv \langle pq|rs|tu \rangle,$$

$$\langle pq|rs|tu \rangle_{\text{robust}} = (pq|\tilde{r}s|\tilde{t}u) + (\tilde{p}q|rs|\tilde{t}u) + (\tilde{p}q|\tilde{r}s|tu) - 2(\tilde{p}q|\tilde{r}s|\tilde{t}u)$$ (11.42)

leads to an integral error that has no term linear in the fitting error.[272] Since each
orbital pair is fitted independently of the others, the density fitting coefficients are the same as for the two-electron case in Eq. (11.38).

In this way, in 2014, Womack and Manby were able to develop formulae for the RDF of all three-electron integrals needed in MP2-F12.[272] In their “DF3” procedure smaller auxiliary basis sets are needed, compared to traditional RI routes, to reach submillihartree errors and it is characterized by the significant advantage of requiring basis sets containing function with up to and including $2L_{\text{occ}}$ units of angular momentum. Last but not least, it is worth mentioning that in the DFT3 formalism 6-center integrals are reduced to 4-center integrals, remarkably decreasing the three-electron integrals formal scaling from $O(N^6)$ to $O(N^4)$.

Therefore, it is clear that robust density fitting is a useful alternative to RI, which can be used to enhance the efficiency and the accuracy of explicitly correlated calculations.

6.11.3 Numerical Quadrature

Numerical-quadrature schemes (QDs) for the evaluation of many-electron integrals were first proposed by Boys and Handy (1969) in the early work on the transcorrelated method.[345] The use QDs as an alternative to RI for the evaluation of many-electron integrals was introduced in explicitly correlated MP2 theory only in 2004 by Ten-no.[271] In this way three-electron integrals can evaluated as

$$\langle ijk| r_{12}^{-1} f_{13} | kml \rangle \approx \sum_g \bar{\psi}_i(r_g) \psi_k(r_g) J_{jm}^g F_{ml}^g$$

with

$$J_{pq}^g = \int \frac{\bar{\psi}_p(r) \psi_q(r)}{|r - r_g|} dr$$

$$F_{pq}^g = \int \bar{\psi}_p(r) f(|r - r_g|) \psi_q(r) dr$$

where $\bar{\psi}_i(r_g) = w_g \psi_i(r_g)$ with the quadrature weights $w_g$, and we have suppressed the integration over the spin coordinate. In this case, the formal computational cost for the elementary objects scales as $O(N^2 G)$, where $G$ is the number of quadrature points, versus the $O(N^3(N + C))$ scaling of CABS-RI, where $C$ is the size of the CAB.[255] The
number of grid points required for good accuracy in the resulting integrals is usually of the order of 10,000 per atom.[271, 339] Hence, QD is more advantageous for large molecules.

For three-electron integrals over cyclic operators such as $r_{12}^{-1}f_{13}f_{23}$, QD cannot be applied to both the electronic coordinates ‘1’ and ‘2’, due to the divergence of $r_{12}^{-1}$ at coincident quadrature points. Therefore, for such integrals and for four-electron integrals, hybrid approaches of QD/RI have been proposed.[339]
Chapter 7

General Notation and Upper Bounds

“There is surely nothing quite so useless as doing with great efficiency what should not be done at all.”


7.1 Introduction

One often reads that the most computationally demanding step in Hartree-Fock (HF) and density-functional theory (DFT) calculations is the computation of $O(N^4)$ two-electron integrals, where $N$ is the size of the basis set. However, if this were true in practice, it would be impossible to perform HF and DFT calculations with thousands of basis functions. That such calculations are now routine is a consequence of the fact that the number of significant (i.e. greater than a threshold $\tau$) integrals grows only quadratically in such systems.

Similarly, explicitly-correlated F12 calculations would be impossible if one had to generate the $O(N^6)$ three-electron and $O(N^8)$ four-electron integrals that arise in the theory. Such calculations are feasible because Kutzelnigg and Klopper proposed a path to circumvent the problem by ingenious use of resolutions.
of the identity (RI)\[252, 262\] to expand these many-electron integrals in terms of efficiently calculated\[20\] two-electron integrals. Nonetheless, RI expansions are not perfect and F12 calculations require large auxiliary basis sets\[354, 355\] to achieve chemically acceptable accuracy.\[332, 333, 356\]

Computing $O(N^6)$ or $O(N^8)$ integrals in a large system would be extremely challenging, but it turns out that the number of significant three- and four-electron integrals is, at worst, $O(N^3)$ and $O(N^4)$, respectively. Indeed, if the correlation factor is a short-range operator (as in modern F12 methods\[245, 246, 247, 248, 271, 249, 250\]), it can be shown that the number drops to only $O(N^2)$. However, to exploit this fully, one must devise rigorous upper bounds (UBs) and then use these to avoid computing vast numbers of negligible integrals. If this can be achieved, it will enable large-scale F12 calculations without the need to introduce RI approximations.

An effective UB should be:

- simple, i.e. much cheaper than the true integral;
- strong, i.e. as close as possible to the true integral;
- consistent, i.e. $N_{\text{sig}} = O(N_{\text{UB}})$, where $N_{\text{sig}}$ is the number of significant integrals and $N_{\text{UB}}$ is the number estimated by the UB.

Many two-electron integral UBs are known\[348, 349, 350, 351, 352\] but few satisfy all three requirements. Moreover, to the best of our knowledge, the only three-electron integral UB that has been proposed is a simple application of the Cauchy-Schwartz inequality.\[338\] In the present Chapter, we construct and test a variety of novel two- and three-electron integral UBs that possess all three properties.

Section 7.2 introduces our general definitions and notation, which will be adopted from here onward. Section 7.3 presents the operators that we will consider in this Chapter. Section 7.4 introduces functions that bound an entire shell (or shell-pair) of Gaussian basis functions of arbitrary angular momentum. In Section 7.5, we discuss strategies to construct UBs and to use them in screening algorithms. Section 7.6 defines bound factors, the building blocks of the UBs developed in this study, and Section 7.7 presents a detailed examination of UBs for one-, two- and three-electron integrals. In Section 7.8 we discuss the contraction problem. Finally in Section 7.9 we discuss the
implementation of the developed UBs into effective screening algorithms.

### 7.2 Gaussians, Shells and Integrals

A primitive Gaussian function (PGF)

\[
|\mathbf{a}|_i \equiv \varphi^A_{\mathbf{a}}(\mathbf{r}) = D^i_{\mathbf{a}} n^i_{\mathbf{a}} \left[ \prod_{t=x,y,z} (t - A_t)^{\alpha_t} \right] e^{-\alpha |\mathbf{r} - \mathbf{A}|^2}
\]  

(2.1)

is defined by its contraction coefficient \(D^i_{\mathbf{a}}\), normalization factor \(n^i_{\mathbf{a}}\), exponent \(\alpha_i\), center \(\mathbf{A} = (A_x, A_y, A_z)\), angular momentum vector \(\mathbf{a} = (a_x, a_y, a_z)\) and its total angular momentum \(\mathbf{a} = a_x + a_y + a_z\). We will usually suppress the primitive index \(i\).

A contracted Gaussian function (CGF)

\[
|\mathbf{a}\rangle \equiv \psi^A_{\mathbf{a}}(\mathbf{r}) = \sum_{i=1}^{K_a} |\mathbf{a}|_i
\]  

(2.2)

is a linear combination of \(K_a\) PGFs, while a CGF-pair

\[
|\mathbf{ab}\rangle \equiv \psi^A_{\mathbf{a}}(\mathbf{r}) \psi^B_{\mathbf{b}}(\mathbf{r}) = \sum_{i=1}^{K_a} \sum_{j=1}^{K_b} |\mathbf{ab}|_{ij}
\]  

(2.3)

is a double sum of PGF-pairs \(|\mathbf{ab}| = \varphi^A_{\mathbf{a}}(\mathbf{r}) \varphi^B_{\mathbf{b}}(\mathbf{r})\).

A primitive shell \(|\mathbf{a}\rangle\) is a set of PGFs sharing the same total angular momentum \(a\), exponent \(\alpha\) and center \(\mathbf{A}\). Similarly, a contracted shell \(|\mathbf{a}\rangle\) is a set of CGFs sharing the same PGFs and total angular momentum. A contracted shell-pair is the set of CGF-pairs obtained by the tensor product \(|\mathbf{ab}\rangle = |\mathbf{a}\rangle \otimes |\mathbf{b}\rangle\). Similarly, a primitive shell-pair \(|\mathbf{ab}\rangle = |\mathbf{a}\rangle \otimes |\mathbf{b}\rangle\) is the set of PGF-pairs. Finally, primitive and contracted shell-quartets are obtained in an analogous way: \(|a_1b_1a_2b_2\rangle = |a_1b_1\rangle \otimes |a_2b_2\rangle\) and \(|a_1a_2b_1b_2\rangle = |a_1b_1\rangle \otimes |a_2b_2\rangle\). For example, \(|1\rangle\) is a set of three \(p\)-type PGFs, a \(|11\rangle \equiv |pp\rangle\) shell-pair is a set of nine PGF-pairs, and a \(|2222\rangle \equiv |dddd\rangle\) shell-quartet is a set of 1296 PGF-quartets.

We write the integrals of CGFs over a one-electron operator \(e_1 = e(r_1)\), a two-electron operator \(f_{12} = f(r_{12})\) (where \(r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|\) is the interelectronic distance) or
In general, we write the integral over a three-electron operator $f_{12g13}$ or $f_{12g13}h_{23}$ as

$$
\langle a_1 | b_1 \rangle = \int \psi_{a_1}^A(r_1) e_1 \psi_{b_1}^B(r_1) dr_1,
$$

$$
\langle a_1 a_2 | b_1 b_2 \rangle = \int \int \psi_{a_1}^A(r_1) \psi_{a_2}^A(r_2) f_{12} \psi_{b_1}^B(r_1) \psi_{b_2}^B(r_2) dr_1 dr_2.
$$

$$
\langle a_1 a_2 a_3 | b_1 b_2 b_3 \rangle = \int \int \int \psi_{a_1}^A(r_1) \psi_{a_2}^A(r_2) \psi_{a_3}^A(r_3) f_{12g13} h_{23} \times
$$

$$
\psi_{b_1}^B(r_1) \psi_{b_2}^B(r_2) \psi_{b_3}^B(r_3) dr_1 dr_2 dr_3.
$$

In general, we write the integral over a $n$-electron operator $f_{1\cdots n}$ of CGFs as

$$
\langle a_1 \cdots a_n | b_1 \cdots b_n \rangle = \langle a_1 \cdots a_n | f_{1\cdots n} | b_1 \cdots b_n \rangle
$$

$$
= \int \cdots \int \psi_{a_1}^A(r_1) \cdots \psi_{a_n}^A(r_n) f_{1\cdots n} \psi_{b_1}^B(r_1) \cdots \psi_{b_n}^B(r_n) dr_1 \cdots dr_n.
$$

If an integral is over PGFs, we use square brackets, e.g.

$$
[a_1 \cdots a_n | b_1 \cdots b_n] = \int \cdots \int \varphi_{a_1}^A(r_1) \cdots \varphi_{a_n}^A(r_n) f_{1\cdots n} \varphi_{b_1}^B(r_1) \cdots \varphi_{b_n}^B(r_n) dr_1 \cdots dr_n.
$$

One-, two- and three-electron fundamentals (i.e. integrals in which all the PGFs are $s$-type functions) are

$$
[0|0] = G_1 \int \varphi_0^{Z_1}(r_1) e_1 dr_1,
$$

$$
[00|00] = G_1 G_2 \int \int \varphi_0^{Z_1}(r_1) \varphi_0^{Z_2}(r_2) f_{12} dr_1 dr_2,
$$

$$
[000|000] = G_1 G_2 G_3
$$

$$
\times \int \int \int \varphi_0^{Z_1}(r_1) \varphi_0^{Z_2}(r_2) \varphi_0^{Z_3}(r_3) f_{12g13} h_{23} dr_1 dr_2 dr_3,
$$

with $0 = (0, 0, 0)$. In general, the $n$-electron fundamental integral (FI) is defined as

$$
[0] = [0 \cdots 0|0 \cdots 0].
$$

The Gaussian product rule reduces it from $2n$ to $n$ centers:

$$
[0] = \left( \prod_{i=1}^{n} G_i \right) \int \cdots \int \varphi_0^{Z_1}(r_1) \cdots \varphi_0^{Z_n}(r_n) f_{1\cdots n} dr_1 \cdots dr_n,
$$

$$
\text{ (2.8)}
$$
where the exponents $\zeta_i$, centroids $Z_i$ and prefactors $G_i$ of the Gaussian products are

\begin{align}
\zeta_i &= \alpha_i + \beta_i, \\
Z_i &= \frac{\alpha_i A_i + \beta_i B_i}{\alpha_i + \beta_i}, \\
G_i &= D_{a_i}n_{a_i}D_{b_i}n_{b_i}\exp\left(-\frac{|AB_i|^2}{\alpha_i^{-1} + \beta_i^{-1}}\right), \\
AB_i &= A_i - B_i.
\end{align}

We also define the inter-centroid vector $Y_{ij} = Z_i - Z_j$, while more general inter-center vectors will be denoted, for example, by $A_iB_j = A_i - B_j$.

Finally, we define an integral class to be the set of all integrals that arise from a shell-pair (one-electron integrals), a shell-quartet (two-electron integrals) or a shell-sextet (three-electron integrals). For example, $\langle 222|111 \rangle$ is a class of 5832 contracted three-electron integrals.

### 7.3 Operators

We will define $f(r_{12})$ to be a short-range (SR) operator if $f$ decays exponentially or faster. Otherwise, it is a long-range (LR) operator. Generic SR and LR operators will be denoted by $S$ and $L$, respectively, and a generic operator will be denoted by $O$. In three-electron operators $f_{12}g_{13}h_{23}$, the SR operators will always precede the LR ones. So, for example, $SSL$ arises but $SLS$ does not.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$N_{\text{sig}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{12}$</td>
<td>$f_{12}g_{13}h_{23}$</td>
</tr>
<tr>
<td>$SS$</td>
<td>$SSS, SSL$</td>
</tr>
<tr>
<td>$SL$</td>
<td>$SLL$</td>
</tr>
<tr>
<td>$LL$</td>
<td>$LLL$</td>
</tr>
</tbody>
</table>

The short- or long-range character of a many-electron operator determines the number of significant integrals that it yields in a large system, as reported in Table 7.1. For this reason, it is essential to consider the range of an operator in order to obtain
strong, consistent UBs for its integrals.

Our methodology is valid for any non-negative operator that yields a bounded potential for a Gaussian, i.e.

\[ \max_{r_2} \int \varphi_0^A(r_1) f_{12} dr_1 < \infty. \] (3.10)

However, because they are ubiquitous in quantum chemistry,[357, 358, 359, 360, 361] we will focus on the LR Coulomb operator

\[ C_{12} = r_{12}^{-1}, \] (3.11)

the SR Slater, Gaussian, Erfc and Delta operators

\[ S_{12} = \exp(-\lambda r_{12}) \] (3.12)
\[ G_{12} = \exp(-\lambda^2 r_{12}^2) \] (3.13)
\[ \mathcal{E}_{12} = r_{12} \text{erfc}(\sqrt{\lambda} r_{12}) \] (3.14)
\[ D_{12} = \delta(r_{12}) \] (3.15)

and their products (e.g. $G_{12}C_{12}$).

When more than one SR operator is involved, they are assumed to have the same exponent. The only exception is for Gaussians, which are ordered by decreasing exponent. For example, in $G_{12}G_{13}G_{23}$, we have $\lambda_{12} \geq \lambda_{13} \geq \lambda_{23}$.

### 7.4 Bounding Gaussians

#### 7.4.1 Shell-bounding Gaussians

The absolute value of the normalized PGF

\[ |\alpha| = D_a \prod_{t=x,y,z} \sqrt{\frac{(2\alpha)^{a_t+1/2}}{\Gamma(a_t + 1/2)}} t^{a_t} e^{-\alpha^2} \] (4.16)
is bounded by the radial function

\[ R_a(r) = |D_a| \sqrt{\frac{(2\alpha)^{a+3/2}}{\gamma_a}} r^a e^{-\alpha r^2}, \]  

(4.17)

where

\[ \gamma_a = \prod_{i=0}^{2} \Gamma \left( \left\lfloor \frac{a+i}{3} \right\rfloor + \frac{1}{2} \right) \]  

(4.18)

and \( \Gamma(x) \) and \( \lfloor x \rfloor \) are the Gamma and floor functions. [362] By combining this with the inequality

\[ \sqrt{(2\alpha)^{a}} r^a e^{-\alpha r^2} \leq \left[ \frac{a}{\sigma e} \right]^{a/2} e^{-(1-\sigma)\alpha r^2} \]  

(4.19)

(which is true for any \( \sigma > 0 \)) one immediately finds that

\[ ||a|| \leq \hat{a}(r) \]  

(4.20)

where

\[ \hat{a}(r) = |D_a| \sqrt{\frac{(2\alpha)^{3/2}}{\gamma_a \gamma_b}} \left[ \frac{a}{\sigma e} \right]^{a} e^{-(1-\sigma)\alpha r^2} \]  

(4.21)

is an \( s \)-type Gaussian that bounds an entire shell of PGFs as tightly as possible. We call it a shell-bounding Gaussian (SBG) and the notation \( \hat{a} \) reminds us that it is spherical.

Figure 7.1 shows the radial part \( R_p(r) \) of a \( p \)-type PGF, and three SBGs \( \hat{p}(r) \) with \( \sigma = 1/3, 2/3 \) and 1. As \( \sigma \) varies, the osculation point between \( R_p(r) \) and \( \hat{p}(r) \) moves. Because they do not vanish at the origin, SBGs overestimate strongly in the core region but less in the valence region.

### 7.4.2 Shell-pair-bounding Gaussians

The product of two SBGs is

\[ \hat{z}(r) = \hat{a}(r) \hat{b}(r) = h e^{-\xi |r-Z|^2}, \]  

(4.22)

where \( \xi = (1-\sigma)\zeta \) and

\[ h = |D_aD_b| \sqrt{\frac{(4\alpha\beta)^{3/2}}{\gamma_a \gamma_b \sigma}} \left[ \frac{a}{\sigma e} \right]^{a} \left[ \frac{b}{\sigma e} \right]^{b} e^{-\frac{(1-\sigma)|AB|^2}{\alpha^1+\beta^2}} \]  

(4.23)
7.5. TYPES OF UPPER BOUND

Figure 7.1: Radial part of a \(p\)-type PGF, \(R_p(r)\), (solid line) and SBGs, \(\hat{\phi}(r)\), for various \(\sigma\) (dashed lines).

We call the \(s\)-type Gaussian \(\hat{\phi}(r)\) a shell-pair-bounding Gaussian (SPBG) because it bounds, in absolute value, an entire shell-pair, i.e.

\[
||ab|| \leq \hat{\phi}(r)
\]  \hspace{1cm} (4.24)

We will discuss an optimal choice for \(\sigma\) in Section 7.9.

7.5 Types of Upper Bound

To construct integral UBs, we will depend heavily on the absolute value inequality

\[
\left| \int \phi(r)dr \right| \leq \int |\phi(r)|dr \hspace{1cm} (5.25)
\]

and the Hölder inequality[362]

\[
\left| \int \phi_1(r)\phi_2(r)dr \right| \leq \left[ \int |\phi_1(r)|^pdr \right]^{1/p} \left[ \int |\phi_2(r)|^qdr \right]^{1/q}, \hspace{1cm} (5.26)
\]

where \(p^{-1} + q^{-1} = 1\) and \(p, q > 1\). Hölder yields the Cauchy-Schwartz inequality[362] if one chooses \(p = q = 2\).
7.5.1 Integral bounds

An integral bound is a number that bounds a particular integral. For example, the Cauchy-Schwartz inequality yields the well-known two-electron integral bound

$$|[a_1 a_2 | b_1 b_2] \leq [a_1 a_1 | b_1 b_1]^{1/2} [a_2 a_2 | b_2 b_2]^{1/2}. \quad (5.27)$$

If one has pre-computed and stored $O(N^2)$ Cauchy-Schwartz factors $[a_1 a_1 | b_1 b_1]$, these yield cheap upper bounds on $O(N^4)$ two-electron integrals.

However, despite their attractive features, integral bounds are poorly suited to modern hardware and software. Bounding every integral before deciding whether or not to compute its exact value places logical branches within inner loops and leads to slow code. Moreover, using bounds to eliminate a few integrals from a class is incompatible with recursive methods for integral generation. This leads naturally to a class strategy.

7.5.2 Class bounds

A class bound is a number that bounds all the integrals in a class. These are particularly effective for large classes because, if the class bound is below $\tau$, a large number of integrals can be skipped on the basis of one test. SBGs lead naturally to class bounds, for example,

$$|[a_1 a_2 | b_1 b_2] \leq [\hat{a}_1 \hat{a}_2 | \hat{b}_1 \hat{b}_2]. \quad (5.28)$$

Non-separable class bounds, such as (5.28), involve quantities that have the same asymptotic scaling as the integrals. Such bounds are therefore always consistent.

Separable class bounds, such as the Cauchy-Schwartz bound derived from (5.28)

$$|[a_1 a_2 | b_1 b_2] \leq [\hat{a}_1 \hat{a}_1 | \hat{b}_1 \hat{b}_1]^{1/2} [\hat{a}_2 \hat{a}_2 | \hat{b}_2 \hat{b}_2]^{1/2} \quad (5.29)$$

involve factors that may scale differently from the integrals themselves. Such bounds may not be consistent.

A specific example may be helpful. The number of significant two-electron integrals over LR and SR operators is $O(N^2)$ and $O(N)$, respectively (see Table 7.1). However, the separable bound (5.29) predicts $O(N^2)$ in both cases and is therefore inconsistent.
for SR operators. In situations when one cannot find a consistent separable bound, one should use a non-separable bound.

### 7.5.3 Shell-$m$tuple bounds

A shell-$m$tuple bound $B_m$ relies only on shell-$m$tuple information, where $m$ is the shell multiplicity: shell-pair ($m = 2$), shell-quartet ($m = 4$), shell-sextet ($m = 6$), etc. If $B_m > \tau$, it indicates that the shell-$m$tuple is significant, i.e. it could yield significant integrals.

A shell-$m$tuple bound is a class bound that depends only on the operator, the basis set and the shell multiplicity $m$, independent of the maximum shell multiplicity $n$ of the integrals. It is also consistent at its specific shell-multiplet level.

Figure 7.2 shows how to use shell-$m$tuple bounds in a primitive three-electron integral screening algorithm. In the first loop, we use a shell-pair bound $[B_2]$ to create a list of significant shell-pairs. In a large system, $O(N^2)$ work generates $O(N)$ significant shell-pairs. In the second loop, we use a shell-quartet bound $[B_4]$ to create a list of significant shell-quartets. Again, in a large system, $O(N^2)$ work generates, at worst, $O(N^2)$ significant shell-quartets. In the third loop, we combine the significant shell-quartets and shell-pairs and use a shell-sextet bound $[B_6]$ to identify significant shell-sextets. In the worst case, this involves $O(N^3)$ work generates $O(N^3)$ significant three-electron integrals. In this way, independent of the operator, shell-$m$tuple bounds reduce the work from $O(N^6)$ to $O(N^3)$ for three-electron integrals.

### 7.6 Intermediates

The use of the Hölder inequality yields intermediate quantities that result from integrating products of SPBGs and operators and then, possibly, maximizing. The quantities are conveniently represented using a square bracket notation $[x]^y$ with the following conventions:

- A digit $n$ inside the brackets indicates the SPBG $\hat{z}_n(r_n)$ and integration over $r_n$. 

146
1: $N_2 = 0$
2: for shells $|a_1| = 1, \ldots, N$ do
3:    for shells $|b_1| = 1, \ldots, N$ do
4:       if $[B_2] > \tau$ then
5:          $N_2 = N_2 + 1$
6:       end if
7:    end for
8: end for
9: This generates $N_2 = \mathcal{O}(N)$ significant shell-pairs
10: $N_4 = 0$
11: for shell-pairs $|a_1 b_1| = 1, \ldots, N_2$ do
12:    for shell-pairs $|a_2 b_2| = 1, \ldots, N_2$ do
13:       if $[B_4] > \tau$ then
14:          $N_4 = N_4 + 1$
15:       end if
16:    end for
17: end for
18: This generates, at worst, $N_4 = \mathcal{O}(N^2)$ significant shell-quartets
19: for shell-quartets $|a_1 b_1 a_2 b_2| = 1, \ldots, N_4$ do
20:    for shell-pairs $|a_3 b_3| = 1, \ldots, N_2$ do
21:       if $[B_6] > \tau$ then
22:          Compute integral class $[a_1 a_2 a_3 | b_1 b_2 b_3]$
23:       end if
24:    end for
25: end for
26: This generates, at worst, $\mathcal{O}(N^3)$ significant three-electron integrals

Figure 7.2: An $\mathcal{O}(N^3)$ algorithm for computing three-electron integrals using shell-mtuplet bounds.
7.7 Upper Bounds

\[ [1] = \int \hat{z}_1 d \mathbf{r}_1 \] 
\[ [12] = \int \int \hat{z}_1 \hat{z}_2 d \mathbf{r}_1 d \mathbf{r}_2 \]

- A superscript indicates a two-electron operator, e.g.

\[ [12]^f = \int \int \hat{z}_1 \hat{z}_2 f_{12} d \mathbf{r}_1 d \mathbf{r}_2 \] 
\[ [123]^{fh} = \int \int \int \hat{z}_1 \hat{z}_2 \hat{z}_3 f_{12} g_{13} h_{23} d \mathbf{r}_1 d \mathbf{r}_2 d \mathbf{r}_3 \]

- We maximize over all coordinates that are in the operators but not in the brackets, e.g.

\[ [1]^f = \max_{r_2} \int \hat{z}_1 f_{12} d \mathbf{r}_1 \] 
\[ [1]^g = \max_{(r_2, r_3)} \int \hat{z}_1 f_{12} g_{13} d \mathbf{r}_1 \] 
\[ [12]^g = \max_{r_3} \int \int \hat{z}_1 \hat{z}_2 f_{12} g_{13} d \mathbf{r}_1 d \mathbf{r}_2 \]

Furthermore, we define \([\hat{1}]^f g\) as the largest \([1]^f g\) factor within a given system and basis set. This implies that

\[ [1]^f g \leq [\hat{1}]^f g. \] 

Since SPBGs are spherically symmetric, \([\hat{1}]^f g\) are basis set dependent, but do not depend on the geometry of the system. A detailed procedure for the computation of the \([\hat{1}]^f g\) factors is presented in Appendix B. Explicit expressions for the remaining intermediates necessary to construct our UBs for various operators can be found in Table 7.2.

7.7 Upper Bounds

In the first part of this section we derive upper bounds for one-, two- and three-electron integrals. We eschewed bounds descending from the Cauchy-Schwartz inequality because they are usually weaker than ours[352] and, for three-electron integrals, they are usually
not simple.

In the second part we discuss their performances for various operators.

7.7.1 Theory

For one-, two- and three-electron integrals, the derivation of our class bounds is described respectively by the diagrams in Figs. 7.3, 7.4 and 7.5.

The bounds are numbered as follows. The first digit indicates the number of electrons, the second the shell-\( m \)-tuplet level (\( m = 6 \) for shell-sextet bounds, \( m = 4 \) for shell-quartet and \( m = 2 \) for shell-pair bounds). The third digit is used to discern bounds that have the same shell-\( m \)-tuplet level and for integrals with the same number of electrons. Thus, 3.6.2 identifies the shell-sextet bound number 2 for three-electron integrals. A shorthand symbol for each bound is reported in the leftmost column of each diagram.

Note that even if there are usually several bounds for each level, we screen the
shell-\textit{m}-tuplet with the cheapest consistent bound. Therefore, their choice depends on the nature of the operator as reported in Table 7.3.

As sketched in Figs. 7.3, 7.4 and 7.5, all the UBs are derived first by using the SPBGs and then, whenever necessary, by repeated applications either of Hölder’s inequality or of Eq. (6.33).

The operators $P_1$ and $P^h$ are only used for sake of conciseness. $P_1$ permutes electron 1 with itself and the other electrons. For example,


In contrast, $P^h$ permutes operator $h$ with itself and the other operators, that is

Symbol

\[ [a_1a_2a_3|fgh|b_1b_2b_3] \]

Level

SPBGs

\[ [B_{6,3}]^{fgh} \]


\[ [B_{6,2}]^{fgh} \]

\[ \min \{ [12]^{fgh}[3], [12]^{fh}[3]^g, [12]^{f}[3]^{gh} \} \]

Hölder

\[ [B_{6,1}]^{fgh} \]

\[ \min \{ [12]^{fgh}[3], [12]^{fh}[3]^g, [12]^{f}[3]^{gh} \} \]

\[ [B_{4,2}]^{fgh} \]

\[ \min \{ [12]^{fgh}[3], [12]^{fh}[3]^g, [12]^{f}[3]^{gh} \} \]

Hölder

\[ [B_{4,1}]^{fgh} \]

\[ \min \{ [12]^{fgh}[3], [12]^{fh}[3]^g, [12]^{f}[3]^{gh} \} \]

\[ [B_{2,2}]^{fgh} \]

\[ \min \{ [12]^{fgh}[3], [12]^{fh}[3]^g, [12]^{f}[3]^{gh} \} \]

\[ [B_{2,1}]^{fgh} \]

\[ \max \mathcal{P}^h \{ [12]^{fgh}[3], [12]^{fh}[3]^g, [12]^{f}[3]^{gh} \} \]

\[ \mathcal{P}^h \] gives rise to bounds which are symmetrized with respect to the exchange of electrons.

In fact, since the operator \( f_{12}g_{13}h_{23} \) is, in general, not symmetric with respect to the exchange of two electrons, \([363]\) such symmetrization becomes necessary for shell-pair
level bounds for three-electron integrals in order to obtain a unique shell-pair list.

Table 7.3: Selected level of shell-ntuplet bounds for two- and three-electron integrals based on the nature of the operator.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Shell-ntuplet bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$m = 2$</td>
</tr>
<tr>
<td>$S$</td>
<td>2.2.1</td>
</tr>
<tr>
<td>$L$</td>
<td>2.2.1</td>
</tr>
<tr>
<td>$SS$</td>
<td>3.2.1</td>
</tr>
<tr>
<td>$SL$</td>
<td>3.2.1</td>
</tr>
<tr>
<td>$LL$</td>
<td>3.2.1</td>
</tr>
<tr>
<td>$SSS$</td>
<td>3.2.1</td>
</tr>
<tr>
<td>$SSL$</td>
<td>3.2.1</td>
</tr>
<tr>
<td>$SLL$</td>
<td>3.2.1</td>
</tr>
<tr>
<td>$LLL$</td>
<td>3.2.1</td>
</tr>
</tbody>
</table>

This lack of symmetry in three electron integrals also gives rise to three different kinds of shell-quartet bounds, that is one for each of the distinct shell-quartets “12”, “13” and “23”. However, unlike for shell-pairs, because i) shell-quartet factors are much more expensive than shell-pair ones, and ii) shell-quartets will never be coupled with themselves, the most efficient procedure is to derive the shell-quartet bound associated with the shell-quartet yielding the shortest list.

In particular, shell-quartets interacting through SR operators will always give shorter lists than the ones interacting through LR ones. Thus, following the conventions of Section 7.3 and without loss of generality, the shell-quarted bounds in Fig. 7.5 are designed for constructing the “12” shell-quartet list, as is it will always be the shortest one.

7.7.2 Performance

One-electron integrals

Undoubtedly the most common one-electron integrals are the overlap integrals ($e = 1$). Additionally, they offer the chance of testing the effectiveness of the SPBGs. Figure 7.6 shows the performance of bound $[B_{2,1}]$ in Fig. 7.3 for various classes of overlap integrals, using a wide range of exponents and distances between centers. We find that our UB rarely overestimates by more than two orders of magnitude and that classes including $p$
shells are the most difficult to bound tightly. We also find that the performance of the UB improves as the total angular momentum of the class increases. This is due to the fact that, since, for same exponent, due to the effect of their polynomial parts, higher momenta GFs tend to be more diffuse than lower angular momenta ones, integrals over these functions are small when the basis functions are centered so far from each other that their tails overlap negligibly. SBGs are much better at reproducing the correct behaviour of the valence region of the GFs than their core structure, hence the observed improved performance.

Two-electron integrals

As reported in Table 7.3, for two-electron integrals over LR operators the separable bound $[B_{4,1}]^L$ is consistent. The upper leftmost graph in Fig. 7.7 illustrates the performance of $[B_{4,1}]^C$ for $\langle 00|C|00 \rangle$ classes. * Similarly to the overlap case discussed above (see Fig. 7.6), $[B_{4,1}]^C$ faithfully estimates the value of two-electron Coulomb integrals in a wide range of situations ($10^{-3} \leq \alpha_i, \beta_i \leq 10^{-3}$, $|AB_1| \leq 15$ and $|Y_{12}| \leq 15$). Except for several outliers, $\langle 00|C|00 \rangle$ is never overestimated by more than two orders of magnitude. This clearly evidences the all-round robustness of the present bound. We note that $[B_{4,1}]^C$ is stronger than the Cauchy-Schwartz-based bound used in most quantum chemistry packages.[352, 20]

For SR two-electron operators, the only consistent bound at the shell-quartet level is $[B_{4,2}]^S$. Because SBGs are $s$-type Gaussians, it reduces the problem to the computation of a single fundamental integral $[12]^S$ to bound the entire class of two-electron integrals.

For $S = \mathcal{G}$, the fundamental integral is known in closed form.[257, 258] The performance of $[B_{4,2}]^\mathcal{G}$ depends on the exponent $\lambda$. The worst case is

$$\lim_{\lambda \to \infty} \sqrt{\frac{\pi}{\lambda}} [\hat{a}_1 \hat{a}_2|G_{12}|\hat{b}_1 \hat{b}_2] = [\hat{a}_1 \hat{a}_2|D_{12}|\hat{b}_1 \hat{b}_2].$$  (7.36)

Thus, the problem reduces to the one-electron overlap case previously studied.

For $S = \mathcal{E}$, the fundamental integral can be obtained in closed form. However, it is computationally cheaper to follow the SBG philosophy (see Section 7.4) and bound $E_{12}$

*We do not consider higher angular momentum classes because they are bound by $s$-type SPBGs (see Fig. 7.4).
Figure 7.6: Performance of $[B_{2.1}]^1$ for $[s|p]$, $[s|d]$, $[p|p]$, $[p|d]$ and $[d|d]$ classes of overlap integrals, where $|\mathbf{A}\mathbf{B}| \leq 15$ and $10^{-3} \leq \alpha, \beta \leq 10^{-3}$. The dashed line indicates overestimation by two orders of magnitude.
Figure 7.7: Performance of shell-quad bound \([B_{4,1}]^C\), \([B_{4,2}]^C\) and \([B_{4,2}]^S\) for \([00\langle 00\rangle]\) classes, where \(10^{-3} \leq \alpha_i, \beta_i \leq 10^{+3}, |AB_4| \leq 15, |Y_{12}| \leq 15\) and \(\lambda = 1\). The dashed line indicates overestimation by two orders of magnitude.

with a Gaussian geminal:

\[
\sqrt{\pi \lambda} \frac{\xi_{12}}{g_{12}} \leq 1. \tag{7.37}
\]

The ratio (7.37) tends to zero when \(r_{12} \to 0\). However, as \(r_{12}\) becomes larger, it quickly reaches values close to unity. This perfect asymptotic behaviour reflects the high accuracy of this UB as shown in Fig. 7.7 where we observe a large density of points near the diagonal, and that the error is always smaller than one order of magnitude.

For \(S = S\), the closed-form expression of \([12]^S\) is computationally too expensive for being used as an UB. [248] Instead we upper bound it using the inequalities reported in Appendix A (Eq. (11.48b)). The lower graph in Fig. 7.7 illustrates the performance of
[B_{4,2}]^G, which is similar to [B_{4,1}]^C.

**Three-electron integrals.**

Of particular interest amongst the three-electron integrals are those over operators of the form SSO, where O can be either LR or a unitary operator, and those over SL, as both kinds arise in F12 methods.\cite{253, 252}

Like for two-electron integrals, the problem of upper-bounding an entire class of three-electron integrals over SSO is reduced, by using the UB [B_{6,3}]^{SSO}, to the computation of a fundamental integral.

When SSO = \{GG, GGC\}, the fundamental integral is known in closed form and is computationally convenient.\cite{257, 258, 245}

For SSO = \{EE, EEC\} a very effective UB for [123]^{SSO} is easily obtained using Eq. (7.37). For SSO = \{SS, SSC\}, consistent UBs have been derived in Appendix A.

For integrals over SL operators, the partially factorized bound [B_{6,2}]^{SL} can be adopted. Whenever SL = \{GC, EC, SC\}, the intermediates necessary for computing [B_{6,2}]^{SL} are reported in Table 7.2. In the case that S = E or S = S these intermediated have been obtained by using either Eq. (7.37) or Eq. (11.48b), respectively. Figure 7.8 illustrates the performance of [B_{6,2}]^{GC}, [B_{6,2}]^{EC} and [B_{6,2}]^{SC} for [000|000] classes. It is particularly important to have a strong UB for these three-electron integrals as they scale quadratically with the system size. From the three shell-sextet bound, [B_{6,2}]^{GC} has been found to be the strongest bound, while [B_{6,2}]^{EC} shows a very similar behavior (as inferred by Eq. (7.37)). Albeit slightly less reliable, the performance of [B_{6,2}]^{SC} is still fairly decent with most of the integrals estimated within a factor of 10^{3/2}.

**7.8 The Contraction Problem**

Suppose that, from the primitive bounds derived in Section 7.7, we want to obtain class bounds for contracted integrals. If each of the CGFs is K-fold contracted, i.e. is a sum of K PGFs (see Eq. (2.2)), a straightforward O(K^2) method to obtain shell-pair contracted bounds is to sum over all its primitive factors. For example

\[
\langle B_{2,1} \rangle^f = \min \left\{ (1)^f \langle \hat{2} \rangle, \langle 1 \rangle \langle \hat{2} \rangle^f \right\},
\] (8.38)

156
with

\[ \langle 1 \rangle^g = \sum_{ij}^{K} |1\rangle_{ij}^{fg}, \quad \text{(8.39)} \]

where \( i \) and \( j \) refer to the PGFs \( |a_i| \) and \( |b_j| \) in the contracted shells \( |a_1| \) and \( |b_1| \), respectively (see Eq. (2.3)). Note that \( \langle 2 \rangle^g \) is the maximum of the contracted factors \( \langle 2 \rangle^g \) yielded by the basis set and their computation is discussed in Appendix B.

Although this idea is very effective at the shell-pair level, it is not suitable for
non-separable shell-quartet and shell-sixtet bounds. For example,

\[ | \langle a_1 a_2 | f | b_1 b_2 \rangle | \leq \sum_{ijkl}^{K} [12ij]_ijkl \]  

(8.40)

requires \( O(K^4) \) computational work. A straightforward extension of (8.40) would lead to \( O(K^6) \) work for three-electron integrals!

To avoid these expensive contraction steps we always contract at the shell-pair level. A detailed algorithm for performing such contraction will be presented in Section 7.9.

The expressions of the contracted bounds are identical to the primitive bounds (see Figs. 7.3, 7.4 and 7.5), except that contracted factors are used. However, the non-separable contracted factors \( \langle 12 \rangle^{fg} \) and \( \langle 123 \rangle^{fg} \) are bound by

\[
\langle 12 \rangle^{fg} \leq \min \left\{ \langle 1 \rangle^{fg} \langle 2 \rangle^{h}, \langle 1 \rangle^{g} \langle 2 \rangle^{f} \right\} \langle 12 \rangle^{fg},
\]  

(8.41a)

\[
\langle 123 \rangle^{fg} \leq \langle 1 \rangle \min \left\{ \langle 2 \rangle^{fh} \langle 3 \rangle^{f}, \langle 2 \rangle^{f} \langle 3 \rangle^{gh} \right\} \langle 123 \rangle^{fg}.
\]  

(8.41b)

The expression in curly braces are derived by using Hölder’s inequality, while the terms \( \langle 12 \rangle^{fg} \) and \( \langle 123 \rangle^{fg} \) are upper bounds of the ratio of the contracted factor \( \langle 123 \rangle^{fg} \) and its Hölder upper bound, that is

\[
\langle 12 \rangle^{fg} \geq \min \left\{ \langle 1 \rangle^{fg} \langle 2 \rangle^{h}, \langle 1 \rangle^{g} \langle 2 \rangle^{f} \right\}
\]  

(8.42a)

\[
\langle 123 \rangle^{fg} \geq \langle 1 \rangle \min \left\{ \langle 2 \rangle^{fh} \langle 3 \rangle^{f}, \langle 1 \rangle \langle 2 \rangle^{f} \langle 3 \rangle^{gh} \right\}
\]  

(8.42b)

and bring the interelectronic coupling. The expression of \( \langle 12 \rangle^{fg} \) and \( \langle 123 \rangle^{fg} \) are given in the two last columns of Table 7.2 for various operators. These are functions of the largest \( \langle \xi_i \rangle \) and/or the smallest \( \langle \xi_i \rangle \) effective exponent in the contracted shell-pair \( |a_i b_i| \), and of the distances

\[ R_{ij} = \max \left\{ 0, \left| \frac{AB_i^+}{2} - \frac{AB_j^+}{2} \right| \right\}
\]  

(8.43)

between two spheres of diameters \( AB_i \) and \( AB_j \) (where \( AB_i^+ = A_i + B_i \)).
CHAPTER 7. GENERAL NOTATION AND UPPER BOUNDS

Algorithm 1 Screening algorithm for primitive one-electron integrals

1: procedure SHELL-PAIR SCREENING
2: for shells \( |a_1| = 1, \ldots, N \) do
3: for shells \( |b_1| = 1, \ldots, N \) do
4: Solve \( \frac{\partial \mathcal{B}_{2,1}}{\partial \sigma_1} = 0 \) to get optimal \( \sigma_1^* \)
5: Compute \( [1]^e \) using \( \sigma_1^* \)
6: if \( [B_{2,1}]^e \geq \tau \) then
7: Compute primitive class \( [a_1|e_1|b_1] \)
8: end if
9: end for
10: end for
11: end procedure

7.9 Algorithms

In the preceding sections we have derived UBs for one-, two- and three-electron integrals, shown that they are efficient and extended them to contracted integrals. In this last Section we discuss how to use these UBs in screening algorithms.

As usual, we begin with the simplest example of primitive one-electron integrals. As sketched in Algorithm 1, we first compute \( \sigma_1^* \) which minimizes the bound \( [B_{2,1}]^e = [1]^e \) itself by solving \( \frac{\partial [1]^e}{\partial \sigma_1} = 0 \). This enables to fully exploit the flexibility of the SPBGs. If \( [B_{2,1}]^e \) is greater or equal to the user-defined threshold \( \tau \), the integral class \( [a_1|e_1|b_1] \) is computed.

Algorithm 2 shows a screening algorithm for primitive two-electron integrals. First, a coarse-grained shell-pair screening is performed with the bound factors \( [1]_0, [1]^f_0 \) computed using

\[
\sigma_1^0 = \frac{a_1 + b_1}{(a_1 + b_1 + 3) + \frac{2|AB|}{a_1^* + \beta_1^*}}
\] (9.44)

which is obtained by approximately solving \( \frac{\partial [1]^e}{\partial \sigma_1} \approx 0 \). The shell-pairs that survive to this coarse-grained screening undergo a fine-grained screening, where the optimal values of the shielding parameter \( \sigma_1 \) and \( \sigma_1^f \) are computed and used to evaluate \( [1] \) and \( [1]^f \), respectively. We point out that the coarse-grained screening performed on all shell-pairs is much cheaper than the fine-grained one for two reasons: i) only a single \( \sigma_1^0 \) is computed and used for all factors, ii) this value of the shielding parameter is computationally cheaper than the optimal one.

Independently of the nature of the operator \( f \), the shell-pair screening generates, in a large system, \( N_2 = \mathcal{O}(N) \) significant shell-pairs. These shell-pairs are then coupled
Algorithm 2 Screening algorithm for primitive two-electron integrals

1: procedure Shell-Pair Screening
2: Compute \( |\mathcal{b}| \) and \( |\mathcal{b}| \)
3: \( N_2 = 0 \)
4: for shells \( |a_1| = 1, \ldots, N \) do
5: for shells \( |b_1| = 1, \ldots, N \) do
6: \( \triangleright \) Coarse-Grained Screening
7: Solve \( \frac{\partial |1\rangle}{\partial \mathcal{a}} = 0 \) (Eq. (9.44)) to get \( \sigma_0^0, |1\rangle, |1\rangle \)
8: if \( |B_2,1| \geq \tau \) then
9: \( \triangleright \) Fine-Grained Screening
10: Solve \( \frac{\partial |\mathcal{b})}{\partial \mathcal{a}} = 0 \) to get \( |1\rangle \) and \( |1\rangle \)
11: if \( |B_2,1| \geq \tau \) then
12: \( N_2 = N_2 + 1 \)
13: Store shell-pair information
14: \( (\text{e.g. } |1\rangle, |1\rangle) \)
15: end if
16: end if
17: end for
18: end for
19: end procedure
20: This yields \( N_2 = \mathcal{O}(N) \) significant shell-pairs
21: procedure Shell-Quartet Screening
22: \( N_4 = 0 \)
23: for shell-pairs \( |a_1,b_1| = 1, \ldots, N_2 \) do
24: for shells \( |a_2,b_2| = 1, \ldots, N_2 \) do
25: switch \( f \) do
26: case \( f = L \)
27: if \( |B_4,1| \geq \tau \) then
28: \( N_4 = N_4 + 1 \)
29: Compute class \( |a_1,a_2| |b_1,b_2| \)
30: end if
31: case \( f = S \)
32: Compute \( |2\rangle |\mathcal{b}) \) using \( \sigma_0^f \) and \( \sigma_0^f \)
33: if \( |B_4,1| \geq \tau \) then
34: \( N_4 = N_4 + 1 \)
35: Compute class \( |a_1,a_2| |b_1,b_2| \)
36: end if
37: end for
38: end for
39: end procedure
40: \( \triangleright \) For \( f = L \) this yields \( N_4 = \mathcal{O}(N^2) \) significant integrals
41: \( \triangleright \) For \( f = S \) this yields \( N_4 = \mathcal{O}(N) \) significant integrals
together to form shell-quartets, which are screened by exploiting different bounds depending on \( f \). If \( f \) is a LR operator then \([B_{4,1}]^f\) is used to generate the significant \( \mathcal{O}(N^2) \) integral. On the contrary, if \( f \) is a SR operator the bound \([B_{4,2}]^f\) yields only \( \mathcal{O}(N) \) significant integrals. Note that if \( f = S \) only the \( \sigma_1^f \) and \( \sigma_2^f \) are used to evaluate the optimal bound factor \([12]^f\).

In the screening algorithm for primitive three-electron integrals, which is sketched in Algorithm 3, shell-pair and shell-quartet screenings are performed along the same lines as for two-electron integrals. Then, the significant shell-pairs and significant shell-quartets are coupled together to form shell-sextets, which are screened with different consistent bound depending on the nature of the operator \( fgh \) (see Table 7.3).

Since contracted shell-quartet and -sextet screenings are straightforward generalization of the primitive ones, the last algorithm that we discuss is the contracted shell-pair screening for three-electron integrals (Algorithm 4). This is by far the most important algorithm for contracted integrals because, as anticipated in Sec. 7.8, in order to avoid expensive contraction steps the contraction is always performed at the shell-pair level. As shown in Algorithm 4, first a coarse-grained screening of the primitive shell pairs \(|a_1 b_1|_{ij}\) arising from the primitives in the contracted shell-pair \(|a_1 b_1\rangle\) is performed by using the non-optimal bound \([B_{2,1}]^{fgh}_{ij}\). The primitive shell-pairs \(|a_1 b_1|_{ij}\) that survive undergo a primitive fine-grained screening, where the optimal values of \( \sigma_{1ij} \) are computed and used to evaluate the optimal primitive level bound \([B_{2,1}]^{fgh}_{ij}\). If \([B_{2,1}]^{fgh}_{ij} \geq \tau\), the optimal primitive factors arising from the specific shell-pair \(|a_1 b_1|_{ij}\) (for example \([1]_{ij}^{fgh}\)) are included in the contraction (for the previous example \(\langle 1 \rangle^{fgh} = \langle 1 \rangle^{fg} + \langle 1 \rangle_{ij}^{fgh}\)). Thus, all the contracted quantities necessary for contracted shell-pair, shell-quartet and shell-sextet bounds (such as \(\xi_1, \xi_2, \xi_3, \hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3\), \(\langle 1 \rangle^{fgh}\)) are constructed in the primitive fine-grained loop. Finally, a contracted fine-grained screening is performed using the optimal \(\langle B_{2,1} \rangle^{fgh}\), yielding \(\mathcal{O}(N)\) significant contracted shell-pairs.

### 7.10 Concluding Remarks

In this Chapter, we have introduced some general notation that will be used throughout the rest of this Part B. Furthermore, we have constructed novel UBs for one-, two- and three-electron integrals, and studied their efficiency. All these bounds are based on a
**Algorithm 3** Screening algorithm for primitive three-electron integrals

1: `procedure SHELL-PAIR SCREENING`
2: Compute $|\{f\}|$, $|\{g\}|$, $|\{h\}|$, $|\{i\}|$, $|\{j\}|$, etc.
3: `for shells $|\{a\}| = 1, \ldots, N$ do`
4: `for shells $|\{b\}| = 1, \ldots, N$ do`
5: `if $|\{b\}| \geq \tau$ then`
6: `if $|\{a\}| \geq \tau$ then`
7: `if $|\{b\}| \geq \tau$ then`
8: `if $|\{a\}| \geq \tau$ then`
9: `for shells $|\{a\}| \geq \tau$ and $|\{b\}| \geq \tau$ do`
10: `Compute $\sigma_1^a$, $\sigma_1^b$, etc. and bound factors`
11: `if $|\{a\}| \geq \tau$ then`
12: `if $|\{b\}| \geq \tau$ then`
13: `if $|\{a\}| \geq \tau$ then`
14: `if $|\{b\}| \geq \tau$ then`
15: `for shells $|\{a\}| \geq \tau$ and $|\{b\}| \geq \tau$ do`
16: `Compute new intermediates for $|\{a\}| \geq \tau$ and $|\{b\}| \geq \tau$`
17: `end for`
18: `end for`
19: `end procedure`
20: `for shell-pairs $|\{a\}|, |\{b\}| = 1, \ldots, N_4$ do`
21: `for shells $|\{a\}| = 1, \ldots, N_4$ do`
22: `for shells $|\{b\}| = 1, \ldots, N_4$ do`
23: `if $|\{a\}| \geq \tau$ then`
24: `if $|\{b\}| \geq \tau$ then`
25: `if $|\{a\}| \geq \tau$ then`
26: `if $|\{b\}| \geq \tau$ then`
27: `for shells $|\{a\}| \geq \tau$ and $|\{b\}| \geq \tau$ do`
28: `Compute $\sigma_1^a$, $\sigma_1^b$, etc. and bound factors`
29: `if $|\{a\}| \geq \tau$ then`
30: `if $|\{b\}| \geq \tau$ then`
31: `if $|\{a\}| \geq \tau$ then`
32: `if $|\{b\}| \geq \tau$ then`
33: `for shells $|\{a\}| \geq \tau$ and $|\{b\}| \geq \tau$ do`
34: `Compute new intermediates for $|\{a\}| \geq \tau$ and $|\{b\}| \geq \tau$`
35: `end for`
36: `end for`
37: `end procedure`

---

**Algorithm 3** Screening algorithm for primitive three-electron integrals (continued)

50: `procedure SHELL-SEXTET SCREENING`
51: $N_4 = 0$
52: for shell-pairs $|\{a\}|, |\{b\}| = 1, \ldots, N_4$ do
53: for shells $|\{a\}| = 1, \ldots, N_4$ do
54: `switch $|\{a\}|, |\{b\}|$`
55: `case $|\{a\}|, |\{b\}| = 1, \ldots, N_4$ do`
56: `Compute new intermediates for $|\{a\}|, |\{b\}| = 1, \ldots, N_4$`
57: `end if`
58: `end for`
59: `end for`
60: `end procedure`

---

7.10. CONCLUDING REMARKS
Algorithm 4: Algorithm for screening contracted shell-pairs of three-electron integrals.

1: procedure CONTRAICTED SHELL-PAIR SCREENING
2: Compute \((\tilde{\mathcal{S}})^f, (\tilde{\mathcal{H}})^h, (\tilde{\mathcal{S}})^h, (\tilde{\mathcal{H}})^f, (\tilde{\mathcal{S}})^{gh}\) and \((\tilde{\mathcal{H}})^{gh}\)
3: \(N_2 = 0\)
4: for contracted shells \([a_1] = 1, \ldots, N\) do
5: for contracted shells \([b_1] = 1, \ldots, N\) do
6: \(K = 0, \xi_1 = 10^{-12}, \xi_1 = 0\)
7: Set to zero contracted bound factors
8: required for \((\tilde{\mathcal{S}})^{gh}\) and \((\tilde{\mathcal{H}})^{gh}\)
9: (e.g. \(\xi_1 = 0, (\xi_1) = 0, (\xi_1) = 0\))
10: Compute maximum primitive factors for
11: primitives in shell pair \([a_1 b_1]\) (e.g. \([\tilde{1}]^{(f)}\))
12: for primitive shells \([a_1, b_1] = 1, \ldots, K_{ab}\) do
13: for primitive shells \([b_1] = 1, \ldots, K_{ab}\) do
14: \(\triangleright\) Primitive Coarse-Grained Screening
15: Solve \(\alpha^{(\tilde{\mathcal{S}})^{gh}} \approx 0\) (Eq. (9.44)) to get \(\sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}}\)
16: Compute initial guess bound factors
17: using \(\sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}}\) (e.g. \([\tilde{1}]^{(f)}_{11,0}\), \([\tilde{1}]^{(f)}_{11,0}\), etc.)
18: if \(\|B_{2,1}\|^{gh} \geq \tau\) then
19: \(\triangleright\) Primitive Fine-Grained Screening
20: Compute optimal \(\sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}}\) and \([\tilde{1}]^{(f)}_{11,0}\)
21: Compute optimal exponents
22: \(\xi_{11,11}^{(\tilde{\mathcal{S}})^{gh}} = (1 - \sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}})\xi_{11,11}\)
23: \(\xi_{11,11}^{(\tilde{\mathcal{H}})^{gh}} = (1 - \sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}})\xi_{11,11}\)
24: \(\delta_{11,11}^{(\tilde{\mathcal{S}})^{gh}} = (1 - \sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}})\delta_{11,11}\)
25: \(\delta_{11,11}^{(\tilde{\mathcal{H}})^{gh}} = (1 - \sigma_{11,11}^{(\tilde{\mathcal{S}})^{gh}})\delta_{11,11}\)
26: if \(\|B_{2,1}\|^{gh} \geq \tau\) then
27: \(K = K + 1\)
28: if \(a, b \not\in \{LL, LLL\}\) then
29: \(\xi_1 = \min(\xi_{11,11}^{(\tilde{\mathcal{S}})^{gh}}, \xi_3); \xi_2 = \max(\xi_{11,11}^{(\tilde{\mathcal{S}})^{gh}}, \xi_3); \xi_3 = \max(\xi_{11,11}^{(\tilde{\mathcal{S}})^{gh}}, \xi_3)\)
30: \(\xi_1 = \min(\xi_{11,11}^{(\tilde{\mathcal{H}})^{gh}}, \xi_3); \xi_2 = \max(\xi_{11,11}^{(\tilde{\mathcal{H}})^{gh}}, \xi_3); \xi_3 = \max(\xi_{11,11}^{(\tilde{\mathcal{H}})^{gh}}, \xi_3)\)
31: if \(\|B_{2,1}\|^{gh} \geq \tau\) then
32: Increment primitive factors
33: end if
34: \(\langle 1, f, s + |1, f, s \rangle\)
35: \(\langle 1, f, s + |1, f, s \rangle\)
36: \(\langle 1, f, s + |1, f, s \rangle\)
37: \(\langle 1, f, s + |1, f, s \rangle\)
38: \(\langle 1, f, s + |1, f, s \rangle\)
39: \(\langle 1, f, s + |1, f, s \rangle\)
40: \(\langle 1, f, s + |1, f, s \rangle\)
41: \(\langle 1, f, s + |1, f, s \rangle\)
42: \(\langle 1, f, s + |1, f, s \rangle\)
43: \(\langle 1, f, s + |1, f, s \rangle\)
44: \(\langle 1, f, s + |1, f, s \rangle\)
45: \(\langle 1, f, s + |1, f, s \rangle\)
46: \(\langle 1, f, s + |1, f, s \rangle\)
47: \(\langle 1, f, s + |1, f, s \rangle\)
48: \(\langle 1, f, s + |1, f, s \rangle\)
49: \(\langle 1, f, s + |1, f, s \rangle\)
50: end if
51: end for
52: end while
53: This yields \(N_2 = O(N)\) significant contracted shell-pairs

new theoretical tool called bounding Gaussians and comply with the following three requirements: they are strong, consistent and simple. We have also shown how the bounds can be easily inserted in integral screening algorithms. Our approach can be
extended to four-electron integrals, and we will report on this in Chapter 11.

7.11 Appendix A: Bounding Slater Geminal Integrals

Many-electron integrals over Slater geminals $S_{12}$ are known to be troublesome. Their UBs are no exception. One of the reasons for this additional complexity can be attributed to their intrinsic inhomogeneity when coupled with PGFs. In this Appendix we introduce some useful inequalities to derive UBs involving $S_{12}$.

7.11.1 Upper bound for the product of two Slater functions

Let us consider the product of two Slater functions with same exponent $\lambda$. Assuming that they are separated by a distance $2u$, we define the origin as the mid-point on the line joining their centers. Therefore, the following piecewise UB is valid:

$$e^{-\lambda|r+u|}e^{-\lambda|r-u|} \leq \begin{cases} e^{-2\lambda u}, & 0 \leq r \leq u, \\ e^{-2\lambda r}, & \text{otherwise}. \end{cases}$$

(11.45)

This bound is exact on the line joining their centers and is asymptotically exact for large $r$.

7.11.2 Primitive factors

We bound the potential of a SPBG with $S_{12}$ as

$$\int \hat{z}_1(r_1)S_{12}dr_1 \leq [1]^2 \Xi_{\lambda}(|r_2 - Z_1|, \xi_1),$$

(11.46)

where the continuous function

$$\Xi_{\lambda}(r, \xi) = \begin{cases} e^{-\xi r^2}, & r \leq \lambda/(2\xi), \\ e^{-\lambda r + \lambda^2/(4\xi)}, & \text{otherwise}, \end{cases}$$

(11.47)
has been obtained by matching the exact behaviour of the potential in Eq. (11.46) for large \( r_2 \) with a tight UB at small \( r_2 \). Using the same procedure, we have

\[
\int \hat{\varepsilon}_1(r_1)S_{12}L_{13}dr_1 \leq [1]^{SL}\xi_\lambda(|r_2 - Z_1|, \xi_1),
\]

\[
[12]^{SL} \leq \min \{[1]^{SL}[2], [1]^{L}[2]^S\} \xi_\lambda(|Y_{12}|, \xi_{12}),
\]

where \( \xi_{12}^{-1} = \xi_1^{-1} + \xi_2^{-1} \). Using Eqs. (11.46) and (11.48a) and assuming (without loss of generality) that \( \xi_2 \leq \xi_3 \), we get

\[
[123]^{SSL} \leq [1][2]^{SL}[3]^S \xi_\lambda[Y_{12} + Y_{13}, (\xi_{12}^{-1} + \xi_{13}^{-1} + 2\xi_1^{-1})^{-1}],
\]

with the help of (11.45).

### 7.11.3 Contracted factors

For contracted integrals the following UBs are adopted

\[
\langle \hat{12} \rangle^{SL} = \xi_\lambda(R_{12}, \hat{\xi}_{12}),
\]

\[
\langle \hat{123} \rangle^{SSL} = \xi_\lambda[R_{12,13}, (\hat{\xi}_{12}^{-1} + \hat{\xi}_{13}^{-1} + 2\hat{\xi}_1^{-1})^{-1}],
\]

where \( R_{ij} \) is defined in Eq. (8.43) and \( R_{12,13} = \sqrt{2R_{12}^2 + 2R_{13}^2 - R_{23}^2} \).

### 7.12 Appendix B: Computation of \([\hat{1}]^{fg}\) and \(\langle \hat{1} \rangle^{fg}\)

As discussed in Section 7.6, the primitive factors \([\hat{1}]^{fg}\) only depend on the basis set adopted for each atomic element (H, He, Li, etc.). In fact, since the SPBGs are spherically symmetric for any \( fg \) which is either monotonically decreasing or constant

\[
[1]^{fg} \equiv [\hat{a}_1|\hat{b}_1]^{fg} \leq \max \{[\hat{a}_1|\hat{a}_1]^{fg}, [\hat{b}_1|\hat{b}_1]^{fg}\},
\]

where

\[
[\hat{a}_1|\hat{a}_1]^{fg} = \max_{(r_2,r_3)} \int \hat{a}_1^2 f_{12} g_{13} dr_1.
\]
Thus, the $\langle \hat{1} \rangle^{fg}$ can be easily evaluated as the largest concentric $[\hat{a}_1 | \hat{a}_1]^{fg}$ determined by the basis set adopted for each distinct atomic element. Hence, these factors can be pre-computed with Algorithm 5, where we have adopted the same notation as in Sec. 7.9.

**Algorithm 5** Computation of $\langle \hat{1} \rangle^{fg}$

1. $\triangleright M$ is the number of distinct atomic elements
2. $\triangleright N_M$ is the number of shells for each atom type
3. Set $\langle \hat{1} \rangle^{fg} = 0$
4. for all distinct atomic elements $1, \ldots, M$
   5. for shells $[a_1] = 1, \ldots, N_M$
      6. $[1]^{fg} = [\hat{a}_1 | \hat{a}_1]^{fg}$
      7. $[1]^{fg}$
      8. Solve $\frac{\partial [1]^{fg}}{\partial a_1} = 0$ to get $\sigma_1^{fg}$
      9. Compute $[1]^{fg}$
     10. $\langle \hat{1} \rangle^{fg} = \max \{ \langle \hat{1} \rangle^{fg}, [1]^{fg} \}$
     11. end for
     12. end for

**Algorithm 6** Computation of $\langle 1 \rangle^{fg}$

1. Set $\langle 1 \rangle^{fg} = 0$
2. for all distinct atomic elements $1, \ldots, M$
   3. for contracted shells $[a_1] = 1, \ldots, N_M$
     4. Set $\langle 1 \rangle^{fg} = 0$
     5. for primitive shells $[a_1] = 1, \ldots, K_{a_1}$
        6. for primitive shells $[a_1] = 1, \ldots, K_{a_1}$
           7. $[1]^{fg} = [\hat{a}_1 | \hat{a}_1]^{fg}$
           8. $[1]^{fg}$
           9. Solve $\frac{\partial [1]^{fg}}{\partial a_1} = 0$ to get $\sigma_1^{fg}$
          10. Compute $[1]^{fg}$
         11. $\langle 1 \rangle^{fg} = \langle 1 \rangle^{fg} + [1]^{fg}$
     12. end for
   13. end for
14. $\langle 1 \rangle^{fg} = \max \{ \langle 1 \rangle^{fg}, [1]^{fg} \}$
15. end for
16. end for

For contracted integrals, firstly, it is necessary to compute $\langle \hat{1} \rangle^{fg}$ (see Algorithm 4). These can be achieved using Algorithm 5. Secondly, the factors $\langle 1 \rangle^{fg}$ are needed for the contracted bounds. Their computation is sketched in Algorithm 6.
Chapter 8

Recurrence Relations for Three-Electron Integrals

“Immortal is the moment when I engendered the recurrence.
For the sake of this moment I bear the recurrence.”

Friedrich Nietzsche — Sämtliche Werke: Kritische Studienausgabe,
Vol. 10, essay # 205* 

8.1 Introduction

In Chapter 7, we presented screening methods which enable to avoid the evaluation
of three-electron integrals with negligible sizes. Here, we present an algorithm to
compute the fraction of significant integrals over three-electron operators using recurrence
relations (RRs). Our recursive approach applies to a general class of multiplicative
three-electron operators and thus generalizes existing schemes that pertain only to
GTGs.[246, 364, 365, 338, 336, 337, 324]

Section 8.2 contains classifications of three-electron operators, and permutational
symmetry considerations. In Section 8.3, we propose a recursive algorithm for the

*Unpublished fragments dating to November 1882-February 1883

167
computation of three-electron integrals. Details of a general scheme for deriving three-electron integral RRs are presented in the Appendix.

8.2 Three-Electron Integrals

All basic definitions for three-electron integrals are consistent with those in Chapter 7, Section 7.2. Furthermore, for conciseness, we will adopt a notation in which missing indices represent s-type gaussians. For example, \( [a_2 a_3] \) is a shorthand for \( [0 a_2 a_3 | 000] \).

We also remind the reader, that the use of unbold indices, e.g. \( \langle a_1 a_2 a_3 | b_1 b_2 b_3 \rangle \), indicates a complete class of integrals from a shell-sextet.

8.2.1 Three-electron operators

We are particularly interested in two types of three-electron operators: “chain” operators of the form \( f_{12} g_{13} \), and “cyclic” operators of the form \( f_{12} g_{13} h_{23} \). In both types, the most interesting cases arise when \([251, 261, 252, 253]\)

\[
\begin{aligned}
f_{12}, g_{12}, h_{12} = & \\
= & \begin{cases}
  r_{12}, & \text{Coulomb operator,} \\
  r_{12}, & \text{anti-Coulomb operator,} \\
  \exp(-\lambda r_{12}), & \text{Slater-type geminal,} \\
  \exp(-\lambda r_{12}^2), & \text{gaussian-type geminal,}
\end{cases}
\end{aligned}
\]

and various combinations of these produce three-electron integrals of practical importance. We note that, by virtue of the identity \( r_{12} \equiv r_{12}^{-1}(r_1^2 + r_2^2 - 2r_1 \cdot r_2) \), integrals involving the anti-Coulomb operator can be reduced to linear combinations of integrals over the Coulomb operator.
Table 8.1: Permutational degeneracy for various operators.

<table>
<thead>
<tr>
<th>Type</th>
<th>Operator</th>
<th>Degeneracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-electron</td>
<td>( f_{12} )</td>
<td>8</td>
</tr>
<tr>
<td>Three-electron chain</td>
<td>( f_{12} g_{13} )</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>( f_{12} f_{13} )</td>
<td>16</td>
</tr>
<tr>
<td>Three-electron cyclic</td>
<td>( f_{12} g_{13} h_{23} )</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>( f_{12} f_{13} h_{23} )</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>( f_{12} f_{13} f_{23} )</td>
<td>48</td>
</tr>
</tbody>
</table>

8.2.2 Permutational symmetry

For real basis functions, it is known\(^{115}\) that two-electron integrals have 8-fold permutational symmetry, meaning that the integrals

\[
\begin{align*}
\langle a_1 a_2 | b_1 b_2 \rangle & \quad \langle b_1 a_2 | a_1 b_2 \rangle & \quad \langle b_1 b_2 | a_1 a_2 \rangle & \quad \langle a_1 b_2 | b_1 a_2 \rangle \\
\langle a_2 a_1 | b_2 b_1 \rangle & \quad \langle b_2 a_1 | a_2 b_1 \rangle & \quad \langle b_2 b_1 | a_2 a_1 \rangle & \quad \langle a_2 b_1 | b_2 a_1 \rangle
\end{align*}
\]

are all equal. Three-electron integrals also exhibit permutational symmetry and, for computational efficiency, it is important that this be fully exploited. The degeneracy depends on the nature of the three-electron operator and the five possible cases are listed in Table 8.1.

8.3 Algorithm

In this Section, we present a recursive algorithm for generating a class of three-electron integrals of arbitrary angular momentum from an initial set of fundamental integrals. The algorithm applies to any three-electron operator of the form \( f_{12} g_{13} h_{23} \) and generalizes the HGP-PRISM algorithm following a OTTTCCCTTT pathway.\(^{270, 366}\) The algorithm is shown schematically in Fig. 8.1.

After selecting a significant shell-sextet, we create a set of generalized fundamental integrals \([0]^{mn}\) (Step O). Next, we build angular momentum on center \( A_3 \) (Step T\(_1\)) and on center \( A_2 \) (Step T\(_2\)) using vertical RRs (VRRs). This choice is motivated by the fact that, for chain operators, the VRR for building momentum on \( A_1 \) is more expensive than that for building on \( A_2 \) and \( A_3 \) (see Appendix). Then, using transfer RRs (TRRs),
Figure 8.1: PRISM representation[366] of a scheme for computing a three-electron integral class. In this work, we consider the (orange) OTTTCCCTTT path.

we transfer momentum onto $A_1$ (Step $T_3$). The primitive $[a_1a_2a_3]$ integrals are then contracted (Step C) and horizontal RRs (HRRs) are used to shift angular momentum from the bra centers $A_3$, $A_2$ and $A_1$ onto the ket centers $B_3$, $B_2$ and $B_1$ (Steps $T_4$, $T_5$ and $T_6$). The number of terms in each of these RRs is summarized in Table 8.2 for cyclic and chain operators. We now describe each step in detail.

### 8.3.1 Construct shell-pairs, -quartets and -sextets

Beginning with a list of shells, a list of significant shell-pairs[20] is constructed. By pairing these shell-pairs, a list of significant shell-quartets is created and then, by pairing the significant pairs and quartets, a list of significant shell-sextets is created. This process, which is critical for the efficiency of the overall algorithm, is achieved by implementing the upper bounds and screening techniques discussed in Chapter 7.

<table>
<thead>
<tr>
<th>Gaussian shells</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[000]^{m}$</td>
</tr>
<tr>
<td>$[00a_3]^{m}$</td>
</tr>
<tr>
<td>$[0a_2a_3]^{000}$</td>
</tr>
<tr>
<td>$[a_1a_2a_3]^{000}$</td>
</tr>
<tr>
<td>$[a_1a_2a_3</td>
</tr>
<tr>
<td>$[a_1a_2a_3</td>
</tr>
<tr>
<td>$[a_1a_2a_3</td>
</tr>
</tbody>
</table>

$\text{CCC} \rightarrow (000)^{m}$

$\text{CCC} \rightarrow (00a_3)^{m}$

$\text{CCC} \rightarrow (0a_2a_3)^{000}$

$\text{CCC} \rightarrow (a_1a_2a_3)^{000}$

$\text{CCC} \rightarrow (a_1a_2a_3|00b_3)$

$\text{CCC} \rightarrow (a_1a_2a_3|0b_2b_3)$

$\text{CCC} \rightarrow (a_1a_2a_3|b_1b_2b_3)$
### Table 8.2: Number of RR terms for cyclic and chain operators.

<table>
<thead>
<tr>
<th>Step</th>
<th>RR type</th>
<th>Expression</th>
<th>Operators</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td>VRR</td>
<td>Eq. (3.9)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
<tr>
<td>$T_2$</td>
<td>VRR</td>
<td>Eq. (3.11)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
<tr>
<td>$T'_3$</td>
<td>VRR</td>
<td>Eq. (5.30)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
<tr>
<td>$T_3$</td>
<td>TRR</td>
<td>Eq. (3.14)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
<tr>
<td>$T_4$</td>
<td>HRR</td>
<td>Eq. (3.15)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
<tr>
<td>$T_5$</td>
<td>HRR</td>
<td>Eq. (3.16)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
<tr>
<td>$T_6$</td>
<td>HRR</td>
<td>Eq. (3.17)</td>
<td>$f_{12}$ $g_{13}$ $h_{23}$</td>
</tr>
</tbody>
</table>

#### 8.3.2 Step O. Form fundamental integrals

Having chosen a significant shell-sextet, we replace the three two-electron operators in its fundamental integral (Chapter 7, Eq. (2.6c)) by their Laplace representations

\[
\begin{align*}
  f(r_{12}) & = \int_0^\infty F(s_1) \exp\left(-s_1 r_{12}^2\right) \, ds_1, \\
  g(r_{13}) & = \int_0^\infty G(s_2) \exp\left(-s_2 r_{13}^2\right) \, ds_2, \\
  h(r_{23}) & = \int_0^\infty H(s_3) \exp\left(-s_3 r_{23}^2\right) \, ds_3.
\end{align*}
\]

\[(3.1a) \quad (3.1b) \quad (3.1c)\]

Table 8.3 contains kernels $F(s)$ for a variety of important two-electron operators $f(r_{12})$.

From the formulae in Table 8.3, one can also easily deduce the Laplace kernels for related functions, such as $f(r_{12})^2$, $f(r_{12})/r_{12}$ and $\nabla^2 f(r_{12})$. Integrating over $r_1$, $r_2$ and $r_3$ then yields

\[
\begin{align*}
  [0] & = G_1 G_2 G_3 \int_0^\infty \int_0^\infty \int_0^\infty F(s_1) G(s_2) H(s_3) w_0(s) \, ds,
\end{align*}
\]

\[(3.2)\]

where

\[
W_0(s) = \left[ \frac{\pi^3}{\xi_1 \xi_2 \xi_3 D(s)} \right]^{3/2} \exp\left[ -\frac{N(s)}{D(s)} \right],
\]

\[(3.3)\]
Table 8.3: Laplace kernels $F(s)$ for various two-electron operators $f(r_{12})$. $n$ is an integer. $\Gamma$ is the gamma function, $H_n$ is a Hermite polynomial, erfc is the complementary error function, and $(\alpha)_j$ is a Pochhammer symbol. $\delta^{(k)}$ and $\theta^{(k)}$ are the kth derivatives of the Dirac delta function and Heaviside step function, respectively.[362]

<table>
<thead>
<tr>
<th>$f(r_{12})$</th>
<th>$F(s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{12}^n \exp(-\lambda r_{12})$</td>
<td>$\frac{2\pi^{-1/2}}{(4\pi)^{n/2+1}} H_n+1 \left(\frac{\lambda}{2s^{1/2}}\right) \exp \left(-\frac{\lambda^2}{4s}\right)$</td>
</tr>
<tr>
<td>$(r_{12}^2 - R^2)^n \exp(-\lambda^2 r_{12}^2)$</td>
<td>$\exp[-R^2(s - \lambda^2)] \delta^{(n)}(s - \lambda^2)$</td>
</tr>
<tr>
<td>$(r_{12}^2 - R^2)^{n-1/2} \exp(-\lambda^2 r_{12}^2)$</td>
<td>$\exp[-R^2(s - \lambda^2)]\theta(s - \lambda^2) \Gamma(-n + 1/2)(s - \lambda^2)^{n+1/2}$</td>
</tr>
<tr>
<td>$r_{12}^{2n} \text{erfc}(\lambda r_{12})$</td>
<td>$\frac{-\pi^{-1/2}}{\Gamma(-n - 1/2)s^{n+1}} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\lambda^2}{s - \lambda^2}\right)^{k+1/2} \frac{\theta(s - \lambda^2)}{k + 1/2}$</td>
</tr>
<tr>
<td>$r_{12}^{2n-1} \text{erfc}(\lambda r_{12})$</td>
<td>$\sum_{k=0}^{n} \binom{n}{k} s^{k-n-1/2} \Gamma(k - n + 1/2) \theta^{(k)}(s - \lambda^2)$</td>
</tr>
<tr>
<td>$r_{12}^{2n} \text{erfc}(\lambda r_{12})^2$</td>
<td>$\frac{2\lambda}{\pi s\sqrt{s - \lambda^2}} \sum_{k=0}^{n} \binom{n}{k} (3/2)<em>k \frac{\theta^{(n-k)}(s - 2\lambda^2)}{(\lambda^2 - s)^k} \sum</em>{j=0}^{k} \frac{(-k)_j}{(3/2)_j} (\lambda^2/s)^j$</td>
</tr>
</tbody>
</table>

and $s = (s_1, s_2, s_3)$. The numerator and denominator are

\[ N(s) = \kappa_{12}s_1 + \kappa_{13}s_2 + \kappa_{23}s_3 + \left(\frac{\kappa_{12}}{\zeta_3} + \frac{\kappa_{13}}{\zeta_2} + \frac{\kappa_{23}}{\zeta_1}\right) (s_1s_2 + s_1s_3 + s_2s_3), \quad (3.4a) \]

\[ D(s) = 1 + \left[\frac{1}{\zeta_1} + \frac{1}{\zeta_2}\right] s_1 + \left[\frac{1}{\zeta_1} + \frac{1}{\zeta_3}\right] s_2 + \left[\frac{1}{\zeta_2} + \frac{1}{\zeta_3}\right] s_3 + \frac{\zeta_1 + \zeta_2 + \zeta_3}{\zeta_1\zeta_2\zeta_3} (s_1s_2 + s_1s_3 + s_2s_3), \quad (3.4b) \]

where

\[ \kappa_{ij} = Y_{ij} \cdot Y_{ij} \quad (3.5) \]

is the squared length of the vector $Y_{ij}$.

For reasons that will become clear later, it is convenient to introduce the generalized
fundamental integral

\[
[0]^\mathbf{m} = G_1 G_2 G_3 \int_0^\infty \int_0^\infty F(s_1) G(s_2) H(s_3) w^{m}(s) \, ds,
\]

where

\[
w^{m}(s) = \frac{s_1^{m_1} s_2^{m_2} s_3^{m_3}}{D(s)^{m_1 + m_2 + m_3}} \left[ \frac{s_1 s_2 + s_1 s_3 + s_2 s_3}{D(s)} \right]^{m_4} w_0(s),
\]

and the auxiliary index vector \(\mathbf{m} = (m_1, m_2, m_3, m_4)\).

To form an \([a_1a_2a_3|b_1b_2b_3]\) class with a cyclic operator, we require all \([0]^\mathbf{m}\) with

\[
0 \leq m_1 \leq a_1 + a_2 + b_1 + b_2,
\]

\[
0 \leq m_2 \leq a_1 + a_3 + b_1 + b_3,
\]

\[
0 \leq m_3 \leq a_2 + a_3 + b_2 + b_3,
\]

\[
0 \leq m_4 \leq a_1 + a_2 + a_3 + b_1 + b_2 + b_3.
\]

To form an \([a_1a_2a_3|b_1b_2b_3]\) class with a chain operator, the ranges of \(m_1, m_2\) and \(m_4\) are as in (3.8), but \(m_3 = 0\).

To construct an \(\langle aa|aa\rangle\) class of two-electron integrals, one needs only \(O(a) [0]^\mathbf{m}\) integrals.\([20]\) However, it follows from (3.8) that, to construct an \(\langle aaa|aaa\rangle\) class of three-electron integrals, we need \(O(a^3)\) (for a chain operator) or \(O(a^4)\) (for a cyclic operator) \([0]^\mathbf{m}\) integrals. This highlights the importance of computing these \([0]^\mathbf{m}\) efficiently. If at least one of the two-electron operators is a GTG, the \([0]^\mathbf{m}\) can be found in closed-form.\([257, 258]\) Otherwise, they can be reduced to one- or two-dimensional integrals, which can then be evaluated by various numerical techniques. This step can consume a significant fraction of the total computation time\([367]\) and a comprehensive treatment of suitable numerical methods merits a detailed discussion which we will be presented in future work.

8.3.3 Step T1. Build momentum on center \(A_3\)

Given a set of \([0]^\mathbf{m}\), integrals of higher angular momentum can be obtained recursively, following Obara and Saika.\([269, 368]\) Whereas VRRs for two-electron integrals have been widely studied, VRRs for three-electron integrals have not, except for GTGs.\([246,\]
The $T_1$ step generates $[a_3]^m$ from $[0]^m$ via the 8-term VRR (see the Appendix for a detailed derivation)

$$
[a_3]^m = (Z_3 - A_3)[a_3]^{(0)} + \zeta_1 \zeta_2 Y_{13}[a_3]^{(2)} + \zeta_1 \zeta_2 Y_{23}[a_3]^{(3)} + (\zeta_1 Y_{13} + \zeta_2 Y_{23})[a_3]^{(4)}
$$

$$
+ \frac{a_3}{2\zeta_3} \left\{ [a_3^{+}]^{(0)} - \zeta_1 \zeta_2 [a_3^{-}]^{(2)} - \zeta_1 \zeta_2 [a_3^{-}]^{(3)} - (\zeta_1 + \zeta_2) [a_3^{-}]^{(4)} \right\}, \quad (3.9)
$$

where the superscript $+$ or $-$ denotes an increment or decrement of one unit of cartesian angular momentum. (Thus, $a_\pm$ is analogous to $a_\pm$ in the notation of Obara and Saika.) The value in the curly superscript indicates which component of the auxiliary index vector $m$ is incremented.

For a chain operator, the $\{3\}$ terms disappear, yielding the 6-term RR

$$
[a_3^+]^m = (Z_3 - A_3)[a_3]^{(0)} + \zeta_1 \zeta_2 Y_{13}[a_3]^{(2)} + (\zeta_1 Y_{13} + \zeta_2 Y_{23})[a_3]^{(4)}
$$

$$
+ \frac{a_3}{2\zeta_3} \left\{ [a_3^{+}]^{(0)} - \zeta_1 \zeta_2 [a_3^{-}]^{(2)} - (\zeta_1 + \zeta_2) [a_3^{-}]^{(4)} \right\}. \quad (3.10)
$$

It is satisfying to note that, by setting $\zeta_2 = 0$ in (3.10), we recover the Obara-Saika two-electron RR.

### 8.3.4 Step $T_2$. Build momentum on center $A_2$

The $T_2$ step forms $[a_2 a_3]$ from $[a_3]^m$ via the 10-term RR

$$
[a_2^+ a_3]^m = (Z_2 - A_2)[a_2 a_3]^{(0)} + \zeta_1 \zeta_3 Y_{12}[a_2 a_3]^{(1)}
$$

$$
- \zeta_1 \zeta_3 Y_{23}[a_2 a_3]^{(3)} + (\zeta_1 Y_{12} - \zeta_3 Y_{23})[a_2 a_3]^{(4)}
$$

$$
+ \frac{a_2}{2\zeta_2} \left\{ [a_2 a_3^{+}]^{(0)} - \zeta_1 \zeta_3 [a_2 a_3^{-}]^{(1)} - \zeta_1 \zeta_3 [a_2 a_3^{-}]^{(3)} - (\zeta_1 + \zeta_3) [a_2 a_3^{-}]^{(4)} \right\}
$$

$$
+ \frac{a_3}{2} \left\{ [a_2 a_3^{+}]^{(3)} + [a_2 a_3^{-}]^{(4)} \right\}. \quad (3.11)
$$
For a chain operator, the \( \{3\} \) terms disappear, yielding the 7-term VRR

\[
\begin{align*}
[a_2^+ a_3]^m &= (Z_2 - A_2)[a_2 a_3]^{(0)} + \zeta_1 \zeta_3 Y_{12}[a_2 a_3]^{(1)} + (\zeta_1 Y_{12} - \zeta_3 Y_{23})[a_2 a_3]^{(4)} \\
&\quad + \frac{a_2}{2 \zeta_2} \left\{ [a_2^+ a_3]^{(0)} - \zeta_1 \zeta_3 [a_2^- a_3]^{(1)} - (\zeta_1 + \zeta_3)[a_2^- a_3]^{(4)} \right\} + \frac{a_3}{2} [a_2 a_3^+]^{(4)}. 
\end{align*}
\]

(3.12)

### 8.3.5 Step T\(_3\). Build momentum on center \( A_1 \)

The \( T_3 \) step generates \( [a_1 a_2 a_3] \) from \( [a_2 a_3] \). There are two possible ways to do this. The first, which we call Step \( T_3' \), is to build angular momentum directly on \( A_1 \) using the 12-term VRR (Eq. (5.30)). A second option exploits the translational invariance

\[
\sum_{j=1}^{3} (\nabla_{A_j} + \nabla_{B_j})[a_1 a_2 a_3] = 0 
\]

(3.13)

to derive the 6-term TRR

\[
[a_1^+ a_2 a_3] = \frac{a_1}{2 \zeta_1} [a_1^- a_2 a_3] + \frac{a_2}{2 \zeta_1} [a_1 a_2^- a_3] + \frac{a_3}{2 \zeta_1} [a_1 a_2 a_3^+] - \frac{\zeta_2}{\zeta_1} [a_1 a_2^+ a_3] \\
- \frac{\zeta_3}{\zeta_1} [a_1 a_2 a_3^+] - \frac{\beta_1 (A_1 - B_1) + \beta_2 (A_2 - B_2) + \beta_3 (A_3 - B_3)}{\zeta_1}[a_1 a_2 a_3],
\]

(3.14)

which transfers momentum between centers that host different electrons.

### 8.3.6 Step C. Contraction

At this stage, following the HGP algorithm,[270] we contract the \( [a_1 a_2 a_3|000] \) to form the \( \langle a_1 a_2 a_3|000 \rangle \). We can perform the contraction at this point because all of the subsequent RRs are independent of the contraction coefficients and exponents. More details about this contraction step can be found in Ref. [20].
8.4 CONCLUDING REMARKS

8.3.7 Steps $T_4$ to $T_6$. Shift momentum to ket centers

We shift momentum to $B_3$, $B_2$ and $B_1$ from $A_3$, $A_2$ and $A_1$, respectively, using the 2-term HRRs

\[
\langle a_1 a_2 a_3 | b_3^+ \rangle = \langle a_1 a_2 a_3^+ | b_3 \rangle + (A_3 - B_3) \langle a_1 a_2 a_3 | b_3 \rangle ,
\]

(3.15)

\[
\langle a_1 a_2 a_3 | b_3^+ b_3 \rangle = \langle a_1 a_2 a_3^+ | b_3 b_3 \rangle + (A_2 - B_2) \langle a_1 a_2 a_3 | b_3 b_3 \rangle ,
\]

(3.16)

\[
\langle a_1 a_2 a_3 | b_2^+ b_2 b_3 \rangle = \langle a_1^+ a_2 a_3^+ | b_2 b_3 \rangle + (A_1 - B_1) \langle a_1 a_2 a_3 | b_2 b_3 \rangle .
\]

(3.17)

8.4 Concluding Remarks

In this Chapter, we have presented a general algorithm to construct three-electron integrals over gaussian basis functions of arbitrary angular momentum from fundamental (momentumless) integrals. The algorithm is based on vertical, transfer and horizontal RRs in the spirit of the Head-Gordon-Pople algorithm. Our approach can be extended to four-electron integrals, and we will report results on this into the next Chapter.

8.5 Appendix: Derivation of VRRs

In this Appendix, we follow the Ahlrichs approach[369] to derive a VRR for the construction of $[a_1]^m$ integrals.

Defining the scaled gradient operator

\[
\hat{D}_{A_1} = \frac{\nabla A_1}{2\alpha_1} ,
\]

(5.18)

we can write the Boys relation[19]

\[
[a_1^+] = \hat{D}_{A_1}[a_1] + \frac{a_1}{2\alpha_1} [a_1^-] ,
\]

(5.19)

which connects an integral of higher momentum to an integral derivative with respect to a coordinate of $A_1$. 

176
In operator form, this can be written as

\[ [a_1^+] = \hat{M}_{a_1} \hat{D}_{A_1} [0] + \frac{a_1}{2\alpha_1} [a_1^+] . \] (5.20)

Substituting the chain rule expression

\[ \hat{D}_{A_1} = \frac{\nabla_{A_1} S_1}{2\alpha_1} \frac{\partial}{\partial S_1} + \frac{\nabla_{A_1} \kappa_{12}}{2\alpha_1} \frac{\partial}{\partial \kappa_{12}} + \frac{\nabla_{A_1} \kappa_{13}}{2\alpha_1} \frac{\partial}{\partial \kappa_{13}} \] (5.21)

into (5.20), using the identities

\[ \frac{\partial [0]^m}{\partial G_1} = G_1^{-1} [0]^{(0)} , \] (5.22)

\[ \frac{\partial [0]^m}{\partial \kappa_{12}} = -\zeta_1 \zeta_2 [0]^{(1)} - \zeta_1 \zeta_2 [0]^{(4)} , \] (5.23)

\[ \frac{\partial [0]^m}{\partial \kappa_{13}} = -\zeta_1 \zeta_2 [0]^{(2)} - \zeta_1 \zeta_3 [0]^{(4)} , \] (5.24)

the commutator property

\[ \hat{M}_{a_1} p = p \hat{M}_{a_1} + a_1^2 \hat{M}_{a_1} , \] (5.25)

(where \( p \) is linear in \( A_1 \) and \( \rho \) is its derivative) and the identities

\[ \hat{D}_{A_1} (Z_1 - A_1) = -\frac{\beta_1}{2\alpha_1 \zeta_1} , \] (5.26)

\[ \hat{D}_{A_1} Y_{12} = \hat{D}_{A_1} Y_{13} = \frac{1}{2\zeta_1} , \] (5.27)

one eventually obtains the 8-term RR for building on \( A_1 \)

\[ [a_1^+]^m = (Z_1 - A_1) [a_1]^{(0)} - \zeta_2 \zeta_3 Y_{12} [a_1]^{(1)} - \zeta_2 \zeta_3 Y_{13} [a_1]^{(2)} - (\zeta_2 Y_{12} + \zeta_3 Y_{13}) [a_1]^{(4)} + \frac{a_1}{2\zeta_1} \left\{ [a_1^+]^{(0)} - \zeta_2 \zeta_3 [a_1^+]^{(1)} - \zeta_2 \zeta_3 [a_1^+]^{(2)} - (\zeta_2 + \zeta_3) [a_1^+]^{(4)} \right\} . \] (5.28)

Equation (3.9), which builds on \( A_3 \), can be derived similarly. However, for chain operators, the RR that builds on \( A_1 \) does not shed terms and it is therefore cheaper to build on \( A_3 \) than on \( A_1 \).
Equation (3.11) can be derived using the relation

\[
[a_2^+ a_3]^m = \hat{M}_{a_3} [a_2^+]^m, \tag{5.29}
\]

and the 12-term VRR for building on \( A_1 \) in Step \( T_3' \) is

\[
[a_1^+ a_2 a_3]^m = \hat{M}_{a_3} \hat{M}_{a_2} [a_1^+]^m \\
= (Z_1 - A_1) [a_1 a_2 a_3]^{(0)} - \zeta_2 \zeta_3 Y_{12} [a_1 a_2 a_3]^{(1)} \\
- \zeta_2 \zeta_3 Y_{13} [a_1 a_2 a_3]^{(2)} - (\zeta_2 Y_{12} + \zeta_3 Y_{13}) [a_1 a_2 a_3]^{(4)} \\
+ \frac{a_1}{2\zeta_1} \{ [a_1^+ a_2 a_3]^{(0)} - \zeta_2 \zeta_3 [a_1^+ a_2 a_3]^{(1)} \} \\
- \zeta_2 \zeta_3 [a_1^+ a_2 a_3]^{(2)} - (\zeta_2 + \zeta_3) [a_1^+ a_2 a_3]^{(4)} \} \\
+ \frac{a_2}{2} \left\{ \zeta_3 [a_1 a_2 a_3]^{(1)} + [a_1 a_2 a_3]^{(4)} \right\} + \frac{a_3}{2} \left\{ \zeta_2 [a_1 a_2 a_3]^{(2)} + [a_1 a_2 a_3]^{(4)} \right\}. \tag{5.30}
\]
Chapter 9

Recurrence Relations for Four-Electron Integrals

“You know my methods. Apply them.”

Arthur Conan Doyle — *The Sign of Four*, Chapter 6, page 112

9.1 Introduction

In the previous Chapter, we reported recurrence relations (RRs) to compute three-electron integrals over gaussian basis functions for multiplicative general operators of the form $f_{12}g_{13}h_{23}$. Here, we generalize our previous study to four-electron integrals.

The present Chapter is organised as follows. In Section 9.2, we introduce the four-electron operators treated in this study. Section 9.3 explains how to calculate fundamental integrals required to start the recursive scheme. In Section 9.4, we report vertical, transfer and horizontal RRs for four-electron integrals. Finally, in Section 9.5, we propose a recursive scheme based on these RRs to calculate classes of four-electron integrals.
9.2 Four-Electron Integrals

All basic definitions for four-electron integrals are consistent with those in Chapter 7, Section 7.2. As for three-electron integrals, we will adopt a notation in which missing indices represent $s$-type Gaussians. For example, $[a_2a_3]$ is a shorthand for $[0a_2a_30000]$. As previously, unbold indices $[a_1a_2a_3a_4|b_1b_2b_3b_4]$ indicate a complete class of integrals from a shell-octet.

9.2.1 Four-electron operators

In the present study, we are particularly interested in the four-electron operators $g_{13}h_{23}i_{34}$ (trident) and $f_{12}h_{23}i_{34}$ (four-electron chain or 4-chain) because they can be required in explicitly-correlated methods such as F12 methods.\[254, 255, 253, 252\] Explicitly-correlated calculations may also require three-electron integrals over the (3-chain) $f_{12}h_{23}$ and (cyclic) $f_{12}g_{13}h_{23}$ operators, as well as two-electron integrals over $f_{12}$. However, we will eschew the study of the two-electron integrals here as they have been extensively studied in the past 25 years.\[261, 370, 246, 332, 344, 371, 372, 248, 271, 250, 373, 339, 324, 374\] Note that the nuclear attraction integrals can be easily obtained by taking the large-exponent limit of a $s$-type shell-pair. We refer the interested reader to Refs. \[366, 20\] for more details about the computation of nuclear attraction integrals.

The structure of these operators is illustrated in Fig. 9.1, where we have adopted a diagrammatic representation. Starting with the “pacman” operator $f_{12}g_{13}h_{23}i_{34}$, we are going to show that one can easily derive all the RRs required to compute two-, three- and four-electron integrals following simple rules. Therefore, in the following, we will focus our analysis on this master “pacman” operator.

9.3 Fundamental Integrals

The first step required to compute integrals of arbitrary angular momentum is the computation of the (momentumless) fundamental integrals $[0]$. These are derived starting from Eq. (2.8), in Chapter 7, using the Gaussian integral representation of each
two-electron operator. For instance, we have

$$f_{12} = \int_0^\infty F(t_{12}) \exp\left(-t_{12}^2 r_{12}^2\right) dt_{12}, \quad (3.1)$$

where $F(t_{12})$ is a Gaussian kernel. Table 9.1 contains kernels $F(t)$ for a variety of important two-electron operators $f_{12}$. More general kernels can be found in Chapter 8, Table 8.3.

Next, the integration over $r_1$, $r_2$, $r_3$ and $r_4$ can be carried out, yielding

$$|0| = G_1 G_2 G_3 G_4 \iiint F(t_{12}) G(t_{13}) H(t_{23}) I(t_{34}) w_0(t) dt,$$  \quad (3.2)
Table 9.1: Kernels $F(t)$ of the Gaussian integral representation for various $f_{12}$ operators. $\delta(x)$ and $\theta(x)$ are respectively the Dirac delta and Heaviside step functions, and $\text{erf}(x)$ and $\text{erfc}(x)$ are the error function and its complement version, respectively.\cite{362}

<table>
<thead>
<tr>
<th>$f_{12}$</th>
<th>$F(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1$</td>
<td>$\delta(t)$</td>
</tr>
<tr>
<td>$r_{12}^{-1}$</td>
<td>$2/\sqrt{\pi}$</td>
</tr>
<tr>
<td>$r_{12}^{-2}$</td>
<td>$2t$</td>
</tr>
<tr>
<td>$(r_{12}^2 + \lambda^2)^{-1/2}$</td>
<td>$(2/\sqrt{\pi}) \exp(-\lambda^2 t^2)$</td>
</tr>
<tr>
<td>$\exp(-\lambda r_{12})$</td>
<td>$(\lambda t^{-2}/\pi) \exp(-\lambda^2 t^{-2}/4)$</td>
</tr>
<tr>
<td>$r_{12}^{-1} \exp(-\lambda r_{12})$</td>
<td>$(2/\sqrt{\pi}) \exp(-\lambda^2 t^{-2}/4)$</td>
</tr>
<tr>
<td>$\exp(-\lambda^2 r_{12}^2)$</td>
<td>$\delta(t - \lambda)$</td>
</tr>
<tr>
<td>$r_{12}^{-1} \text{erf}(\lambda r_{12})$</td>
<td>$(2/\sqrt{\pi}) \theta(t - \lambda)$</td>
</tr>
<tr>
<td>$r_{12}^{-1} \text{erfc}(\lambda r_{12})$</td>
<td>$(2/\sqrt{\pi}) [1 - \theta(t - \lambda)]$</td>
</tr>
</tbody>
</table>

where $t = (t_{12}, t_{13}, t_{23}, t_{34})$ and

$$w_0(t) = \left( \frac{\pi^4}{D(t)} \right)^{3/2} \exp \left[ -\frac{N(t)}{D(t)} \right].$$ \hfill (3.3)

Defining the following polynomials

\begin{align*}
s_1 &= t_{12}^2, & s_2 &= t_{13}^2, & s_3 &= t_{23}^2, \\
s_4 &= t_{34}^2, & s_5 &= t_{12}^2 t_{13}^2 + t_{12}^2 t_{23}^2 + t_{13}^2 t_{23}^2, & s_6 &= t_{12}^2 t_{34}^2, \\
s_7 &= t_{13}^2 t_{34}^2, & s_8 &= t_{23}^2 t_{34}^2, & s_9 &= t_{12}^2 t_{13}^2 t_{34}^2 + t_{12}^2 t_{23}^2 t_{34}^2 + t_{13}^2 t_{23}^2 t_{34}^2, \hfill (3.4a) & \hfill (3.4b) & \hfill (3.4c)
\end{align*}

we have

$$D(t) = \zeta_1 \zeta_2 \zeta_3 \zeta_4 + (\zeta_1 + \zeta_2) \zeta_3 \zeta_4 s_1 + \zeta_2 \zeta_3 (\zeta_1 + \zeta_3) s_2 + \zeta_1 \zeta_4 (\zeta_2 + \zeta_3) s_3 + \zeta_1 \zeta_2 (\zeta_3 + \zeta_4) s_4 + (\zeta_1 + \zeta_2 + \zeta_3) s_5 + (\zeta_1 + \zeta_2) (\zeta_3 + \zeta_4) s_6 + \zeta_2 (\zeta_1 + \zeta_3 + \zeta_4) s_7 + \zeta_1 (\zeta_2 + \zeta_3 + \zeta_4) s_8 + (\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4) s_9, \hfill (3.5a)$$

$$N(t) = \zeta_3 \zeta_1 \kappa_{12} s_1 + \zeta_2 \zeta_3 \kappa_{13} s_2 + \zeta_1 \zeta_4 \kappa_{23} s_3 + \zeta_1 \zeta_2 \kappa_{34} s_4 + \zeta_4 (\kappa_{12} + \kappa_{13} + \kappa_{23}) s_5 + [ (\zeta_3 + \zeta_4) \kappa_{12} + (\zeta_1 + \zeta_2) \kappa_{34} ] s_6 + \zeta_2 (\kappa_{13} + \kappa_{14} + \kappa_{34}) s_7 + \zeta_1 (\kappa_{23} + \kappa_{24} + \kappa_{34}) s_8 + (\kappa_{12} + \kappa_{13} + \kappa_{14} + \kappa_{23} + \kappa_{24} + \kappa_{34}) s_9, \hfill (3.5b)$$

182
where $\kappa_{ij} = \zeta_i \zeta_j |Z_i - Z_j|^2$.

Following Obara and Saika,[269, 368] vertical RRs (VRRs) are obtained by differentiation of Eq. (3.2) with respect to the center coordinates. Therefore, one can show that (3.2) has to be generalized to the following form:

$$
[0]^m = G_1 G_2 G_3 G_4 \iiint F(t_{12}) G(t_{13}) H(t_{23}) I(t_{34}) w_m(t) \, dt,
$$

where $m = (m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8, m_9)$ is, for the pacman operator, a nine-dimensional auxiliary index, and

$$
w_m(t) = w_0(t) \prod_{k=1}^{\dim m} \left[ \frac{s_k}{D(t)} \right]^{m_k}.
$$

As reported in Fig. 9.1, while the fundamental integrals of the pacman operator contains 9 auxiliary indices (i.e. $\dim m = 9$), the two interesting four-electron operators (trident and chain) contains 7. The cyclic and chain three-electron operators have only four and three, respectively, while the two-electron chain operator has a single $m$ component (i.e. $\dim m = 1$). In Chapters 10 and 11, we will show that these numbers of auxiliary indices are drastically reduced if one uses Gaussian Geminals due to their factorization properties.

### 9.4 Recurrence Relations

In this Section, we report vertical, transfer and horizontal RRs for the computation of four-electron integrals of arbitrary angular momentum. In particular, we refer the interested readers to the appendix of Chapter 8[363] for more details about how to derive these VRRs following the late R. Ahlrichs.[369]
9.4. RECURRENCE RELATIONS

9.4.1 Vertical recurrence relations

To build angular momentum over center $A_1$, we have derived the following 24-term VRR:

$$[a_1^+ a_2 a_3 a_4]^m = Z_1 A_1 [a_1 a_2 a_3 a_4]^m - \zeta_2 \zeta_4 Z_{12} [a_1 a_2 a_3 a_4]^{(1)} - \zeta_2 \zeta_4 Z_{13} [a_1 a_2 a_3 a_4]^{(2)}$$

$$- \zeta_4 (\zeta_2 Z_{12} + \zeta_2 Z_{13}) [a_1 a_2 a_3 a_4]^{(5)} - \zeta_2 (\zeta_3 + \zeta_4) Z_{12} [a_1 a_2 a_3 a_4]^{(6)}$$

$$- \zeta_2 (\zeta_3 Z_{13} + \zeta_4 Z_{14}) [a_1 a_2 a_3 a_4]^{(7)} - (\zeta_2 Z_{12} + \zeta_3 Z_{13} + \zeta_4 Z_{14}) [a_1 a_2 a_3 a_4]^{(9)}$$

$$+ \frac{a_1}{2 \zeta_1} \left\{ [a_1 a_2 a_3 a_4]^m - \zeta_2 \zeta_4 Z_{12} [a_1 a_2 a_3 a_4]^{(1)} - \zeta_2 \zeta_4 Z_{13} [a_1 a_2 a_3 a_4]^{(2)}$$

$$- \zeta_4 (\zeta_2 + \zeta_4) [a_1 a_2 a_3 a_4]^{(5)} - \zeta_2 (\zeta_3 + \zeta_4) [a_1 a_2 a_3 a_4]^{(6)}$$

$$- \zeta_2 (\zeta_3 + \zeta_4) [a_1 a_2 a_3 a_4]^{(7)} - (\zeta_2 + \zeta_4) [a_1 a_2 a_3 a_4]^{(9)} \right\}$$

$$+ \frac{a_2}{2} \left\{ \zeta_3 \zeta_4 [a_1 a_2 a_3 a_4]^{(1)} + \zeta_1 [a_1 a_2 a_3 a_4]^{(5)}$$

$$+ (\zeta_3 + \zeta_4) [a_1 a_2 a_3 a_4]^{(6)} + [a_1 a_2 a_3 a_4]^{(9)} \right\}$$

$$+ \frac{a_3}{2} \left\{ \zeta_3 \zeta_4 [a_1 a_2 a_3 a_4]^{(2)} + \zeta_1 [a_1 a_2 a_3 a_4]^{(5)}$$

$$+ \zeta_2 [a_1 a_2 a_3 a_4]^{(7)} + [a_1 a_2 a_3 a_4]^{(9)} \right\}$$

$$+ \frac{a_4}{2} \left\{ \zeta_2 [a_1 a_2 a_3 a_4]^{(7)} + [a_1 a_2 a_3 a_4]^{(9)} \right\}.$$  

(4.8)

where, consistently with Chapter 8, the superscript $+$ or $-$ denotes an increment or decrement of one unit of Cartesian angular momentum (thus, $a^\pm$ is analogous to $a \pm 1_i$ in the notation of Obara and Saika), and the value in the curly superscript indicates which component of the auxiliary index vector $m$ is incremented.

Because Eq. (4.8) builds angular momentum over $A_1$ and all four bra centers have non-zero angular momentum, we will call this expression VRR$^{A_1}_4$. The VRRs used to obtain $[a_1^+ a_2 a_3]^m$, $[a_1^+ a_2]^m$ and $[a_1^+]^m$ can be easily derived from Eq. (4.8) by setting successively $a_4 = 0$, $a_3 = 0$ and $a_2 = 0$. These are respectively named VRR$^{A_1}_3$, VRR$^{A_1}_2$, and VRR$^{A_1}_1$.

One can easily derive VRR$^4$ for the trident and 4-chain operators following the simple rules given in Fig. 9.1. We obtain the VRRs for the trident operator by removing the terms {1} and {6}. Similarly, 4-chain VRRs are obtained by removing the terms {2} and {7}. This yields a 18- and 17-term VRR$^{A_1}_4$ for the trident and 4-chain operators, respectively.
\[ a_{1} a_{2} a_{3} a_{4} \]
Again, the corresponding expressions for VRR	extsubscript{1}, VRR	extsubscript{2}, VRR	extsubscript{3} can be easily derived from Eqs. (4.9), (4.10) and (4.11). The number of terms for each of these VRRs is reported in Fig. 9.2 for the 3-chain \(f_{12}h_{23}\) (top left), cyclic \(f_{12}g_{13}h_{23}\) (top right), 4-chain \(f_{12}h_{23}i_{34}\) (bottom left) and trident \(g_{13}h_{23}i_{34}\) (bottom right) operators.

9.4.2 Transfer recurrence relations

TRRs redistribute angular momentum between centers hosting to different electrons. Using translational invariance, one can derive

\[
[a_1^+ a_2 a_3 a_4]^{m} = Z_4 A_4 [a_1 a_2 a_3 a_4]^m + \zeta_1 Z_3 [a_1 a_2 a_3 a_4]^4 + \zeta_3 [a_1 a_2 a_3 a_4]^6
+ \zeta_2 Z_4 [a_1 a_2 a_3 a_4]^7 + \zeta_1 Z_4 [a_1 a_2 a_3 a_4]^8
+ \frac{a_4}{2\zeta_1} [a_1 a_2 a_3 a_4]^m - \zeta_1 [a_1 a_2 a_3 a_4]^4
- \zeta_3 [a_1 a_2 a_3 a_4]^6 - \zeta_2 [a_1 a_2 a_3 a_4]^7 - \zeta_1 [a_1 a_2 a_3 a_4]^8
+ \frac{a_1}{2} \left\{ \zeta_2 [a_1 a_2 a_3 a_4]^7 + [a_1^{-} a_2 a_3 a_4]^9 \right\}
+ \frac{a_2}{2} \left\{ \zeta_1 [a_1 a_2 a_3 a_4]^8 + [a_1 a_2 a_3 a_4]^9 \right\}
+ \frac{a_3}{2} \left\{ \zeta_1 [a_1 a_2 a_3 a_4]^4 + (\zeta_1 + \zeta_2) [a_1 a_2 a_3 a_4]^6 \right\}
+ \zeta_2 [a_1 a_2 a_3 a_4]^{-7} + \zeta_1 [a_1 a_2 a_3 a_4]^{-8} + [a_1 a_2 a_3 a_4]^9 \right\}
\]

(4.11)

9.4. RECURRENCE RELATIONS
Figure 9.2: Graph representation of the VRRs for the 3-chain \( f_{12} h_{23} \) (top left), cyclic \( f_{12} g_{13} h_{23} \) (top right), 4-chain \( f_{12} h_{23} i_{34} \) (bottom left) and trident \( g_{13} h_{23} i_{34} \) (bottom right) operators. The edge label gives the number of terms in the corresponding VRR. The red path corresponds to the algorithm generating the smallest number of intermediates.

### 9.4.3 Horizontal recurrence relations

The so-called HRRs enable to shift momentum between centers over the same electronic coordinate:

\[
\langle a_1 a_2 a_3 a_4 | b_i^+ \rangle = \langle a_1 a_2 a_3 a_4 | b_i \rangle + A_4 B_4 \langle a_1 a_2 a_3 a_4 | b_i \rangle, \quad (4.13a)
\]

\[
\langle a_1 a_2 a_3 a_4 | b_i^+ b_j \rangle = \langle a_1 a_2 a_3 a_4 | b_i b_j \rangle + A_3 B_3 \langle a_1 a_2 a_3 a_4 | b_i b_j \rangle, \quad (4.13b)
\]

\[
\langle a_1 a_2 a_3 a_4 | b_i^+ b_j b_i b_i \rangle = \langle a_1 a_2 a_3 a_4 | b_i b_j b_i b_i \rangle + A_2 B_2 \langle a_1 a_2 a_3 a_4 | b_i b_j b_i b_i \rangle, \quad (4.13c)
\]

\[
\langle a_1 a_2 a_3 a_4 | b_i^+ b_j b_i b_i \rangle = \langle a_1 a_2 a_3 a_4 | b_i b_j b_i b_i \rangle + A_1 B_1 \langle a_1 a_2 a_3 a_4 | b_i b_j b_i b_i \rangle. \quad (4.13d)
\]
Note that HRRs can be applied to contracted integrals because they are independent of the contraction coefficients and exponents.

Table 9.2: Number of intermediates required to compute various integral classes for two-, three- and four-electron operators. The path generating the minimum number of intermediates is highlighted in bold. The number of terms in the RRs and the associated incremental center are also reported.

<table>
<thead>
<tr>
<th>Integral type</th>
<th>operator</th>
<th>path</th>
<th>number of terms</th>
<th>centers</th>
<th>integral class</th>
</tr>
</thead>
<tbody>
<tr>
<td>two-electron</td>
<td>$f_{12}$</td>
<td>VV</td>
<td>(4,6)</td>
<td>$A_2A_1$</td>
<td>4 13 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VT</td>
<td>(4,4)</td>
<td>$A_2A_1$</td>
<td>7 19 37</td>
</tr>
<tr>
<td>three-electron chain</td>
<td>$f_{12}h_{23}$</td>
<td>VVV</td>
<td>(6,10,9)</td>
<td>$A_3A_1A_2$</td>
<td>34 230 873</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVV</td>
<td>(8,8,9)</td>
<td>$A_3A_1A_2$</td>
<td>32 209 778</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVT</td>
<td>(6,7,6)</td>
<td>$A_3A_1A_2$</td>
<td>38 246 873</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVT</td>
<td>(6,10,6)</td>
<td>$A_3A_2A_1$</td>
<td>43 314 1,256</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVT</td>
<td>(8,8,6)</td>
<td>$A_2A_3A_1$</td>
<td>40 260 923</td>
</tr>
<tr>
<td>cyclic</td>
<td>$f_{12}g_{13}h_{23}$</td>
<td>VVV</td>
<td>(8,10,12)</td>
<td>$A_3A_2A_1$</td>
<td>52 469 2,216</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVT</td>
<td>(8,10,6)</td>
<td>$A_1A_2A_1$</td>
<td>61 539 2,426</td>
</tr>
<tr>
<td>four-electron chain</td>
<td>$f_{12}h_{23}h_{34}$</td>
<td>VVV</td>
<td>(10,11,20,24)</td>
<td>$A_4A_3A_2A_3$</td>
<td>465 13,781 150,961</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,18,20,17)</td>
<td>$A_4A_3A_2A_3$</td>
<td>436 12,535 133,891</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(14,18,16,17)</td>
<td>$A_4A_3A_2A_3$</td>
<td>433 12,704 138,913</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(11,10,20,8)</td>
<td>$A_4A_1A_3A_2$</td>
<td>435 12,863 141,679</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVT</td>
<td>(10,18,20,8)</td>
<td>$A_4A_3A_2A_1$</td>
<td>532 16,295 181,178</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVT</td>
<td>(14,14,20,8)</td>
<td>$A_4A_3A_2A_1$</td>
<td>560 17,029 188,242</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVT</td>
<td>(14,18,16,8)</td>
<td>$A_4A_3A_2A_1$</td>
<td>559 17,487 199,050</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,12,14,28)</td>
<td>$A_4A_2A_1A_3$</td>
<td>445 13,139 143,619</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,12,24,18)</td>
<td>$A_4A_2A_3A_1$</td>
<td>447 13,381 148,911</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,20,16,18)</td>
<td>$A_4A_3A_2A_1$</td>
<td>418 12,447 133,853</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(16,14,16,18)</td>
<td>$A_4A_3A_2A_1$</td>
<td>418 12,054 129,322</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,12,14,8)</td>
<td>$A_4A_2A_1A_3$</td>
<td>470 13,306 136,584</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,12,24,8)</td>
<td>$A_4A_2A_3A_1$</td>
<td>546 16,917 191,171</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(10,20,16,8)</td>
<td>$A_4A_3A_2A_1$</td>
<td>521 15,515 168,958</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VVVV</td>
<td>(16,14,16,8)</td>
<td>$A_4A_3A_2A_1$</td>
<td>499 13,969 142,264</td>
</tr>
</tbody>
</table>

9.5 Algorithm

In this Section, we describe a recursive scheme for the computation of three- and four-electron integrals based on a late contraction scheme à la Head-Gordon-Pople (HGP). [270] The general skeleton of the algorithm is shown in Fig. 9.3 for two representative examples: the 3-chain operator $f_{12}g_{23}$ (left) and the 4-chain operator $f_{12}h_{23}h_{34}$ (right).

First, let us focus on the 3-chain operator.

To compute a class of three-electron integrals $\langle a_1a_2a_3|b_1b_2b_3\rangle$, starting from the fundamental integrals $[000]^m$, we first build up angular momentum over center $A_3$ with the 6-term VRR$_1$ to obtain $[00a_3]$. Then, we use the 10-term VRR$_2$ over $A_2$ to obtain
Finally, we build up momentum over the last bra center $A_1$ using the 9-term VRR$_3$ to get $[a_1 a_2 a_3]$. 

Note that, in our previous Chapter, we stated that one could use either the 6-term TRR instead or the 12-term VRR$_3$. However, we have found here that the number of intermediates (i.e. the number of pre-computed classes needed to calculate a given class) required by the paths involving the TRR is much larger (see Table 9.2). This is readily understood: note that a unit increase in momentum on the last center requires the same increase on all the other centers (as evidenced by the second term in the right-hand side of (4.12)). Hence, the TRR is computationally expensive for three- and four-electron integrals due to the large number of centers. As illustrated in the top left graph of Fig. 9.2, other paths, corresponding to different VRRs, are possible. However, we have found that they do generate a larger number of intermediates, as reported in Table 9.2.

Similarly to the 3-chain operator, for the 4-chain operator, we get $[a_1 a_2 a_3 a_4]$ by successively building up momentum over $A_4$, $A_3$, $A_2$ and $A_1$. The number of intermediates required by the other paths are gathered in Table 9.2. Again, the paths involving the 8-term TRR (reported in Eq. (4.12)) are much more expensive.
The last two steps of the algorithm are common to the three- and four-electron integral schemes. Following the HGP algorithm, [270] we contract the integrals: 

\[ a_1 a_2 a_3 a_4 |0000 \] to form \( \langle a_1 a_2 a_3 a_4 |0000 \rangle \) in the four-electron case, or \[ a_1 a_2 a_3 |000 \] to form \( \langle a_1 a_2 a_3 |000 \rangle \) in the three-electron case. More details about the contraction step can be found in Ref. [20]. The final step of the algorithm shifts momentum to the ket centers from the bra centers with the help of the 2-term HRRs reported in Section 9.4.3.

9.6 Concluding Remarks

In this Chapter, we have reported recurrence relations (RRs) for the efficient and accurate computation of four-electron integrals over Gaussian basis functions and a general class of multiplicative four-electron operators of the form \( f_{12} g_{13} h_{23} i_{34} \). Starting from this master operator, one can easily derive the RRs for various operators arising in explicitly-correlated methods following simple diagrammatic rules (see Fig. 9.1).

Here, we have derived three types of RRs: i) starting from the fundamental integrals, vertical RRs (VRRs) allow to increase the angular momentum over the bra centers; ii) the transfer RR (TRR) redistributes angular momentum between centers hosting different electrons, and can be used instead of the VRR on the last bra center; iii) the horizontal RRs (HRRs) enable to shift momentum from the bra to the ket centers corresponding to the same electronic coordinate. Importantly, HRRs can be applied to contracted integrals.

Finally, after carefully studying the different paths one can follow to build up angular momentum (see Fig. 9.2), we have proposed a late-contraction recursive scheme which minimizes the number of intermediates to be computed (see Fig. 9.3). We believe our approach represents a major step towards a fast and accurate computational scheme for three- and four-electron integrals within explicitly-correlated methods. It also paves the way to contraction-effective methods for these types of integrals. In particular, an early contraction scheme would have significant computational benefits.

The next two Chapters will present analogous schemes, including the screening steps, for two-, three- and four-electron integrals arising in explicitly correlated methods when the correlation factor is a linear combination of GGs. In this case, we will show that,
due to the factorization properties of GGs, the computational cost is drastically reduced.
Chapter 10

Two-Electron Integrals Over Gaussian Geminals

“It is to Gauss that we owe our deliverance from that absurd method of estimating forces by a variable standard which prevailed so long even among men of science.”

James Clerk Maxwell — Introductory Lecture on Experimental Physics

10.1 Introduction

Fashions come and go in quantum chemistry, but Gaussian functions are here to stay. Their unique combination of desirable properties – strong localization, infinite differentiability and closure under multiplication – have cemented their status as near-ideal computational building blocks and they enjoy an almost unchallenged pre-eminence in most of the popular molecular orbital software packages.

However, they have not always been so admired. Following Boys’ landmark proposal[19] that they be used as basis functions for molecular orbitals, many of the field’s leading exponents were skeptical and continued to persevere with exponential functions of various types, either devoting years of effort to the multicenter
two-electron integrals that such functions entail or approximating those integrals by ingenious combinations of fitting and neglect.

It is less well known that, ten years after the Boys paper and long before Gaussian basis functions had garnered universal approval, Boys[257] and Singer[258] proposed that a Gaussian in the distance $r_{12}$ between two electrons, the so-called Gaussian Geminal

$$G_{12} = \exp(-\lambda r_{12}^2)$$

(1.1)

could be a similarly potent two-electron basis function because (quoting Boys[257]) “there are explicit formulas for all the necessary many-dimensional integrals” which arise.

Over the years, a number of other Geminal functions have been contemplated, primarily because the lack of cusps[17] in Gaussians has been perceived as a serious deficiency. Ten-no showed that, although the Slater Geminal $\exp(-\lambda r_{12})$ is more difficult to handle than its Gaussian cousin, the integrals that it generates can be computed, albeit with some effort.[248, 339] Soon afterwards, in an elegant comparative study,[249] Tew and Klopper concluded that the Slater Geminal is superior to three other non-Gaussian alternatives. However, although one-electron Gaussians are inherently incapable of modeling nuclear-electron cusps, and two-electron Gaussian Geminals are likewise unable to capture electron-electron cusps, it has been established both theoretically[239, 375, 376, 377] and empirically that a “brute force” saturation of function space, including Gaussians with large exponents, can reduce the cusp-related errors to any desired level. As a consequence, the Boys-Singer idea has taken root and flourished over the years in a number of groups.[378, 319, 320, 325, 326, 327, 328, 379, 380, 381, 317, 382, 383, 384, 385, 245, 246, 364, 365, 335, 386, 324]

Twenty years ago, Gill and Adamson suggested[360] that the short-range part of the Coulomb operator could be systematically improved toward the full operator by adding Gaussians, i.e.

$$\frac{1}{r_{12}} \approx \frac{\text{erfc}(\omega r_{12})}{r_{12}} + \frac{2\omega}{\sqrt{\pi}} \sum_{j=1}^{m} c_j \exp(-\omega^2 \alpha_j^2 r_{12}^2)$$

(1.2)

where erfc is the complementary error function[362] and the coefficients $c_j$ and exponents $\alpha_j$ are dimensionless constants. If the $c_j$ and $\alpha_j$ are chosen to match the first $4m - 2$
derivatives at $r_{12} = 0$, one obtains a family of Coulomb-attenuated potentials termed CAP($m$).[387] A few years later, Sirbu and King[388] proposed a single-Gaussian expansion in which the parameters were instead tailored for a subsequent perturbation theory. In 2004, Toulouse et al.[389] adopted the CAP(1) potential, renaming it “erfgau” and using it to explain the accuracy of the Local Density Approximation (LDA) in density functional theory (DFT). The erfgau potential was adopted by Song and co-workers,[390, 391] who later discovered problems for which it is useful to drop the erfc term [392, 393, 394] and subsequently to add a second Gaussian,[395] to form the “2Gau” potential. In 2Gau, the more diffuse Gaussian ($\alpha = 0.006$) is used to estimate the medium-range part of the exchange energy.

Boys-Singer, CAP($m$), erfgau, 2Gau and F12 calculations all require the evaluation of two-electron integrals over the Geminal operator (1.1) and, for efficient applications to large chemical systems, it is desirable to be able to compute these fast. If the molecular orbitals are expanded in a Gaussian basis, i.e. the basis functions and the two-electron operator are both Gaussian, the resulting integrals factorize immediately into their Cartesian components[267] and can be formed without the need for an “auxiliary index” $m$. This offers important computational advantages and the usual recurrence relations (RRs) that are effective for forming traditional Coulomb integrals[266, 267, 264, 268, 269, 270, 396, 397, 398, 366, 399, 20, 400, 369, 401, 363] must be appropriately simplified and optimized in order to be maximally effective.

In addition to constructing optimal RRs and using them intelligently, one must also fully exploit the fact that the Gaussian Geminal (1.1) is a short-range operator whose matrix elements between spatially well-separated charge distributions are negligible and should be systematically avoided. It is easy to show that, if the basis set contains $N$ functions, there are $O(N^4)$ two-electron integrals but only $O(N)$ of these are non-negligible in a large molecule. Adamson et al.[361] and others[402] have discussed special techniques for identifying this tiny subset of worthwhile integrals and avoiding the others.

In the following Sections, we present a near-optimal algorithm for constructing two-electron integrals over (1.1) in a basis set of contracted Gaussian functions. After defining our notation and Rys integrals in Section 10.2, we discuss a detailed algorithm in
10.2 ADDITIONAL NOTATION AND RYS INTEGRALS

Section 10.3 and its application to the formation of a \( pp|pp \) class in Section 10.4. Finally, computational costs are examined in Section 10.5.

The study in this Chapter paves the way to Chapter 11, that will discuss the evaluation of three- and four-electron integrals arising in explicitly correlated methods when Gaussian-Geminal correlation factors are used.

### 10.2 Additional Notation and Rys Integrals

Definitions for Gaussian basis functions, shells, fundamental integrals, and integral classes have been already provided in Chapter 7, Section 7.2. Additionally, in this Chapter we will suppress the operator in the braket notation, that is \( [a_1 a_2|G_{12}|b_1 b_2] \) and \( \langle a_1 a_2|G_{12}|b_1 b_2 \rangle \equiv \langle a_1 a_2|b_1 b_2 \rangle \).

It is easy to show that

\[
[a_1 a_2|b_1 b_2] = G_1 G_2 T_{a_1 a_2 b_1 b_2} T_{a_1 b_1 a_2 b_2} T_{a_1 b_1 a_2 b_2} T_{a_1 b_1 a_2 b_2} (2.3)
\]

that is the primitive two-electron integral \( [a_1 a_2|b_1 b_2] \) factorises into the product of three Rys integrals of the form

\[
T_{a_1 b_1 a_2 b_2} = \frac{\sqrt{\xi_1 \xi_2}}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (s - A_1)^{a_1} (s - B_1)^{b_1} (t - A_2)^{a_2} (t - B_2)^{b_2} e^{-\zeta_1 (s-z_1)^2} e^{-\zeta_2 (t-z_2)^2} ds dt
\]

and the simplest of these is

\[
T_{0000} = \left[ \frac{\lambda^{-1}}{\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1}} \right]^{1/2} \exp \left[ -\frac{(Z_1 - Z_2)^2}{\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1}} \right] (2.5)
\]

When discussing RRs, we use a compact vector notation that we introduced in Chapter 8[363] wherein \( a^\pm \) represents the three angular momentum vectors formed by incrementing or decrementing the \( x, y \) or \( z \) components of \( a \). For example, Boys’ famous formula[19] for the derivatives of the primitive Gaussian, \( [a] \), with respect to
the components of its center is

\[ \nabla A[a] = 2a[a^+] - a[a^-] \quad (2.6) \]

This notation avoids the need for the \( a \pm 1 \) motifs which arise in the Obara-Saika notation.\[269, 368\]

As usual, we will use non-bold indices to indicate shells of Gaussians. For example, \([11]\) denotes a primitive Gaussian shell-pair arising from two p shells. Similarly, \((22)10\) denotes a class of \(6 \times 6 \times 3 \times 1 = 108\) contracted integrals that arise from two d shells, a p shell and an s shell.

### 10.3 Algorithm

In this Section, we present an efficient algorithm for generating an \(a_1a_2|b_1b_2\) class from shell data. We use \(L\) to denote the maximum angular momentum in the basis set. For example, if the basis contains only s, p and d functions, then \(L = 2\).

#### 10.3.1 Significant shell-pairs

In Chapter 7, we have derived the primitive class bound

\[ ||ab|cd|| \leq \min\{|1|^G|2|, |1|2|^G\} \quad (3.7) \]

where the primitive bound factor is

\[ |1|^G = |1| \left( \frac{\xi_1}{\lambda + \xi_1} \right)^{3/2} \quad (3.8) \]

and

\[ |1| = |D_{a_1}D_{b_1}| \sqrt{\frac{(4\alpha_1\beta_1)^{3/2}}{\gamma_{a_1}\gamma_{b_1}}} \left[ \frac{a_1}{\sigma e} \right]^{a_1} \left[ \frac{b_1}{\sigma e} \right]^{b_1} e^{-\frac{(1-\sigma_1)AB_1^2}{\alpha_1^{1+\beta_1}}} \left( \frac{\pi}{\xi_1} \right)^{3/2} \quad (3.9) \]

with \(\xi_1 = (1 - \sigma_1)(\alpha_1 + \beta_1)\), and \(\sigma_1\) is an adjustable parameter that we set to

\[ \sigma_1^G = \frac{(a_1 + b_1)(\alpha_1 + \beta_1 + \lambda)}{(a_1 + b_1 + 3)(\alpha_1 + \beta_1) + 2\frac{|AB_1|^2}{\alpha_1^{1+\beta_1}}(\alpha_1 + \beta_1 + \lambda)} \quad (3.10) \]
when used for $[1]^G$ and to

$$\sigma_1 = \frac{(a_1 + b_1)}{(a_1 + b_1 + 3) + 2 \frac{|AB|^2}{a_1^{i+1} + b_1^{i+1}}} \quad (3.11)$$

when used for $[1]$.

**Algorithm 7** Construction of significant shell-pairs for computing integrals over $G_{12}$

1: Compute $[1]$ and $[1]^G$
3: nSigShellPairs = 0
4: for contracted shell $\langle a \rangle$ do
5: for contracted shell $\langle b \rangle$ do
6: $K_{ab} = 0$; $\xi_{\min} = 10^6$; $(1) = 0$; $(1)^G = 0$
7: for primitive shell $[a_1]$ in $\langle a \rangle$ do
8: for primitive shell $[b_1]$ in $\langle b \rangle$ do
9: Compute both $\sigma_{1j}$ and $\sigma_{1j}^G$ values using Eqs. 3.10 and 3.11
10: Compute primitive bound factor $[1]_{ij}$ using Eq. 3.9
11: Compute primitive bound factor $[1]_{ij}^G$ using Eq. 3.8
12: if $\min\{[1]_{ij}^G, [1]_{ij}\} > \tau$ then
13: $K_{ab} = K_{ab} + 1$
14: $\xi_{\min} = \min\{(1 - \sigma_{1j})/(\alpha_i + \beta_j), \xi_{\min}\}$
15: $(1) = (1) + [1]_{ij}$
16: $(1)^G = (1)^G + [1]_{ij}^G$
17: end if
18: end for
19: end for
20: if $K_{ab} = 0$ then
21: Discard contracted shell-pair $\langle ab \rangle$
22: else
23: nSigShellPairs = nSigShellPairs + 1
24: Store degree of contraction $K_{ab}$
25: Store minimum exponent $\xi_{\min}$
26: Store contracted bound factors $(1)$ and $(1)^G$
27: end if
28: end for
29: end for

Bound factors $[2]$ and $[2]^G$ are defined similarly and are bounded by their maxima in the basis set, $\widehat{[2]}$ and $\widehat{[2]}^G$, respectively. We remind the reader that the $\widehat{[2]}$ and $\widehat{[2]}^G$ factors do not depend on the geometry of the system and can be computed by looping only over shells (and not shell-pairs), as described in Chapter 7, Section 7.12 (Appendix B). Analogous considerations apply for the contracted factors $\langle 2 \rangle$ and $\langle 2 \rangle^G$, and their maxima over the basis set, $\langle 2 \rangle$ and $\langle 2 \rangle^G$.

Thus, to build a list of “significant” contracted shell-pairs, i.e. those that could yield integrals over $G$ that exceed a user-specified threshold $\tau$, we follow the scheme in Algorithm 7. In an extended system with $N$ shells, this generates only $O(N)$ significant shell-pairs.
10.3.2 Significant shell-quartets

Shell-quartets are produced by pairing significant shell-pairs. However, because $G_{12}$ is short-ranged, most of the quartets in an extended system are not significant and we therefore require a strong upper bound to identify and discard these as cheaply as possible. Bounds[349, 350, 20] for Coulomb integrals are weak when applied to integrals over $G_{12}$, but, in Chapter 7, we have derived the strong bound

$$|\langle a_1 a_2 | b_1 b_2 \rangle| \leq \min\{\langle 1 \rangle^{\mathcal{G}} \langle 2 \rangle, \langle 1 \rangle \langle 2 \rangle^{\mathcal{G}} \} \langle \tilde{I} 2 \rangle^{\mathcal{G}}$$

(3.12)

where

$$\langle \tilde{I} 2 \rangle^{\mathcal{G}} = \exp \left[ -\frac{R_{12}^2}{\xi_{1\min}^{-1} + \lambda^{-1} + \xi_{2\min}^{-1}} \right]$$

(3.13)

$\xi_{1\min}$ is the smallest effective exponent in the $\langle a_1 b_1 |$ shell-pair (see Algorithm 7) and $R_{12}$ is the distance between a sphere with diameter $\overline{AB}_1$ and another with diameter $\overline{AB}_2$, as defined in Chapter 7, Eq. (8.43). Thus, to identify the significant contracted shell-quartets, we follow the scheme in Algorithm 8. In an extended system with $N$ shells, this generates only $O(N)$ significant quartets.

**Algorithm 8** Identifying significant shell-quartets for computing integrals over $\mathcal{G}$

1: for significant shell-pair $\langle a_1 b_1 |$ do
2: for significant shell-pair $| a_2 b_2 \rangle$ do
3: if $\min\{\langle 1 \rangle^{\mathcal{G}} \langle 2 \rangle, \langle 1 \rangle \langle 2 \rangle^{\mathcal{G}} \} < \tau$ then
4: $\langle a_1 a_2 | b_1 b_2 \rangle$ is not significant
5: else
6: Compute $\langle \tilde{I} 2 \rangle^{\mathcal{G}}$ using Eqs. 3.13
7: if $\min\{\langle 1 \rangle^{\mathcal{G}} \langle 2 \rangle, \langle 1 \rangle \langle 2 \rangle^{\mathcal{G}} \} \langle \tilde{I} 2 \rangle^{\mathcal{G}} < \tau$ then
8: $\langle a_1 a_2 | b_1 b_2 \rangle$ is not significant
9: else
10: $\langle a_1 a_2 | b_1 b_2 \rangle$ is significant
11: end if
12: end if
13: end for
14: end for

10.3.3 Construct [00|00]

For each of the primitive quartets in a significant contracted shell-quartet, we form

$$f_{\xi_1} = \frac{\xi_1^{-1}}{\xi_1^{-1} + \lambda^{-1} + \xi_2^{-1}} \quad f_\lambda = \frac{\lambda^{-1}}{\xi_1^{-1} + \lambda^{-1} + \xi_2^{-1}} \quad f_{\xi_2} = \frac{\xi_2^{-1}}{\xi_1^{-1} + \lambda^{-1} + \xi_2^{-1}}$$

(3.14)
It then follows from Eqs. (2.3) and (2.5) that

\[ [00|00] = G_1 \, G_2 \, f_\lambda^{3/2} \exp \left[ -\frac{|Z_1 - Z_2|^2}{\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1}} \right] \]  

(3.15)

For optimal efficiency, the exponential should be computed via a Chebyshev interpolation.\(^{[367]}\)

If the class angular momentum \(a_1 + a_2 + b_1 + b_2 > 0\), we need a pathway to form the required \(\langle c_1 c_2 | 00 \rangle\) classes. If the degree of contraction \(K_{\text{tot}} = K_{a_1} K_{a_2} K_{b_1} K_{b_2}\) is small, we use the late-contraction path (Section 10.3.4). If \(K_{\text{tot}}\) is large, the early-contraction path (Section 10.3.5) is more economical.

### 10.3.4 Construct \(\langle c_1 c_2 | 00 \rangle\) by late contraction

When \(K_{\text{tot}}\) is small, it is best to build angular momentum and then contract, in the spirit of the Head-Gordon-Pople approach.\(^{[270]}\) We form the Gaussian centroids

\[ R_{A_1} = Z_1 A_1 + f_{\zeta_1} Z_2 Z_1 \quad \quad R_{A_2} = Z_2 A_2 + f_{\zeta_2} Z_1 Z_2 \]  

(3.16)

and the inverse exponents

\[ g_{\zeta_1} = \frac{1 - f_{\zeta_1}}{2\zeta_1} \quad \quad g_{\lambda} = \frac{f_{\zeta_2}}{2\zeta_2} = \frac{f_{\zeta_1}}{2\zeta_1} \quad \quad g_{\zeta_2} = \frac{1 - f_{\zeta_2}}{2\zeta_2} \]  

(3.17)

and then use RRs to transform \([00|00]\) into the required \([c_1 c_2 | 00]\) classes. We call the following algorithm the late path.

To form an \([c_1 c_2 | 00]\) class in which \(c_1 \geq c_2\), we use the 3-term RR\(^{[324]}\)

\[ [c_1^+ c_2 | 00] = R_{A_1} [c_1 c_2 | 00] + c_1 g_{\zeta_1} [c_1^+ c_2 | 00] + c_2 g_{\lambda} [c_1 c_2^+ | 00] \]  

(3.18)

To form an \([c_1 c_2 | 00]\) class in which \(c_1 < c_2\), we use the 3-term RR\(^{[324]}\)

\[ [c_1 c_2^+ | 00] = R_{A_2} [c_1 c_2 | 00] + c_2 g_{\zeta_2} [c_2 c_1^- | 00] + c_2 g_{\zeta_2} [c_1 c_2^- | 00] \]  

(3.19)

Finally, the \([c_1 c_2 | 00]\) are contracted into \(\langle c_1 c_2 | 00 \rangle\) using the simple sum

\[ \langle c_1 c_2 | 00 \rangle = \sum_{i=1}^{K_{a_1}} \sum_{j=1}^{K_{b_1}} \sum_{k=1}^{K_{a_2}} \sum_{l=1}^{K_{b_2}} [c_1 c_2 | 00] \]  

(3.20)
Table 10.1: Cost (multiplies + additions) of forming an \([c_1c_2|00]\) class using Eqs. 3.18 and 3.19 in terms of the primitive shell pairs \([c_2|0]\) (ket) and \([c_1|0]\) (bra).

<table>
<thead>
<tr>
<th></th>
<th>60</th>
<th>50</th>
<th>40</th>
<th>30</th>
<th>20</th>
<th>10</th>
<th>00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>41 + 13</td>
<td>138 + 54</td>
<td>303 + 135</td>
<td>537 + 257</td>
<td>840 + 420</td>
<td>1212 + 624</td>
<td>1653 + 869</td>
</tr>
<tr>
<td></td>
<td>30 + 9</td>
<td>102 + 39</td>
<td>225 + 99</td>
<td>399 + 189</td>
<td>624 + 309</td>
<td>900 + 459</td>
<td>1212 + 624</td>
</tr>
<tr>
<td></td>
<td>21 + 6</td>
<td>72 + 27</td>
<td>159 + 69</td>
<td>282 + 132</td>
<td>441 + 216</td>
<td>624 + 309</td>
<td>840 + 420</td>
</tr>
<tr>
<td></td>
<td>13 + 3</td>
<td>48 + 18</td>
<td>105 + 45</td>
<td>185 + 85</td>
<td>282 + 132</td>
<td>399 + 189</td>
<td>537 + 257</td>
</tr>
<tr>
<td></td>
<td>9 + 3</td>
<td>30 + 12</td>
<td>66 + 30</td>
<td>105 + 45</td>
<td>159 + 69</td>
<td>225 + 99</td>
<td>303 + 135</td>
</tr>
<tr>
<td></td>
<td>3 + 0</td>
<td>12 + 3</td>
<td>30 + 12</td>
<td>48 + 18</td>
<td>72 + 27</td>
<td>102 + 39</td>
<td>138 + 54</td>
</tr>
<tr>
<td></td>
<td>0 + 0</td>
<td>3 + 0</td>
<td>9 + 3</td>
<td>13 + 3</td>
<td>21 + 6</td>
<td>30 + 9</td>
<td>41 + 13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are many ways to use Eqs. 3.18 and 3.19 because, in general, an \([c_1c_2|00]\) integral can be formed in three ways, viz. by incrementing the \(x\), \(y\) or \(z\) component of its angular momentum. The pursuit of an optimal strategy to form the required \([c_1c_2|00]\) classes therefore leads to a optimization problem related to the McMurchie-Davidson tree-search problem.\[398\] Optimal flop costs (decomposed into multiplies and additions) for forming various \([c_1c_2|00]\) classes are shown in Table 10.1.

### 10.3.5 Construct \(\langle c_1c_2|00\rangle\) by early contraction

When \(K_{\text{tot}}\) is large, it is better to contract the \([00|00]\) and then build angular momentum, in the spirit of the Pople-Hehre axis switch method\[265\] and the CCTTT path in the PRISM algorithm.\[20\] To achieve this, we substitute the identities\[20\]

\[
Z_1A_1 = B_1A_1 \frac{2\beta_1}{2\zeta_1} \quad (3.21a)
\]
\[
Z_2A_2 = B_2A_2 \frac{2\beta_2}{2\zeta_2} \quad (3.21b)
\]
\[
Z_1Z_2 = B_1A_1 \frac{2\beta_1}{2\zeta_1} + A_2B_2 \frac{2\beta_2}{2\zeta_2} + A_1A_2 \quad (3.21c)
\]

201
into the primitive RRs (Eqs. 3.18 and 3.19) to obtain the “contraction-ready” RRs

\[
[c_1^+ c_2^+ |00] = \left[ B_1 A_1 \frac{2\beta_1}{2\xi_1} (1 - f_{\xi_1}) + B_2 A_2 \frac{2\beta_2}{2\xi_2} f_{\xi_1} + A_2 A_1 f_{\xi_1} \right] [c_1 c_2 |00]
+ c_1 \frac{1 - f_{\xi_1}}{2\xi_1} [c_1^+ c_2 |00] + c_2 f_{\xi_1} [c_1 c_2^+ |00]
\]

(3.22)

\[
[c_1 c_2^+ |00] = \left[ B_1 A_1 \frac{2\beta_2}{2\xi_2} (1 - f_{\xi_2}) + B_1 A_1 \frac{2\beta_1}{2\xi_1} f_{\xi_2} + A_1 A_2 f_{\xi_2} \right] [c_1 c_2 |00]
+ c_1 f_{\xi_2} [c_1 c_2^+ |00] + c_2 \frac{1 - f_{\xi_2}}{2\xi_2} [c_1 c_2^+ |00]
\]

(3.23)

Subtracting Eq. 3.23 from Eq. 3.22 yields a symmetrical RR

\[
[c_1^+ c_2 |00] - [c_1 c_2^+ |00] = \left[ B_1 A_1 \frac{2\beta_1}{2\xi_1} f_{\lambda} + A_2 B_2 \frac{2\beta_2}{2\xi_2} f_{\lambda} + A_2 A_1 (1 - f_{\lambda}) \right] [c_1 c_2 |00]
+ c_1 f_{\lambda} [c_1^+ c_2 |00] - c_2 f_{\lambda} [c_1 c_2^+ |00]
\]

(3.24)

that relates an integral to its four neighbours and involves \(f_{\lambda}\), rather than \(f_{\xi_1}\) or \(f_{\xi_2}\).

Contracted 7-term RRs emerge from Eqs. (3.22), (3.23) and (3.24) by substituting the identities \(f_{\xi_1} = (2\lambda / 2\xi_1)f_{\lambda}\) and \(f_{\xi_2} = (2\lambda / 2\xi_2)f_{\lambda}\) and summing over primitives[397]. This leads to the following algorithm which we call the early path.

First, \([00 |00]|00\) are half-contracted into \([00 |00]|00\) using the scaled ket contraction[396]

\[
[00 |00]|00\rangle_{dq} = \sum_{k=1}^{K_{a_2}} \sum_{l=1}^{K_{b_2}} \frac{(2\beta_2)^d}{(2\xi_2)^q} f_{\lambda} [00 |00]\] (3.25)

Then, \([00 |00]|00\) are fully contracted into \([00 |00]|00\) using the scaled bra contraction[396]

\[
bp [00 |00]|00\rangle_{dq} = \sum_{k=1}^{K_{a_1}} \sum_{j=1}^{K_{b_1}} \frac{(2\beta_1)^b}{(2\xi_1)^p} [00 |00]|00\rangle_{dq} (3.26)
\]

Then, \([00 |00]|00\) are transformed into \(c_1^+ |00|00\) using the 6-term RR (from Eq. 3.22)

\[
bp (c_1^+ |00|00)_{dq} = B_1 A_1 (b+1)(p+1) (c_1 |00|00)_{dq} - 2\lambda B_1 A_1 (b+1)(p+2) (c_1 |00|00)_{dq}^{r+1}
+ 2\lambda B_2 A_1 b(p+1) (c_1 |00|00)_{dq}^{r+1}(d+1)(q+1) + 2\lambda A_2 A_1 b(p+1) (c_1 |00|00)_{dq}^{r+1}
+ c_1 b(p+1) (c_1 |00|00)_{dq} - 2\lambda c_1 b(p+2) (c_1^+ |00|00)_{dq}^{r+1}
\]

(3.27)
CHAPTER 10. TWO-ELECTRON INTEGRALS OVER GAUSSIAN GEMINALS

Table 10.2: Cost (multiplies + additions) of forming an \( c_1 c_2 |00\) class using Eqs. 3.27 and 3.28 in terms of the contracted shell pairs \(|c_20\rangle\) (ket) and \(|c_10\rangle\) (bra).

<table>
<thead>
<tr>
<th></th>
<th>(00)</th>
<th>(10)</th>
<th>(20)</th>
<th>(30)</th>
<th>(40)</th>
<th>(50)</th>
<th>(60)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>118 + 125</td>
<td>354 + 396</td>
<td>723 + 813</td>
<td>1231 + 1383</td>
<td>1884 + 2106</td>
<td>2682 + 2982</td>
<td>3626 + 4011</td>
</tr>
<tr>
<td>50</td>
<td>87 + 93</td>
<td>261 + 297</td>
<td>537 + 609</td>
<td>915 + 1035</td>
<td>1401 + 1575</td>
<td>1995 + 2229</td>
<td>2697 + 2997</td>
</tr>
<tr>
<td>40</td>
<td>63 + 66</td>
<td>189 + 207</td>
<td>387 + 429</td>
<td>657 + 732</td>
<td>1002 + 1116</td>
<td>1425 + 1581</td>
<td>1926 + 2127</td>
</tr>
<tr>
<td>30</td>
<td>43 + 43</td>
<td>129 + 138</td>
<td>261 + 285</td>
<td>445 + 485</td>
<td>681 + 738</td>
<td>969 + 1044</td>
<td>1310 + 1403</td>
</tr>
<tr>
<td>20</td>
<td>24 + 27</td>
<td>72 + 84</td>
<td>147 + 174</td>
<td>249 + 297</td>
<td>381 + 453</td>
<td>543 + 642</td>
<td>735 + 864</td>
</tr>
<tr>
<td>10</td>
<td>12 + 12</td>
<td>36 + 39</td>
<td>75 + 81</td>
<td>129 + 138</td>
<td>198 + 210</td>
<td>282 + 297</td>
<td>381 + 399</td>
</tr>
<tr>
<td>00</td>
<td>0 + 0</td>
<td>12 + 9</td>
<td>27 + 24</td>
<td>46 + 36</td>
<td>69 + 57</td>
<td>96 + 81</td>
<td>131 + 110</td>
</tr>
</tbody>
</table>

Finally, \( c_1|00\rangle\) are transformed into \( c_1 c_2 |00\rangle\) using the 7-term RR (from Eq. 3.24)

\[
bp \langle c_1 c_2^+ |00\rangle_d = bp \langle c_1^+ c_2 |00\rangle_d + A_1 B_1 (b+1)(p+1) \langle c_1 c_2 |00\rangle_d^{r+1}
+ B_2 A_1 bp \langle c_1 c_2 |00\rangle_d^{(d+1)(q+1)} + A_1 A_2 bp \langle c_1 c_2 |00\rangle_d^{r+1}
- A_1 A_2 bp \langle c_1 c_2 |00\rangle_d^{r+1} - c_1 b(p+1) \langle c_1^- c_2 |00\rangle_d^{r+1}
+ c_2 bp \langle c_1 c_2^- |00\rangle_d^{r+1}
\]

Optimal flop costs for forming various \( c_1 c_2 |00\rangle\) classes in this way are shown in Table 10.2.

10.3.6 Construct \( a_1 a_2 |b_1 b_2\rangle\)

Finally, we transform \( c_1 c_2 |00\rangle\) into \( a_1 c_2 |b_1 0\rangle\) and thence into \( a_1 a_2 |b_1 b_2\rangle\) using the 2-term RRs[268, 270, 399]

\[
\langle a_1 c_2 | b_1^+ 0 \rangle = \langle a_1^+ c_2 | b_1 0 \rangle + A_1 B_1 \langle a_1 c_2 | b_1 0 \rangle
\]

\[
\langle a_1 a_2 | b_1 b_2^+ \rangle = \langle a_1 a_2^+ | b_1 b_2 \rangle + A_2 B_2 \langle a_1 a_2 | b_1 b_2 \rangle
\]

10.4 Examples: Forming a \( pp |pp\rangle\) Class

We now consider in detail the steps required to form a \( pp|pp\rangle\) class from shell-pair data. We assume, as usual, that each shell has degree of contraction \( K\). We measure the cost of each step by the number of flops (floating-point operations) that it requires. Although it is known[403, 399] that the number of memory operations (mops) is often just as important as the number of flops, the flop cost remains a useful comparator between different algorithms. We assume that all useful shell-pair data (e.g. \( 2\beta_2/2\zeta_2\),

\[203\]
$G_1, Z_1 A_1$, etc.) have been pre-computed.

### 10.4.1 The late path

Figure 10.1 shows how the late path builds angular momentum and Table 10.3 lists all of the steps needed. In addition to $K^4$ divides, square roots and exponentials, the \( \langle pp|pp \rangle \) flop cost is

$$C_{pppp}^{\text{late}} = 329K^4 + 324 \quad (4.31)$$

Figure 10.1 shows that the \([20|00]\) class can be avoided for \( \langle pp|pp \rangle \) construction, illustrating a general feature of the strategy in Section 10.3.4, which avoids the construction of unnecessary intermediate classes. This saving grows with the angular momentum of the target class, as can be seen in Figures 10.2 and 10.3 in the Appendix.

![Diagram](image)

Figure 10.1: The \([c_1 c_2|00]\) classes needed to form a \( \langle pp|pp \rangle \) class on the late path.

Table 10.3 reveals that building angular momentum on $A_2$ (eq 3.19) represents two-thirds of the primitive work and contraction (Eq. 3.20) only a quarter. We also see that the horizontal recurrence relations (HRRs), Eqs. 3.29 and 3.30, perform half of the work to form a contracted \( \langle pp|pp \rangle \) class but that, as $K$ increases, their contribution becomes a minor component.

### 10.4.2 The early path

Table 10.4 shows the \( b_p \langle 00|00 \rangle_{dq}^r \) required on the early path and Table 10.5 lists all of the steps required. In addition to $K^4$ divides, square roots and exponentials, the \( \langle pp|pp \rangle \) flop cost is

$$C_{pppp}^{\text{early}} = 59K^4 + 356K^2 + 6455K^0 \quad (4.32)$$
Table 10.3: Steps and flop costs required to form a $<pp|pp>$ class on the late path. A, M, D, S, E are add, multiply, divide, square root and exponential, respectively.

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.14</td>
<td>$1/(\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1})$</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3.14</td>
<td>$f_{\zeta_1}, f_3, f_{\zeta_2}$</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3.15</td>
<td>$f_{3/2}$</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3.15</td>
<td>$Z_1Z_2$</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>3.15</td>
<td>$</td>
<td>Z_1Z_2</td>
</tr>
<tr>
<td>6</td>
<td>3.15</td>
<td>$[00</td>
<td>00]$</td>
</tr>
<tr>
<td>7</td>
<td>3.16</td>
<td>$R_{A_1}, R_{A_2}$</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>3.17</td>
<td>$g_{\zeta_1}, g_3, g_{\zeta_2}$</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>3.18</td>
<td>$[10</td>
<td>00]$</td>
</tr>
<tr>
<td>10</td>
<td>3.19</td>
<td>$[01</td>
<td>00], [11</td>
</tr>
<tr>
<td>11</td>
<td>3.19</td>
<td>$[02</td>
<td>00], [12</td>
</tr>
<tr>
<td>12</td>
<td>3.20</td>
<td>$\langle 22</td>
<td>00 \rangle, \langle 12</td>
</tr>
</tbody>
</table>

**Total $K^4$ work** 173 156 1 1 1

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>3.29</td>
<td>$\langle 12</td>
<td>10 \rangle, \langle 11</td>
</tr>
<tr>
<td>14</td>
<td>3.30</td>
<td>$\langle 11</td>
<td>11 \rangle$</td>
</tr>
</tbody>
</table>

**Total $K^0$ work** 162 162

Table 10.4 shows that, whereas $<pp|pp>$ requires 24 distinct bra scalings (rows), it needs only 8 ket scalings (columns). This illustrates a general feature of the strategy in Section 10.3.5, which first builds on $A_1$ using Eq. 3.27 and later on $A_2$ using Eq. 3.28, creating a strong asymmetry in the number of required bra and ket scalings. This decreases the cost of the ket contraction (Eq. 3.25) but increases the cost of the bra contraction (eq 3.26). These changes create the tiny $K^4$ coefficient, significantly larger $K^2$ coefficient and huge $K^0$ coefficient in eq. 4.32.

An $<aa|aa>$ class consists of $O(a^8)$ integrals and yet one can show that, on the early path, there are only $(a + 1)(3a + 1)$ ket scalings and the number of ket-contracted $[00|00]\_{dq}$ is only

$$\text{No. of } |00|00\rangle_{dq} = \frac{(a + 1)(4a + 3)(7a + 2)}{6}$$  \hspace{1cm} (4.33)$$

For this reason, the early path does remarkably little work at the primitive level (Table 10.5). As each primitive $|00|00\rangle$ is formed by Eq. 3.15, it is immediately contracted into the required $|00|00\rangle_{dq}$ integrals using Eq. 3.25. If done efficiently (Algorithm 9), each ket-contraction requires only a single multiply and add. The subsequent bra-contraction
Table 10.4: The 178 $v_p \langle 00 \rangle_{dq}^r$ integrals required to form a $\langle pp|pp \rangle$ class on the early path. The row index is $(b,p)$, the column index is $(d,q)$, and the entries in the body are the required $r$ values. The final row shows the 21 half-contracted $\langle 00 \rangle_{dq}^r$ integrals that are required.

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(1,1)</th>
<th>(2,2)</th>
<th>(3,3)</th>
<th>(4,4)</th>
<th>(0,1)</th>
<th>(1,2)</th>
<th>(2,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td>0,1,2</td>
<td>1,2</td>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2,2)</td>
<td>0,1,2</td>
<td>1,2</td>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,3)</td>
<td>0,1,2</td>
<td>1,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4,4)</td>
<td>0,1,2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0,1)</td>
<td>0,1,2,3</td>
<td>1,2,3</td>
<td>2,3</td>
<td>3</td>
<td>1,2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,2)</td>
<td>0,1,2,3</td>
<td>1,2,3</td>
<td>2,3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2,3)</td>
<td>0,1,2,3</td>
<td>1,2,3</td>
<td>2,3</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,4)</td>
<td>1,2,3</td>
<td>1,2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4,5)</td>
<td>1,2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0,2)</td>
<td>0,1,2,3,4</td>
<td>1,2,3,4</td>
<td>2,3,4</td>
<td>3,4</td>
<td>4</td>
<td>2,3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(1,3)</td>
<td>1,2,3,4</td>
<td>1,2,3,4</td>
<td>2,3,4</td>
<td>3,4</td>
<td></td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>(2,4)</td>
<td>1,2,3,4</td>
<td>2,3,4</td>
<td>2,3,4</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,5)</td>
<td>2,3,4</td>
<td>2,3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4,6)</td>
<td>2,3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0,3)</td>
<td>1,2,3,4</td>
<td>2,3,4</td>
<td>2,3,4</td>
<td>3,4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,4)</td>
<td>2,3,4</td>
<td>2,3,4</td>
<td>3,4</td>
<td>3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2,5)</td>
<td>2,3,4</td>
<td>3,4</td>
<td>3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,6)</td>
<td>3,4</td>
<td>3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4,7)</td>
<td>3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0,4)</td>
<td>2,3,4</td>
<td>3,4</td>
<td>3,4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,5)</td>
<td>3,4</td>
<td>3,4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2,6)</td>
<td>3,4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3,7)</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4,8)</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>0,1,2,3,4</td>
<td>1,2,3,4</td>
<td>2,3,4</td>
<td>3,4</td>
<td>4</td>
<td>1,2,3</td>
<td>2,3</td>
<td>3</td>
</tr>
</tbody>
</table>
Algorithm 9 Forming the \( |00\rangle|00\rangle \) integrals required for \( \langle pp|pp \rangle \) on the early path.

1. for bra primitive shell-pair do
2.   Initialize all \(|00\rangle|00\rangle_{dq}\) to zero
3.   for ket primitive shell-pair do
4.     \( s_{01} = 1/2\xi_2 \) and \( s_{11} = 2\beta_2/2\zeta_2 \) (pre-computed)
5.     Compute \( f_\lambda \) using Eq. 3.14
6.     Compute \(|00\rangle|00\rangle\) using Eq. 3.15
7.     \( T_{00} = |00\rangle|00\rangle_0 \); Add \( T_{00} \) to \(|00\rangle|00\rangle_0\)
8.     \( T_{10} = T_{00} * f_\lambda \); Add \( T_{10} \) to \(|00\rangle|00\rangle_{10}\)
9.     \( T_{20} = T_{00} * f_\lambda \); Add \( T_{20} \) to \(|00\rangle|00\rangle_{20}\)
10. \( T_{30} = T_{00} * f_\lambda \); Add \( T_{30} \) to \(|00\rangle|00\rangle_{30}\)
11. \( T_{40} = T_{00} * f_\lambda \); Add \( T_{40} \) to \(|00\rangle|00\rangle_{40}\)
12. \( T_{11} = T_{10} * s_{11} \); Add \( T_{11} \) to \(|00\rangle|00\rangle_{11}\)
13. \( T_{12} = T_{11} * f_\lambda \); Add \( T_{12} \) to \(|00\rangle|00\rangle_{22}\)
14. \( T_{13} = T_{12} * f_\lambda \); Add \( T_{13} \) to \(|00\rangle|00\rangle_{33}\)
15. \( T_{14} = T_{13} * f_\lambda \); Add \( T_{14} \) to \(|00\rangle|00\rangle_{44}\)
16. \( T_{22} = T_{21} * s_{11} \); Add \( T_{22} \) to \(|00\rangle|00\rangle_{22}\)
17. \( T_{23} = T_{22} * f_\lambda \); Add \( T_{23} \) to \(|00\rangle|00\rangle_{33}\)
18. \( T_{24} = T_{23} * f_\lambda \); Add \( T_{24} \) to \(|00\rangle|00\rangle_{44}\)
19. \( T_{33} = T_{32} * s_{11} \); Add \( T_{33} \) to \(|00\rangle|00\rangle_{33}\)
20. \( T_{34} = T_{33} * f_\lambda \); Add \( T_{34} \) to \(|00\rangle|00\rangle_{44}\)
21. \( T_{44} = T_{43} * s_{11} \); Add \( T_{44} \) to \(|00\rangle|00\rangle_{44}\)
22. \( T_{21} = T_{00} * s_{01} \); Add \( T_{21} \) to \(|00\rangle|00\rangle_{21}\)
23. \( T_{22} = T_{00} * s_{01} \); Add \( T_{22} \) to \(|00\rangle|00\rangle_{22}\)
24. \( T_{23} = T_{00} * s_{01} \); Add \( T_{23} \) to \(|00\rangle|00\rangle_{23}\)
25. \( T_{24} = T_{00} * s_{01} \); Add \( T_{24} \) to \(|00\rangle|00\rangle_{24}\)
26. \( T_{33} = T_{00} * s_{01} \); Add \( T_{33} \) to \(|00\rangle|00\rangle_{33}\)
27. \( T_{34} = T_{00} * s_{01} \); Add \( T_{34} \) to \(|00\rangle|00\rangle_{34}\)
28. \( T_{44} = T_{00} * s_{01} \); Add \( T_{44} \) to \(|00\rangle|00\rangle_{44}\)
29. end for
30. end for

10.5 Computational Costs

As shown in the Examples above, the flop cost to form a given class by a given algorithm is

\[
\text{Cost} = x K^4 + y K^2 + z K^0
\]  

(5.34)

The \( x \), \( y \) and \( z \) parameters have been discussed in many previous papers and, in Table 10.6, we give their values for forming \( \langle pp|pp \rangle \), \( \langle dd|dd \rangle \) and \( \langle ff|ff \rangle \) classes by the
10.5. COMPUTATIONAL COSTS

Table 10.5: Steps and flop costs required to form a \(pp|pp\) class on the early path.

A, M, D, S, E are add, multiply, divide, square root and exponential, respectively.

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.14</td>
<td>(1/(\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1}))</td>
<td>M 2</td>
</tr>
<tr>
<td></td>
<td>3.14</td>
<td>(f_{\zeta_1}, f_{\zeta_2})</td>
<td>A 1</td>
</tr>
<tr>
<td>3</td>
<td>3.15</td>
<td>(f_{\lambda}^{3/2})</td>
<td>D 1</td>
</tr>
<tr>
<td>4</td>
<td>3.15</td>
<td>(Z_1Z_2)</td>
<td>S 1</td>
</tr>
<tr>
<td>5</td>
<td>3.15</td>
<td>(</td>
<td>Z_1Z_2</td>
</tr>
<tr>
<td>6</td>
<td>3.15</td>
<td>([00][00])</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3.25</td>
<td>([00][00]r_{dq})</td>
<td>21</td>
</tr>
</tbody>
</table>

Total \(K^4\) work \(31\ 28\ 1\ 1\ 1\)

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3.26</td>
<td>(b_p \langle 00\rangle [00]_{dq})</td>
<td>M 178</td>
</tr>
</tbody>
</table>

Total \(K^2\) work \(178\ 178\)

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>3.27</td>
<td>(b_p \langle a0\rangle [00]_{dq})</td>
<td>M 2411</td>
</tr>
</tbody>
</table>

Total \(K^0\) work \(3413\ 3042\)

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3.28</td>
<td>(\langle 22\rangle [00], \langle 12\rangle [00], \langle 21\rangle [00], \langle 11\rangle [00])</td>
<td>M 840</td>
</tr>
<tr>
<td>11</td>
<td>3.29</td>
<td>(\langle 12\rangle [10], \langle 11\rangle [10])</td>
<td>M 81</td>
</tr>
<tr>
<td>12</td>
<td>3.30</td>
<td>(\langle 11\rangle [11])</td>
<td>M 81</td>
</tr>
</tbody>
</table>

Total \(K^0\) work \(3413\ 3042\)

Table 10.6: Flop-cost parameters for forming integral classes by present and previous algorithms.

<table>
<thead>
<tr>
<th>Class</th>
<th>Flop-cost parameter</th>
<th>Present algorithms</th>
<th>Previous algorithms[352]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early</td>
<td>Late</td>
<td>PH</td>
</tr>
<tr>
<td>(pp</td>
<td>pp)</td>
<td>x</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>356</td>
<td>2,300</td>
</tr>
<tr>
<td></td>
<td>z</td>
<td>6,455</td>
<td>4,000</td>
</tr>
<tr>
<td>(dd</td>
<td>dd)</td>
<td>x</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>4,130</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>z</td>
<td>273,736</td>
<td>11,256</td>
</tr>
<tr>
<td>(ff</td>
<td>ff)</td>
<td>x</td>
<td>477</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>20,986</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>z</td>
<td>3,574,460</td>
<td>135,024</td>
</tr>
</tbody>
</table>

208
early and late paths described above, comparing these with Pople-Hehre (PH), Head-Gordon-Pople (HGP) and McMurchie-Davidson (MD) algorithms.

For all classes studied, the late path is systematically cheaper than the HGP and MD algorithms. For $K_{tot} = 1$, the late path is roughly twice as fast as HGP and, for large $K_{tot}$, the late path is between 2.75 and 4.5 times faster. The early path is uncompetitive for small $K_{tot}$, but quickly overtakes all other algorithms as $K_{tot}$ increases. It outperforms the late path if $K_{tot} \geq 30$ for $\langle pp|pp\rangle$, if $K_{tot} \geq 81$ for $\langle dd|dd\rangle$, or if $K_{tot} \geq 158$ for $\langle ff|ff\rangle$.

The efficiency of the early path for strongly contracted classes is even more attractive for calculations in which Eq. 1.1 is replaced by the contracted Gaussian Geminal

$$G = \sum_{m=1}^{K_G} c_m \exp(-\lambda_m r_{12}^2)$$

The total degree of contraction is now $K_{tot} = K_a^1 K_a^2 K_{b_1} K_{b_2} \times K_G$ and the ket contraction (Eq. 3.25) naturally generalizes to include the Geminal contraction, becoming

$$|00\rangle_{dq} = \sum_{k=1}^{K_{a_2}} \sum_{l=1}^{K_{a_2}} \sum_{m=1}^{K_G} (2\beta_2)^d (2\zeta_2)^p f_{dq}^k |00\rangle$$

It is therefore immediately clear that calculations using contracted Gaussian Geminals with $K_G \gg 1$ will benefit greatly from the early path.

### 10.6 Concluding Remarks

We have presented an efficient algorithm to construct two-electron integrals over a Gaussian Geminal operator. Our method employs vertical and horizontal RRs in the spirit of the Head-Gordon-Pople approach and offers late- and early-contraction paths in the PRISM style. Flop-cost parameters reveal that the new algorithm is much cheaper computationally than previous schemes. We therefore expect that efficient implementations of the new approach will significantly reduce the cost of calculations involving such integrals. These include the CAP($n$), LCgau, LC2gau and F12 calculations where Gaussian Geminals are used.
10.7 Appendix: Forming $\langle dd|dd \rangle$ & $\langle ff|ff \rangle$ on the Late Path

![Diagram showing the process of forming $\langle dd|dd \rangle$ and $\langle ff|ff \rangle$ classes on the late path.]

Figure 10.2: The [$c_1c_2|00$] classes needed to form a $\langle dd|dd \rangle$ class on the late path.

![Diagram showing the process of forming $\langle ff|ff \rangle$ classes on the late path.]

Figure 10.3: The [$c_1c_2|00$] classes needed to form an $\langle ff|ff \rangle$ class on the late path.
Table 10.7: Steps and flop-costs required to form a \( \langle dd|dd \rangle \) class on the late path.
A, M, D, S, E are add, multiply, divide, square root and exponential, respectively.

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>M</th>
<th>A</th>
<th>D</th>
<th>S</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.14</td>
<td>( 1/(\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1}) )</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3.14</td>
<td>( f_{\zeta_1}, f_{\lambda}, f_{\zeta_2} )</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.15</td>
<td>( f_{\lambda}^{3/2} )</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.15</td>
<td>( Z_1Z_2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>3.15</td>
<td>(</td>
<td>Z_1Z_2</td>
<td>^2/(\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1}) )</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.15</td>
<td>( [00</td>
<td>00] )</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>3.16</td>
<td>( R_{A1}, R_{A2} )</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3.17</td>
<td>( g_{\zeta_1}, g_{\lambda}, g_{\zeta_2} )</td>
<td></td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3.18</td>
<td>( {0,1,2}0</td>
<td>00 )</td>
<td>12</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3.19</td>
<td>( {0,1,2,3}1</td>
<td>00 )</td>
<td>93</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3.19</td>
<td>( {0,1,2,3,4}2</td>
<td>00 )</td>
<td>369</td>
<td>159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3.19</td>
<td>( {1,2,3,4}3</td>
<td>00 )</td>
<td>620</td>
<td>280</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>3.19</td>
<td>( {2,3,4}4</td>
<td>00 )</td>
<td>882</td>
<td>417</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>3.20</td>
<td>( \langle 22</td>
<td>00 \rangle, \ldots, \langle 44</td>
<td>00 \rangle )</td>
<td>961</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Total ( K^4 ) work</strong></td>
<td>1996</td>
<td>1868</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>3.29</td>
<td>( \langle 2,3\rangle{2,3,4}</td>
<td>10 )</td>
<td>1488</td>
<td>1488</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>3.29</td>
<td>( \langle 2{2,3,4}</td>
<td>20 )</td>
<td>1116</td>
<td>1116</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>3.30</td>
<td>( \langle 2{2,3}</td>
<td>21 )</td>
<td>1728</td>
<td>1728</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>3.30</td>
<td>( \langle 22</td>
<td>22 \rangle )</td>
<td>1296</td>
<td>1296</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Total ( K^0 ) work</strong></td>
<td>5628</td>
<td>5628</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
10.7. APPENDIX: FORMING $\langle DD|DD \rangle$ & $\langle FF|FF \rangle$ ON THE LATE PATH

Table 10.8: Steps and flop-costs required to form a $\langle ff|ff \rangle$ class on the late path. A, M, D, S, E are add, multiply, divide, square root and exponential, respectively.

<table>
<thead>
<tr>
<th>Step</th>
<th>Equation</th>
<th>Computed quantity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>1</td>
<td>3.14</td>
<td>$1/(\zeta_1^{-1} + \lambda^{-1} + \zeta_2^{-1})$</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3.14</td>
<td>$f_{\zeta_1}, f_{\lambda}, f_{\zeta_2}$</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3.15</td>
<td>$f_{\lambda}^{3/2}$</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3.15</td>
<td>$\mathbf{Z}_1 \mathbf{Z}_2$</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>3.15</td>
<td>$</td>
<td>\mathbf{Z}_1 \mathbf{Z}_2</td>
</tr>
<tr>
<td>6</td>
<td>3.15</td>
<td>$[00</td>
<td>00]$</td>
</tr>
<tr>
<td>7</td>
<td>3.16</td>
<td>$\mathbf{R}<em>{A_1}, \mathbf{R}</em>{A_2}$</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>3.17</td>
<td>$g_{\zeta_1}, g_{\lambda}, g_{\zeta_2}$</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>3.18</td>
<td>${0,1,2,3}0</td>
<td>00$</td>
</tr>
<tr>
<td>10</td>
<td>3.19</td>
<td>${0,1,2,3,4,5}1</td>
<td>00$</td>
</tr>
<tr>
<td>11</td>
<td>3.19</td>
<td>${0,1,2,3,4,5}2</td>
<td>00$</td>
</tr>
<tr>
<td>12</td>
<td>3.19</td>
<td>${0,1,2,3,4,5,6}3</td>
<td>00$</td>
</tr>
<tr>
<td>13</td>
<td>3.19</td>
<td>${1,2,3,4,5,6}4</td>
<td>00$</td>
</tr>
<tr>
<td>14</td>
<td>3.19</td>
<td>${2,3,4,5,6}5</td>
<td>00$</td>
</tr>
<tr>
<td>15</td>
<td>3.19</td>
<td>${3,4,5,6}6</td>
<td>00$</td>
</tr>
<tr>
<td>16</td>
<td>3.20</td>
<td>$\langle 33</td>
<td>00 \rangle, \ldots, \langle 66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>3.29</td>
<td>$\langle 3,4,5\rangle, {3,4,5,6}</td>
<td>10$</td>
</tr>
<tr>
<td>18</td>
<td>3.29</td>
<td>$\langle 3,4\rangle, {3,4,5,6}</td>
<td>20$</td>
</tr>
<tr>
<td>19</td>
<td>3.29</td>
<td>${3,4,5,6}</td>
<td>30$</td>
</tr>
<tr>
<td>20</td>
<td>3.30</td>
<td>$\langle 33</td>
<td>33 \rangle$</td>
</tr>
<tr>
<td>21</td>
<td>3.30</td>
<td>$\langle 3,4\rangle</td>
<td>32$</td>
</tr>
<tr>
<td>22</td>
<td>3.30</td>
<td>${33</td>
<td>33 \rangle$</td>
</tr>
</tbody>
</table>

**Total $K^4$ work** | 12393 | 11567 | 1 | 1 | 1 |

**Total $K^0$ work** | 67512 | 67512 |
Chapter 11

Three- and Four-Electron Integrals Involving Gaussian Geminals

"Out of intense complexities, intense simplicities emerge."

Winston Churchill — *The World Crisis*, Chapter XLII, page 623

11.1 Introduction

In Chapter 8 and 9, we have presented algorithms for the evaluation of three- and four-electron integrals arising in explicitly correlated methods when generic multiplicative operators of the form $f_{12} = f(|r_1 - r_2|)$ are adopted as correlation factors. In Chapter 10, we have shown how two-electron integrals over Gaussian Geminals, $G_{12}$, can be computed via a combination of consistent screening (that is, only the significant integrals are evaluated) and recursion techniques, much more cheaply than in any other previous scheme. Such an enhancement is possible for two main reasons: i) the short-range nature of $G_{12}$, ii) the unique factorisation properties of the primitive integrals $[a_1a_2|G_{12}|b_1b_2]$ over Gaussian basis functions. In this last Chapter, we show that, because of i) and
ii), an analogous improvement is obtained in the evaluation of three- and four-electron integrals if the correlation factor is a Gaussian Geminal, that is \( f_{12} = g_{12} \).

As mentioned in Chapter 6, methods for evaluating many-electron integrals involving GGs have already been developed. Persson and Taylor\[246\] derived recurrence relations based on Hermite Gaussians, analogously to the work of McMurchie and Davidson for two-electron integrals.[264] These recurrence relations were implemented by Dahle.[338, 336, 337] Saito and Suzuki\[365\] also proposed an approach based on the work by Obara and Saika.[269, 368] More recently, a general formulation using Rys polynomials\[266, 267, 268\] was published by Komornicki and King.[324] Even if limited to the three-center case, it is worth mentioning that May has also developed recurrence relations for two types of three-electron integrals.[404] These recurrence relations were implemented by Womack using automatically-generated code.[405]

A major limitation of all these approaches is that they do not include any integral screening.* In fact, as discussed in Chapter 7, a remarkable consequence of the short-range nature of the Slater and Gaussian correlation factors is that, even if formally scaling as \( \mathcal{O}(N^6) \) and \( \mathcal{O}(N^8) \), there are only \( \mathcal{O}(N^2) \) significant three- and four-electron integrals in a large system. Therefore, it is paramount to devise rigorous upper bounds to avoid computing the large number of negligible integrals.

As we will show later, another, usually neglected, important feature of many-electron integrals involving GGs, is that they exhibit a common mathematical structure such that they can be all cast in terms of a unified theory. This enables considerable simplifications both in the formulae and possibly in the implementation.

The present Chapter is organized as follows. In Section 11.2, we discuss the structure of the three- and four-electron operators considered here. The next three sections contain the main ingredients for the efficient computation of three- and four-electron integrals involving GGs: i) fundamental integrals (FIs) in Section 11.3, ii) upper bounds (UBs) in Section 11.4, and iii) recurrence relations (RRs) in Section 11.5. In Section 11.6, we give an overall view of our algorithm which is an extension of the late-contraction path of PRISM (see Refs. [270, 20] and references therein). Note that the RRs developed in this study differ from the ones reported in our previous Chapters, as they are specifically

---

*Komornicki and King mentioned the crucial importance of an effective integral screening in Ref. [324].
11.2 Three- and Four-Electron Integrals

11.2.1 Three- and four-electron operators

In this study, we are particularly interested in the “master” four-electron operator $C_{12}G_{13}G_{14}G_{23}G_{34}$ (where $C_{12} = r_{12}^{-1}$ is the Coulomb operator) because the three types of three-electron integrals and the three types of four-electron integrals that can be tailored for the unique factorization properties brought by the association of Gaussian basis functions and GGs.
required in F12 calculations can be easily generated from it (see Fig. 11.1). These three types of three-electron integrals are composed by a single type of integrals over the cyclic operator $C_{12}G_{13}G_{23}$, and two types of integrals over the three-electron chain (or 3-chain) operators $C_{12}G_{23}$ and $G_{13}G_{23}$. F12 calculations may also require three types of four-electron integrals: two types of integrals over the 4-chain operators $C_{12}G_{14}G_{23}G_{34}$ and $C_{12}G_{13}G_{34}$, as well as one type over the trident operator $C_{12}G_{13}G_{14}$. Explicitly-correlated methods also requires two-electron integrals. However, their computation has been thoroughly studied in the literature.\[261, 370, 246, 332, 344, 371, 372, 248, 271, 250, 373, 339, 324, 254, 255, 374, 253, 252\] Similarly, the nuclear attraction integrals can be easily obtained by taking the large-exponent limit of a $s$-type shell-pair.

Starting with the “master” operator $C_{12}G_{13}G_{14}G_{23}G_{34}$, we will show that one can easily obtain all the FIs as well as the RRs required to compute three- and four-electron integrals within F12 calculations. This is illustrated in Fig. 11.1 where we have used a diagrammatic representation of the operators. The number $N_{\text{sig}}$ of significant integrals in a large system with $N$ CGFs is also reported.

### 11.3 Fundamental Integrals

Following Persson and Taylor,\[245\] the $[\mathbf{0}]^m$ are derived starting from the momentumless integral in Chapter 7, Eq. (2.8), using the following Gaussian integral representation for the Coulomb operator

$$C_{12} = \frac{2}{\sqrt{\pi}} \int_0^\infty \exp(-u^2r_{12}^2) du.$$ (3.1)
After a lengthy derivation, one can show that the closed-form expression of the FIs is

\[
[0]^m = \frac{2}{\sqrt{\pi}} [0]_G \sqrt{\frac{\delta_0}{\delta_1 - \delta_0}} \left( \frac{\delta_1}{\delta_1 - \delta_0} \right)^m F_m \left[ \frac{\delta_1 (Y_1 - Y_0)}{\delta_1 - \delta_0} \right],
\]

(3.2)

where \( m \) is an auxiliary index, \( F_m(t) \) is the generalized Boys function, and

\[
[0]_G = \left( \prod_{i=1}^{4} G_i \right) \left( \frac{\pi^{3/2}}{\delta_0} \right) \exp(-Y_0)
\]

(3.3)

is the FI of the “pure” GG operator \( G_{13}G_{14}G_{23}G_{34} \) from which one can easily get the FI of the 3-chain operator \( G_{13}G_{23} \) by setting \( \lambda_{14} = \lambda_{34} = 0 \). While the FIs involving a Coulomb operator contain an auxiliary index \( m \), the FIs over “pure” GG operators (like \( G_{13}G_{23} \)) do not, thanks to the factorization properties of GGs.

The various quantities required to compute (3.2) are

\[
\delta_u = \zeta + \lambda_u = \zeta + G + u^2 C,
\]

(3.4)

where

\[
\zeta = \begin{pmatrix} \zeta_1 & 0 & 0 & 0 \\ 0 & \zeta_2 & 0 & 0 \\ 0 & 0 & \zeta_3 & 0 \\ 0 & 0 & 0 & \zeta_4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\]

(3.5a)

\[
G = \begin{pmatrix} \lambda_{13} + \lambda_{14} & 0 & -\lambda_{13} & -\lambda_{14} \\ 0 & \lambda_{23} & -\lambda_{23} & 0 \\ -\lambda_{13} & -\lambda_{23} & \lambda_{13} + \lambda_{23} + \lambda_{34} & -\lambda_{34} \\ -\lambda_{14} & 0 & -\lambda_{34} & \lambda_{14} + \lambda_{34} \end{pmatrix},
\]

(3.5b)
and

$$\Delta_u = \zeta \cdot \delta_u^{-1} \cdot \zeta,$$

$$Y^k = \begin{pmatrix} 0 & Y^k_{12} & Y^k_{13} & Y^k_{14} \\ 0 & 0 & 0 & Y^k_{24} \\ 0 & 0 & 0 & Y^k_{34} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

(3.6a)

$$\delta_u = \det(\delta_u),$$

$$Y_u = \text{Tr}(\Delta_u \cdot Y^2).$$

(3.6b)

The generalized Boys function $F_m(t)$ in Eq. (3.2) can be computed efficiently using well-established algorithms.[367, 406, 407]

11.4 Upper Bounds

In this section, we report UBs for primitive and contracted three- and four-electron integrals. As reported in Chapter 7, our UBs are required to be simple (i.e. significantly computationally cheaper than the true integral), strong (i.e. as close as possible to the true integral in the threshold region $10^{-14} - 10^{-8}$), and consistent (i.e. the number of significant integrals $N_{\text{sig}} = \mathcal{O}(N_{\text{UB}})$, where $N_{\text{UB}}$ is the number of integrals estimated by the UB). We also remind the reader that our screening algorithms are based on primitive, $[B_m]$, and contracted, $\langle B_m \rangle$, shell-$nt$uplet bounds. These are based on shell-$nt$uplet information only: shell-pair ($m = 2$), shell-quartet ($m = 4$), shell-sextet ($m = 6$) and shell-octet ($m = 8$). Thus, for each category of three- and four-electron integrals, we will report from shell-pair to shell-sextet (or shell-octet) bounds.

Figure 11.2 is a schematic representation of the overall screening scheme for contracted four-electron integrals. First, we use a primitive shell-pair bound $[B_2]$ to create a list of significant primitive shell-pairs. For a given contracted shell-pair, if at least one of its primitive shell-pairs has survived, a contracted shell-pair bound $\langle B_2 \rangle$ is used to decide whether or not this contracted shell-pair is worth keeping. The second step consists in using a shell-quartet bound $\langle B_4 \rangle$ to create a list of significant contracted shell-quartets by pairing the contracted shell-pairs with themselves. Then, we combine the significant shell-quartets and shell-pairs, and a shell-sextet bound $\langle B_6 \rangle$ identifies the significant contracted shell-sextets. Finally, the shell-sextets are paired with the
shell-pairs. If the resulting shell-octet quantity is found to be significant, the contracted integral class \( \langle a_1 a_2 a_3 a_4 | b_1 b_2 b_3 b_4 \rangle \) must be computed via RRs, as discussed in the next section. The number of significant shell-\(m\)tuplets generated at each step is given in Table 11.1. As one can see, the size of any shell-\(m\)tuplet list is, at worst, quadratic in a large system.

During the shell-pair screening, either a contracted or a primitive path is followed depending on the degree of contraction of the integral class \( K_{\text{tot}} = \prod_{i=1}^{n} K_{a_i} K_{b_i} \). If \( K_{\text{tot}} > 1 \), the contracted path is enforced, otherwise the primitive path is followed. This enables to adopt the more effective primitive bounds for primitive integral classes which are usually associated with medium and high angular momentum PGFs and, therefore, are more expensive to evaluate via RRs. The scheme for primitive four-electron integrals differs only by the use of primitive bounds instead of contracted ones. The three-electron integrals screening scheme can be easily deduced from Fig. 11.2.

Note that we bound an entire class of integrals with a single UB. This is a particularly desirable feature, especially when dealing with three- or four-electron integrals where the size of a class can be extremely large. For example, the simple \([ppp|ppp]\) and \([pppp|pppp]\) classes are made of 729 and 4,096 integrals!

### 11.4.1 Primitive bounds

In this section we present UBs for primitive three- and four-electron integrals. Without loss of generality, we assume that the Geminal operators are ordered by decreasing exponent, i.e. \( \lambda_{13} \leq \lambda_{14} \leq \lambda_{23} \leq \lambda_{34} \).

All the required primitive bounds have the form

\[
[B_m] = \begin{cases} 
\max [I_m], & m = 2, \\
\min [I_m], & m > 2, 
\end{cases}
\]  

(4.7)

where \( m \) is the shell multiplicity.

The bound sets \([I_m]\) are reported in Table 11.1.

They also require the bound factors in Table 11.2, which are easily computed with
Table 11.1: Primitive bound sets \([I_m]\) for three- and four-electron integrals and number \(N_{\text{sig}}\) of significant shell-\(m\)tuplets in a large system with \(N\) PFGs.

<table>
<thead>
<tr>
<th>Integral type operator</th>
<th>(N_{\text{sig}})</th>
<th>(N_{\text{sig}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>three-electron chain (C_{\text{trident}}) ({[1]^{12}<em>{13}, [\bar{1}]^{13}</em>{14}, [\bar{1}]^{12}_{14}}) &amp; (\Omega(N)) &amp; (\Omega(2N))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>four-electron chain (C_{\text{tetrahedron}}) ({\bar{1}}^{12}<em>{13}, [\bar{1}]^{12}</em>{13}, [\bar{1}]^{13}<em>{14}, [\bar{1}]^{14}</em>{13}}) &amp; (\Omega(N)) &amp; (\Omega(N))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trident (C_{\text{trident}}) ({[12]^{12}<em>{13}, [\bar{1}]^{13}</em>{14}, [\bar{1}]^{12}_{14}}) &amp; (\Omega(N)) &amp; (\Omega(N))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11.2: Bounds factors for three- and four-electron integrals. \(\xi_i = (1 - \sigma_i)\zeta_i\), and \(\delta_0\) and \(Y\) are given by (4.8a) and (4.8b). The \(h_i\) factors are defined in Chapter 7, Eq. (4.21).

<table>
<thead>
<tr>
<th>Bound factor</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>([1]_{13,14}^{13,14})</td>
<td>(h_1 \left( \frac{\pi}{\xi_1 + \lambda_{13} + \lambda_{14}} \right) )</td>
</tr>
<tr>
<td>([12,13,14]^{12,13,14})</td>
<td>(h_1 \left( \frac{\pi}{\xi_1 + \lambda_{13} + \lambda_{14}} \right) )</td>
</tr>
<tr>
<td>([13]^{13})</td>
<td>(h_1 h_3 \left( \frac{\pi^2}{\delta_0^{3/2}} \right) \exp(-Y_{13}) )</td>
</tr>
<tr>
<td>([13]_{12,13}^{12,13})</td>
<td>(\frac{2}{\sqrt{\pi}} \left( \frac{\delta_{13}}{\delta_{0}^{3/2}} \right) \exp(-Y_{13}) )</td>
</tr>
<tr>
<td>([134]_{12,13,14}^{12,13,14})</td>
<td>(h_1 h_3 h_4 \left( \frac{\pi^3}{\delta_0^{3/2}} \right) \exp(-Y_{13,14}) )</td>
</tr>
<tr>
<td>([134]_{12,13,14}^{13,14})</td>
<td>(2 \sqrt{\pi} \left( \frac{\delta_{13}^{13,14}}{\delta_{0}^{13,14}} \right) \right. )</td>
</tr>
</tbody>
</table>

220
the following quantities:

\[
\delta_{u}^{13,14,23,34} = \det(\delta_{u}^{13,14,23,34}),
\]

(4.8a)

\[
\hat{Y}^{13,14,23,34} = \Tr\left(\Delta_{0}^{13,14,23,34} \cdot Y^{2}\right),
\]

(4.8b)

where

\[
\delta_{u}^{13,14,23,34} = \xi + \lambda_{u} = \begin{pmatrix}
\xi_1 & 0 & 0 & 0 \\
0 & \xi_2 & 0 & 0 \\
0 & 0 & \xi_3 & 0 \\
0 & 0 & 0 & \xi_4 \\
\end{pmatrix} + \lambda_{u},
\]

(4.9a)

\[
\Delta_{u}^{13,14,23,34} = \xi \cdot \left(\delta_{u}^{13,14,23,34}\right)^{-1} \cdot \xi,
\]

(4.9b)

where \(\xi_i = (1 - \sigma_i)\zeta_i\). We point out that bound factors \([1]^{12}\) or \([1]^{12,13}\) reported in Table 11.2 can be obtained from \([1]^{12,13,14}\) by setting \(\lambda_{15} = \lambda_{14} = 0\) or \(\lambda_{14} = 0\), respectively. As described in Chapter 7, the parameter \(\sigma_1\) is ultimately obtained by solving the quadratic equation

\[
\frac{\partial \hat{[1]}^{13,14,23,34}}{\partial \sigma_1} = 0.
\]

(4.10)

Factors of the kind \([1]^{13,14,23,34}\) are also required for the bounds in Table 11.1. They are defined as the largest factor \([1]^{13,14,23,34}\) within a given system and basis set, and can be pre-computed and stored with the remaining basis set information.

### 11.4.2 Contracted bounds

Contracted integral bounds are straightforward variations of primitive ones. While contracting at the shell-pair level \((m = 2)\) only requires \(O(K^2)\) computational work, contracting at the shell-quartet, -sextet or -octet level would require \(O(K^4), O(K^6)\) or \(O(K^8)\) work, respectively.

Therefore, as sketched in Fig. 11.2, we use a primitive bound for a first screening of the shell-pairs, then contracted bounds are used to screen shell-pairs, -quartets, -sextets and -octets. Considering \(K\)-fold CGFs, the contraction step never exceeds
\( O(K^2) \) computational cost. Bound factors such as

\[
\langle 1 \rangle_{13,14}^{13,14,1} = \sum_{ij}^{K} \left[ 1 \right]_{ij}^{13,14,1},
\]  

(4.11)

are computed within the shell-pair loop, while their maximum within the basis set \( \langle 1 \rangle_{13,14}^{13,14} \) can be pre-computed. As in Chapter 7, in Eq. (4.11), \( i \) and \( j \) refer to the PGFs \( |a_1|_i \) and \( |b_1|_j \) in the contracted shells \( |a_1 \rangle \) and \( |b_1 \rangle \), respectively.

The expressions of the contracted bounds are identical to the primitive bounds, with the only exception that the contracted factors \( \langle 13 \rangle_{12,13,23}^{12,13,23} \), \( \langle 123 \rangle_{12,13,23}^{12,13,23} \), and \( \langle 134 \rangle_{12,13,23}^{12,13,23} \) are bound by

\[
\begin{align*}
\langle 13 \rangle_{12,13,23}^{12,13,23} &\leq \min \left\{ \langle 1 \rangle_{12,13}^{12,13} \langle 3 \rangle_{23}^{23}, \langle 1 \rangle_{12}^{12} \langle 3 \rangle_{13,23}^{13,23} \right\} \exp(-\hat{Y}_{13}), \\
\langle 123 \rangle_{12,13,23}^{12,13,23} &\leq \langle 2 \rangle \min \left\{ \langle 1 \rangle_{12,13}^{12,13} \langle 3 \rangle_{23}^{23}, \langle 1 \rangle_{12}^{12} \langle 3 \rangle_{13,23}^{13,23} \right\} \exp(-\hat{Y}_{13,23}), \\
\langle 134 \rangle_{12,13,23}^{12,13,23} &\leq \min \left\{ \langle 1 \rangle_{12,13}^{12,13} \langle 3 \rangle_{23}^{23} \langle 4 \rangle_{34}^{44}, \langle 1 \rangle_{12,14}^{12,14} \langle 3 \rangle_{13,23}^{13,23} \right\} \exp(-\hat{Y}_{13,14,23}),
\end{align*}
\]

(4.12)

where

\[
\hat{Y}_{13,14,23} = \text{Tr} \left( \Delta_0^{13,14,23,34} \cdot \hat{R}^2 \right),
\]  

(4.13)

can be evaluated with the following expressions

\[
\begin{align*}
\tilde{\delta}_{u}^{13,14,23,34} &= \tilde{\xi} + \lambda_u = \begin{pmatrix}
\xi_1 & 0 & 0 & 0 \\
0 & \xi_2 & 0 & 0 \\
0 & 0 & \xi_3 & 0 \\
0 & 0 & 0 & \xi_4
\end{pmatrix} + \lambda_u, \\
\tilde{\Delta}_{u}^{13,14,23,34} &= \tilde{\xi} \cdot (\tilde{\delta}_{u}^{13,14,23,34})^{-1} \cdot \tilde{\xi}, \\
\tilde{R}^k &= \begin{pmatrix}
0 & R^k_{12} & R^k_{13} & R^k_{14} \\
0 & 0 & 0 & R^k_{24} \\
0 & 0 & 0 & R^k_{34} \\
0 & 0 & 0 & 0
\end{pmatrix},
\end{align*}
\]

(4.14)

where \( \tilde{\xi}_i \) is the smallest effective exponent \( \xi_i \) in the contracted shell-pair \( |a_i b_i \rangle \), and
and \( R_{ij} \) is the distance between two spheres of diameters \( A_i B_i \) and \( A_j B_j \) as defined in Chapter 7, Eq. (8.43).

### 11.5 Recurrence Relations

![Graph representation of the VRRs for the 3-chain operator \( C_{12} G_{23} \) (left) and trident operator \( C_{12} G_{13} G_{14} \) (right). The edge label gives the number of terms in the corresponding VRR. The red path corresponds to the algorithm generating the smallest number of intermediates.](image)

#### 11.5.1 Vertical recurrence relations

Following Obara and Saika,[269, 368] vertical RRs (VRRs) are obtained by differentiation of Eq. (3.2) with respect to the centers coordinates.[369, 363] For the integrals considered in this study, one can show that

\[
[\cdots a_i^+ \cdots]^m = \left( Z_i A_i - \hat{D}_i Y_0 \right) [\cdots a_i \cdots]^m - \left( \hat{D}_i Y_1 - \hat{D}_i Y_0 \right) [\cdots a_i \cdots]^{m+1} \\
+ \sum_{j=1}^{n} a_j \left\{ \left( \frac{\delta_{ij}}{2\xi_i} - \hat{D}_{ij} Y_0 \right) [\cdots a_j^- \cdots]^m - \left( \hat{D}_{ij} Y_1 - \hat{D}_{ij} Y_0 \right) [\cdots a_j^- \cdots]^{m+1} \right\},
\]

(5.15)
where $\delta_{ij}$ is the Kronecker delta,[362]
\[
\hat{D}_i = \frac{\nabla A_i}{2\alpha_i}, \quad \hat{D}_{ij} = \hat{D}_i \hat{D}_j,
\]
and
\[
\hat{D}_i Y_u = \text{Tr} \left( \Delta_u \cdot \hat{D}_i Y^2 \right), \quad (\hat{D}_i Y^2)_{kl} = \kappa_{ikl}(Y)_{kl}, \quad (5.17a)
\]
\[
\hat{D}_{ij} Y_u = \text{Tr} \left( \Delta_u \cdot \hat{D}_{ij} Y^2 \right), \quad (\hat{D}_{ij} Y^2)_{kl} = \frac{\kappa_{ikl}\kappa_{jkl}}{2}, \quad (5.17b)
\]
with
\[
\varepsilon_{ij} = \begin{cases} 
1, & \text{if } i \leq j, \\
0, & \text{otherwise}, 
\end{cases} \quad \kappa_{ijk} = \frac{\varepsilon_{ij} \delta_{ki} - \delta_{ij} \varepsilon_{ki}}{\zeta_i}. \quad (5.18)
\]

One can easily derive VRRs for other three- and four-electron operators following the simple rules given in Fig. 11.1. The number of terms for each of these VRRs is reported in Table 11.3 for various two-, three- and four-electron operators.

Note that for a pure GG operator, we have $m = 0$ and $Y_1 = Y_0$. Therefore, Eq. (5.15) reduces to a simpler expression:
\[
[\cdots a_i^+ \cdots] = \left( Z_i A_i - \hat{D}_i Y_0 \right) [\cdots a_i \cdots] + \sum^n_{j=1} a_j \left( \frac{\delta_{ij}}{2\zeta_i} \hat{D}_j Y_0 \right) [\cdots a_j^- \cdots]. \quad (5.19)
\]

### 11.5.2 Transfer recurrence relations

Transfer RRs (TRRs) redistribute angular momentum between centers referring to different electrons.[363] Using the translational invariance, one can derive
\[
[\cdots a_i^+ \cdots] = \sum^n_{j=1} a_j \left[ \cdots a_j^- \cdots \right] - \sum^n_{j \neq i} \frac{\zeta_j}{\zeta_i} \left[ \cdots a_j^+ \cdots \right] - \sum^n_{j=1} \beta_j A_j B_j \left[ \cdots a_j \cdots \right]. \quad (5.20)
\]

Note that Eq. (5.20) can only be used to build up angular momentum on the last center. Moreover, to increase the momentum by one unit on this last center, one must increase the momentum by the same amount on all the other centers (as evidenced by the second term in the right-hand side of (5.20)). Therefore, the TRR is computationally expensive.
for three- and four-electron integrals due to the large number of centers (see below). As mentioned by Ahlrichs,[369] the TRR can be beneficial for very high angular momentum two-electron integral classes.

### 11.5.3 Horizontal recurrence relations

The so-called horizontal RRs (HRRs) enable to shift momentum between centers over the same electronic coordinate:[363]

\[
\langle \cdots a_i \cdots | b_i^+ \cdots \rangle = \langle \cdots a_i^+ \cdots | b_i \cdots \rangle + A_i B_i \langle \cdots a_i \cdots | b_i \cdots \rangle.
\]

(5.21)

Note that HRRs can be applied to contracted integrals because they are independent of the contraction coefficients and exponents.

### 11.6 Algorithm

In this Section, we propose a recursive algorithm for the computation of a class of three- or four-electron integrals of arbitrary angular momentum. The present recursive algorithm is based on a late-contraction scheme inspired by the Head-Gordon-Pople algorithm[270] following a BOVVVCCCHHHH path. The general skeleton of the algorithm is shown in Fig. 11.4 for the trident operator \( C_{12}G_{13}G_{14} \). We will use this example to illustrate each step.

Based on the shell data, the first step of the algorithm (step B) is to decide whether or not a given class of integrals is significant or negligible. If the integral class is found to be significant by the screening algorithm presented in Section 11.4 and depicted in Fig. 11.2, an initial set of FIs is computed (step O) via the formulae gathered in Section 11.3.

Starting with these FIs, angular momentum is then built up over the different bra centers \( A_1, A_2, A_3 \) and \( A_4 \) using the VRRs derived in Section 11.5.1. To minimize the computational cost, one has to think carefully how to perform this step. Indeed, the cost depends on the order in which this increase in angular momentum is performed. This is illustrated in Fig. 11.3, where we have represented the various possible pathways for the 3-chain operator \( C_{12}G_{23} \) (left) and the trident operator \( C_{12}G_{13}G_{14} \) (right). The red path
corresponds to the path generating the least intermediates (i.e. requiring the smallest number of classes in order to compute a given class). Different paths are compared in Table 11.3 for various two-, three- and four-electron operators, where we have reported the number of intermediates generated by each path for various integral classes.

Taking the 3-chain operator $C_{12}G_{23}$ as an example, one can see that, to compute a [ppp] class, it is more advantageous to build momentum over center $A_3$, then over centers $A_2$, and finally over center $A_1$ using VRRs with 4, 6 and 6 terms, respectively. The alternative path corresponding to building momentum over $A_3$, $A_1$, and then $A_2$ with 4-, 5- and 7-term VRRs is slightly more expensive for a [ppp] class but becomes affordable for high angular momentum classes. For both paths, using the TRR instead of the last VRR implies a large increase in the number of intermediates.

For the trident operator, we successively build angular momentum over $A_4$, $A_3$, $A_1$...
Table 11.3: Number of intermediates required to compute various integral classes for two-, three- and four-electron operators. The path generating the minimum number of intermediates is highlighted in bold. The number of terms in the RRs and the associated incremental center are also reported.

<table>
<thead>
<tr>
<th>Integral type</th>
<th>operator path</th>
<th>number of terms</th>
<th>centers</th>
<th>integral class</th>
</tr>
</thead>
<tbody>
<tr>
<td>two-electron chain</td>
<td>$G_{12}$</td>
<td>(2,3)</td>
<td>$A_2 A_1$</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>VV</td>
<td>VTT</td>
<td>(2,4)</td>
<td>$A_2 A_1$</td>
</tr>
<tr>
<td></td>
<td>$C_{12}$</td>
<td>(4,6)</td>
<td>$A_2 A_1$</td>
<td>4</td>
</tr>
<tr>
<td>three-electron chain</td>
<td>$G_{13}G_{23}$</td>
<td>(2,3,4)</td>
<td>$A_3 A_2 A_1$</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>VV</td>
<td>VTT</td>
<td>(2,3,6)</td>
<td>$A_3 A_2 A_1$</td>
</tr>
<tr>
<td>cyclic</td>
<td>$C_{12}G_{23}$</td>
<td>(4,5,7)</td>
<td>$A_2 A_1$</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>VVV</td>
<td>VVV</td>
<td>(4,6,6)</td>
<td>$A_2 A_1$</td>
</tr>
<tr>
<td></td>
<td>VVT</td>
<td>(4,5,6)</td>
<td>$A_2 A_1$</td>
<td>16</td>
</tr>
<tr>
<td>four-electron chain</td>
<td>$C_{12}G_{14}G_{23}$</td>
<td>(4,5,7,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>VVVV</td>
<td>VVVV</td>
<td>(4,6,6,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
</tr>
<tr>
<td></td>
<td>VVVT</td>
<td>VVVT</td>
<td>(4,5,7,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
</tr>
<tr>
<td></td>
<td>VVV</td>
<td>VVV</td>
<td>(4,6,6,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
</tr>
<tr>
<td></td>
<td>VVT</td>
<td>VVT</td>
<td>(4,6,6,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
</tr>
<tr>
<td>trident</td>
<td>$C_{12}G_{13}G_{23}$</td>
<td>(4,6,8,7)</td>
<td>$A_3 A_2 A_1 A_4$</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>VVVV</td>
<td>VVVV</td>
<td>(4,6,6,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
</tr>
<tr>
<td></td>
<td>VVVT</td>
<td>(4,6,8,8)</td>
<td>$A_3 A_2 A_1 A_4$</td>
<td>34</td>
</tr>
</tbody>
</table>

and $A_2$ using VRRs with 4, 6, 8 and 7 terms. The pathway using VRRs with 4, 6, 6, and 9 terms is more expensive due to the large number of terms of the VRR building up momentum over the last center. Again, using the TRR instead of the last VRR significantly increases the number of intermediates.

The path involving the minimal number of intermediates is given in Table 11.3 for various two-, three- and four-electron operators. It is interesting to point out that it is never beneficial to use the TRR derived in Eq. (5.20) (see Sec. 11.5.2).

One can easily show that, for operators involving the Coulomb operator, the number of intermediates required to compute a $n$-electron integral class $[a \ldots a]$ increases as $O(a^{n+1})$ for the VRR-only paths (see Table 11.3). This number is reduced to $O(a^n)$ if one uses the TRR to build up angular momentum on the last center. However, the prefactor is much larger and the crossover happens for extremely high angular momentum for three- and four-electron integrals. For “pure” GG operators, such as $G_{12}$ or $G_{13}G_{23}$, the number of intermediates required to compute a class $[a \ldots a]$ increases as
$O(n^a)$ for any type of paths.

Finally, we note that the optimal path for the trident $C_{12}G_{13}G_{14}$ and the 4-chain $C_{12}G_{13}G_{34}$ is similar, thanks to their similar structure. Indeed, these two operators can be seen as two “linked” GGs ($G_{13}G_{14}$ or $G_{13}G_{34}$) interacting with the Coulomb operator $C_{12}$ (see Fig. 11.1), while the other 4-chain operator $C_{12}G_{14}G_{23}$ can be seen as two “unlinked” GGs ($G_{14}$ and $G_{23}$) interacting with the Coulomb operator.

When angular momentum has been built over all the bra centers, following the HGP algorithm,[270] we contract $[a_1a_2a_3a_4|0000]$ to form $(a_1a_2a_3a_4|0000)$ (step CCCC). We can perform the contraction at this point because all of the subsequent RRs are independent of the contraction coefficients and exponents. More details about this contraction step can be found in Ref. [20].

The last step of the algorithm (step HHHH) shifts momentum from the bra center $A_1, A_2, A_3$ and $A_4$ to the ket centers $B_1, B_2, B_3$ and $B_4$ using the two-term HRRs given by Eq. (5.21) in Section 11.5.3.

It is certainly of interest to compare the number of intermediates in Table 11.3 with that required to construct a given three- and four-electron integral class when general multiplicative correlation factors of the form $f_{12} = f(|r_1 - r_2|)$ are adopted, which is reported in Table 9.2 (Chapter 9). For three-electron integrals, using GGs yields a reduction of the number of intermediates which ranges from one order of magnitude for $[ppp|ppp]$ classes to almost two orders of magnitude for $[ddd|ddd]$ classes. For four-electron integrals such reduction becomes extreme, ranging from one order of magnitude for $[pppp|pppp]$ classes to three orders of magnitude for $[dddd|dddd]$ classes.

The computational convenience of GGs is even more striking if one considers that the $[0]^m$ required for each class are much fewer and incredibly cheaper to evaluate for GGs than for more general $f_{12}$ factors. In fact, in the first case the $[0]^m$ bring at most one auxiliary index and only one evaluation of the generalized Boys function is required, while, as discussed in Chapter 9, the $[0]^m$ for more general correlation factors might involve up to 7 auxiliary indices and usually numerical quadrature techniques are required for their evaluation.
11.7 Concluding Remarks

We have presented the three main ingredients to compute three- and four-electron integrals involving GGs. Firstly, a straightforward method to compute the FIs is given. Secondly, scaling-consistent UBs are reported, as they allow to evaluate only the $O(N^2)$ significant integrals in a large system. Finally, the significant integrals are computed via a recursive scheme based on vertical and horizontal RRs, which can be viewed as an extension of the PRISM late-contraction path to three- and four-electron integrals.

We have also shown that using Gaussian Geminals (or linear combinations of them), within the scheme presented, yields a dramatic reduction of the computational complexity of these integrals.

We believe our approach represents a major step towards an accurate and efficient computational scheme for three- and four-electron integrals. It also paves the way to even more contraction-effective methods for these types of integrals. In particular, based on the study performed in Chapter 10, we believe that an early-contraction scheme would have additional significant computational benefits.
Chapter 12

Conclusions

In this thesis we contributed to the development of two exciting areas of theoretical quantum chemistry.

In Part A we developed and assessed the performance of single-determinant models for electronic excited states.

Chapter 2 provided a compact description of some of the most relevant and employed quantum chemical methods for electronic excited states, with a particular emphasis on those approaches encountered in the remainder of the thesis.

In Chapter 3, after introducing the Maximum Overlap Method (MOM) developed by Gilbert et al.[131] we examined the extent to which single-determinant HF solutions are reasonable approximations to the ground state and ten singly excited states of the H$_2$ molecule. We found that, with only one exception, HF yields surprisingly accurate models for the low-lying excited states of this molecule.

In Chapter 4, by introducing a modified version of the MOM, the Initial Maximum Overlap Method (IMOM), we applied the single-determinant approach to three of the most challenging topics in the chemistry of excited states: double excitations, conical intersections and charge-transfer states. We concluded that the IMOM protocol provides a straightforward and reliable method for obtaining these solutions, and we have shown that they may be preferable to other low-cost excited-state methods.

In Chapter 5, by exploiting the structural simplicity of the orbital picture emerging from the single-determinant scheme we defined the excitation number, $\eta$, as a rigorous metric for the characterisation of multiply excited states. Furthermore, we have shown
that \( \eta \) values are pleasingly close to integers, which allows an unambiguous classification of multiple excitations. We have used this to correctly describe several multiply excited states that have been misassigned in the past.

In Part B we presented effective methods for the accurate evaluation of many-electron integrals arising in the explicitly correlated electronic structure theory.

Chapter 6 introduced briefly explicitly correlated method, with a particular emphasis on the integrals’ problem.

In Chapter 7 we constructed and tested the performance of a variety of novel many-electron integral upper bounds. Furthermore, we discussed the implementation of such bounds in effective screening algorithms which enable to compute only the significant fraction of integrals.

In Chapter 8 and 9 we presented recurrence-relations-based algorithms, in the PRISM[20] style, for the evaluation of the non-negligible integrals over a general class of multiplicative three- and four-electron operators.

In Chapter 10 we devised a new computational scheme for two-electron integrals over Gaussian Geminals which exploits the unique factorisation properties of these integrals. A detailed FLOP count revealed that our scheme is computationally much cheaper than any other previous scheme.

Finally, in Chapter 11 we extended the “late-contraction path” and the screening techniques developed in Chapter 10 to three- and four-electron integrals involving Gaussian Geminals. The developed computational scheme enables to evaluate, in a large system, only the \( O(N^2) \) significant integrals arising in F12 calculations. The computational cost of classes of such integrals, evaluated by counting the number of intermediate classes, revealed that our Gaussian-Geminal-based scheme yields a dramatic reduction of the computational complexity of three- and four-electron integrals.

In conclusion, a single-determinant framework provides a simple and accurate alternative for modelling excited states in cases where other low-cost methods, such as CIS and TD-DFT, either perform poorly or fail completely. Moreover, we found that, in a non-negligible number of cases, single-determinant energies obtained in this way can even be extremely accurate, rivalling with far more expensive methods such as CASSCF and MRMP.
Furthermore, we conclude that the developed computational schemes for many-electron integrals arising in explicitly correlated methods provide an accurate and effective alternative to the RI strategy. In fact, such schemes become more and more convenient as the size of the system increases because i) the RI larger auxiliary basis is no more required, ii) our screening techniques enable to compute only the significant fraction of integrals, iii) the integrals are evaluated exactly or in a nearly-exact fashion. The effectiveness of our approach is dramatically enhanced if Gaussian Geminals, or linear combinations of Gaussian Geminals, are adopted as correlations factor. Based on the results of Chapter 10, we also believe that an early-contraction scheme for three- and four-electron integrals involving Gaussian Geminals would have additional significant computational benefits.
Bibliography


        45, 64, 66, 69

        Springer, 2012. 45, 48, 63, 69

        68, 81, 82, 89

        58, 68, 82, 107

        2004. 45, 58, 68, 82, 107

        2004. 45, 58, 68, 82, 107

        1051, 2006. 45, 52, 58, 68, 82, 94, 107


        108:4439, 1998. 45, 58, 68, 82

        45, 58, 68, 82


        45, 58, 68, 82, 89


[116] F. Jensen. *Introduction to Computational Chemistry*. John Wiley & Sons Ltd, Chichester, 2007. 63, 64, 70


242


246


247


BIBLIOGRAPHY


252


253


256
BIBLIOGRAPHY


257


258


