

LONG-RANGE EFFECTS IN QM/MM
CALCULATIONS: EWALD SUMMATION IN
NON-MINIMAL BASIS SETS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Zachary Conner Holden, M.S., B.S.

Graduate Program In Chemistry

The Ohio State University

2015

Dissertation Committee:

John M. Herbert, Advisor

James V. Coe

Marcos Sotomayor

© Copyright by
Zachary Conner Holden
2015

ABSTRACT

An implementation of Ewald summation for use in mixed quantum mechanics/molecular mechanics (QM/MM) calculations is presented, which builds upon previous work by others that was limited to semi-empirical electronic structure for the QM region. Unlike previous work, our implementation describes the wave function’s periodic images using “ChEIPG” atomic charges, which are determined by fitting to the QM electrostatic potential evaluated on a real-space grid. This implementation is stable even for large Gaussian basis sets with diffuse exponents, and is thus appropriate when the QM region is described by a correlated wave function. Derivatives of the ChEIPG charges with respect to the QM density matrix are a potentially serious bottleneck in this approach, so we introduce a ChEIPG algorithm based on atom-centered Lebedev grids. The ChEIPG charges thus obtained exhibit good rotational invariance even for sparse grids, enabling significant cost savings. Upon further examination new digestion routines were created to enable an ever more significant cost savings. Detailed analysis of the optimal choice of user-selected Ewald parameters, as well as timing breakdowns, are presented.

To my parents, Eric and Karen.

VITA

1987	Born, Chambersburg, Pennsylvania
2005	Chambersburg Area Senior High School
2009	B.S. Chemistry, B.S. Physics, University of Pittsburgh
2009–2012	Graduate Teaching Associate, The Ohio State University
2012	M.S. Chemistry, The Ohio State University
2012–2015	Graduate Research Associate, The Ohio State University
2015–present	Graduate Teaching Associate, The Ohio State University

PUBLICATIONS

- (4) Z. C. Holden and J. M. Herbert. **Periodic Boundary Conditions for QM/MM Calculations: *Ab Initio* Molecular Dynamics in Extended Gaussian Basis Sets.** *In Preparation.*
- (3) Z. C. Holden, R. M. Richard, and J. M. Herbert. **Erratum: Periodic Boundary Conditions for QM/MM Calculations: Ewald Summation for Extended Gaussian Basis Sets.** *J. Chem. Phys.* **142**, 059901 (2015).
- (2) Y. Shao, Z. Gan, et al. **Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package.** *Mol. Phys.* **113**, 184-215 (2015).

(1) Z. C. Holden, R. M. Richard, and J. M. Herbert. **Periodic Boundary Conditions for QM/MM Calculations: Ewald Summation for Extended Gaussian Basis Sets.** *J. Chem. Phys.* **139**, 244108 (2013).

FIELDS OF STUDY

Major Field: Chemistry
Theoretical Physical Chemistry

TABLE OF CONTENTS

ABSTRACT	ii
DEDICATION	iii
VITA	iv
LIST OF FIGURES	xi
LIST OF TABLES	xiv

CHAPTER	PAGE
1 Introduction	1
2 Ewald Summation Theory	7
2.1 Electrostatic Energy	7
2.2 Gradients	11
3 QM/MM and PBC	14
3.1 Energy corrections	15
3.2 Fock matrix corrections	19
3.3 Position derivative of the Ewald energy and potential	21
3.4 Position derivative of the SCF energy	23
3.5 Mulliken Charge Derivatives	26
4 Parameters for the Ewald sums	28

5	ChEIPG charges	32
5.1	Basic theory	32
5.2	Lebedev grid implementation	34
5.3	Density derivative of the charges	36
5.4	Position derivatives of ChEIPG Charges	37
5.4.1	Atom-independent grids	39
5.4.2	Atom-dependent grids	41
5.4.3	Weights of rectangular ChEIPG charges	43
6	Results	44
6.1	Charge schemes	44
6.2	Lebedev ChEIPG charges	48
6.2.1	Gas phase	48
6.2.2	Condensed phase	52
6.3	Timings	55
6.4	A new digestion routine	63
6.5	Gradient results	65
7	Conclusion	70

APPENDICES

A	Ewald Summation	72
A.1	Transformation	72
A.2	E_{cell}	73
A.3	E_{real}	74
A.4	E_{recip} and E_{self}	75
A.4.1	E_{self}	75
A.4.2	E_{recip}	76
A.5	E_{dipole}	79
A.6	E_{charge}	82
A.7	E_{corr}	86
A.8	Orthorhombic Boxes vs Cubic Boxes	87

B	Ewald Derivatives	88
B.1	Further Energies	88
B.2	Charge-Charge Gradients	89
B.2.1	$\nabla_k E_{self}$ and $\nabla_k E_{charge}$	89
B.2.2	$\nabla_k E_{dipole}$	90
B.2.3	$\nabla_k E_{real}$	90
B.2.4	$\nabla_k E_{recip}$	91
B.2.5	$\nabla_k E_{corr}$	94
C	MM Dynamics	96
C.1	van der Waal's Energy	96
C.1.1	vdW Energy	96
C.1.2	Switching Function	96
C.1.3	Smooth Energy	97
C.2	van der Waal's Gradient	98
C.2.1	vdW Gradient	98
C.2.2	Derivative of Switching Function	99
C.2.3	Smooth vdW Gradient	100
C.3	Minimum Image Convention	100
C.4	Confining Potential	101
D	Ewald Energy Correction	104
D.1	$\Delta E_{QM/QM}^{PI}$	104
D.2	$\Delta E_{QM/MM}^{PI}$	107
D.2.1	Real Space Term	107
D.2.2	Reciprocal Term	108
D.3	Forming Pair Potential	109
D.4	Maximum Vectors and Timings	111
D.4.1	Real Space	111
D.4.2	Reciprocal Space	113
D.4.3	Finding η_{min}	114
D.4.4	Orthorhombic Cells	115
E	Ewald Fock Matrix Correction	120

F	Spatial Derivatives	122
	F.1 Normal HF Derivative	122
	F.2 Ewald HF Derivative	124
	F.2.1 w.r.t. QM Atom Position	124
	F.2.2 w.r.t. MM Atom Position	128
	F.2.3 Derivative of the Pair Potential	130
	F.3 Logistics in Q-Chem	133
	F.3.1 Outputs and Relationships	133
	F.3.2 QM Derivatives	134
	F.3.3 QM/MM Derivatives	135
	F.3.4 QM/MM Ewald Derivatives	136
	F.3.5 Scratch Files	136
	F.3.6 Outline of Code for Computing $\frac{\Delta E}{\Delta P_{\mu\nu}}$ Directly	137
G	Charge Schemes	139
	G.1 Mulliken	139
	G.2 QM Potential at a Grid Point	139
	G.3 ChEIP-G	140
	G.3.1 Charges	140
H	Charge Derivatives	145
	H.1 Density Derivative for Mulliken Charges	145
	H.1.1 Non-Symmetric Form	145
	H.1.2 Symmetric Form	145
	H.2 Density Derivative for ChEIP-G Charges	146
	H.2.1 Simplification of Derivative	146
	H.2.2 Alternate Simplification of Derivative	148
	H.2.3 Application of Derivative to Fock Matrix	149
	H.3 Spatial Derivative for ChEIP-G Charges	149
	H.3.1 Simplify Terms 1 and 6	151
	H.3.2 Terms 2 and 4	152
	H.3.3 Terms 3 and 5	153
	H.3.4 Recombine	153
	H.3.5 Alternate Derivation	154
	H.3.6 Programmed Versions	158

H.4	Needed Derivatives for q_A^∇	159
H.4.1	$(\mathbf{G}^{-1})_{CD}^{\nabla A}$	159
H.4.2	$e_C^{\nabla A}$	161
H.5	M Derivatives	164
H.5.1	Mulliken Charges	164
H.5.2	ChELP-G Charges	165
H.6	ChELPG Rectangular Weights	167
H.6.1	ChELPG Weighting Scheme	167
H.6.2	Weight Derivatives	168
I	Vectors Inside a Supersphere	170
I.1	All Vectors Inside a Supersphere, N	170
I.1.1	$n = 0$	170
I.1.2	$n = 1$	170
I.1.3	$n = 2$	171
I.1.4	$n = 3$	171
I.2	All Positive Vectors in Sphere (First Octant), M	172
I.2.1	$n = 0$	172
I.2.2	$n = 1$	172
I.2.3	$n = 2$	173
I.2.4	$n = 3$	173
	Bibliography	174

LIST OF FIGURES

FIGURE	PAGE
1.1 Radial distribution functions for chloride-water(n) systems. For system I: $n = 31$, II: $n = 63$, and $n = 127$. The top graph is for the chloride-oxygen distribution and the bottom graph is for the chloride-hydrogen distribution. This figure is reproduced from ref. ¹	2
1.2 Velocity autocorrelation function for chloride-water(n) systems. For system I: $n = 31$, II: $n = 63$, and $n = 127$. This figure is reproduced from ref. ¹	3
1.3 The potential of mean force (PMF) for ammonium chloride with 1034 TIP3P water molecules. This figure is reproduced from ref. ²	4
1.4 The potential of mean force (PMF) for methyl phosphate with 1034 TIP3P water molecules. This figure is reproduced from ref. ²	5
2.1 A figure showing periodic replication.	8
3.1 A color coded pictorial representation, where the orange oval represents the QM region, for all interactions included in E_{total}	16
3.2 A pictorial representation of a replicated QM/MM simulation cell, where the blue oval represents the QM region (second image), and its collapse onto point charges (third image). Note that any circle not surrounded by the QM region is a point charge.	17

5.1	Plot of the Lebedev grid for H ₂ in plane containing the internuclear axis. (a) Actual grid using $N_p = 302$ points per radial shell and nine shells per atom. (b) Number of grid points contained in each 0.25 Å × 0.25 Å cell, equivalent to the contribution of each cell to the least-squares fit when all the weights w_k are identical. (c) Each cell's contribution when a simple weighting scheme ($w_k = 1/n_k$) is used.	35
6.1	Converged partial charges for the oxygen atom of a single QM water molecule in an MM water box, plotted against the number of basis functions used to describe the QM region, for 6-31(x+,y+)G* and 6-311(x+,y+)G* basis sets. In (a), the QM/MM-Ewald method uses Mulliken image charges whereas in (b) it uses ChEIPG image charges. In the latter case, Mulliken charges were also computed upon SCF convergence.	45
6.2	Total SCF energy versus the number of basis functions, for a single QM water molecule in a box of MM water molecules, using either (a) Mulliken image charges or (b) ChEIPG image charges.	47
6.3	Convergence of the Cartesian ChEIPG charges (atomic units) computed at the HF/aug-cc-pVDZ level, as a function of the Cartesian grid spacing Δx . The reference values were computed using $\Delta x = 0.05$ Å. The vertical axis plots the mean unsigned error (MUE) for all of the charges in the molecule.	49
6.4	Convergence of the Lebedev ChEIPG charges (atomic units) computed at the HF/aug-cc-pVDZ level, as a function of the grid spacing Δx . The reference values were computed using a Cartesian grid with $\Delta x = 0.05$ Å. The vertical axis plots the mean unsigned error (MUE) for all of the charges in the molecule. In (a), the weighting scheme discussed in Section 5.2 is employed ($w_k = 1/n_k$), whereas in (b) the weights are all equal.	50
6.5	Rotational invariance (a.u.) of the Lebedev and Cartesian ChEIPG charges on the hydroxyl oxygen of glycine as the molecule is rotated by 180° about an axis. Charges are computed at the HF/6-31G* level of theory.	51

6.6	The total CPU time required to calculate $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ [Eq. (3.9)] for (a) 11 QM water molecules in a large MM water box, and (b) 27 water molecules plus cytidine in the QM region, surrounded by a larger box of MM water molecules.	56
6.7	Percentage of the total single-point energy calculation time that is spent in calculating derivatives $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ and $\partial Q_\alpha/\partial P_{\mu\nu}$, for a water test system (small QM region) (top graph) and an aqueous cytidine test system (large QM region) (bottom graph). The time required to calculate $\partial Q_\alpha/\partial P_{\mu\nu}$ is independent of η but becomes a smaller percentage as the time to compute $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ increases.	58
6.8	Conservation of energy for QM/MM calculations of aqueous cytidine in which the QM region consists of a region of the cytidine molecule, described at the HF/6-31G level. The y-axis is the energy of the time step minus the average energy over the last 100 time steps.	67
C.1	Graph of the hyperbolic tangent function with parameters of interest shown.	102

LIST OF TABLES

TABLE	PAGE
<p>6.1 SCF energies and Lebedev ChEIPG charges as the simulation cell (containing 5 QM and 211 MM water molecules) is rotated with respect to the axes that define the Lebedev unit spheres. The final column is the difference between the maximum and minimum values for the various quantities in each row. The SCF convergence threshold was set to 10^{-5} hartree.</p>	54
<p>6.2 Magnitude of the force on a single oxygen atom as the simulation cell (containing 5 QM and 211 MM water molecules) is rotated with respect to the axes that define the Lebedev unit spheres. The final column is the difference between the maximum and minimum values for the various quantities in each row. The SCF convergence threshold was set to 10^{-7} hartree and the integral threshold to 10^{-10}, with the corresponding value of C determined from Eq. (4.1).</p>	54
<p>6.3 Timings (in seconds) for a single-point QM/MM calculation of cytidine (QM) in water (MM), with periodic boundary conditions using $\eta = 0.04 \text{ \AA}^{-1}$. The QM region consists of 30 atoms and 349 basis functions (B3LYP/6-31+G*) and the MM region consists of 12,840 point charges (4,160 TIP3P water molecules). The Lebedev and Cartesian ChEIPG grids consist of 885 and 32,598 points, respectively. For the steps that must be repeated at each SCF calculation, two columns of timing data are provided, corresponding to the second (left column) and ninth (right column) SCF cycles. The row labeled “other” includes the XC quadrature step and any remaining overhead associated with the Fock build.</p>	59

- 6.4 Timing data (rounded to the nearest second) for QM/MM calculations of aqueous cytidine in which the QM region consists of a region of specified radius, R , around the cytidine molecule, described at the B3LYP/6-31+G* level. All calculations were performed with periodic boundary conditions using $\eta = 0.04 \text{ \AA}^{-1}$ and ChElPG charges. For the steps that must be repeated at each SCF calculation, two columns of timing data are provided, corresponding to the second (left column) and ninth (right column) SCF cycles. Timings labeled “other” includes the XC quadrature step and any remaining overhead associated with the Fock build. All calculations were run in serial on a single Intel Xeon x5650 processor with 48 GB RAM with no competing processes on the node. 61
- 6.5 Timing data (rounded to the nearest tenth of a second) for QM/MM calculations of aqueous cytidine in which the QM region consists of a region of the cytidine molecule, described at the HF/6-31G level. All calculations were performed with periodic boundary conditions using $\eta = 0.12 \text{ \AA}^{-1}$ and ChElPG charges. All calculations were run in serial on a single Intel Xeon x5650 processor with 48 GB RAM with no competing processes on the node. 66

CHAPTER 1

Introduction

As computers have become more powerful, scientists have developed new methods to condensed phase calculations. Unfortunately, the best theory, CCSD(T) (coupled cluster singles, doubles, and perturbative triples), can only be used to calculate small systems, about 200 basis functions. Obviously, this method is not appropriate for condensed phase calculations. Originally, quantum mechanics could only be performed on single, small molecules with small basis sets in the gas phase.

The original way that scientists attempted to model condensed phase systems was through cluster calculations. This entails taking a molecule of interest and surrounding it by a single or several solvation layers. Calculations are then performed on this system in the hope that it will resemble the bulk solvated behavior of the system of interest. Unfortunately, these calculations usually leave something to be desired.

Figure 1.1 show the radial distribution functions for a chloride-water system of various sizes. The long range behavior of a radial distribution function should be constant because the probability of finding a certain atom from another atom is a constant value. As can be seen in the figure, the long range behavior decreases for smaller cluster sizes. None of the clusters show a constant value at long range. The

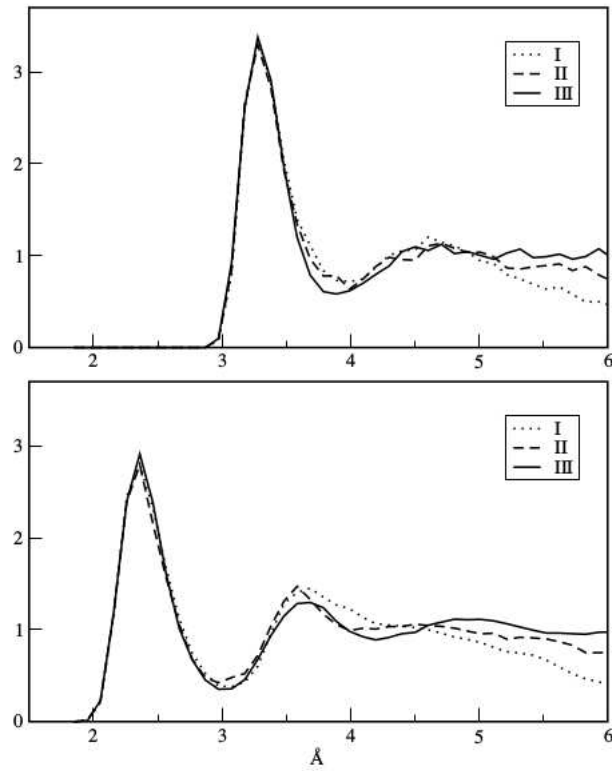


Figure 1.1: Radial distribution functions for chloride-water(n) systems. For system I: $n = 31$, II: $n = 63$, and $n = 127$. The top graph is for the chloride-oxygen distribution and the bottom graph is for the chloride-hydrogen distribution. This figure is reproduced from ref.¹.

lack of smoothness in the curves amounts from the small sample size in the numerical integration.

Also in ref. ??, the autovelocity correlation function was calculated for the chloride-water systems. This figure is shown in fig. 1.2. At long times, the autovelocity correlation function should be zero because the velocity at the given time

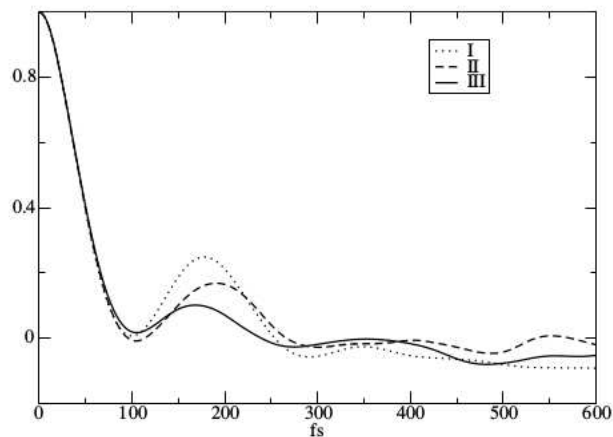


Figure 1.2: Velocity autocorrelation function for chloride-water(n) systems. For system I: $n = 31$, II: $n = 63$, and $n = 127$. This figure is reproduced from ref.¹.

should be decoupled from the velocity at time zero. It can be argued that the simulation was not run for long enough for the velocities to decouple. The autovelocity correlation function varies greatly with the system size. Unfortunately, there is no way to experimentally measure this function, so it raises the question, when has the bulk limit be reached? With the wide variance in this function with respect to cluster size, it is not reasonable to assume that the bulk limit has been reached even in the largest cluster.

The next step is to use a QM/MM (quantum mechanics/molecular mechanics) calculation. This entails surrounding the quantum mechanical (QM) region with a molecular mechanics (MM) region. The MM region is computed using Newton's laws. It adds long range behavior, but only as far as the region extends. In order to get proper long range behavior either a very large number of MM atoms must be used

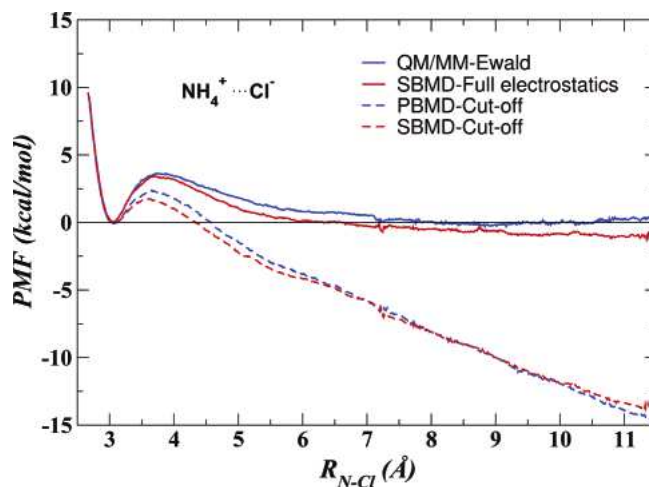


Figure 1.3: The potential of mean force (PMF) for ammonium chloride with 1034 TIP3P water molecules. This figure is reproduced from ref.².

or a periodic boundary condition must be used. Periodic boundary conditions are the cheapest way to incorporate these long range behaviors. The QM/MM/Ewald method was originally implemented by Nam, Gao, and York² for semi-empirical methods.

As can be seen in fig. 1.3, periodic boundary conditions are needed to get the long range properties correct. Even when a cutoff of 11.5 Å is used for the electrostatics, the long range behavior of the PMF is incorrect. This figure shows the obvious need for periodic boundary conditions for long range electrostatic interactions.

Nam, Gao, and York also computed the PMF for methyl phosphate, which is the simplest example of a phosphoryl in a dissociative phosphoryl transfer mechanism. Once again a cut-off shows the incorrect long range behavior; however, there is an accuracy issue as well. In their paper the authors state that the free energy calculated from this PMF profile produces a reaction free energy that is in error by as much

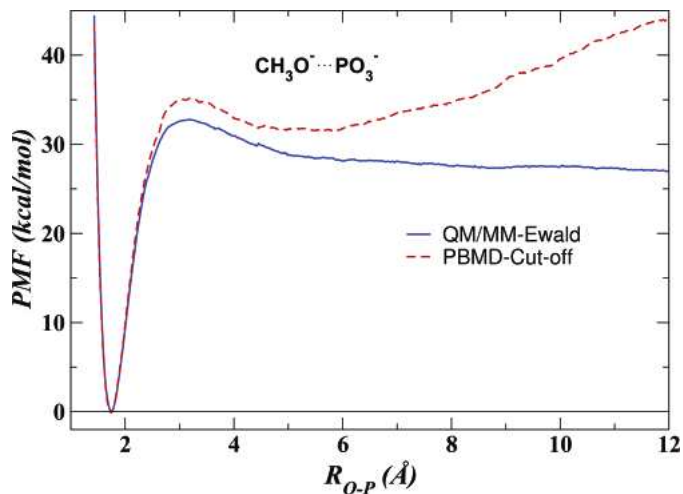


Figure 1.4: The potential of mean force (PMF) for methyl phosphate with 1034 TIP3P water molecules. This figure is reproduced from ref.².

as 10 kcal/mol. The authors suggest that better semi-empirical methods or better parametrization are needed. In reality, these measures are not required if one could actually use a Hartree-Fock based method.

Nam *et al.*² implemented their Ewald method by using Mulliken atomic charges to represent the periodic images of the QM wave function. With an appropriate correction to the Fock matrix, the self-consistent field (SCF) procedure remains variational in these periodic QM/MM simulations.² The method in Ref. 2 was designed for semi-empirical QM calculations in minimal basis sets, for which Mulliken charges are well-behaved, but experiences serious convergence problems in larger basis sets.³ To circumvent this problem, and to facilitate periodic QM/MM calculations in arbitrary basis sets, we recently reformulated the QM/MM-Ewald method of Ref. 2 to use charges derived from the electrostatic potential⁴ (ChElPG) to represent the image

wave functions.³

ChEIPG charges are more expensive to compute than Mulliken charges, and their derivatives even more so. The derivatives originally posed a problem because they were overly expensive and the bottleneck of the calculation. Originally an attempt was made to use Lebedev centered atomic grids (as opposed to the convention rectangular grids) to reduce the cost of the ChEIPG density derivatives which are needed for single point energy calculations. Lebedev grids did reduce the computation time as compared to the rectangular grids; however, they did not eliminate the bottleneck. Upon development of the gradients, it was discovered that the bottleneck in these calculations were unsurprisingly found to be the ChEIPG position derivatives. Unfortunately, due to the non-analytical weighting scheme of the Lebedev algorithm, position gradients are not possible for these charges. This meant that the gradients must be computed with the rectangular grid. At that point rewriting of the integral digestion routine was investigated. This new digestion routine for both density derivatives and position derivatives of ChEIPG charges sped up these routine to the point where they were no longer the bottlenecks of the calculation.

In this document, a derivation of the analytical expression of the QM/MM Ewald method for energies and gradients is presented. Their performance is then documented with timing breakdowns, parallelization across multiple processors, and a discussion of the optimal user-selected Ewald parameters.

CHAPTER 2

Ewald Summation Theory

2.1 Electrostatic Energy

Charge–charge interactions decay very slowly with distance and may not become negligible in a calculation until the distance is on the order of hundreds of nanometers. As such, the pairwise sum over such interactions is slowly convergent. In fact, it is only conditionally convergent in a periodically-replicated simulation cell,⁵ shown in fig. 2.1, which is the problem that Ewald summation is designed to overcome. This section provides a brief overview of the Ewald summation technique, which also serves to introduce the notation that we will use.

Traditional charge-charge Ewald summation splits the pairwise summation into two parts: a real-space portion, based on a short-range interaction potential whose pairwise sum converges quickly; and a long-range portion based on a slowly-varying interaction potential whose pairwise sum converges relatively quickly in reciprocal space. The Coulomb potential is partitioned using the error function (erf) and complementary error function (erfc), according to

$$\frac{1}{r} = \frac{\text{erf}(\eta r)}{r} + \frac{\text{erfc}(\eta r)}{r} . \quad (2.1)$$

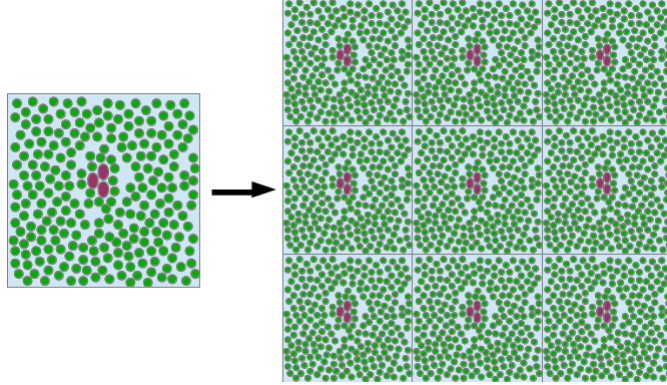


Figure 2.1: A figure showing periodic replication.

Note that $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$. The Ewald parameter η controls the length scale ($\sim \eta^{-1}$) on which the short-range function $\operatorname{erfc}(\eta r)/r$ decays, and thus controls how much of the pairwise Coulomb sum is performed in real space. As η increases, more of the summation is performed in reciprocal space, whereas setting $\eta = 0$ is the same as performing the pairwise sum entirely in real space.

In the context of Ewald summation, the Coulomb energy in the simulation cell, E_{cell} , is traditionally partitioned as

$$E_{\text{cell}} = E_{\text{real}} + E_{\text{self}} + E_{\text{recip}} + E_{\text{charge}} + E_{\text{dipole}}. \quad (2.2)$$

The energy components include the real-space energy,

$$E_{\text{real}} = \sum_{\mathbf{n}} \sum_j^{N_{\text{MM}}} \sum_{k>j}^{N_{\text{MM}}} q_j q_k \frac{\operatorname{erfc}(\eta |\mathbf{r}_{jk} + \mathbf{n}L|)}{|\mathbf{r}_{jk} + \mathbf{n}L|} \quad (2.3)$$

(where $\mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k$), the Coulomb self-energy,

$$E_{\text{self}} = -\frac{\eta}{\sqrt{\pi}} \sum_j^{N_{\text{MM}}} q_j^2, \quad (2.4)$$

and the reciprocal-space energy,

$$E_{\text{recip}} = \sum_{\mathbf{m} \neq \mathbf{0}} \frac{|S(\mathbf{m})|^2}{2L\pi |\mathbf{m}|^2} \exp\left(\frac{-\pi^2 |\mathbf{m}|^2}{\eta^2 L^2}\right). \quad (2.5)$$

In these expressions, L denotes the length of the (cubic) simulation cell, and \mathbf{n} and \mathbf{m} are real-space and reciprocal-space lattice vectors, respectively, sums over which extend to infinity. The quantity $S(\mathbf{m})$ in Eq. (2.17) is known as the *structure factor*,⁶ and is discussed below. The quantities q_j and q_k are point charges, the total number of which is denoted N_{MM} . For brevity, we have presented these equations for the case that the unit cell is cubic. This is not a fundamental limitation of the Ewald formalism, but the non-cubic case would require, splitting the sum over lattice vectors \mathbf{n} in Eq. (2.13) into separate sums over n_x , n_y , and n_z , with lengths L_x , L_y , and L_z for each side of the simulation cell.

The final two terms in Eq. (2.2) warrant some additional comments. The quantity

$$E_{\text{charge}} = -\frac{Q_{\text{tot}}^2 \pi}{2L^3 \eta^2}, \quad (2.6)$$

is known as the *surface charge term*, where

$$Q_{\text{tot}} = \sum_j^{N_{\text{MM}}} q_j \quad (2.7)$$

represents the total charge of the simulation cell. Since the Coulomb energy is divergent if $Q_{\text{tot}} \neq 0$, Ewald summation can only be used to compute the Coulomb energy for a neutral simulation cell, and E_{charge} represents the energy required to surround a charged cell with a charge-compensating membrane of opposite charge. (If the cell is electrically neutral, then $E_{\text{charge}} = 0$.) Artifacts due to Ewald simulation of a charged unit cell have been noted,⁷⁻⁹ but will not concern us here.

The final component of E_{cell} is the *surface dipole term*,¹⁰

$$E_{\text{dipole}} = -\frac{\pi}{(2\varepsilon + 1)L^3} \sum_{j,k}^{N_{\text{MM}}} q_j q_k |\mathbf{r}_{jk}|^2, \quad (2.8)$$

If $Q_{\text{tot}} = 0$ (which is often assumed when this term is discussed in the literature, as for example in Ref. 11), the E_{dipole} is proportional to the square of the dipole moment of the simulation cell. In Eq. (2.8), we imagine placing the *supercell* (the simulation cell along with all of its periodic images) into a dielectric medium, whose dielectric constant is denoted by ε in Eq. (2.8).^{5,10,12} Often, one assumes “tin foil” boundary conditions ($\varepsilon = \infty$, corresponding to placing the supercell inside of a conductor), in which case $E_{\text{dipole}} = 0$.

The last bit of notation to explain is the quantity $|S(\mathbf{m})|^2$ in Eq. (2.9). In its most general form, this quantity is defined as

$$|S(\mathbf{m})|^2 = \sum_{j,k}^{N_{\text{MM}}} q_j q_k \exp\left(\frac{2\pi i}{L}(\mathbf{m} \cdot \mathbf{r}_{jk})\right). \quad (2.9)$$

Despite the appearance of $i = \sqrt{-1}$, the quantity $|S(\mathbf{m})|^2$ is real, as the squared-modulus notation indicates, and can be rewritten in a way that makes the real value obvious:

$$|S(\mathbf{m})|^2 = \sum_{j,k}^{N_{\text{MM}}} q_j q_k \cos\left(\frac{2\pi}{L}(\mathbf{m} \cdot \mathbf{r}_{jk})\right). \quad (2.10)$$

More often, this quantity is further simplified by separating \mathbf{r}_j and \mathbf{r}_k , which allows the double summation to be recast as two identical single summations, with the result⁶

$$|S(\mathbf{m})|^2 = \left[\sum_j^{N_{\text{MM}}} q_j \cos\left(\frac{2\pi}{L}(\mathbf{m} \cdot \mathbf{r}_j)\right) \right]^2. \quad (2.11)$$

The latter form requires fewer operations to compute, and is therefore preferred. For QM/MM applications, however, some of the \mathbf{r}_j vectors correspond to QM atoms and sum to MM atoms, hence the simplifications leading to Eq. (2.11) will not be possible and Eq. (2.10) must be used instead.

2.2 Gradients

In the presence of long-range Ewald summation, the mutual Coulomb interaction of the simulation cell with itself and its periodic images can be partitioned into five terms. Letting $\hat{\nabla}_k$ denote the nuclear gradient operator with respect to atom k , we have

$$\begin{aligned}\hat{\nabla}_k E_{\text{cell}} &= \hat{\nabla}_k (E_{\text{real}} + E_{\text{self}} + E_{\text{recip}} + E_{\text{charge}} + E_{\text{dipole}}) \\ &= \hat{\nabla}_k (E_{\text{real}} + E_{\text{recip}} + E_{\text{dipole}}).\end{aligned}\tag{2.12}$$

The second equality follows because $\hat{\nabla}_k E_{\text{self}} = 0 = \hat{\nabla}_k E_{\text{charge}}$, since neither E_{self} nor E_{charge} depends on the atomic positions.

The real space energy is³

$$E_{\text{real}} = \sum_{\mathbf{n}} \sum_j^{N_{\text{MM}}} \sum_{k>j}^{N_{\text{MM}}} q_j q_k \frac{\text{erfc}(\eta|\mathbf{r}_{jk} + \mathbf{n}L|)}{|\mathbf{r}_{jk} + \mathbf{n}L|},\tag{2.13}$$

where $\mathbf{r}_{jk} = |\mathbf{r}_j - \mathbf{r}_k|$. The derivative of this energy is

$$\begin{aligned}\hat{\nabla}_k E_{\text{real}} &= - \sum_{\mathbf{n}} \sum_{j \neq k}^{N_{\text{MM}}} q_j q_k \left(\frac{\text{erfc}(\eta|\mathbf{r}_{jk} + \mathbf{n}L|)}{|\mathbf{r}_{jk} + \mathbf{n}L|} \right. \\ &\quad \left. + \frac{2\eta}{\sqrt{\pi}} \exp(-\eta^2|\mathbf{r}_{jk} + \mathbf{n}L|^2) \right) \frac{\mathbf{r}_{jk} + \mathbf{n}L}{|\mathbf{r}_{jk} + \mathbf{n}L|^2}.\end{aligned}\tag{2.14}$$

The restriction that $j \neq k$ arises from the fact that the case $j = k$ is already excluded for $\mathbf{n} = \mathbf{0}$ (lest we count an atom's Coulomb interaction with itself), and for $\mathbf{n} \neq \mathbf{0}$

and $j = k$ the distance $\mathbf{r}_{jk} = \mathbf{n}L$, which is no longer dependent on \mathbf{r}_k , hence its gradient in Eq. (2.14) vanishes.

The dipole energy is³

$$E_{\text{dipole}} = -\frac{\pi}{(2\varepsilon + 1)L^3} \sum_{j,k}^{N_{\text{MM}}} q_j q_k |\mathbf{r}_{jk}|^2 \quad (2.15)$$

and its gradient is

$$\hat{\nabla}_k E_{\text{dipole}} = \frac{4\pi}{(2\varepsilon + 1)L^3} \sum_{j \neq k}^{N_{\text{MM}}} q_j q_k \mathbf{r}_{jk} \cdot \quad (2.16)$$

Both the dipole energy and its gradient vanish under tin-foil boundary conditions ($\varepsilon \rightarrow 0$).

The final term in Eq. (2.12) is $\hat{\nabla}_k E_{\text{recip}}$. The quantity E_{recip} can be written in the compact form³

$$E_{\text{recip}} = \frac{1}{2} \sum_{\mathbf{m} \neq \mathbf{0}} w(\mathbf{m}) |S(\mathbf{m})|^2 \quad (2.17)$$

where

$$w(\mathbf{m}) = \left(\frac{L^2}{V\pi|\mathbf{m}|^2} \right) \exp\left(\frac{-\pi^2|\mathbf{m}|^2}{\eta^2 L^2} \right) \quad (2.18)$$

is independent of the atomic positions but

$$|S(\mathbf{m})|^2 = \sum_{j,k} q_j q_k \exp\left(\frac{2\pi i}{L} \mathbf{m} \cdot \mathbf{r}_{jk} \right) \quad (2.19)$$

depends on the atomic positions via \mathbf{r}_{jk} . Therefore

$$\hat{\nabla}_k E_{\text{recip}} = \frac{1}{2} \sum_{\mathbf{m} \neq \mathbf{0}} w(\mathbf{m}) \hat{\nabla}_k |S(\mathbf{m})|^2 \cdot \quad (2.20)$$

The gradient of the structure factor in this equation is

$$\hat{\nabla}_k |S(\mathbf{m})|^2 = \sum_j^{N_{\text{MM}}} \frac{4\pi q_j q_k}{L} \sin\left(\frac{2\pi}{L} \mathbf{m} \cdot \mathbf{r}_{jk} \right) \mathbf{m} \quad (2.21)$$

The restriction $j \neq k$ does not appear because this derivative vanishes identically for $j = k$.

The form of the gradient in Eq. (2.21) can always be used, *i.e.*, regardless of whether q_k is an MM point charge or an atomic point charge representing the image of the QM wave function. If the indices j and k are both MM point charges (so that summations over these indices run over the same set of atoms, then a more computationally efficient form of Eq. (2.21) is

$$\hat{\nabla}_k |S(\mathbf{m})|^2 = \frac{4\pi q_k}{L} \left\{ \left[\sum_j^{N_{\text{MM}}} q_j \sin\left(\frac{2\pi}{L} \mathbf{m} \cdot \mathbf{r}_j\right) \right] \cos\left(\frac{2\pi}{L} \mathbf{m} \cdot \mathbf{r}_k\right) - \left[\sum_j^{N_{\text{MM}}} q_j \cos\left(\frac{2\pi}{L} \mathbf{m} \cdot \mathbf{r}_j\right) \right] \sin\left(\frac{2\pi}{L} \mathbf{m} \cdot \mathbf{r}_k\right) \right\} \mathbf{m} \quad (2.22)$$

The terms in square brackets are sums over all MM atoms and have no dependence on \mathbf{r}_k . These sums can be computed once and stored, eliminating the need for a double sum over atomic indices. This makes Eq. (2.22) preferable to Eq. (2.21) for MM/MM interactions. For the interactions between a QM atom k and all MM atoms j , Eq. (2.21) will be used instead.

CHAPTER 3

QM/MM and PBC

The QM/MM-Ewald technique introduced by Nam *et al.*² is based upon the reasonable assumption that the simulation cell is large compared to the spatial extent of the QM wave function. As such, a large MM “buffer” screens the interaction between the electron density and its periodic images, so collapsing this density onto point charges for the purpose of computing these long-range Coulomb interactions should not engender serious error. Once the density is reduced to point charges, classical Ewald summation can be applied. In this section, we describe the basic theory behind obtaining PBC corrections to the SCF energy and Fock matrix. In developing this theory, we are concerned only with electrostatic interactions, as other QM/MM interactions such as non-bonded Lennard-Jones interactions operate on shorter length scales and PBC implementations based on smooth cutoffs should be fine. Thus, “total” energy will refer to the QM electronic structure energy plus all MM and QM/MM electrostatic interactions; other MM interactions can simply be tacked on to the formulas appearing below.

3.1 Energy corrections

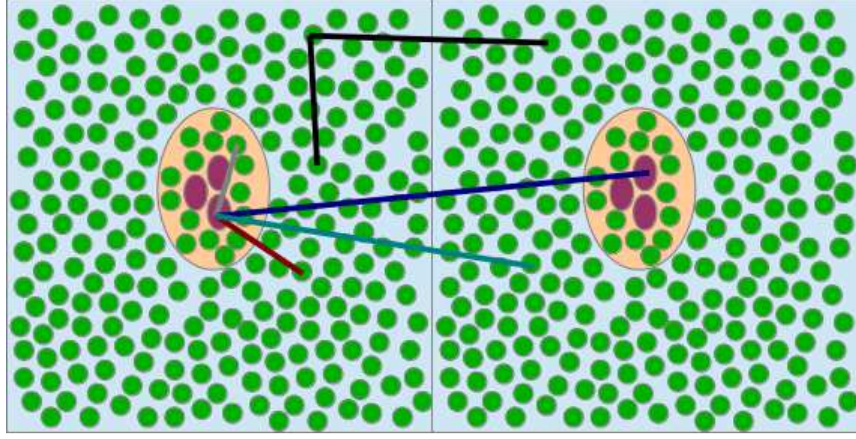
We first write the total QM/MM supersystem (SS) energy, which includes the interactions between all periodic images, as

$$E_{\text{total}} = E_{\text{QM-QM}}^{\text{SS}} + E_{\text{QM-MM}}^{\text{SS}} + E_{\text{MM-MM}}^{\text{SS}}. \quad (3.1)$$

The final term, $E_{\text{MM-MM}}^{\text{SS}}$, can be evaluated using a standard, classical Ewald summation and need not be discussed further. It is helpful to partition the other two terms into interactions between atoms in the simulation cell with other atoms in the simulation cell, which we will call the real-space (RS) interactions, and also interactions between the simulation cell and atoms contained in the periodic images (PI). The SS energies in Eq. (3.1) can thus be broken down into RS and PI parts, shown pictorially in fig. 3.1:

$$\begin{aligned} E_{\text{total}} = & E_{\text{QM-QM}}^{\text{RS}} + \Delta E_{\text{QM-QM}}^{\text{PI}} + E_{\text{QM-MM}}^{\text{RS}} \\ & + \Delta E_{\text{QM-MM}}^{\text{PI}} + E_{\text{MM-MM}}^{\text{SS}}. \end{aligned} \quad (3.2)$$

The term $E_{\text{QM-QM}}^{\text{RS}}$ (interaction between QM atoms in the simulation cell with other QM atoms in the simulation cell) is simply the result of some QM electronic structure calculation. The term $E_{\text{QM-MM}}^{\text{RS}}$ results from some QM/MM interaction scheme; note that Eq. (3.1) tacitly assumes an “additive” QM/MM scheme, as opposed to a “subtractive” scheme such as ONIOM.¹³ For the latter, there are no QM periodic images so Ewald summation involves the MM system only, and is therefore straightforward.



$$E_{QM-QM}^{RS} + \Delta E_{QM-QM}^{PI} + E_{QM-MM}^{RS} + \Delta E_{QM-MM}^{PI} + E_{MM-MM}^{SS}$$

Figure 3.1: A color coded pictorial representation, where the orange oval represents the QM region, for all interactions included in E_{total} .

The remaining terms in Eq. (3.2) are calculated as differences between a SS calculation and a RS calculation,

$$\Delta E^{PI} = E^{SS} - E^{RS} . \quad (3.3)$$

In particular, ΔE_{QM-MM}^{PI} is obtained using a QM region embedded in a periodically-replicated supercell of MM regions, but without replication of the QM region. This interaction energy can be decomposed into real- and reciprocal-space parts, the latter of which will involve only MM atoms provided that the QM region is fully enveloped by the short-range part of the Coulomb potential in Eq. (2.1). The term ΔE_{QM-QM}^{PI} in Eq. (3.2) is obtained from a periodic array point charges obtained from collapsing the QM electron density onto atom-centered charges, as described below and shown in fig. 3.2.

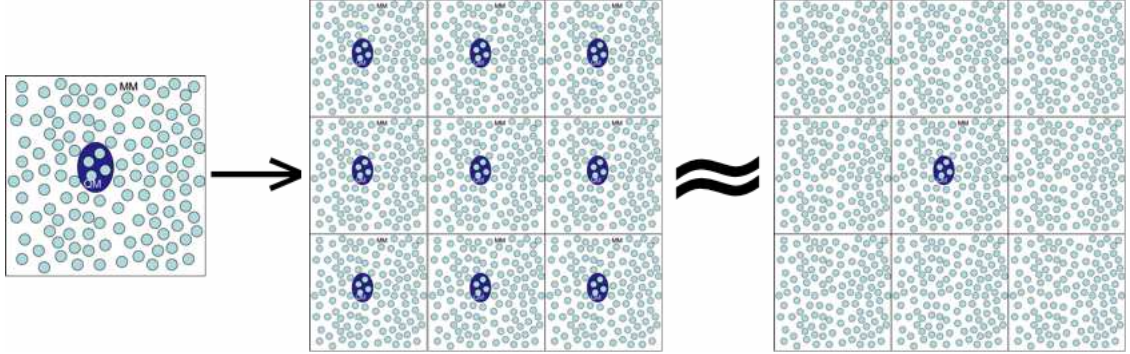


Figure 3.2: A pictorial representation of a replicated QM/MM simulation cell, where the blue oval represents the QM region (second image), and its collapse onto point charges (third image). Note that any circle not surrounded by the QM region is a point charge.

Applying the fundamental assumption that the QM images are far apart and screened by a wide buffer of MM charges, the calculation effectively reduces to a series of pairwise Coulomb interactions between the atoms in the simulation cell and those contained in the periodic images. It is therefore expected that $\Delta E_{\text{QM-QM}}^{\text{PI}}$ and $\Delta E_{\text{QM-MM}}^{\text{PI}}$ will have similar forms:

$$\Delta E_{\text{QM-QM}}^{\text{PI}} = \frac{1}{2} \sum_{\alpha, \beta}^{N_{\text{QM}}} Q_{\alpha} Q_{\beta} \omega(\mathbf{r}_{\alpha\beta}) \quad (3.4)$$

and

$$\Delta E_{\text{QM-MM}}^{\text{PI}} = \sum_{\alpha}^{N_{\text{QM}}} \sum_j^{N_{\text{MM}}} Q_{\alpha} q_j \omega(\mathbf{r}_{\alpha j}) . \quad (3.5)$$

For clarity, we use Q_{α} to denote the partial charge on a QM atom and q_j to denote the partial charge on an MM atom. As in standard Ewald summation, both quantities can be described by a potential function, $\omega(\mathbf{r})$. For a neutral simulation cell ($Q_{\text{tot}} = 0$)

with tin-foil boundary conditions ($\varepsilon = \infty$), this potential is

$$\begin{aligned} \omega(\mathbf{r}_{\alpha\beta}) = & \sum_{\mathbf{m} \neq \mathbf{0}} \frac{e^{-\pi^2 |\mathbf{m}|^2 / \eta^2 L^2}}{L\pi |\mathbf{m}|^2} \cos\left(\frac{2\pi}{L} (\mathbf{m} \cdot \mathbf{r}_{\alpha\beta})\right) \\ & + \sum_{\mathbf{n} \neq \mathbf{0}} \frac{\operatorname{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{n}L|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{n}L|} - \frac{\operatorname{erf}(\eta |\mathbf{r}_{\alpha\beta}|)}{|\mathbf{r}_{\alpha\beta}|} \end{aligned} \quad (3.6)$$

In obtaining this result, the form of $|S(\mathbf{m})|^2$ in Eq. (2.10) has been used in the reciprocal (first) term in Eq. (3.6). The reason to prefer this form, as opposed to that in Eq. (2.11), is that the latter requires that each of the indices runs over the same sum, which is not the case for QM/MM interactions where one of the summation indices in Eq. (2.9) represents QM atoms while the other represents MM atoms.

The potential in Eq. (3.6) warrants some comments. First, the term containing a sum over $\mathbf{m} \neq \mathbf{0}$ is directly analogous to the reciprocal term in Eq. (2.17), whereas the erf and erfc terms are analogous to the real space term in Eq. (2.13). The different appearance of the erf term (representing the $\mathbf{n} = \mathbf{0}$ vector) is due to the fact that an energy $E_{\text{real}}^{\text{RS}}$, with a Coulomb potential of $1/r$, has been subtracted out of $E_{\text{real}}^{\text{SS}}$ with a Coulomb potential of $\operatorname{erfc}(\eta r)/r$, to afford $\Delta E_{\text{real}}^{\text{PI}}$, with a Coulomb potential of $-\operatorname{erf}(\eta r)/r$.

It should also be noted that nothing analogous to the self-energy is immediately apparent in Eq. (3.6). In the QM-QM PI correction [Eq. (3.4)], there is no restriction on the sum and so $\alpha = \beta$ is allowed. For $\alpha = \beta$, the Coulomb interaction is given by

$$\lim_{r \rightarrow 0} \frac{\operatorname{erf}(\eta r)}{r} = \frac{2\eta}{\sqrt{\pi}}, \quad (3.7)$$

which is in fact the self term. Furthermore, there is no self energy corresponding to the QM-MM PI correction [Eq. (3.5)], since the atom types in the two summations

are different.

Finally, it is worth noting that the potential in Eq. (3.6) differs from the Ewald potential given by Nam *et al.*,² insofar as the term in Eq. (3.6) containing the sum over $\mathbf{n} \neq \mathbf{0}$ is absent in Ref. 2. The authors of Ref. 2 assume that the Ewald parameter η has been chosen such that only the simulation cell must be considered in the real-space portion of the Ewald sum. This is a reasonable assumption but is not assumed *a priori* in this work, on the basis that cost considerations for more general QM/MM calculations might favor a different partition of the effort. In the case of a charged system, the charge term [Eq. (2.6)] is included in the MM Ewald summation. We henceforth assume tin-foil boundary conditions and therefore omit the dipole term in Eq. (2.2).

3.2 Fock matrix corrections

The corrections above must now be incorporated into the Fock matrix, which is computed by taking the derivative of the energy with respect to the density matrix.

Using the chain rule, this correction can be expressed as

$$\Delta F_{\mu\nu}^{\text{PI}} \equiv \frac{\partial \Delta E^{\text{PI}}}{\partial P_{\mu\nu}} = \sum_{\alpha}^{N_{\text{QM}}} \frac{\partial \Delta E^{\text{PI}}}{\partial Q_{\alpha}} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \quad (3.8)$$

where $\Delta E^{\text{PI}} = E_{\text{QM-QM}}^{\text{PI}} + E_{\text{QM-MM}}^{\text{PI}}$. The energy derivative with respect to an atomic point charge can be evaluated directly from Eqs. (3.4) and (3.5):

$$\frac{\partial \Delta E^{\text{PI}}}{\partial Q_{\alpha}} = \sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega(\mathbf{r}_{\alpha\beta}) + \sum_j^{N_{\text{MM}}} q_j \omega(\mathbf{r}_{\alpha j}). \quad (3.9)$$

The cost of evaluating Eq. (3.9) can be significantly reduced by recognizing that the Ewald potential depends upon the positions of the atoms (both QM and MM),

but not on any details of the electronic structure. Those details are encoded into the QM charges Q_β , which are the only quantities in Eq. (3.9) that change from one SCF cycle to the next. Thus, we can pre-compute the Ewald potential at the relevant interatomic distances prior to entering the SCF iterations. In anticipation of doing this, let us define a column vector

$$\boldsymbol{\omega}_\alpha = \left[\omega(\mathbf{r}_{\alpha 1}) \quad \omega(\mathbf{r}_{\alpha 1}) \quad \cdots \quad \omega(\mathbf{r}_{\alpha N_{\text{QM}}}) \quad \sum_j^{N_{\text{MM}}} q_j \omega(\mathbf{r}_{\alpha j}) \right]^\dagger. \quad (3.10)$$

The final entry in this vector is identical to the second term in Eq. (3.9). Next, define another column vector

$$\mathbf{Q} = [Q_1 \quad Q_2 \quad \cdots \quad Q_{N_{\text{QM}}} \quad 1]^\dagger. \quad (3.11)$$

(Save for the final entry, the vector \mathbf{Q} consists simply of the QM atomic charges.)

Using this new notation, we can rewrite Eq. (3.9) as

$$\Delta F_{\mu\nu}^{\text{PI}} = \sum_\alpha^{N_{\text{QM}}} \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \mathbf{Q}^\dagger \boldsymbol{\omega}_\alpha. \quad (3.12)$$

It remains to evaluate the charge derivatives $\partial Q_\alpha / \partial P_{\mu\nu}$. The form of these derivatives depends upon the charge scheme that is used (Mulliken, Löwdin, ChEIPG⁴, etc.). For Mulliken or Löwdin charges, these derivatives are quite simple. The Mulliken atomic charges, for example, are defined as

$$Q_\alpha = Z_\alpha - \sum_{\substack{\mu \in \alpha \\ \nu}} P_{\mu\nu} S_{\mu\nu}, \quad (3.13)$$

hence the requisite derivatives are nothing more than overlap matrix elements:

$$\frac{\partial Q_\alpha}{\partial P_{\mu\nu}} = -S_{\mu\nu} \delta_{\mu \in \alpha}. \quad (3.14)$$

Here, $\delta_{\mu\in\alpha}$ is a Kronecker delta-type symbol signifying that both atomic orbital (AO) basis function μ must be centered on atom α , else the derivative is zero by definition. It is not terribly surprising to discover that the Mulliken charges are unstable in extended basis sets, and because in this context these charges make their way into the Fock matrix, we find that the Mulliken-based QM/MM-Ewald scheme is difficult or impossible to converge in extended basis sets. (Data to this effect are provided in Section 6; we encounter similar difficulties in attempting to use Mulliken or Löwdin charges in the context of the self-consistent XPol charge-embedding procedure.¹⁴) ChElPG charges (shown in Chapter 5), on the other hand, appear to be stable and robust, but the derivatives $\partial Q_\alpha/\partial P_{\mu\nu}$ are far more costly in the ChElPG case. These derivatives will be given explicitly below, following a discussion of the basic theory behind ChElPG charges.

3.3 Position derivative of the Ewald energy and potential

The energy of a periodically replicated QM/MM system can be expressed as³

$$E_{\text{total}} = E_{\text{QM-QM}}^{\text{RS}} + E_{\text{QM-MM}}^{\text{RS}} + E_{\text{MM-MM}}^{\text{SS}} + \Delta E^{\text{PI}} \quad (3.15)$$

where E_{total} is total energy of the simulation cell, $E_{\text{QM-QM}}^{\text{RS}}$ and $E_{\text{QM-MM}}^{\text{RS}}$ are the simulation cell QM and QM/MM energies, respectively, in real space (RS), and $E_{\text{MM-MM}}^{\text{SS}}$ is the MM energy for the supersystem (SS), meaning the simulation cell and its periodic images. Lastly, ΔE^{PI} is the energy of interaction between the simulation cell and the periodic images. Assuming a neutral system with tin-foil boundary conditions, this

final term can be written

$$\Delta E^{\text{PI}} = \sum_{\alpha}^{N_{\text{QM}}} Q_{\alpha} \left[\frac{1}{2} \sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega(\mathbf{r}_{\alpha\beta}) + \sum_j^{N_{\text{MM}}} q_j \omega(\mathbf{r}_{\alpha j}) \right] \quad (3.16)$$

where Q_{α} is the charge on a QM atom and q_j is the charge on an MM atom. This results in a Fock matrix correction, $\Delta F_{\mu\nu}^{\text{PI}} \equiv \partial(\Delta E^{\text{PI}})/\partial P_{\mu\nu}$. This correction is^{2,3}

$$\Delta F_{\mu\nu}^{\text{PI}} = \sum_{\alpha}^{N_{\text{QM}}} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right). \quad (3.17)$$

The quantity $\omega_{\alpha k} \equiv \omega(\mathbf{r}_{\alpha k})$ is the real-space Ewald potential,³

$$\begin{aligned} \omega(\mathbf{r}_{\alpha k}) = & \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{\cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_{\alpha k})}{V \pi |\mathbf{a}_m|^2} e^{-\pi^2 |\mathbf{a}_m|^2 / \eta^2} \\ & + \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha k} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha k} + \mathbf{b}_n|} - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha k}|)}{|\mathbf{r}_{\alpha k}|}. \end{aligned} \quad (3.18)$$

The gradient of ΔE^{PI} with respect to a nuclear displacement is different depending upon whether it is a QM or an MM atom that is displaced. We first consider a displacement, x_{γ} . Note that the MM charges are fixed but the QM charges are not, and we have

$$\begin{aligned} \frac{\partial \Delta E^{\text{PI}}}{\partial x_{\gamma}} = & \sum_{\alpha}^{N_{\text{QM}}} Q_{\alpha} Q_{\gamma} \frac{\partial \omega_{\alpha\gamma}}{\partial x_{\gamma}} (1 - \delta_{\alpha\gamma}/2) + \sum_j^{N_{\text{MM}}} q_j Q_{\gamma} \frac{\partial \omega_{\gamma j}}{\partial x_{\gamma}} + \sum_{\alpha, \beta}^{N_{\text{QM}}} \frac{\partial Q_{\alpha}}{\partial x_{\gamma}} Q_{\beta} \omega_{\alpha\beta} \\ & + \sum_j^{N_{\text{MM}}} \sum_{\alpha}^{N_{\text{QM}}} q_j \frac{\partial Q_{\alpha}}{\partial x_{\gamma}} \omega_{\alpha j}. \end{aligned} \quad (3.19)$$

In practice, we can eliminate the $\delta_{\alpha\gamma}$ term since the dependence on nuclear coordinates vanishes for $\alpha = \gamma$, since $\mathbf{r}_{\alpha\alpha} = 0$. Thus

$$\frac{\partial \Delta E^{\text{PI}}}{\partial x_{\gamma}} = \sum_{\alpha}^{N_{\text{QM}}} Q_{\alpha} Q_{\gamma} \frac{\partial \omega_{\alpha\gamma}}{\partial x_{\gamma}} + \sum_j^{N_{\text{MM}}} q_j Q_{\gamma} \frac{\partial \omega_{\gamma j}}{\partial x_{\gamma}} + \sum_{\alpha}^{N_{\text{QM}}} \frac{\partial Q_{\alpha}}{\partial x_{\gamma}} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right). \quad (3.20)$$

To obtain the derivative of ΔE^{PI} with respect to an MM coordinate x_k , the MM charges come from a force field and are independent of the atom position. The QM charges are dependent on the MM atom positions because the charges are dependent on the density matrix. Thus, yielding

$$\frac{\partial \Delta E^{\text{PI}}}{\partial x_k} = \sum_{\alpha}^{N_{\text{QM}}} q_k Q_{\alpha} \frac{\partial \omega_{\alpha k}}{\partial x_k} + \sum_{\alpha}^{N_{\text{QM}}} \frac{\partial Q_{\alpha}}{\partial x_k} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right). \quad (3.21)$$

Expressions for the charge derivatives $\partial Q_{\alpha}/\partial x_{\gamma}$ are not universal and depend upon how the QM charges are obtained from the wave function. These derivatives will be discussed later. The derivative of the Ewald potential is universal, and we have

$$\hat{\nabla}_{\gamma} \omega_{\gamma k} = \left[\frac{\text{erf}(\eta |\mathbf{r}_{\gamma k}|)}{|\mathbf{r}_{\gamma k}|} - \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma k}|^2} \right] \frac{\mathbf{r}_{\gamma k}}{|\mathbf{r}_{\gamma k}|^2} - 2\pi \sum_{\mathbf{a}_m \neq \mathbf{0}} \omega(\mathbf{a}_m) \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_{\gamma k}) \mathbf{a}_m - \sum_{\mathbf{b}_n \neq \mathbf{0}} \left(\frac{\text{erfc}(\eta |\mathbf{r}_{\gamma k} + \mathbf{b}_n|)}{|\mathbf{r}_{\gamma k} + \mathbf{b}_n|} + \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma k} + \mathbf{b}_n|^2} \right) \frac{\mathbf{r}_{\gamma k} + \mathbf{b}_n}{|\mathbf{r}_{\gamma k} + \mathbf{b}_n|^2} \quad (3.22)$$

where we have opted for the use of Eq. (2.21) in the reciprocal term. The derivative with respect to a MM nuclear coordinate is obtained easily by recognizing that

$$\hat{\nabla}_{\gamma} \omega_{\gamma k} = -\hat{\nabla}_k \omega_{\gamma k}. \quad (3.23)$$

3.4 Position derivative of the SCF energy

The derivative of the Hartree-Fock energy ε_{HF} is

$$\frac{\partial \varepsilon_{\text{HF}}}{\partial x} = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{\text{nuc}}}{\partial x} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} F_{\mu\nu} \quad (3.24)$$

where the Fock matrix elements are

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle P_{\lambda\sigma}, \quad (3.25)$$

and $P_{\mu\nu}$ represents a density matrix. Equation (3.24) is adaptable to DFT in a straightforward way, so for brevity we do not include the exchange-correlation term in the derivation that follows. As shown long ago,¹⁵ Eq. (3.24) is needlessly expensive because the density matrix derivatives $\partial P_{\mu\nu}/\partial x$ can be eliminated in favor of the energy-weighted density matrix,

$$W_{\mu\nu} = \sum_i \varepsilon_i c_{i\mu}^* c_{\nu i}, \quad (3.26)$$

so that

$$\frac{\partial \varepsilon_{\text{HF}}}{\partial x} = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{\text{nuc}}}{\partial x} - \sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x}. \quad (3.27)$$

This transformation relies on two facts:

$$\sum_{\nu} F_{\mu\nu} c_{\nu i} = \sum_{\nu} \varepsilon_i S_{\mu\nu} c_{\nu i} \quad (3.28)$$

and

$$\sum_{\mu\nu} c_{i\mu}^* S_{\mu\nu} c_{\nu j} = \delta_{ij}. \quad (3.29)$$

Let us use the chain rule to write the charge derivatives $\partial Q_{\alpha}/\partial x_{\gamma}$ as

$$\frac{\partial Q_{\alpha}}{\partial x_{\gamma}} = \sum_{\mu\nu} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \frac{\partial P_{\mu\nu}}{\partial x_{\gamma}} + \sum_M \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial M}{\partial x_{\gamma}}. \quad (3.30)$$

Here, M represents any quantity on which Q_{α} depends, save for the density matrix. Mulliken charges, for example, depend on the overlap matrix elements $S_{\mu\nu}$, hence $M = S_{\mu\nu}$ in that case. These ‘‘M-derivatives’’ will be evaluated below, for both Mulliken and ChEIPG charges. In the general case, we can use Eq. (3.30) to rewrite

Eq. (3.20) as

$$\begin{aligned} \frac{\partial \Delta E^{\text{PI}}}{\partial x_\gamma} = & \sum_{\alpha}^{N_{\text{QM}}} Q_{\alpha} Q_{\gamma} \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma} + \sum_j^{N_{\text{MM}}} q_j Q_{\gamma} \frac{\partial \omega_{\gamma j}}{\partial x_\gamma} + \sum_{\alpha}^{N_{\text{QM}}} \sum_{\mu\nu} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} \right. \\ & \left. + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right) + \sum_{\alpha}^{N_{\text{QM}}} \sum_M \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial M}{\partial x_\gamma} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right). \end{aligned} \quad (3.31)$$

Note that the third term in this equation looks like the Ewald Fock matrix correction [Eq. (3.17)] contracted with the density matrix derivative $\partial P_{\mu\nu}/\partial x_\gamma$:

$$\sum_{\alpha}^{N_{\text{QM}}} \sum_{\mu\nu} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right) = \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \Delta F_{\mu\nu}^{\text{PI}}. \quad (3.32)$$

In the gradient of the *total* energy, $\partial(\varepsilon_{\text{HF}} + \Delta E^{\text{PI}})/\partial x_\gamma$, this term is simply added to the final term in Eq. (3.24), and is eventually folded into the term containing the energy-weighted density matrix, since the Fock matrix for the periodic calculations has matrix elements $F_{\mu\nu} + \Delta F_{\mu\nu}^{\text{PI}}$. Therefore, the final gradient expression is obtained by combining Eqs. (3.24) and (3.31) and using Eq. (3.32):

$$\begin{aligned} \frac{\partial (\varepsilon_{\text{HF}} + \Delta E^{\text{PI}})}{\partial x_\gamma} = & \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_\gamma} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x_\gamma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{\text{nuc}}}{\partial x_\gamma} - \\ & \sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} + \sum_{\alpha}^{N_{\text{QM}}} Q_{\alpha} Q_{\gamma} \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma} + \sum_j^{N_{\text{MM}}} q_j Q_{\gamma} \frac{\partial \omega_{\gamma j}}{\partial x_\gamma} + \\ & \sum_{\alpha}^{N_{\text{QM}}} \sum_M \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial M}{\partial x_\gamma} \left(\sum_{\beta}^{N_{\text{QM}}} Q_{\beta} \omega_{\alpha\beta} + \sum_j^{N_{\text{MM}}} q_j \omega_{\alpha j} \right). \end{aligned} \quad (3.33)$$

There are three extra terms as compared to the traditional Hartree-Fock gradient. Derivatives of the Ewald potential can be evaluated analytically [Eq. (3.22)]. Derivatives of the QM atomic charges, which appear in the final term, are discussed below.

The MM derivative is simpler. When an MM atomic coordinate x_k is displaced, the analogue of Eq. (3.30) for the QM charge derivatives reads

$$\frac{\partial Q_\alpha}{\partial x_k} = \sum_{\mu\nu} \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \frac{\partial P_{\mu\nu}}{\partial x_k}. \quad (3.34)$$

The ‘‘M-derivative’’ term in Eq. (3.30) is absent, because quantities such as $S_{\mu\nu}$ (and other quantities required to evaluate the ChEIPG charges) do not depend on the MM atomic positions. Displacement of x_k appears in $d\varepsilon_{\text{HF}}/dx_k$ in the form $dH_{\mu\nu}/dx_k$ [see Eq. (3.27)] This results in a derivative

$$\frac{\partial (\varepsilon_{\text{HF}} + \Delta E^{\text{PI}})}{\partial x_k} = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_k} + \frac{\partial V_{\text{nuc}}}{\partial x_k} + \sum_{\alpha}^{N_{\text{QM}}} q_k Q_\alpha \frac{\partial \omega_{\alpha k}}{\partial x_k}. \quad (3.35)$$

3.5 Mulliken Charge Derivatives

Mulliken atomic charges were used in the original, minimal-basis QM/MM-Ewald method of Ref. 2, but perform poorly (sometimes leading to SCF convergence failure) in larger basis sets, or large QM regions.³ For completeness, however, we consider the Mulliken case here, as it provides a gentle introduction to the more difficult ChEIPG case that is considered in Section 5.4.

What is needed is to evaluate the ‘‘M-derivatives’’ $\sum_M (\partial Q_\alpha / \partial M) (\partial M / \partial x_\gamma)$, for all independent variables upon which Q_α depends *except* $P_{\mu\nu}$, as that dependence is already folded into the energy-weighted density matrix. Mulliken charges are defined as

$$Q_\alpha = Z_\alpha - \sum_{\substack{\mu \in \alpha \\ \nu}} P_{\mu\nu} S_{\mu\nu}, \quad (3.36)$$

so the independent variables are the $P_{\mu\nu}$ and the $S_{\mu\nu}$. Therefore the M-derivative is

$$\begin{aligned} \sum_M \frac{\partial Q_\alpha}{\partial M} \frac{\partial M}{\partial x_\gamma} &= \sum_{\mu\nu} \frac{\partial Q_\alpha}{\partial S_{\mu\nu}} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} \\ &= - \sum_{\substack{\mu \in \alpha \\ \nu}} P_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma}. \end{aligned} \tag{3.37}$$

CHAPTER 4

Parameters for the Ewald sums

In addition to the charge scheme, the user-controlled Ewald parameter, η , can greatly influence the calculation time. This parameter controls how much of the pairwise sum is performed in real space, and thus controls how many vectors are required to converge the vector summations in the real-space term [Eq. (2.13)] and the reciprocal-space term [Eq. (2.17)]. Both summations converge as Gaussian functions.⁵ Following Ref. 5, we thus choose a constant, C , such that $\exp(C^2)$ is within a specified convergence threshold. We take this to be the same as the threshold (drop tolerance) used for the one- and two-electron integrals,

$$C = \sqrt{-\ln(\text{Integral Threshold})} . \quad (4.1)$$

Unless otherwise stated, the integral threshold will be set to 10^{-8} here, the default integral threshold for a single point energy calculation in Q-Chem.

In real space, the argument of the complimentary error function [see Eq. (2.13)] controls the convergence, hence we want

$$\exp(C^2) \leq \exp(\eta^2 |\mathbf{r}_{\alpha\beta} + \mathbf{n}_{\max} L|^2) , \quad (4.2)$$

where the vector $\mathbf{n}_{\max} = (n_{\max}, 0, 0)$ specifies how many periodic boxes one must use in the calculation to achieve a required level of accuracy. This is equivalent to figuring out how far away two atoms must be before their pairwise interaction contributes less than the integral threshold. Thus we obtain

$$C \leq \eta |\mathbf{r}_{\alpha\beta} + \mathbf{n}_{\max} L|. \quad (4.3)$$

Each of the components of $\mathbf{r}_{\alpha\beta}$ must be less than the box length, L . Replacing $\mathbf{n}_{\max} + 1$ with \mathbf{n}_{\max} for convenience, one obtains

$$C \leq \eta |\mathbf{n}_{\max} L| = \eta n_{\max} L \quad (4.4)$$

and therefore

$$n_{\max} = \text{ceiling} \left(\frac{C}{\eta L} \right). \quad (4.5)$$

Equation (4.5) specifies the largest vector that must be included in the real-space sum in order to achieve a certain drop tolerance. If the integers n_x, n_y, n_z are run from $-n_{\max}$ to n_{\max} , however, there are unnecessary vectors that are included in this “supercube”. The farthest *distance* that needs to be considered is actually $|\mathbf{n}_{\max}|$, so we need include only those lattice vectors satisfying the condition $|\mathbf{n}| \leq |\mathbf{n}_{\max}|$. Enforcing the condition creates a “supersphere” where some lattice vectors from the corners of the supercube have been excluded. Note from Eq. (4.5) that $n_{\max} = 0$ when $C/\eta L < 1/2$. This condition leads to a cutoff radius, $R_c = C/\eta < L/2$, so that all significant interactions are included using the minimum-image convention (cutoff at half the box length) in the real-space sum. In this case, the real-space sum is calculated only within the simulation cell, with the result that the $\mathbf{n} \neq \mathbf{0}$ term in

Eq. (3.6) is zero. This is in accordance with the assumption made by Nam *et al.*² and is often the case for large simulation cells.

The requisite number of reciprocal-space vectors \mathbf{m} is calculated in the same manner. Since the argument of the exponential function in Eq. (2.17) is $\pi^2|\mathbf{m}|^2/\eta^2L^2$, this quantity replaces $(\eta|\mathbf{r}_{\alpha\beta} + \mathbf{n}_{\max}L|)^2$ in an inequality similar to Eq. (4.2), with the result

$$m_{\max} = \text{ceiling} \left(\frac{CL\eta}{\pi} \right) . \quad (4.6)$$

Each of the elements in \mathbf{m} runs from $-m_{\max}$ to m_{\max} with $\mathbf{0}$ excluded and subject to a constraint that $|\mathbf{m}| < |\mathbf{m}_{\max}|$.

Now it is possible to determine the exact number of vectors that will be needed for the calculation, given a particular value of η . The number of total vectors in the supercube is

$$v_{\text{tot}} = (2n_{\max} + 1)^3 + (2m_{\max} + 1)^3 . \quad (4.7)$$

The number of total vectors that satisfy the constraints (*i.e.*, the supersphere) is not so easily computed but can be determined through recursion relations. In order to find these numbers the reader should consult Sloane's handbook of integer sequences, (specifically, series A000605).^{16,17}

In this work, we employ standard Ewald summation as opposed to the particle-mesh Ewald technique^{18,19} that is more common in strictly classical simulations. As such, the cost of both the real- and reciprocal-space sums scales quadratically with the number of vectors. As such, the minimum number of vectors leads to the fastest

calculation and the value of η that affords this minimum is the optimal Ewald parameter. This value can be determined by numerical solution of the equation

$$\frac{2CL^3\eta^3}{\sqrt{\pi^3}} + \frac{L^2\eta^2}{\sqrt{\pi}} - L\eta - 2C = 0 . \quad (4.8)$$

Alternatively, one may build a table where one finds the number of total vectors in the supercube and then chooses the η value corresponding to the smallest number of vectors.

CHAPTER 5

ChEIPG charges

5.1 Basic theory

By construction, the ChEIPG atomic charges minimize, in a least-squares sense, the difference between the QM electrostatic potential (evaluated on a grid) and the electrostatic potential derived from a set of atom-centered point charges (evaluated on the same grid), subject to the constraint that the atomic charges must sum to the molecular charge.⁴ A complete discussion of the ChEIPG formalism, using the same notation that is used here, can be found in Ref. 20. Briefly, the ChEIPG charges are given by

$$Q_A = \sum_B^{N_{\text{MM}}} e_B (\mathbf{G}^{-1})_{BA} - \left(\frac{\sum_{BC}^{N_{\text{MM}}} e_B (\mathbf{G}^{-1})_{BC} - Q_{\text{tot}}}{\sum_{BC}^{N_{\text{MM}}} (\mathbf{G}^{-1})_{BC}} \right) \sum_B^{N_{\text{MM}}} (\mathbf{G}^{-1})_{BA} \quad (5.1)$$

where

$$e_B = \sum_k^{N_{\text{grid}}} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_B|} \left(\sum_J^{N_{\text{MM}}} \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu}^{N_{\text{basis}}} (\mathbf{I}_k)_{\mu\nu} P_{\mu\nu} \right) \quad (5.2)$$

and

$$G_{AB} = \sum_k^{N_{\text{grid}}} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_A| |\mathbf{r}_k - \mathbf{r}_B|} . \quad (5.3)$$

The indices A, B, \dots index nuclei and k indexes the grid points on which the electrostatic potential is evaluated; the quantity w_k is the weight assigned to the k th grid point. We have previously introduced a weighting scheme that ensures that the charges are continuous functions of the atomic coordinates,^{14,20} although a different weighting scheme, a first attempt at a speedup, will be used in this section, as described below. Finally, the quantity

$$(\mathbf{I}_k)_{\mu\nu} = \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle_{\mathbf{r}} \quad (5.4)$$

is a charge–density Coulomb integral. (The subscript \mathbf{r} indicates that the electron coordinate \mathbf{r} is the integration variable.)

Historically, the nature of the grid on which to evaluate the electrostatic potential was a source of debate, with various incarnations of the least-square fitting algorithm using different types of grids. One early algorithm²¹ (originally called “ChEIP”) used a set of concentric, atom-centered spherical grids. However, the charges thus obtained from were shown to be sensitive to molecular conformation,⁴ which was problematic because a main goal was to use ChEIP charges to parameterize force fields. The ChEIPG algorithm⁴ (so called to distinguish it from ChEIP) consists of replacing these spherical grids with a Cartesian grid, deleting points within the van der Waals region in order to fit to the long-range parts of the electrostatic potential. Although this reduced the conformational dependence of the charges, it was later demonstrated that atom-centered grids (including those with icosahedral symmetry) lead to far better rotational invariance of the charges, as compared to Cartesian grids.²²

In the present context, contraction of the integrals $(\mathbf{I}_k)_{\mu\nu}$ with certain quantities that arise in the construction of $\partial Q_\alpha/\partial P_{\mu\nu}$ proves to be a serious bottleneck, and it therefore behooves us to reduce the number of grid points (N_{grid}) as much as possible. Since atom-centered Lebedev grids of octahedral symmetry are already ubiquitous in Gaussian-orbital-based DFT codes, this was the natural choice to explore.

5.2 Lebedev grid implementation

This section documents our implementation of Eq. (5.1) using atom-centered Lebedev grids. (Once the grid is constructed, this works like any other ChEIPG algorithm, but there are some numerical aspects worthy of discussion.) It could be argued that the charges thus obtained should no longer be called “ChEIPG” charges, since the only difference between the ChEIP and ChEIPG algorithms is how the grid is constructed. However, the ChEIPG acronym is widely known and emphasizes the fact that there is a grid-based aspect to the calculation. Thus, we refer to our algorithm as a Lebedev grid-based implementation of the ChEIPG charges. We will retain some of the terminology from the original paper on ChEIPG charges.⁴ Namely, the *head space* refers to the distance from the van der Waals (vdW) surface to the outermost radial shells that constitute the grid. (There are no points within the vdW surface.) Also, let Δx denote the spacing between radial shells. These two parameters, along with the number of Lebedev points per shell (N_p), serve to define the ChEIPG grid. The Lebedev grid with N_p points on the unit sphere is constructed on each atom and

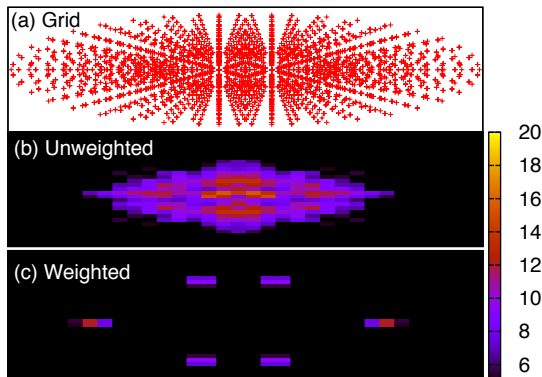


Figure 5.1: Plot of the Lebedev grid for H_2 in plane containing the internuclear axis. (a) Actual grid using $N_p = 302$ points per radial shell and nine shells per atom. (b) Number of grid points contained in each $0.25 \text{ \AA} \times 0.25 \text{ \AA}$ cell, equivalent to the contribution of each cell to the least-squares fit when all the weights w_k are identical. (c) Each cell's contribution when a simple weighting scheme ($w_k = 1/n_k$) is used.

then its radius is scaled by a factor

$$\alpha_i = \mathbf{r}_{\text{vdW}} + (i - 1)\Delta x \quad (5.5)$$

for the i th shell. Radial shells are constructed from \mathbf{r}_{vdW} out to the head space distance. Aside from the symmetry of the grids, this procedure is similar to that used by Spackman²² to evaluate Cartesian versus atom-centered grids, and also to one of the original ChEIP algorithms.²¹

Although it is possible to perform a weighted least squares fit of the electrostatic potential using the weights w_k in Eq. (5.1), the original ChEIPG paper of Breneman and Wiberg⁴ sets all $w_k = 1$, and the authors in fact emphasize the importance of using an isotropic grid to reduce conformational dependence of the charges. The use of atom-centered Lebedev grids leads to a highly anisotropic coverage of real space, as

is evident from the Lebedev grid for H₂, which is shown in Fig. 5.1(a). In Fig. 5.1(b), we plot the number of grid points contained in each 0.25 Å × 0.25 Å cell, which shows how each such cell contributes to the ChEIPG fit when all $w_k = 1$. Given the cylindrical nature of the point density we expect radial anisotropy, which is readily apparent in Fig. 5.1(b), but what is perhaps less intuitive is the fact that there is also anisotropy orthogonal to the bonding plane. The latter arises from the presence of “seams” where the two atom-centered grids meet.

To ameliorate this anisotropy, we propose a simple weighting scheme in which $w_k = 1/n_k$, where n_k is the number of grid points contained within the cell where the point k resides. Figure 5.1(c) shows that this scheme significantly reduces the anisotropy of the grid.

5.3 Density derivative of the charges

As compared to the Mulliken or Löwdin charge scheme, the derivatives $\partial Q_\alpha/\partial P_{\mu\nu}$ are significantly more complicated in the case that Q_α is a ChEIPG atomic charge. To evaluate these derivatives in the ChEIPG case, note that e_B is the only quantity in Eq. (5.1) that is dependent on $P_{\mu\nu}$. Using the notation of Herbert *et al.*,²⁰

$$\begin{aligned} \frac{\partial Q_A}{\partial P_{\mu\nu}} = & - \sum_B^{N_{\text{MM}}} (\mathbf{G}^{-1})_{BA} \sum_k^{N_{\text{grid}}} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_B|} (\mathbf{I}_k)_{\mu\nu} \\ & + \gamma_A \sum_{BC}^{N_{\text{MM}}} (\mathbf{G}^{-1})_{CB} \sum_k^{N_{\text{grid}}} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} (\mathbf{I}_k)_{\mu\nu} \end{aligned} \quad (5.6)$$

where

$$\gamma_A = \frac{\sum_B^{N_{\text{MM}}} (\mathbf{G}^{-1})_{BA}}{\sum_{BC}^{N_{\text{MM}}} (\mathbf{G}^{-1})_{BC}}. \quad (5.7)$$

The quantity $(\mathbf{I}_k)_{\mu\nu}$ in Eq. (5.6) contains the integrals in this derivative; therefore, it is desirable to rearrange the derivative in such a way so that it needs to be calculated only once for each atom. Defining

$$\lambda_{Ak} = \sum_B^{N_{\text{MM}}} (\mathbf{G}^{-1})_{BA} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_B|}, \quad (5.8)$$

we obtain the following compact result:

$$\frac{\partial Q_A}{\partial P_{\mu\nu}} = \sum_B^{N_{\text{MM}}} (\boldsymbol{\Omega}_B)_{\mu\nu} (\gamma_A - \delta_{BA}). \quad (5.9)$$

The quantity

$$(\boldsymbol{\Omega}_B)_{\mu\nu} = \sum_k^{N_{\text{grid}}} \lambda_{Bk} (\mathbf{I}_k)_{\mu\nu} \quad (5.10)$$

consists of charge–density integrals in the AO basis, with “charges” λ_{Bk} located at points \mathbf{r}_k . Combining this with Eq. (3.12) and rearranging the order of summations, one obtains a correction to the Fock matrix in which $(\boldsymbol{\Omega}_B)_{\mu\nu}$ is evaluated just once for each B , and thus the integrals $(\mathbf{I}_k)_{\mu\nu}$ are calculated exactly once per QM atom.

5.4 Position derivatives of ChEIPG Charges

The ChEIPG charge derivatives are much more complicated. Instead of finding the M-derivative directly as we did for Mulliken charges, it is easier to find the full derivative of the ChEIPG charges, and then subtract out the density derivative term. The

ChEIPG charges can be expressed as²⁰

$$q_A = \sum_B e_B (\mathbf{G}^{-1})_{BA} - \left(\frac{\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \right) \sum_B (\mathbf{G}^{-1})_{BA} . \quad (5.11)$$

In this representation, uppercase Roman letters stand for QM atoms, the index k will stand for grid points, Q is the total charge on the system. Therefore, q_A in this section is equivalent to Q_α in other sections. The matrix \mathbf{G} is defined by

$$G_{CD} = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|} \quad (5.12)$$

where w_k is a weight for the k th grid point, and

$$e_C = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right) . \quad (5.13)$$

Here

$$(\mathbb{I}_k)_{\mu\nu} = \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle_{\mathbf{r}} \quad (5.14)$$

is a one-electron charge–density integral. (The subscript “ \mathbf{r} ” is intended to indicate that this is the integration variable.)

The ChEIPG charge derivatives can be expressed in a relatively compact form:

$$q_C^\nabla = \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla + \Xi (\mathbf{G}^{-1})_{MN}^\nabla \right] (\delta_{NC} - \gamma_C) . \quad (5.15)$$

Here we have switched to a superscript notation to make the equations easier to read; however, it is important to remember that every derivative is still taken with respect to a nuclear coordinate, even though this is omitted from the notation for brevity.

This equation introduces the notation

$$\Xi = \frac{Q - \sum_{KL} e_K (\mathbf{G}^{-1})_{KL}}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \quad (5.16)$$

and

$$\gamma_A = \frac{\sum_B (\mathbf{G}^{-1})_{AB}}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} . \quad (5.17)$$

Equation (5.15) is deceptively compact, because although both $(\mathbf{G}^{-1})_{MN}^\nabla$ and e_M^∇ can be expressed in analytic form, these derivatives are fairly involved. One complexity is the appearance of the grid points \mathbf{r}_k , which may be independent of x_γ in the case of a regular Cartesian grid (as in the original formulation of the ChEIPG algorithm⁴), or they might depend upon the nuclear positions (as in our Lebedev grid-based formulation³). Each of these cases will be examined in turn. Before doing so, note that

$$(\mathbf{G}^{-1})^\nabla = (\mathbf{G}^{-1})(\mathbf{G}^\nabla)(\mathbf{G}^{-1}) , \quad (5.18)$$

which follows by differentiating the condition $\mathbf{G}\mathbf{G}^{-1} = \mathbf{1}$. In view of this, the derivative $(\mathbf{G}^{-1})^\nabla$ in Eq. (5.15) is readily obtained once \mathbf{G}^∇ is known, so we will only discuss the latter.

5.4.1 Atom-independent grids

In this section, we limit our discussion to the case where the positions of the ChEIPG grid points are independent of the positions of the nuclei. (Most such implementations use a rectangular grid, but the gradient formalism presented below does not require this.) Some implementations use a uniform weighting scheme ($w_k \equiv 1$), in which case the derivatives of w_k vanishes. A non-uniform weighting function, intended to insure smoothness despite the use of a fixed Cartesian grid, was introduced in our previous implementation of ChEIPG charges,²⁰ and will be discussed below.

The derivative of the \mathbf{G} matrix with respect to a nuclear coordinate on atom A is

$$G_{CD}^{\nabla A} = \sum_k \frac{w_k^{\nabla A}}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|} + \sum_k (\mathbf{r}_k - \mathbf{r}_A) \left(\frac{w_k \delta_{CA}}{|\mathbf{r}_k - \mathbf{r}_A|^3 |\mathbf{r}_k - \mathbf{r}_D|} + \frac{w_k \delta_{DA}}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_A|^3} \right). \quad (5.19)$$

This part of the gradient has a simple analytic form that is easy to program.

The derivative of e_C is more complicated and we will break it into pieces for convenience, writing

$$e_C^{\nabla A} = R_C^{\nabla A} + S_C^{\nabla A} + T_C^{\nabla A} + U_C^{\nabla A} \quad (5.20)$$

where

$$R_C^{\nabla A} = \sum_k \left(\frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \right)^{\nabla A} \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right] \quad (5.21)$$

$$S_C^{\nabla A} = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} \right)^{\nabla A} \quad (5.22)$$

$$T_C^{\nabla A} = - \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu}^{\nabla A} P_{\mu\nu} \quad (5.23)$$

$$U_C^{\nabla A} = - \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu}^{\nabla A} \quad (5.24)$$

The $R_C^{\nabla A}$ and $S_C^{\nabla A}$ terms are straightforward to derive and inexpensive to evaluate:

$$R_C^{\nabla A} = \sum_k \left(\frac{w_k^{\nabla A}}{|\mathbf{r}_k - \mathbf{r}_C|} + \frac{w_k (\mathbf{r}_k - \mathbf{r}_A)}{|\mathbf{r}_k - \mathbf{r}_A|^3} \delta_{CA} \right) \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right] \quad (5.25)$$

$$S_C^{\nabla A} = \sum_k \frac{w_k Z_A (\mathbf{r}_k - \mathbf{r}_A)}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_A|^3}. \quad (5.26)$$

The $T_C^{\nabla A}$ term is the most difficult and time-consuming part of the ChEIPG M-derivative, as it involves derivatives of one-electron integrals:

$$T_C^{\nabla A} = - \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} P_{\mu\nu} \left(\left\langle \mu^{\nabla A} \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle_{\mathbf{r}} + \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu^{\nabla A} \right\rangle_{\mathbf{r}} \right) \quad (5.27)$$

There is no derivative over the distance because the grid point is independent of the atom position. Finally,

$$U_C^{\nabla A} = \sum_{\mu\nu} \frac{\partial Q_C}{\partial P_{\mu\nu}} P_{\mu\nu}^{\nabla A} . \quad (5.28)$$

If the term $U_C^{\nabla A}$ is excluded from the gradient, the charge derivative is identical to the M-derivative. With this in mind, define

$$\tilde{e}_C^{\nabla A} = e_C^{\nabla A} - U_C^{\nabla A} . \quad (5.29)$$

The M-derivative then assumes the form

$$\begin{aligned} \sum_M \frac{\partial Q_C}{\partial M} \frac{\partial M}{\partial x_A} &= q_C^{\nabla A} - \sum_{\mu\nu} \frac{\partial Q_C}{\partial P_{\mu\nu}} P_{\mu\nu}^{\nabla A} \\ &= \sum_{MN} \left[\tilde{e}_M^{\nabla A} (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^{\nabla A} + \Xi (\mathbf{G}^{-1})_{MN}^{\nabla A} \right] (\delta_{NC} - \gamma_C) \end{aligned} \quad (5.30)$$

Using Eqs. (5.19), (5.25), (5.26), (5.27), and (5.30), one can assemble the appropriate M-derivative to be used in Eq. (3.33).

5.4.2 Atom-dependent grids

Atom-dependent grids add a new level of complexity to the derivatives. We assume, as in our previous Lebedev grid-based implementation of ChElPG,³ that such grids consist of atom-centered radial shells. In such a case, the locations of the grid points can be expressed as

$$\mathbf{r}_k = \mathbf{r}_A + \mathbf{r}_{d,n} , \quad (5.31)$$

where the vector $\mathbf{r}_{d,n}$ depends upon the radial spacing (d) and the number of angular grid points (n). Thus, we assume in what follows that the grid point locations (\mathbf{r}_k)

depend linearly on the atomic positions (\mathbf{r}_A). Equations (5.19), (5.25), (5.26), and (5.27) derived above are thereby replaced by Eqs. (5.32)–(5.35) that are presented below, and used to assemble the derivatives in Eq. (5.30). Contained inside each of Eqs. (5.32)–(5.35) is an analogous equation, introduced above for the case of an atom-independent grid.

The new gradient equations are

$$\begin{aligned}
G_{CD}^{\nabla A} = & \sum_k \frac{w_k^{\nabla A}}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|} \\
& + \sum_{k \notin A} (\mathbf{r}_k - \mathbf{r}_A) \left(\frac{w_k \delta_{CA}}{|\mathbf{r}_k - \mathbf{r}_A|^3 |\mathbf{r}_k - \mathbf{r}_D|} + \frac{w_k \delta_{DA}}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_A|^3} \right) \\
& - \sum_{k \in A} \frac{w_k (1 - \delta_{CD})}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|} \left(\frac{(\mathbf{r}_k - \mathbf{r}_C) \delta_{CA}}{|\mathbf{r}_k - \mathbf{r}_C|^2} + \frac{(\mathbf{r}_k - \mathbf{r}_D) \delta_{DA}}{|\mathbf{r}_k - \mathbf{r}_D|^2} \right) \quad (5.32)
\end{aligned}$$

$$\begin{aligned}
R_C^{\nabla A} = & \sum_k \frac{w_k^{\nabla A}}{|\mathbf{r}_k - \mathbf{r}_C|} \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right) \\
& + \sum_{k \notin A} \frac{w_k (\mathbf{r}_k - \mathbf{r}_A) \delta_{CA}}{|\mathbf{r}_k - \mathbf{r}_A|^3} \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right) \\
& - \sum_{k \in A} \frac{w_k (\mathbf{r}_k - \mathbf{r}_C) (1 - \delta_{CA})}{|\mathbf{r}_k - \mathbf{r}_C|^3} \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right) \quad (5.33)
\end{aligned}$$

$$S_C^{\nabla A} = \sum_{k \notin A} \frac{w_k Z_A (\mathbf{r}_k - \mathbf{r}_A)}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_A|^3} - \sum_{k \in A} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_J \frac{Z_J (\mathbf{r}_k - \mathbf{r}_J)}{|\mathbf{r}_k - \mathbf{r}_J|^3} (1 - \delta_{JA}) \quad (5.34)$$

$$\begin{aligned}
T_C^{\nabla A} = & - \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} P_{\mu\nu} \left[\left\langle \mu^{\nabla A} \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle_{\mathbf{r}} + \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu^{\nabla A} \right\rangle_{\mathbf{r}} \right. \\
& \left. - \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|^3} \right| \nu \right\rangle_{\mathbf{r}} \right] \quad (5.35)
\end{aligned}$$

5.4.3 Weights of rectangular ChEIPG charges

In Q-Chem’s implementation of the ChEIPG charges,²⁰ the weights

$$w_k(\mathbf{r}_k, \{\mathbf{r}_A\}) = w_k^{\text{LR}}(\mathbf{r}_k, \{\mathbf{r}_A\}) \prod_J^{N_{\text{atoms}}} A_k^J(\mathbf{r}_k, \mathbf{r}_J) \quad (5.36)$$

associated with the grid points are chosen to consist of a long-range weighting function, w_k^{LR} , and a product of atomic switching functions A_k^J . The aim of this approach is to ensure that the ChEIPG charges are smooth, continuous functions of the nuclear coordinates, despite the use of a fixed Cartesian grid to evaluate the electrostatic potential. Gradients of Eq. (5.36) are straightforward:

$$\hat{\nabla}_L w_k = (\hat{\nabla}_L w_k^{\text{LR}}) \prod_J^{N_{\text{atoms}}} A_k^J + w_k^{\text{LR}} (\hat{\nabla}_L A_k^L) \prod_{J \neq L}^{N_{\text{atoms}}} A_k^J. \quad (5.37)$$

The gradient $\hat{\nabla}_L w_k$ goes to zero if grid point k is not in the region where the switching functions are significant.

CHAPTER 6

Results

The Lebedev ChEIPG method and the QM/MM-Ewald method described above have been implemented in a locally-modified version of Q-CHEM v. 4.0.^{23,24} Here, we describe various numerical tests designed to evaluate the numerical performance of the method.

6.1 Charge schemes

In attempting to implement the algorithm in Ref. 2, we encountered serious SCF convergence problems that we suspected were due to the use of extended basis sets in conjunction with Mulliken image charges. Since no such difficulties have been reported in previous minimal-basis implementations of the algorithm,^{2,25} we first wanted to verify that the Mulliken version [with charge derivatives given in Eq. (3.14)] does indeed work in a minimal basis set. To test this, calculations were run at the QM = Hartree-Fock (HF)/STO-3G level of theory, for a single QM water molecule in a box of 215 TIP3P water molecules,²⁶ with $L = 18.643$ Å corresponding to ambient liquid density. This calculation converged rapidly using Mulliken charges, and in comparison to the corresponding calculation using a Cartesian ChEIPG grid (head

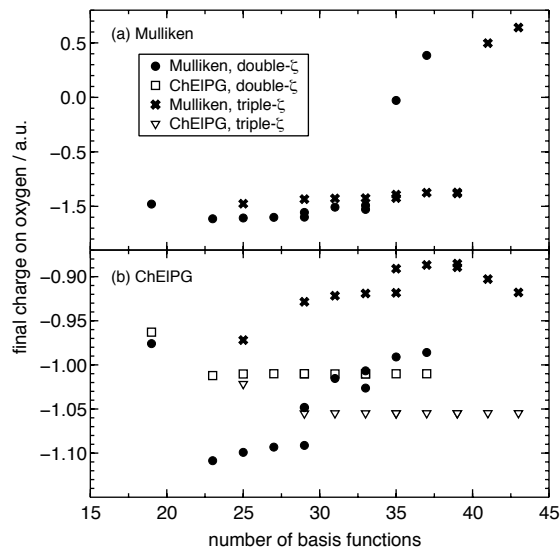


Figure 6.1: Converged partial charges for the oxygen atom of a single QM water molecule in an MM water box, plotted against the number of basis functions used to describe the QM region, for $6-31(x+,y+)G^*$ and $6-311(x+,y+)G^*$ basis sets. In (a), the QM/MM-Ewald method uses Mulliken image charges whereas in (b) it uses ChEIPG image charges. In the latter case, Mulliken charges were also computed upon SCF convergence.

spacing of 5 \AA and $\Delta x = 1 \text{ \AA}$), essentially the same energy is obtained. Moreover, if we compute Mulliken charges using the density matrix obtained from the ChEIPG Ewald calculation, we obtain values within 0.003 a.u. of the Mulliken charges obtained from the Mulliken Ewald calculation. This confirms that various charge schemes work equally well in minimal basis sets.

In order to test extended basis sets, the same calculation was performed using the $6-31(x+,y+)G^*$ and $6-311(x+,y+)G^*$ basis sets, where x and y range from 0 to 3, except that the $6-31(3+,+)G^*$, $6-31(3+,2+)G^*$, and $6-311(3+,+)G^*$ basis sets were

excluded because in these cases the Mulliken-based Ewald procedure fails to converge after 50 SCF cycles. Figure 6.1 shows the final, converged point charge on the oxygen atom of the QM water molecule as a function of basis set size. We converge the SCF Ewald calculation using either Mulliken image charges [Fig. 6.1(a)] or else ChEIPG image charges [Fig. 6.1(b)], and in the latter case we also compute Mulliken charges using the final, converged SCF density matrix. From panel (a) we see that the use of Mulliken image charges—when the calculation can be converged—often leads to a positive partial charge on the oxygen atom in larger basis sets. Not only is this behavior not observed with ChEIPG charges, but if we use ChEIPG image charges to converge the SCF calculation (*i.e.*, the ChEIPG charges are used to construct the Fock matrix correction $\Delta F_{\mu\nu}^{\text{PI}}$), then the Mulliken charges obtained upon convergence are reasonable [see Fig. 6.1(b)]. This suggests that the problem lies with instabilities in the Mulliken charge derivatives as the basis set is expanded, which are exacerbated when these charges are included as part of the self-consistent iteration procedure.

These instabilities are borne out by the SCF energies, plotted as a function of basis size in Fig. 6.2. When Mulliken image charges are employed, the correct SCF energy of ≈ -76 hartree is obtained only in small basis sets; in larger basis sets, the “converged” SCF energy differs from this value by as much as 1.5 hartree. For ChEIPG image charges, the SCF energy is stable with respect to basis-set expansion.

Clearly, Mulliken charges cannot be used for QM/MM-Ewald calculations in non-minimal basis sets. The remainder of this work explores the use of ChEIPG images charges. In that case, one must determine electrostatic grid parameters to ensure that

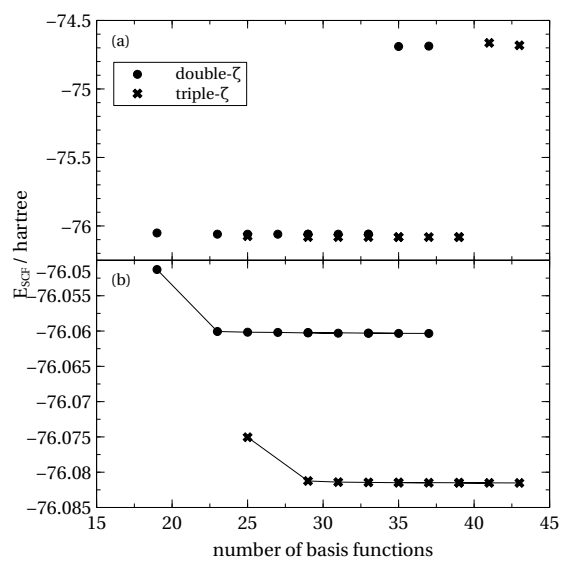


Figure 6.2: Total SCF energy versus the number of basis functions, for a single QM water molecule in a box of MM water molecules, using either (a) Mulliken image charges or (b) ChEIPG image charges.

the charges are converged. Tests of how the ChEIPG charges converge with respect to grid parameters are presented in the next section.

6.2 Lebedev ChEIPG charges

6.2.1 Gas phase

Our Lebedev grid-based implementation of the ChEIPG algorithm is new, and here we seek to test it against the original Cartesian grid-based version of Breneman and Wiberg.⁴ The Lebedev version is inherently much more efficient, as it uses far fewer grid points for the same head space and grid spacing, so we seek to understand how sparse we can make the Lebedev grid without adversely affecting the charges that are obtained. In these tests, we leave the head space set to 2.8 Å (the value recommended by Breneman and Wiberg⁴), and use Bondi radii²⁷ to define the vdW surface. Atom-centered radial Lebedev shells with $N_p = 590$ points per shell extend from the atomic Bondi radius out to 2.8 Å away from that surface, in radial increments of Δx . (The value of N_p has previously been shown to provide good rotational invariance, in the context of polarizable continuum model calculations where the vdW cavity is discretized with atom-centered Lebedev grids.²⁸) Choosing bins of volume $(\Delta x)^3$, this leaves only Δx as a parameter to test convergence of the ChEIPG charges.

We first aim to determine whether the Lebedev ChEIPG charges provided are a reasonable representation of the electrostatic potential. To that end, we first examine the convergence behavior of the Cartesian ChEIPG charges, in order to establish a

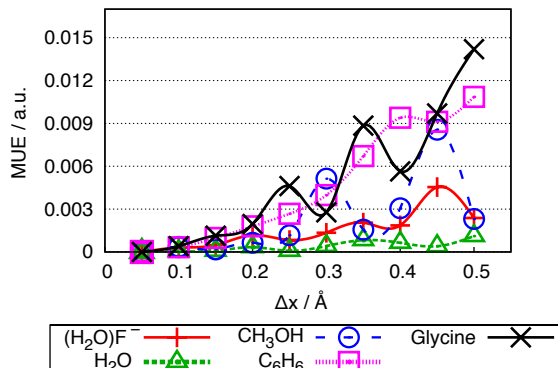


Figure 6.3: Convergence of the Cartesian ChEIPG charges (atomic units) computed at the HF/aug-cc-pVDZ level, as a function of the Cartesian grid spacing Δx . The reference values were computed using $\Delta x = 0.05$ Å. The vertical axis plots the mean unsigned error (MUE) for all of the charges in the molecule.

baseline. Note that the Cartesian ChEIPG charges provide the best possible representation of the electrostatic potential, in a least-squares sense, in the limit that $\Delta x \rightarrow 0$, and we will take Cartesian ChEIPG charges computed using $\Delta x = 0.05$ Å to be the “true” ChEIPG charges. (This choice is justified by the fact that the charges change by only $\sim 10^{-3}$ a.u. when Δx is increased to 0.10 Å.) Convergence of the Cartesian ChEIPG charges towards these “true” values, as a function of Δx , is plotted in Fig. 6.3 for several small molecules. Even for $\Delta x = 0.5$ Å, the charges are already converged to about two decimal places.

The convergence of the weighted and unweighted Lebedev ChEIPG charges, as a function of Δx , is shown in Fig. 6.4. Use of the weighting scheme tends to afford better agreement with the Cartesian ChEIPG charges, suggesting that an approximately isotropic grid is indeed important for reproducing Cartesian ChEIPG charges.

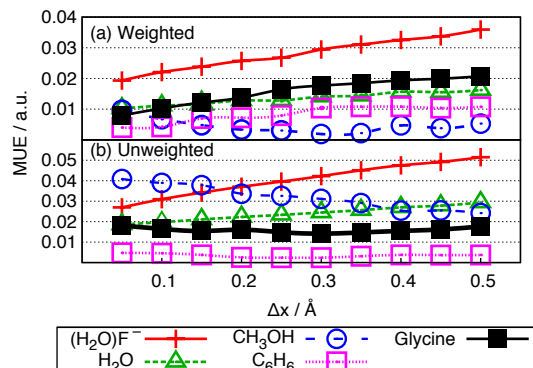


Figure 6.4: Convergence of the Lebedev ChEIPG charges (atomic units) computed at the HF/aug-cc-pVDZ level, as a function of the grid spacing Δx . The reference values were computed using a Cartesian grid with $\Delta x = 0.05$ Å. The vertical axis plots the mean unsigned error (MUE) for all of the charges in the molecule. In (a), the weighting scheme discussed in Section 5.2 is employed ($w_k = 1/n_k$), whereas in (b) the weights are all equal.

Interestingly, the slope of errors with respect to Δx is about the same regardless of whether the weighting scheme is used or not. We take this to mean that the charges converge at about the same rate with respect to Δx , but converge to different values depending on whether the weighting is used. The data in Fig. 6.4 suggest that it is reasonable to expect errors of the same order of magnitude, or maybe only slightly larger, as those seen for Cartesian ChEIPG charges when using Lebedev ChEIPG charges.

Given that we can converge the Lebedev ChEIPG charges to about the same values as their Cartesian counterparts, we now turn our attention to the rotational invariance of the ChEIPG charges. To this end, we have computed the HF/6-31G* ChEIPG charges of a glycine molecule in the standard nuclear orientation (principle

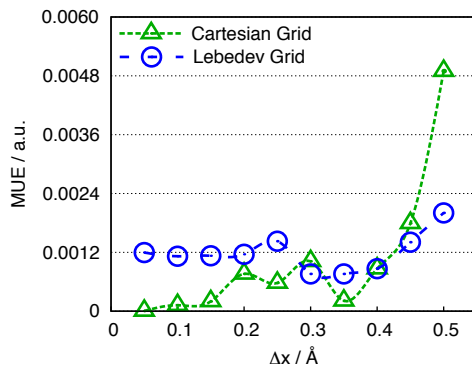


Figure 6.5: Rotational invariance (a.u.) of the Lebedev and Cartesian ChEIPG charges on the hydroxyl oxygen of glycine as the molecule is rotated by 180° about an axis. Charges are computed at the HF/6-31G* level of theory.

axes of nuclear charge) and also after rotation around the x axis from 0 to π radians, in increments of $\pi/12$ radians. We examine the ChEIPG charge on the hydroxyl oxygen atom as a function of this rotation angle, as compared to the value obtained at 0° . The mean unsigned error (MUE), which is a measure of rotational invariance, is plotted in Fig. 6.5 as a function of Δx .

For comparison, the actual ChEIPG charge on the hydroxyl oxygen ranges from -0.78 to -0.80 a.u., whereas the data in Fig. 6.5 suggest that both the Cartesian and Lebedev grids afford charges that are rotationally invariant to within 0.01 a.u. or better, even for $\Delta x = 0.5 \text{ \AA}$. For grid spacings $\Delta x < 0.15 \text{ \AA}$, no further improvement in the rotational invariance is observed. One interesting point is that the Lebedev grid exhibits better rotational invariance when Δx is large, but (slightly) worse invariance when Δx is small. This is seemingly at odds with Spackman’s results for spherical and icosahedral grids.²² However, Spackman took care to only compare grids of

relatively the same density.

In this particular glycine example, the Cartesian grid has sides of $\approx 21 \text{ \AA}$, meaning the Cartesian grids occupy a volume of $\approx 9300 \text{ \AA}^3$ and the Lebedev grid (assuming it is constructed from spheres) has a volume of $\approx 4800 \text{ \AA}^3$. At $\Delta x = 0.5 \text{ \AA}$, the Cartesian grid contains about 15,800 points and the Lebedev grid about 9,000 grid points, or point densities of 1.7 \AA^{-3} and 1.9 \AA^{-3} , respectively. On the other hand, at $\Delta x = 0.05 \text{ \AA}$ the Cartesian grid has about 1.5×10^7 points for a point density of about $1,600 \text{ \AA}^{-3}$, while the Lebedev grid has 84,000 points for a density of 17.5 \AA^{-3} . In other words, the Cartesian grid is about 100 times more dense than the Lebedev grid for $\Delta x = 0.05 \text{ \AA}$. We conclude that for Lebedev and Cartesian grids of similar densities (*e.g.*, the $\Delta x = 0.5 \text{ \AA}$ case), the Lebedev grids exhibit better rotational invariance. For cases where the point densities are very different (*e.g.*, the $\Delta x = 0.5 \text{ \AA}$ case), the more dense grid exhibits the better rotational invariance. This is consistent with Spackman's results.²²

6.2.2 Condensed phase

The results above show that use of weighted Lebedev grids affords ChEIPG charges that are nearly identical to those obtained using Cartesian grids, but can do so with far fewer grid points. However, the benchmarks above use $N_p = 590$ points per radial shell, which will be expensive in calculations with larger QM regions. Reducing this number to $N_p = 50$, we have tested the rotational invariance of the ChEIPG charges in the context of a QM/MM-Ewald calculation, taking as a test system a QM region composed of five water molecules (B3LYP/6-31+G* level) in a periodic cell containing

211 MM water molecules ($L = 18.643 \text{ \AA}$). The ChElPG charges were computed using a head space of 3.0 \AA with $\Delta x = 0.5 \text{ \AA}$. We carried out single-point energy calculations after rotating the entire simulation cell in increments of 90° , leaving fixed the axis system that defines the ChElPG unit spheres. (That is, the axes of the simulation cell are rotated with respect to the axes that define the grid.)

	0°	90°	180°	270°	$ \Delta_{\max-\min} $
$E_{\text{SCF}}/\text{a.u.}$	-382.2489807663	-382.2489808032	-382.2489806988	-382.2489807549	0.000000
$\epsilon_{\text{HOMO}}/\text{a.u.}$	-0.300498	-0.300498	-0.300498	-0.300498	0.000000
$Q_{\text{oxy}}/\text{a.u.}$	-1.026878	-1.030694	-1.035725	-1.027292	0.008847
$Q_{\text{hyd1}}/\text{a.u.}$	0.465895	0.468642	0.467734	0.466975	0.001839
$Q_{\text{hyd2}}/\text{a.u.}$	0.559032	0.560517	0.564779	0.557386	0.007393

Table 6.1: SCF energies and Lebedev ChEIPG charges as the simulation cell (containing 5 QM and 211 MM water molecules) is rotated with respect to the axes that define the Lebedev unit spheres. The final column is the difference between the maximum and minimum values for the various quantities in each row. The SCF convergence threshold was set to 10^{-5} hartree.

54

	force / a.u.				$ \Delta_{\max-\min} $
	0°	90°	180°	270°	
$ \mathbf{F}_{\text{total}} $	0.0318286779	0.0318296463	0.0318290238	0.0318289497	0.0000010
$ \mathbf{F}_{\text{QM/MM}} $	0.0318488287	0.0318484627	0.0318490519	0.0318486728	0.0000006
$ \mathbf{F}_{\text{Ewald}} $	0.0000201508	0.0000188163	0.0000200282	0.0000197231	0.0000013

Table 6.2: Magnitude of the force on a single oxygen atom as the simulation cell (containing 5 QM and 211 MM water molecules) is rotated with respect to the axes that define the Lebedev unit spheres. The final column is the difference between the maximum and minimum values for the various quantities in each row. The SCF convergence threshold was set to 10^{-7} hartree and the integral threshold to 10^{-10} , with the corresponding value of C determined from Eq. (4.1).

The results, which are summarized in Tables 6.1 and 6.2, use a sparser grid than was used for the gas-phase calculations, yet good rotational invariance of both energies (Table 6.1) and forces (Table 6.2) is observed. The variation in the SCF energy as a function of rotation angle is smaller than the SCF convergence threshold (10^{-5} hartree, for the calculations in Table 6.1). The convergence threshold was tightened to 10^{-7} hartree for the gradient calculations (Table 6.2), yet the variation in different components of the force is no larger than 1.3×10^{-7} a.u. These results suggest rotational invariance can be achieved in condensed-phase systems using grids that are far sparser than those used in the gas-phase calculations presented above.

6.3 Timings

One drawback to the use of ChEIPG charges is the expense associated with computing the charge derivatives $\partial Q_\alpha / \partial P_{\mu\nu}$, especially the tensor $\mathbf{\Omega}_B$ in Eq. (5.10). The choice of the Ewald parameter η can also make a large difference in calculation time, as it controls the number of vectors used in the real- and reciprocal-space sums. A poor choice for η can double the calculation time, in our experience.

To understand how the Ewald parameter affects the calculation time, two systems were analyzed. The first is intended to be indicative of a fairly small QM region, consisting of 11 QM water molecules (B3LYP/6-31+G* level of theory, for a total of 253 basis functions) in a $L = 18.643$ Å simulation cell containing 205 TIP3P water molecules.²⁶ The second calculation is much larger, and consists of a QM region containing an aqueous cytidine molecule and all 27 water molecules that reside within

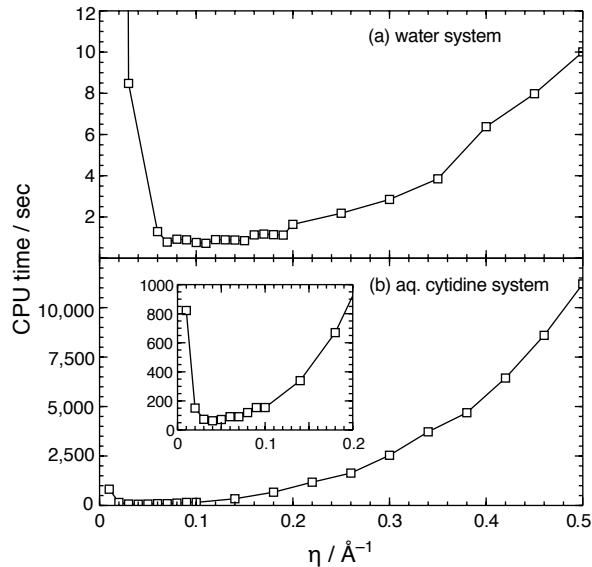


Figure 6.6: The total CPU time required to calculate $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ [Eq. (3.9)] for (a) 11 QM water molecules in a large MM water box, and (b) 27 water molecules plus cytidine in the QM region, surrounded by a larger box of MM water molecules.

6 \AA of the cytidine molecule. The QM region is described at the B3LYP/6-31+G* level (970 basis functions) and placed in a $L = 50.0$ \AA simulation cell containing 4,122 TIP3P water molecules. In both cases, the ChEIPG parameters are set to 3.0 \AA for the head space, $N_p = 50$, and $\Delta x = 0.5$ \AA , since these values afford good rotational invariance for the test case in Section 6.2.2. We compute the SCF energy of both systems as a function of η . For the first system, all values of η afford the same energy to within 10^{-10} hartree, while for the larger system the variation is no greater than 2×10^{-7} hartree. This implies that we have indeed converged both the real- and the reciprocal-space sums for each value of η , which should be the case if one follows the recommendations in Section 4.

Figure 6.6 shows the CPU time required to compute the $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ [Eq. (3.9)], as a function of η , for these two test systems. Note that in a single-point calculation, this is essentially a one-time cost insofar as the main cost is in calculating the Ewald potential, which is done outside of the SCF iterations. Unless the Ewald parameter is chosen poorly, the cost of this step is small in comparison to the time required to compute the ChELPG charge derivatives, $\partial Q_\alpha/\partial P_{\mu\nu}$. This is demonstrated in Fig. 6.7, where we compare (as a function of η) the fraction of the total job time that is consumed in computing derivatives $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ versus the fraction required to compute the derivatives $\partial Q_\alpha/\partial P_{\mu\nu}$. Note that all of these calculations are exact (within the integral drop tolerance), insofar as we use the criteria given in Section 4 to decide how many vectors are necessary to converge the real- and reciprocal-space sums.

As can be seen in Figs. 6.6 and 6.7, a poor choice for η can make a large difference in the calculation time. This issue is less important in smaller systems where the time to compute the Ewald potential is small; however, in a large system (such as cytidine in 27 QM water molecules), this step can become the bottleneck if η is chosen too large. This point has not been emphasized previously in the context of Ewald summation for QM/MM calculations.

Most classical implementations of Ewald summation are based on the particle-mesh Ewald method,^{18,19} in which the reciprocal-space summation is faster (scaling as $N_{\text{vec}} \log N_{\text{vec}}$ with respect to the number of reciprocal lattice vectors, N_{vec}) than the real-space summation (which scales as N_{MM}^2 with respect to the number of point

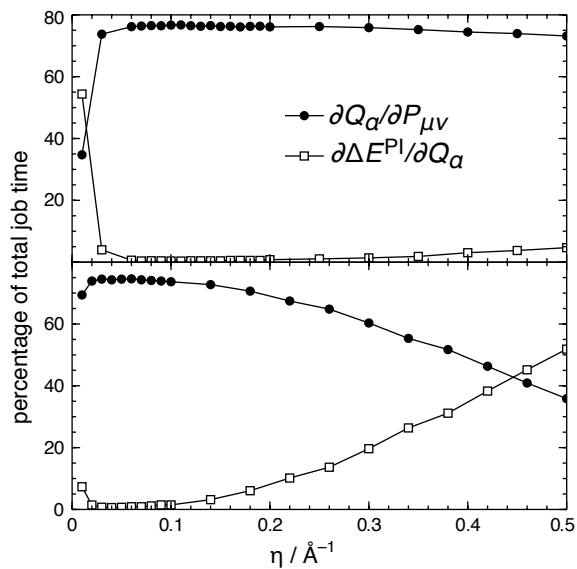


Figure 6.7: Percentage of the total single-point energy calculation time that is spent in calculating derivatives $\partial \Delta E^{\text{PI}} / \partial Q_\alpha$ and $\partial Q_\alpha / \partial P_{\mu\nu}$, for a water test system (small QM region) (top graph) and an aqueous cytidine test system (large QM region) (bottom graph). The time required to calculate $\partial Q_\alpha / \partial P_{\mu\nu}$ is independent of η but becomes a smaller percentage as the time to compute $\partial \Delta E^{\text{PI}} / \partial Q_\alpha$ increases.

	Lebedev		Cartesian	
One-electron integrals	6.4		6.4	
Ewald potential ($\partial\Delta E^{\text{PI}}/\partial Q_\alpha$)	31.5		31.4	
Two-electron integrals	17.6	12.2	17.7	12.4
ChEIPG charges	0.4	0.4	3.9	3.8
$\partial Q_\alpha/\partial P_{\mu\nu}$	16.2	16.1	176.4	176.2
Other	6.1	4.8	6.1	4.7
Total Fock build	71.7	33.4	235.4	197.1
Total SCF	348.6		1658.9	

Table 6.3: Timings (in seconds) for a single-point QM/MM calculation of cytidine (QM) in water (MM), with periodic boundary conditions using $\eta = 0.04 \text{ \AA}^{-1}$. The QM region consists of 30 atoms and 349 basis functions (B3LYP/6-31+G*) and the MM region consists of 12,840 point charges (4,160 TIP3P water molecules). The Lebedev and Cartesian ChEIPG grids consist of 885 and 32,598 points, respectively. For the steps that must be repeated at each SCF calculation, two columns of timing data are provided, corresponding to the second (left column) and ninth (right column) SCF cycles. The row labeled “other” includes the XC quadrature step and any remaining overhead associated with the Fock build.

charges). For this reason, a larger Ewald parameter is generally selected, in order to perform more of the summation in reciprocal space, which may not be an effective strategy for the present implementation, where the cost of the reciprocal-space sum scales as N_{vec}^2 . Although a particle-mesh implementation of QM/MM-Ewald may be interesting to consider (especially in view of the recent quantum Ewald mesh for evaluation of electron repulsion integrals²⁹), at present the $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ term is often not the bottleneck of the calculation, as can be seen in Fig. 6.7. As such, there seems to be little need to accelerate this part of the calculation at present.

It is beneficial to analyze the complete timings of the QM/MM-Ewald calculations,

and we will present timings for a variety of aqueous cytidine calculations performed at the B3LYP/6-31+G* level, using ChEIPG charges with a head space of 3.0 Å, $N_p = 26$, and $\Delta x = 0.5$ Å. (Although the convergence tests of the ChEIPG charges reported in Section 6.2.1 used a much larger number of Lebedev grid points, numerical tests of QM/MM-Ewald calculations, comparing $N_p = 26$ to $N_p = 590$, demonstrate that the converged SCF energies differ by less than the convergence threshold of 10^{-5} hartree.) Table 6.3 compares timing data for Lebedev and Cartesian ChEIPG grids in the QM/MM-Ewald procedure, for a calculation in which only the cytidine molecule is treated at a QM level. The use of Lebedev grids reduces the number of grid points from 32,598 points to just 885 points, which substantially reduces the cost of computing the charge derivatives $\partial Q_\alpha / \partial P_{\mu\nu}$. At the same time, the difference between the SCF energies in the two calculations is only 6.663×10^{-6} hartree, which is smaller than the SCF convergence threshold of 10^{-5} hartree, so there is every reason to prefer the Lebedev-based approach.

	$R = 6 \text{ \AA}$		$R = 7 \text{ \AA}$		$R = 8 \text{ \AA}$		$R = 9 \text{ \AA}$	
No. QM atoms	108		174		249		345	
No. MM atoms	12,402		12,336		12,231		12,165	
No. basis functions	947		1,453		2,028		2,764	
No. ChEIPG grid points	2,700		4,194		5,683		7,642	
One-electron integrals	27		44		67		93	
Ewald potential ($\partial\Delta E^{\text{PI}}/\partial Q_\alpha$)	104		166		236		324	
Two-electron integrals	291	122	943	317	2,169	735	4,546	1,288
ChEIPG charges	4	4	11	11	24	24	48	47
$\partial Q_\alpha/\partial P_{\mu\nu}$	762	763	3,476	3,477	10,761	10,767	29,930	29,925
Other	45	29	85	63	137	102	207	153
Total Fock build	1,206	917	4,681	3,868	13,327	11,629	35,054	31,417
Total SCF	8,363		29,795		86,921		231,248	

Table 6.4: Timing data (rounded to the nearest second) for QM/MM calculations of aqueous cytidine in which the QM region consists of a region of specified radius, R , around the cytidine molecule, described at the B3LYP/6-31+G* level. All calculations were performed with periodic boundary conditions using $\eta = 0.04 \text{ \AA}^{-1}$ and ChEIPG charges. For the steps that must be repeated at each SCF calculation, two columns of timing data are provided, corresponding to the second (left column) and ninth (right column) SCF cycles. Timings labeled “other” includes the XC quadrature step and any remaining overhead associated with the Fock build. All calculations were run in serial on a single Intel Xeon x5650 processor with 48 GB RAM with no competing processes on the node.

Timing data are provided in Table 6.4 for a sequence of related calculations in which the QM region consists of the cytidine molecule plus all water molecules containing an atom within some specified distance, R , of the glycosidic nitrogen. (All calculations contained 4,160 QM + MM water molecules in the simulation cell.) These calculations were performed at the B3LYP/6-31+G* level with $\eta = 0.04 \text{ \AA}^{-1}$. ChEIPG grid parameters are the same as those for the cytidine-only QM region discussed above.

The data in Table 6.4 reveal that the time needed to calculate the one-electron integrals is almost negligible. (Note that the one-electron integral timings quoted the table includes only the QM-MM interactions within the simulation cell. The time required to calculate the QM-MM image interactions is included in the $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ term.) Given the data in Table 6.4, it seems that there is little motivation at this point to work on accelerating the one-electron integral evaluation, *e.g.*, using asymptotic expansions.³⁰

The data also reveal that the cost of computing $\partial\Delta E^{\text{PI}}/\partial Q_\alpha$ is an order-of-magnitude less than the cost of computing electron repulsion integrals, except for the smallest QM regions. As such, the particle-mesh Ewald technique,¹⁸ which is generally regarded as the method of choice for implementing PBC in classical simulations, does *not* appear to be a promising way forward in the present context, since the most expensive step in our QM/MM-Ewald algorithm (by a very wide margin, especially for large QM regions) is calculation of the charge derivatives $\partial Q_\alpha/\partial P_{\mu\nu}$. In particular, the matrix $\mathbf{\Omega}_B$ in Eq. (5.10) must be computed N_{QM} times (once for each

QM atom, B) in order to calculate the charge derivatives $\partial Q_\alpha/\partial P_{\mu\nu}$. Each of the $\mathbf{\Omega}_B$ matrices is independent of one another so this step can be trivially parallelized across N_{MM} processors, and further parallelism will be as good as the parallelization of the one-electron integrals $(\mathbf{I}_k)_{\mu\nu}$ [Eq. (5.4)]. Even a factor of two reduction in the time to calculate $\partial Q_\alpha/\partial P_{\mu\nu}$ would reduce the total SCF time for the calculations in Table 6.4 by a minimum of 45%, and by 70% in the case of the smallest ($R = 6 \text{ \AA}$) QM region.

6.4 A new digestion routine

Upon examination of the results from the energy calculations and the derivative calculations, the time required by the $\partial Q_\alpha/\partial P_{\mu\nu}$ and $\partial Q_\alpha/\partial x_\gamma$ were deemed unacceptable. Reason tells us that two electron integrals should dominate the calculation except in specific cases. The number of two electrons integrals scales as $\mathcal{O}(N_{\text{basis}}^4)$, nominally. In practice, a screening procedure is used to determine pairs of basis functions that are significant. These are known as functions pairs. In reality, the number of two electron integrals scales as $\mathcal{O}(N_{\text{fp}}^2)$, where N_{fp} is the number of function pairs.

In the case of ChELPG density derivatives, there are $N_{\text{fp}}N_{\text{grid}}N_{\text{atoms}}$ integrals that need to be computed. The N_{atoms} dependence arises because of the weights on the charge-density integrals and the sum over the number of grid points, see eqns. 5.9 and 5.10. Normally when integrals are computed they are combined with density matrix using a routine known as a digestion routine. The digestion routine performs a sum over the index of the coordinates. This sum keeps the arrays in memory to a small

size, namely N_{fp} . In the ChElPG charges there is an outer sum over the number of atoms causes us to recompute $N_{fp}N_{grid}$ integrals N_{atoms} times if the digestion routine is to be used. If we do not use the conventional digestion routine, a new routine can be written that will perform $N_{fp}N_{grid}$ integrals and store them in core memory. These integrals can then be combined with the $N_{atoms}N_{grid}$ weights. By doing this new routine, the number of integrals computed can be reduced by a factor of N_{atoms} ; therefore, we expect a roughly N_{atoms} speedup in the computation of this term.

In a 15.0 Å cubic box containing one cytidine molecule (30 atoms) as the QM region at the HF/6-31G level ($N_{basis} = 179$ and $N_{fp} = 13329$) and 110 TIP3P MM waters using a rectangular grid from ChElPG charges with $\Delta x = 0.75\text{Å}$ and the head space being 2.8 Å, there is a significant speedup in the amount of time spent in the $\partial Q_\alpha/\partial P_{\mu\nu}$ and $\partial Q_\alpha/\partial x_\gamma$ terms. The results can be seen in table 6.5. The classic routine consists of doing the digestion routine with the sum over grid points done internally. The new routine consists of doing the digestion using the new scheme described earlier in this section. It is obvious that the charge derivatives are the bottlenecks for both the energy and the gradient calculations. The new routine no longer makes the ChElPG derivatives the bottleneck on one processor. The speed up between the new routine and the classic routine is a factor of $N_{atoms}/2$ for this calculation. While this is not as high as the value predicted, it is expected that the speed up will approach N_{atoms} when the system becomes large because there is a larger difference between the time to compute the number of integrals and the linear algebra routine needed to combine the integral values with the weights.

On one processor the computation of the two-electron integrals is the bottleneck process. The reader will note that the computation of ChEIPG charges are also a large portion of the calculation. There are three terms that are unique to a QM/MM Ewald calculation: ChEIPG charges, Ewald potential, and $\partial Q_\alpha/\partial P_{\mu\nu}$. The computation of the ChEIPG charges is the dominant term because the charges must be recomputed at every SCF cycle in order to respond to the change in the density matrix. The Ewald potential and $\partial Q_\alpha/\partial P_{\mu\nu}$ are both only dependent on the position of the nuclei and not on the density matrix, so these terms only have to be calculated once per geometry.

Table 6.5 also showcases the speed up of the calculations across multiple processors using OpenMP. The best speed up is shown in the two-electron parts of the calculation. The most significant speed up for all terms is between 1 and 5 processors. Very little speed up is obtained going from 5 processors to 20 processors. This tells us that this job is not appropriate to look at parallel speed up because the overhead is a significant portion of the calculation.

6.5 Gradient results

If the gradient for a calculation is correct, an NVE AIMD simulation should show a conservation of energy. Such behavior is shown in figure 6.8. This figure hides the fact that there is an upward trend in the total energy. This is believed to be caused by a flaw in the algorithm when the AIMD calculation is restarted. When running an AIMD simulation over long time periods, the total calculation for all time steps must

No. of Processors	1	5	10	15	20
One-electron integrals	0.2	0.1	0.1	0.1	0.1
Two-electron integrals	19.9	4.5	2.6	2.2	2.4
ChEIPG charges	17.4	5.3	5.0	4.8	5.0
Ewald potential	1.4	0.6	0.6	0.6	0.6
$\partial Q_\alpha/\partial P_{\mu\nu}$ (New routine)	2.2	2.1	2.6	3.6	4.4
$\partial Q_\alpha/\partial P_{\mu\nu}$ (Classic routine)	33.4	8.9	8.4	7.9	8.0
Derivative of one-electron integrals	0.5	0.1	0.1	0.1	0.1
Derivative of two-electron integrals	7.5	1.5	1.5	1.4	1.4
Derivative of Ewald potential	0.4	0.6	1.0	1.5	8.0
Derivative of ChEIPG charges (New routine)	7.3	2.3	2.7	3.3	3.8
Derivative of ChEIPG charges (Classic routine)	127.7	34.0	37.6	40.2	40.7

Table 6.5: Timing data (rounded to the nearest tenth of a second) for QM/MM calculations of aqueous cytidine in which the QM region consists of a region of the cytidine molecule, described at the HF/6-31G level. All calculations were performed with periodic boundary conditions using $\eta = 0.12 \text{ \AA}^{-1}$ and ChEIPG charges. All calculations were run in serial on a single Intel Xeon x5650 processor with 48 GB RAM with no competing processes on the node.

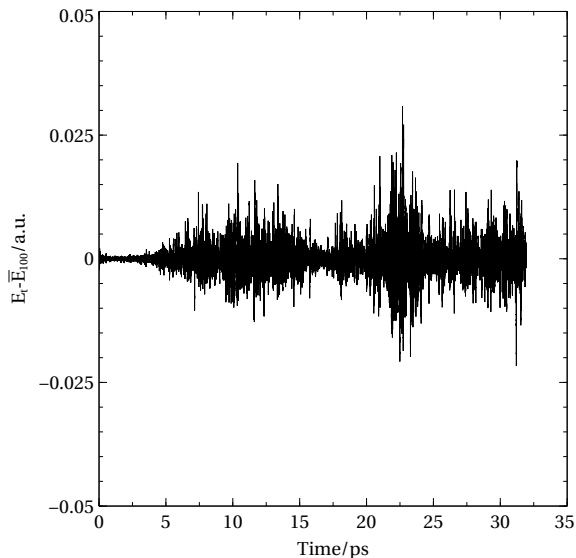


Figure 6.8: Conservation of energy for QM/MM calculations of aqueous cytidine in which the QM region consists of a region of the cytidine molecule, described at the HF/6-31G level. The y-axis is the energy of the time step minus the average energy over the last 100 time steps.

be split into a series of calculations over some of the time steps. After one calculation with a fraction of the total time steps is completed a new calculation is started from the scratch files (coordinates and velocities) of the previous AIMD calculation in order to calculate time steps.

The analytical gradients agree with a three point finite difference routine with a step size of 10^{-3} Å to 10^{-6} for QM gradients and 10^{-8} for MM gradients. Thus, it is expected that force calculations are correct and that the problem solely comes from the AIMD portion of the calculation. There are two routines that are important only in dynamics calculations. The first is a routine that enforces periodic boundary

conditions. When a periodic boundary condition is employed, it is expected that there are no molecules outside of the simulation cell. In a dynamics simulation it is not unreasonable for a force to be applied to a molecule that causes the molecule to be outside the edge of the simulation cell. In order to keep the simulation cell “intacted”, the molecule must be translated by the cell length (in the appropriate direction) so that it “enter” from the opposite edge of the cell. One complexity is that the connectivity information must be known and stored so that all atoms in a molecule are translated together. Currently this is done based on the position of the center of mass of the molecule.

The other routine that must be done is to center the simulation cell. In order to tell if a molecule has strayed past the edge of the simulation cell, it is necessary to know where the center of the simulation cell is. Currently, this is taken care of at the beginning of the simulation by measuring the minimum and maximum atom coordinates and subtracting to find the center of the simulation cell as input by the user. All atoms are then translated so that the center of the cell occurs at the origin. While this routine makes sense for a large group of small molecules, such as a water box, it may not prove to be the best method for a large solute molecule in a solvent. In this case it may be better to recenter the box at every time step where the center of the simulation cell is the center of mass of the solute molecule. The reason for this is because one can imagine a large solute molecule, which still has its center of mass inside the simulation cell, but where a significant portion of the molecule is outside of the edge of the cell. In this case, the simulation cell may not provide a

good representation of the solvation of the solute molecule.

CHAPTER 7

Conclusion

Although the theory of Ewald summation for QM/MM calculations has been described before,^{2,25} in the context of semi-empirical QM methods, we have provided a robust and general way to extend this technique to extended basis sets, where earlier implementations based on Mulliken image charges for the QM electron density experience stability problems. These are alleviated by using ChEIPG image charges instead, and the relatively high cost of computing such charges is mitigated somewhat by an implementation of the ChEIPG algorithm using atom-centered Lebedev grids for evaluation of the QM electrostatic potential. These Lebedev ChEIPG charges exhibit good rotational invariance and reproduce the QM/MM-Ewald results using traditional ChEIPG charges, even for very sparse grids. This is important, because for large QM regions the cost of evaluating derivatives of the ChEIPG charges with respect to the density matrix becomes the overwhelming bottleneck in the calculation. (For a more realistic QM region of 349 basis functions, this cost is comparable to the cost of building the Coulomb and exchange matrices for a hybrid density functional.)

After gradients were implemented, new integral digestion routines were specifically written for ChEIPG charge derivatives, so as to minimize the number of integrals

computed. These routines decrease the time needed to compute derivatives by at least an order of magnitude for small systems. These new routines have eliminated the need to use Lebedev grids and have sped up the use of these routines to the point where the QM/MM-Ewald part of the calculation is no longer the bottle neck of the calculation. This allows one to add long range effects to QM/MM calculations and has led to a QM/MM-Ewald procedure that is a promising way to perform periodic QM/MM calculations in a Gaussian-orbital-based SCF electronic structure code. The method works for both HF and DFT calculations, including functionals of arbitrary complexity. Post-HF correlated wave functions can be built upon HF molecular orbitals and eigenvalues that are polarized by the PBC, and the fact that large basis sets can be used means that QM/MM calculations with correlated wave functions are possible.

This document shows the method required to achieve long-range electrostatics in QM/MM calculations for non-minimal basis sets using the Hartree-Fock method. The theory for this method is developed in a general way so that if desired, it could be implemented for post-HF methods and any charge scheme the user desires.

APPENDIX A

Ewald Summation

Note:

$$\Gamma(y) = \int_0^{\infty} x^{y-1} e^{-x} dx \quad (\text{A.1})$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad (\text{A.2})$$

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-x^2} dx \quad (\text{A.3})$$

σ is introduced because of a singularity at 0

The prime on the ij sum denotes that $i \neq j$ iff $\mathbf{n} = 0$

Definitions: $\mathbf{b}_n = \mathbf{n}L$; $\mathbf{a}_m = \frac{\mathbf{m}}{L}$; V is Volume; D is the number of Dimensions of replication

A.1 Transformation

$$\frac{1}{\Gamma(s)} \int_0^{\infty} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} dt \quad (\text{A.4})$$

Let: $u = t |\mathbf{r}_{ij} + \mathbf{b}_n|^2$; $du = |\mathbf{r}_{ij} + \mathbf{b}_n|^2 dt$; $t = \frac{u}{|\mathbf{r}_{ij} + \mathbf{b}_n|^2}$; $dt = \frac{du}{|\mathbf{r}_{ij} + \mathbf{b}_n|^2}$

$$\frac{1}{\Gamma(s)} \int_0^\infty \left(\frac{u}{|\mathbf{r}_{ij} + \mathbf{b}_n|^2} \right)^{s-1} e^{-u} \left(\frac{du}{|\mathbf{r}_{ij} + \mathbf{b}_n|^2} \right) \quad (\text{A.5})$$

$$\frac{1}{\Gamma(s)} |\mathbf{r}_{ij} + \mathbf{b}_n|^{-2} |\mathbf{r}_{ij} + \mathbf{b}_n|^{-2(s-1)} \int_0^\infty u^{s-1} e^{-u} du \quad (\text{A.6})$$

$$\frac{1}{\Gamma(s)} |\mathbf{r}_{ij} + \mathbf{b}_n|^{-2s} \Gamma(s) \quad (\text{A.7})$$

$$|\mathbf{r}_{ij} + \mathbf{b}_n|^{-2s} \quad (\text{A.8})$$

A.2 E_{cell}

$$E_{cell} = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} q_i q_j |\mathbf{r}_{ij} + \mathbf{b}_n|^{-2s} \quad (\text{A.9})$$

$$E_{cell} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^\infty t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt \quad (\text{A.10})$$

$$E_{cell} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\Gamma(s)} \left[\underbrace{\int_0^{\eta^2} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt}_{E_{k-space}} + \underbrace{\int_{\eta^2}^\infty t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt}_{E_{real}} \right] \quad (\text{A.11})$$

A.3 E_{real}

$$E_{real} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\Gamma(s)} \int_{\eta^2}^{\infty} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt \quad (\text{A.12})$$

$$\text{True Form: } E_{real}(s) = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\Gamma(s)} \int_{\eta^2}^{\infty} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} dt \quad (\text{A.13})$$

$$E_{real}\left(s = \frac{1}{2}\right) = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\Gamma\left(\frac{1}{2}\right)} \int_{\eta^2}^{\infty} t^{-\frac{1}{2}} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} dt \quad (\text{A.14})$$

Let: $u^2 = t|\mathbf{r}_{ij} + \mathbf{b}_n|^2$; $2u du = dt|\mathbf{r}_{ij} + \mathbf{b}_n|^2$; $t^{-\frac{1}{2}} = \frac{|\mathbf{r}_{ij} + \mathbf{b}_n|}{u}$; $dt = |\mathbf{r}_{ij} + \mathbf{b}_n|^{-2} 2u du$

$$E_{real}\left(s = \frac{1}{2}\right) = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\sqrt{\pi}} \int_{\eta|\mathbf{r}_{ij} + \mathbf{b}_n|}^{\infty} \frac{|\mathbf{r}_{ij} + \mathbf{b}_n|}{u} e^{-u^2} \frac{2u}{|\mathbf{r}_{ij} + \mathbf{b}_n|^2} du \quad (\text{A.15})$$

$$E_{real}\left(s = \frac{1}{2}\right) = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\sqrt{\pi}} \int_{\eta|\mathbf{r}_{ij} + \mathbf{b}_n|}^{\infty} \frac{2}{|\mathbf{r}_{ij} + \mathbf{b}_n|} e^{-u^2} du \quad (\text{A.16})$$

$$E_{real}\left(s = \frac{1}{2}\right) = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{b}_n| \sqrt{\pi}} \int_{\eta|\mathbf{r}_{ij} + \mathbf{b}_n|}^{\infty} e^{-u^2} du \quad (\text{A.17})$$

$$E_{real}\left(s = \frac{1}{2}\right) = \frac{1}{2} \sum_{\mathbf{b}_n} \sum'_{ij} q_i q_j \frac{\text{erfc}(\eta|\mathbf{r}_{ij} + \mathbf{b}_n|)}{|\mathbf{r}_{ij} + \mathbf{b}_n|} \quad (\text{A.18})$$

A.4 E_{recip} and E_{self}

$$E_{k-space} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum'_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt \quad (\text{A.19})$$

$$E_{k-space} = \underbrace{\frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt}_{E_{k-other}} - \underbrace{\frac{1}{2} \lim_{\sigma \rightarrow 0} \delta_{\mathbf{b}_n \mathbf{0}} \delta_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt}_{E_{self}} \quad (\text{A.20})$$

A.4.1 E_{self}

$$E_{self} = -\frac{1}{2} \lim_{\sigma \rightarrow 0} \delta_{\mathbf{b}_n \mathbf{0}} \sum_i \delta_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt \quad (\text{A.21})$$

$$\text{True Form: } E_{self} = -\frac{1}{2} \sum_i \frac{q_i^2}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} dt \quad (\text{A.22})$$

$$E_{self} \left(s = \frac{1}{2} \right) = -\frac{1}{2} \sum_i \frac{q_i^2}{\Gamma(\frac{1}{2})} \int_0^{\eta^2} t^{-\frac{1}{2}} dt \quad (\text{A.23})$$

$$E_{self} \left(s = \frac{1}{2} \right) = -\frac{1}{2} \sum_i \frac{q_i^2}{\sqrt{\pi}} 2t^{\frac{1}{2}} \Big|_{t=0}^{\eta^2} \quad (\text{A.24})$$

$$E_{self} \left(s = \frac{1}{2} \right) = -\frac{\eta}{\sqrt{\pi}} \sum_i q_i^2 \quad (\text{A.25})$$

A.4.2 E_{recip}

$$E_{k-other} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t|\mathbf{r}_{ij} + \mathbf{b}_n|^2} e^{-\sigma \mathbf{b}_n} dt \quad (\text{A.26})$$

See Appendix ??

$$E_{k-other} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_{\mathbf{b}_n} \sum_{ij} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} e^{-(\sigma+t)|\mathbf{b}_n + \frac{t}{\sigma+t} \rho_{ij}|^2} dt \quad (\text{A.27})$$

$$\text{Jacobi Sum Formula: } \sum_{\mathbf{b}} e^{-y|\mathbf{x} + \mathbf{b}|^2} = \frac{1}{V} \left(\frac{\pi}{y} \right)^{\frac{D}{2}} \sum_{\mathbf{a}} e^{-\frac{\pi^2}{y} \mathbf{a}^2} e^{2\pi i \mathbf{a} \cdot \mathbf{x}}$$

$$\text{Let: } y = \sigma + t \text{ and } \mathbf{x} = \frac{t}{\sigma+t} \rho_{ij}$$

$$E_{k-other} = \lim_{\sigma \rightarrow 0} \sum_{\mathbf{a}_m} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} \frac{1}{V} \left(\frac{\pi}{\sigma+t} \right)^{\frac{D}{2}} e^{-\frac{\pi^2}{\sigma+t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \frac{t}{\sigma+t} \rho_{ij}} dt \quad (\text{A.28})$$

$$E_{k-other} = \lim_{\sigma \rightarrow 0} \sum_{\mathbf{a}_m} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} e^{-\frac{\pi^2}{\sigma+t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \frac{t}{\sigma+t} \rho_{ij}} dt \quad (\text{A.29})$$

$$\begin{aligned} E_{k-other} &= \lim_{\sigma \rightarrow 0} \underbrace{\sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} e^{-\frac{\pi^2}{\sigma+t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \frac{t}{\sigma+t} \rho_{ij}} dt}_{E_{recip}} \\ &+ \lim_{\sigma \rightarrow 0} \underbrace{\delta_{\mathbf{a}_m \mathbf{0}} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} e^{-\frac{\pi^2}{\sigma+t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \frac{t}{\sigma+t} \rho_{ij}} dt}_{E_{other}} \end{aligned} \quad (\text{A.30})$$

$$E_{recip} = \lim_{\sigma \rightarrow 0} \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} e^{-\frac{\pi^2}{\sigma+t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \frac{t}{\sigma+t} \rho_{ij}} dt \quad (\text{A.31})$$

$$\text{True Form: } E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j \pi^{\frac{D}{2}}}{2V\Gamma(s)} \int_0^{\eta^2} t^{s-1-\frac{D}{2}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\pi^2}{t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \rho_{ij}} dt \quad (\text{A.32})$$

For $s = \frac{1}{2}$ and $D = 3$; the box is replicated in 3 out of 3 dimensions so $\mathbf{z}_{ij} = \mathbf{0}$ and

$$\rho_{ij} = \mathbf{r}_{ij}$$

$$E_{recip} \left(s = \frac{1}{2}, D = 3 \right) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j \pi^{\frac{3}{2}}}{2V\Gamma\left(\frac{1}{2}\right)} \int_0^{\eta^2} t^{\frac{1}{2}-1-\frac{3}{2}} e^{-\frac{\pi^2}{t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} \frac{1}{\pi^2 |\mathbf{a}_m|^2} dt \quad (\text{A.33})$$

$$E_{recip} \left(s = \frac{1}{2}, D = 3 \right) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j}{2V\pi |\mathbf{a}_m|^2} \int_0^{\eta^2} t^{-2} e^{-\frac{\pi^2}{t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} dt \quad (\text{A.34})$$

$$\text{Let: } v = \frac{-\pi^2 \mathbf{a}_m^2}{t}; \quad dv = \frac{\pi^2 \mathbf{a}_m^2}{t^2} dt; \quad t^{-2} dt = \frac{dv}{\pi^2 \mathbf{a}_m^2}$$

$$E_{recip} \left(s = \frac{1}{2}, D = 3 \right) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j}{2V\pi |\mathbf{a}_m|^2} \int_0^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} e^v e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} dv \quad (\text{A.35})$$

$$E_{recip} \left(s = \frac{1}{2}, D = 3 \right) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j}{2V\pi |\mathbf{a}_m|^2} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} e^v \Big|_0^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \quad (\text{A.36})$$

$$E_{recip} \left(s = \frac{1}{2}, D = 3 \right) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} \quad (\text{A.37})$$

$$\text{Let: } w(\mathbf{a}_m) = \frac{1}{V\pi \mathbf{a}_m^2} e^{\frac{-\pi^2 |\mathbf{a}_m|^2}{\eta^2}} \text{ and } |S(\mathbf{a}_m)|^2 = \sum_{ij} q_i q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}}$$

$$E_{recip} \left(s = \frac{1}{2}, D = 3 \right) = \frac{1}{2} \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \quad (\text{A.38})$$

$$|S(\mathbf{a}_m)|^2 = \sum_{ij} q_i q_j e^{2\pi i \mathbf{a}_m \cdot (\mathbf{r}_i - \mathbf{r}_j)} \quad (\text{A.39})$$

$$|S(\mathbf{a}_m)|^2 = \sum_{ij} q_i q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_i} e^{-2\pi i \mathbf{a}_m \cdot \mathbf{r}_j} \quad (\text{A.40})$$

$$|S(\mathbf{a}_m)|^2 = \sum_i q_i e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_i} \sum_j q_j e^{-2\pi i \mathbf{a}_m \cdot \mathbf{r}_j} \quad (\text{A.41})$$

$$|S(\mathbf{a}_m)|^2 = \sum_i q_i e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_i} \sum_i q_i e^{-2\pi i \mathbf{a}_m \cdot \mathbf{r}_i} \quad (\text{A.42})$$

$$|S(\mathbf{a}_m)|^2 = \left| \sum_i q_i e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_i} \right|^2 \quad (\text{A.43})$$

$$|S(\mathbf{a}_m)|^2 = \left| \sum_i q_i [\cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) + i \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_i)] \right|^2 \quad (\text{A.44})$$

$$|S(\mathbf{a}_m)|^2 = \left[\sum_i q_i \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right]^2 + \left[\sum_i q_i \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right]^2 \quad (\text{A.45})$$

A.5 E_{dipole}

$$E_{other} = \lim_{\sigma \rightarrow 0} \delta_{\mathbf{a}_m \mathbf{0}} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} e^{-\frac{\pi^2}{\sigma+t} \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \frac{t}{\sigma+t} \rho_{ij}} dt \quad (\text{A.46})$$

$$E_{other} = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} dt \quad (\text{A.47})$$

$\lim_{\sigma \rightarrow 0} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} = t^{s-1-\frac{D}{2}}$ if $s + \frac{D}{2} > 1$ then there is a divergence at $t = 0$

$$E_{other} = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \left[\underbrace{\int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt}_{E_{div}} + \underbrace{\int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} \left[e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} - 1 \right] dt}_{E_{dipole}} \right] \quad (\text{A.48})$$

$$\text{True Form: } E_{dipole} = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} \left[e^{-t\mathbf{z}_{ij}^2} e^{-\frac{\sigma t}{\sigma+t} \rho_{ij}^2} - 1 \right] dt \quad (\text{A.49})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma\left(\frac{1}{2}\right)} \frac{\pi^{\frac{3}{2}}}{V} \int_0^{\eta^2} \frac{t^{-\frac{1}{2}}}{(\sigma+t)^{\frac{3}{2}}} \left[e^{-\frac{\sigma t}{\sigma+t} \mathbf{r}_{ij}^2} - 1 \right] dt \quad (\text{A.50})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{q_i q_j \pi}{2V} \int_0^{\eta^2} \frac{t^{-\frac{1}{2}}}{(\sigma+t)^{\frac{3}{2}}} \left[\left(\sum_{\xi=0}^{\infty} \frac{1}{\xi!} \left(-\frac{\sigma t}{\sigma+t} \mathbf{r}_{ij}^2 \right)^{\xi} \right) - 1 \right] dt \quad (\text{A.51})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{q_i q_j \pi}{2V} \int_0^{\eta^2} \frac{t^{-\frac{1}{2}}}{(\sigma + t)^{\frac{3}{2}}} \left[\sum_{\xi=1}^{\infty} \frac{1}{\xi!} \left(-\frac{\sigma t}{\sigma + t} \mathbf{r}_{ij}^2 \right)^\xi \right] dt \quad (\text{A.52})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \sum_{\xi=1}^{\infty} \frac{q_i q_j \pi}{2V} \frac{\sigma^\xi \mathbf{r}_{ij}^{2\xi}}{\xi!} (-1)^\xi \int_0^{\eta^2} t^{\xi-\frac{1}{2}} (\sigma + t)^{-(\xi+\frac{3}{2})} dt \quad (\text{A.53})$$

$$\begin{aligned} \text{Use: } \int x^c (a + bx)^d dx &= \frac{x^{c+1} (a + bx)^d}{c + d + 1} + \frac{ad}{c + d + 1} \int x^c (a + bx)^{d-1} dx \\ &= \frac{1}{a(d+1)} \left[-x^{c+1} (a + bx)^{d+1} \right. \\ &\quad \left. + (c + d + 2) \int x^c (a + bx)^{d+1} \right] dx \end{aligned} \quad (\text{A.54})$$

$$\text{Let: } a = \sigma; b = 1; c = \xi - \frac{1}{2}; d = -\left(\xi + \frac{3}{2}\right); x = t$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \sum_{\xi=1}^{\infty} \frac{q_i q_j \pi}{2V} \frac{\sigma^\xi \mathbf{r}_{ij}^{2\xi}}{\xi!} (-1)^\xi \left[\frac{-1}{\sigma \left(\xi + \frac{1}{2}\right)} \left[-t^{\xi+\frac{1}{2}} (\sigma + t)^{-(\xi+\frac{1}{2})} + 0 \right] \right] \Bigg|_{t=0}^{\eta^2} \quad (\text{A.55})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \sum_{\xi=1}^{\infty} \frac{q_i q_j \pi}{2V} \frac{\sigma^{\xi-1} \mathbf{r}_{ij}^{2\xi}}{\xi! \left(\xi + \frac{1}{2}\right)} (-1)^{\xi+2} \eta^{2\xi+1} (\sigma + \eta^2)^{-(\xi+\frac{1}{2})} \quad (\text{A.56})$$

The only term that survives is $\xi = 1$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{q_i q_j \pi \mathbf{r}_{ij}^2}{2V^{\frac{3}{2}}} (-1)^3 \eta^3 (\sigma + \eta^2)^{-\frac{3}{2}} \quad (\text{A.57})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = \sum_{ij} -\frac{q_i q_j \pi}{3V} \mathbf{r}_{ij}^2 \quad (\text{A.58})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{\pi}{3V} \sum_{ij} q_i q_j \mathbf{r}_{ij}^2 \quad (\text{A.59})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{\pi}{3V} \sum_{ij} q_i q_j |\mathbf{r}_i - \mathbf{r}_j|^2 \quad (\text{A.60})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{\pi}{3V} \sum_{ij} q_i q_j [|\mathbf{r}_i|^2 - 2\mathbf{r}_i \cdot \mathbf{r}_j + |\mathbf{r}_j|^2] \quad (\text{A.61})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{\pi}{3V} \left[\sum_i q_i |\mathbf{r}_i|^2 \sum_j q_j - \sum_{ij} 2q_i q_j \mathbf{r}_i \cdot \mathbf{r}_j + \sum_j q_j |\mathbf{r}_j|^2 \sum_i q_i \right] \quad (\text{A.62})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{2\pi}{3V} \left[Q \sum_i q_i |\mathbf{r}_i|^2 - \sum_i q_i \mathbf{r}_i \cdot \sum_j q_j \mathbf{r}_j \right] \quad (\text{A.63})$$

$$E_{dipole} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{2\pi}{3V} \left[Q \sum_i q_i |\mathbf{r}_i|^2 - \left(\sum_i q_i \mathbf{r}_i \right)^2 \right] \quad (\text{A.64})$$

A.6 E_{charge}

$$\text{True Form: } E_{div} = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt \quad (\text{A.65})$$

$$E_{div} \left(s = \frac{1}{2}, D = 3 \right) = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{q_i q_j \pi}{2V} \int_0^{\eta^2} t^{-\frac{1}{2}} (\sigma+t)^{-\frac{3}{2}} dt \quad (\text{A.66})$$

Replicated/Image Charge Density (charge density of macrosystem (\mathfrak{R}) minus charge density of simulation cell):

$$\rho(\mathbf{r}) = \sum_i \sum_{\mathbf{b}_n} q_i \delta(\mathbf{r} - \mathbf{r}_i - \mathbf{b}_n) - \frac{Q}{V} \quad (\text{A.67})$$

Energy of Simulation Cell:

Note: Term in brackets removes self interactions. \tilde{N} = the total number of cells.

$$E_{cell} = \frac{1}{2\tilde{N}} \iint_{\mathfrak{R}} \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} [1 - \delta(\mathbf{r} - \mathbf{r}')] d\mathbf{r} d\mathbf{r}' \quad (\text{A.68})$$

$$E_{cell} = \frac{1}{2\tilde{N}} \iint_{\mathfrak{R}} \frac{1 - \delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \left[\sum_{ij} \sum_{\mathbf{b}_n \mathbf{b}'_n} ' q_i q_j \delta(\mathbf{r} - \mathbf{r}_i - \mathbf{b}_n) \delta(\mathbf{r}' - \mathbf{r}_j - \mathbf{b}'_n) - \frac{2Q}{V} \sum_i \sum_{\mathbf{b}_n} q_i \delta(\mathbf{r} - \mathbf{r}_i - \mathbf{b}_n) + \frac{Q^2}{V^2} \right] d\mathbf{r} d\mathbf{r}' \quad (\text{A.69})$$

$$E_{cell} = \frac{1}{2\tilde{N}} \sum_{ij} \sum_{\mathbf{b}_n \mathbf{b}'_n} ' \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{b}_n - \mathbf{b}'_n|} - \frac{1}{\tilde{N}} \frac{Q}{V} \sum_i \sum_{\mathbf{b}_n} \int_{\mathfrak{R}} \frac{q_i [1 - \delta(\mathbf{r} - \mathbf{r}_i - \mathbf{b}_n)]}{|\mathbf{r} - \mathbf{r}_i - \mathbf{b}_n|} d\mathbf{r} + \frac{1}{2\tilde{N}} \frac{Q^2}{V^2} \iint_{\mathfrak{R}} \frac{1 - \delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (\text{A.70})$$

Let: $\mathbf{c}_n = \mathbf{b}_n - \mathbf{b}'_n$

Redefine: 2nd Term: $\mathbf{r} = \mathbf{r} - \mathbf{r}_i - \mathbf{b}_n$ and 3rd Term: $\mathbf{r} = \mathbf{r} - \mathbf{r}'$

$$E_{cell} = \frac{1}{2\tilde{N}} \sum_{ij} \sum_{\mathbf{c}_n \mathbf{b}'_n} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{c}_n|} - \frac{1}{\tilde{N}} \frac{Q}{V} \sum_i q_i \sum_{\mathbf{b}_n} \int_{\mathfrak{R}} \frac{1 - \delta(\mathbf{r})}{|\mathbf{r}|} d\mathbf{r} + \frac{1}{2\tilde{N}} \frac{Q^2}{V^2} \iint_{\mathfrak{R}} \frac{1 - \delta(\mathbf{r})}{|\mathbf{r}|} d\mathbf{r} d\mathbf{r}' \quad (\text{A.71})$$

$$E_{cell} = \frac{1}{2\tilde{N}} \sum_{ij} \sum_{\mathbf{c}_n} \tilde{N} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{c}_n|} - \frac{1}{\tilde{N}} \frac{Q}{V} Q \tilde{N} \int_{\mathfrak{R}} \frac{1 - \delta(\mathbf{r})}{|\mathbf{r}|} d\mathbf{r} + \frac{1}{2\tilde{N}} \frac{Q^2}{V^2} \int_{\mathfrak{R}} \tilde{N} V \frac{1 - \delta(\mathbf{r})}{|\mathbf{r}|} d\mathbf{r} \quad (\text{A.72})$$

$\delta(\mathbf{r}) = 0$ as long as $D > 1$

$$E_{cell} = \underbrace{\frac{1}{2} \sum_{ij} \sum_{\mathbf{c}_n} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \mathbf{c}_n|}}_{E_{total}} - \underbrace{\frac{1}{2} \frac{Q^2}{V} \int_{\mathfrak{R}} \frac{1}{|\mathbf{r}|} d\mathbf{r}}_{E_{period}} \quad (\text{A.73})$$

E_{total} is the energy of the macrosystem and E_{period} is the energy of periodic images

$$E_{cell} = E_{total} - E_{period} \quad (\text{A.74})$$

$$E_{cell} = E_{real} + E_{recip} + E_{self} + E_{dipole} + E_{div} \quad (\text{A.75})$$

$$E_{total} = E_{real} + E_{recip} + E_{self} + E_{dipole} + E_{charge} \quad (\text{A.76})$$

$$E_{charge} = E_{div} + E_{period} \quad (\text{A.77})$$

Back Transform using $\mathbf{r} = \mathbf{r} - \mathbf{r}_i$:

$$E_{period} = -\frac{1}{2V} \sum_i q_i \int_{\Re} |\mathbf{r} - \mathbf{r}_i|^{-2s} d\mathbf{r} \quad (\text{A.78})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \sum_{ij} \frac{1}{2} \frac{q_i q_j}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt - \frac{1}{2V} \sum_i q_i \int_{\Re} |\mathbf{r} - \mathbf{r}_i|^{-2s} d\mathbf{r} \quad (\text{A.79})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \frac{1}{2} \frac{Q^2}{\Gamma(s)} \frac{\pi^{\frac{D}{2}}}{V} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt - \lim_{\sigma \rightarrow 0} \frac{1}{2V} \sum_i q_i \int_{\Re} |\mathbf{r} - \mathbf{r}_i|^{-2s} e^{-\sigma|\mathbf{r}-\mathbf{r}_i|^2} d\mathbf{r} \quad (\text{A.80})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \frac{Q}{2V} \left[\frac{Q\pi^{\frac{D}{2}}}{\Gamma(s)} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt - \sum_i q_i \int_{\Re} |\mathbf{r} - \mathbf{r}_i|^{-2s} e^{-\sigma|\mathbf{r}-\mathbf{r}_i|^2} d\mathbf{r} \right] \quad (\text{A.81})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \frac{Q}{2V} \left[\frac{Q\pi^{\frac{D}{2}}}{\Gamma(s)} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt - \sum_i q_i \int_0^{\infty} \int_{\Re} \frac{1}{\Gamma(s)} t^{s-1} e^{-t|\mathbf{r}-\mathbf{r}_i|^2} e^{-\sigma|\mathbf{r}-\mathbf{r}_i|^2} d\mathbf{r} dt \right] \quad (\text{A.82})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \frac{Q}{2V} \left[\frac{Q\pi^{\frac{D}{2}}}{\Gamma(s)} \int_0^{\eta^2} \frac{t^{s-1}}{(\sigma+t)^{\frac{D}{2}}} dt - \frac{1}{\Gamma(s)} \sum_i q_i \int_0^{\infty} \int_{\Re} t^{s-1} e^{-(\sigma+t)|\mathbf{r}-\mathbf{r}_i|^2} d\mathbf{r} dt \right] \quad (\text{A.83})$$

$$\text{Use: } \int_{\Re} a e^{-|\mathbf{x}+\mathbf{b}|/c^2} d\mathbf{x} = a |c|^D \pi^{\frac{D}{2}}$$

$$\text{Let: } a = t^{s-1} \text{ and } c = (\sigma+t)^{-1/2}$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \frac{Q}{2V} \left[\frac{Q\pi^{\frac{D}{2}}}{\Gamma(s)} \int_0^{\eta^2} t^{s-1} (\sigma + t)^{-\frac{D}{2}} dt - \frac{1}{\Gamma(s)} \sum_i q_i \int_0^{\infty} t^{s-1} (\sigma + t)^{-\frac{D}{2}} \pi^{\frac{D}{2}} dt \right] \quad (\text{A.84})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} \frac{Q^2\pi^{\frac{D}{2}}}{2V\Gamma(s)} \left[\int_0^{\eta^2} t^{s-1} (\sigma + t)^{-\frac{D}{2}} dt - \int_0^{\infty} t^{s-1} (\sigma + t)^{-\frac{D}{2}} dt \right] \quad (\text{A.85})$$

$$E_{charge} = \lim_{\sigma \rightarrow 0} -\frac{Q^2\pi^{\frac{D}{2}}}{2V\Gamma(s)} \int_{\eta^2}^{\infty} t^{s-1} (\sigma + t)^{-\frac{D}{2}} dt \quad (\text{A.86})$$

$$\text{True Form: } E_{charge} = -\frac{Q^2\pi^{\frac{D}{2}}}{2V\Gamma(s)} \int_{\eta^2}^{\infty} t^{s-1-\frac{D}{2}} dt \quad (\text{A.87})$$

$$E_{charge} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{Q^2\pi^{\frac{3}{2}}}{2V\Gamma\left(\frac{1}{2}\right)} \int_{\eta^2}^{\infty} t^{\frac{1}{2}-1-\frac{3}{2}} dt \quad (\text{A.88})$$

$$E_{charge} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{Q^2\pi}{2V} \int_{\eta^2}^{\infty} t^{-2} dt \quad (\text{A.89})$$

$$E_{charge} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{Q^2\pi}{2V} (-t^{-1}) \Big|_{\eta^2}^{\infty} \quad (\text{A.90})$$

$$E_{charge} \left(s = \frac{1}{2}, D = 3 \right) = -\frac{Q^2\pi}{2V\eta^2} \quad (\text{A.91})$$

A.7 E_{corr}

This is a correction term that is needed to remove double counting when performing MD on bonded systems. Eq. A.92 is the equation for E_{cell} when Ewald summation is used only to include non-1-2, non-1-3, and non-1-4 interactions; hence, a correction term is needed. Eq. A.93 states that the energy from the complete Ewald summation is the energy of the cell.

$$E_{cell} = E_{bond} + E_{angle} + E_{torsion} + E_{ewald} + E_{corr} \quad (\text{A.92})$$

$$E_{cell} = E_{ewald} \quad (\text{A.93})$$

$$-E_{corr} = E_{bond} + E_{angle} + E_{torsion} \quad (\text{A.94})$$

$$E_{corr} = - \sum_{\substack{i \in 1 \\ j \in 2,3,4}} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} = -\frac{1}{2} \sum_{i \neq j \in 1,2,3,4} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} \quad (\text{A.95})$$

The sums are to indicate that all 1-2, 1-3, and 1-4 interactions should be excluded. Note that for certain force fields 1-5 interactions should also be included. w_{ij} is a weight that is used in some force fields w_{12} and w_{13} are often 1, but w_{14} is usually smaller, something around 0.5.

A.8 Orthorhombic Boxes vs Cubic Boxes

- For a primitive cubic cell: $a = b = c$, $\alpha = \beta = \gamma = 90$, L will be used for the box length
- For a primitive orthorhombic cell: $a \neq b \neq c$, $\alpha = \beta = \gamma = 90$, $[L_x \ L_y \ L_z]$ will be used as the box length vector

Variables	P-Cubic Definition	P-Orthorhombic Definition	Energies Affected
\mathbf{b}_n	$[n_x L \ n_y L \ n_z L]$	$[n_x L_x \ n_y L_y \ n_z L_z]$	E_{real}
\mathbf{a}_m	$[\frac{m_x}{L} \ \frac{m_y}{L} \ \frac{m_z}{L}]$	$[\frac{m_x}{L_x} \ \frac{m_y}{L_y} \ \frac{m_z}{L_z}]$	E_{recip}
V	L^3	$L_x L_y L_z$	$E_{recip}, E_{dipole}, E_{charge}$

APPENDIX B

Ewald Derivatives

Note:

$$\nabla_i f(|\mathbf{r}_{ij}|) = f'(|\mathbf{r}_{ij}|) \frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|} = -\nabla_j f(|\mathbf{r}_{ij}|) \quad (\text{B.1})$$

$$\nabla_i \nabla_j f(|\mathbf{r}_{ij}|) = \left(\frac{f'(|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|} - f''(|\mathbf{r}_{ij}|) \right) \frac{\mathbf{r}_{ij} \mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^2} - \frac{f'(|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|} \mathbf{1} \quad (\text{B.2})$$

$$\nabla_i |\mathbf{r}_{ij}|^{-1} = -\frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^3} \quad (\text{B.3})$$

$$\nabla_i \nabla_j f(|\mathbf{r}_{ij}|) = \frac{1}{|\mathbf{r}_{ij}|^3} \left(\mathbf{1} - \frac{\mathbf{r}_{ij} \mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^2} \right) \quad (\text{B.4})$$

$$\frac{d}{dx} \text{erfc}(x) = -\frac{2e^{-x^2}}{\sqrt{\pi}} \quad (\text{B.5})$$

B.1 Further Energies

$$E^{cd} = \sum_{ij} \mu_i \cdot \nabla_i \sum_{\mathbf{b}_n} q_j |\mathbf{r}_{ij} + \mathbf{b}_n|^{-1} \quad (\text{B.6})$$

$$E^{dd} = \frac{1}{2} \sum_{ij} \mu_i \mu_j : \nabla_i \nabla_j \sum_{\mathbf{b}_n}' |\mathbf{r}_{ij} + \mathbf{b}_n|^{-1} \quad (\text{B.7})$$

B.2 Charge-Charge Gradients

For $s = \frac{1}{2}$ and $D = 3$

$$E_{real} = \frac{1}{2} \sum_{\mathbf{b}_n} \sum_{ij}' q_i q_j \frac{\text{erfc}(\eta |\mathbf{r}_{ij} + \mathbf{b}_n|)}{|\mathbf{r}_{ij} + \mathbf{b}_n|} \quad (\text{B.8})$$

$$E_{self} = -\frac{\eta}{\sqrt{\pi}} \sum_i q_i^2 \quad (\text{B.9})$$

$$E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_{ij} \frac{q_i q_j}{2V\pi |\mathbf{a}_m|^2} e^{-\frac{\pi^2 \mathbf{a}_m^2}{\eta^2}} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} \quad (\text{B.10})$$

$$E_{dipole} = -\frac{\pi}{3V} \sum_{ij} q_i q_j \mathbf{r}_{ij}^2 \quad (\text{B.11})$$

$$E_{charge} = -\frac{Q^2 \pi}{2V\eta^2} \quad (\text{B.12})$$

$$E_{corr} = \sum_{\substack{i \in 1 \\ j \in 2,3,4}} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} \quad (\text{B.13})$$

B.2.1 $\nabla_k E_{self}$ and $\nabla_k E_{charge}$

$$\nabla_k E_{self} = 0 \quad (\text{B.14})$$

$$\nabla_k E_{charge} = 0 \quad (\text{B.15})$$

B.2.2 $\nabla_k E_{dipole}$

$$\nabla_k E_{dipole} = -\frac{\pi}{3V} \nabla_k \sum_{ij} q_i q_j \mathbf{r}_{ij}^2 \quad (\text{B.16})$$

$$\nabla_k E_{dipole} = -\frac{\pi}{3V} \nabla_k \left(\sum_{\substack{i \neq k \\ j}} q_i q_j \mathbf{r}_{ij}^2 + \sum_j q_k q_j \mathbf{r}_{kj}^2 \right) \quad (\text{B.17})$$

$$\nabla_k E_{dipole} = -\frac{\pi}{3V} \nabla_k \left(\sum_{ij \neq k} q_i q_j \mathbf{r}_{ij}^2 + \sum_{j \neq k} q_k q_j \mathbf{r}_{kj}^2 + \sum_{i \neq k} q_i q_k \mathbf{r}_{ik}^2 + q_k \mathbf{r}_{kk}^2 \right) \quad (\text{B.18})$$

$$\nabla_k E_{dipole} = -\frac{\pi}{3V} \nabla_k \left(\sum_{j \neq k} q_k q_j \mathbf{r}_{kj}^2 + \sum_{i \neq k} q_i q_k \mathbf{r}_{ik}^2 \right) \quad (\text{B.19})$$

$$\nabla_k E_{dipole} = -\frac{2\pi}{3V} \nabla_k \sum_{i \neq k} q_i q_k \mathbf{r}_{ik}^2 \quad (\text{B.20})$$

$$\nabla_k E_{dipole} = \frac{4\pi}{3V} \sum_{i \neq k} q_i q_k \mathbf{r}_{ik} \quad (\text{B.21})$$

B.2.3 $\nabla_k E_{real}$

$$x_{ij} = \frac{\text{erfc}(\eta|\mathbf{r}_{ij} + \mathbf{b}_n|)}{|\mathbf{r}_{ij} + \mathbf{b}_n|}$$

$$\nabla_k E_{real} = \frac{1}{2} \sum_{\mathbf{b}_n} \nabla_k \sum'_{ij} q_i q_j x_{ij} \quad (\text{B.22})$$

$$\nabla_k E_{real} = \frac{1}{2} \sum_{\mathbf{b}_n} \nabla_k \left(\sum_{\substack{i \neq k \\ j}}' q_i q_j x_{ij} + \sum_j' q_k q_j x_{kj} \right) \quad (\text{B.23})$$

$$\nabla_k E_{real} = \frac{1}{2} \sum_{\mathbf{b}_n} \nabla_k \left(\sum_{ij \neq k}' q_i q_j x_{ij} + \sum_{j \neq k} q_k q_j x_{kj} + \sum_{i \neq k} q_i q_k x_{ik} \right) + \frac{1}{2} \sum_{\mathbf{b}_n \neq \mathbf{0}} \nabla_k q_k q_k x_{kk} \quad (\text{B.24})$$

$$\nabla_k E_{real} = \sum_{\mathbf{b}_n} \nabla_k \sum_{i \neq k} q_i q_k x_{ik} + \frac{1}{2} \sum_{\mathbf{b}_n \neq \mathbf{0}} \nabla_k q_k q_k x_{kk} \quad (\text{B.25})$$

$$\nabla_k E_{real} = \sum_{\mathbf{b}_n} \nabla_k \sum_{i \neq k} q_i q_k \frac{\text{erfc}(\eta |\mathbf{r}_{ik} + \mathbf{b}_n|)}{|\mathbf{r}_{ik} + \mathbf{b}_n|} + \frac{1}{2} \sum_{\mathbf{b}_n \neq \mathbf{0}} \nabla_k q_k q_k \frac{\text{erfc}(\eta |\mathbf{b}_n|)}{|\mathbf{b}_n|} \quad (\text{B.26})$$

$$\nabla_k E_{real} = \sum_{\mathbf{b}_n} \sum_{i \neq k} q_i q_k \nabla_k \frac{\text{erfc}(\eta |\mathbf{r}_{ik} + \mathbf{b}_n|)}{|\mathbf{r}_{ik} + \mathbf{b}_n|} \quad (\text{B.27})$$

$$\nabla_k E_{real} = \sum_{\mathbf{b}_n} \sum_{i \neq k} q_i q_k \left[-\frac{\text{erfc}(\eta |\mathbf{r}_{ik} + \mathbf{b}_n|)}{|\mathbf{r}_{ik} + \mathbf{b}_n|^2} + -\frac{2\eta}{\sqrt{\pi}} \frac{e^{-\eta^2 |\mathbf{r}_{ik} + \mathbf{b}_n|^2}}{|\mathbf{r}_{ik} + \mathbf{b}_n|} \right] \frac{\mathbf{r}_{ik} + \mathbf{b}_n}{|\mathbf{r}_{ik} + \mathbf{b}_n|} \quad (\text{B.28})$$

$$\nabla_k E_{real} = -\sum_{\mathbf{b}_n} \sum_{i \neq k} q_i q_k \left[\frac{\text{erfc}(\eta |\mathbf{r}_{ik} + \mathbf{b}_n|)}{|\mathbf{r}_{ik} + \mathbf{b}_n|} + \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{ik} + \mathbf{b}_n|^2} \right] \frac{\mathbf{r}_{ik} + \mathbf{b}_n}{|\mathbf{r}_{ik} + \mathbf{b}_n|^2} \quad (\text{B.29})$$

B.2.4 $\nabla_k E_{recip}$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \sum_{ij} q_i q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} \quad (\text{B.30})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \sum_{ij} q_i q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} \quad (\text{B.31})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \left(\sum_{\substack{i \neq k \\ j}} q_i q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} + \sum_j q_k q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{kj}} \right) \quad (\text{B.32})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \left(\sum_{ij \neq k} q_i q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ij}} + \sum_{i \neq k} q_i q_k e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ik}} \right. \\ \left. + \sum_{j \neq k} q_k q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{kj}} + q_k^2 e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{kk}} \right) \quad (\text{B.33})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \left(\sum_{i \neq k} q_i q_k e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ik}} + \sum_{j \neq k} q_k q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{kj}} \right) \quad (\text{B.34})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \left(\sum_{i \neq k} q_i q_k e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ik}} + \sum_{i \neq k} q_i q_k e^{2\pi i \mathbf{a}_m \cdot -\mathbf{r}_{ik}} \right) \quad (\text{B.35})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \sum_{i \neq k} q_i q_k (e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ik}} + e^{-2\pi i \mathbf{a}_m \cdot \mathbf{r}_{ik}}) \quad (\text{B.36})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \nabla_k \sum_{i \neq k} q_i q_k 2 \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_{ik}) \quad (\text{B.37})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \sum_{i \neq k} q_i q_k \nabla_k \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_{ik}) \quad (\text{B.38})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V \pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \sum_{i \neq k} q_i q_k \cdot -\sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_{ik}) \cdot -2\pi \mathbf{a}_m \quad (\text{B.39})$$

Note: the restriction is removed because $\mathbf{r}_{kk} = \mathbf{0}$ which will cause the argument of the sine function to be 0 and therefore the sine function itself to be 0.

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \sum_i \frac{2q_i q_k}{V |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_{ik}) \mathbf{a}_m \quad (\text{B.40})$$

OR

$$\nabla_k E_{recip} = \nabla_k \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V \pi \mathbf{a}_m^2} e^{\frac{-\pi^2 |\mathbf{a}_m|^2}{\eta^2}} \left(\left[\sum_i q_i \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right]^2 + \left[\sum_i q_i \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right]^2 \right) \quad (\text{B.41})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V \pi \mathbf{a}_m^2} e^{\frac{-\pi^2 |\mathbf{a}_m|^2}{\eta^2}} \nabla_k \left(\left[\sum_i q_i \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right]^2 + \left[\sum_i q_i \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right]^2 \right) \quad (\text{B.42})$$

$$\nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V \pi \mathbf{a}_m^2} e^{\frac{-\pi^2 |\mathbf{a}_m|^2}{\eta^2}} \left(2 \left[\sum_i q_i \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right] \nabla_k \left[\sum_i q_i \cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right] + 2 \left[\sum_i q_i \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right] \nabla_k \left[\sum_i q_i \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_i) \right] \right) \quad (\text{B.43})$$

$$\begin{aligned} \nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{2V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2|\mathbf{a}_m|^2}{\eta^2}} & \left(2 \left[\sum_i q_i \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_i) \right] \right. \\ & \left. [-q_k \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_k) 2\pi\mathbf{a}_m] + 2 \left[\sum_i q_i \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_i) \right] [q_k \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_k) 2\pi\mathbf{a}_m] \right) \end{aligned} \quad (\text{B.44})$$

$$\begin{aligned} \nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{4\pi}{2V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2|\mathbf{a}_m|^2}{\eta^2}} & \left(\left[\sum_i q_i \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_i) \right] [-q_k \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_k)] \right. \\ & \left. + \left[\sum_i q_i \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_i) \right] [q_k \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_k)] \right) \mathbf{a}_m \end{aligned} \quad (\text{B.45})$$

$$\begin{aligned} \nabla_k E_{recip} = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{2}{V\mathbf{a}_m^2} e^{-\frac{\pi^2|\mathbf{a}_m|^2}{\eta^2}} & \left(\left[\sum_i q_i \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_i) \right] [q_k \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_k)] \right. \\ & \left. - \left[\sum_i q_i \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_i) \right] [q_k \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_k)] \right) \mathbf{a}_m \end{aligned} \quad (\text{B.46})$$

B.2.5 $\nabla_k E_{corr}$

$$\nabla_k E_{corr} = \nabla_k \sum_{\substack{i \in 1 \\ j \in 2,3,4}} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} \quad (\text{B.47})$$

$$\nabla_k E_{corr} = -\frac{1}{2} \nabla_k \sum_{i \neq j \in 1,2,3,4} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} \quad (\text{B.48})$$

$$\nabla_k E_{corr} = -\frac{1}{2} \nabla_k \left(\sum_{\substack{i \neq k \\ i \neq j}} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} + \sum_{j \neq k \in 1,2,3,4} \frac{w_{kj} q_k q_j}{|\mathbf{r}_{kj}|} \right) \quad (\text{B.49})$$

$$\nabla_k E_{corr} = -\frac{1}{2} \nabla_k \left(\sum_{\substack{i \neq k \\ i, j \in \{1, 2, 3, 4\} \\ j \neq k}} \frac{w_{ij} q_i q_j}{|\mathbf{r}_{ij}|} + \sum_{i \neq k \in \{1, 2, 3, 4\}} \frac{w_{ik} q_i q_k}{|\mathbf{r}_{ik}|} + \sum_{j \neq k \in \{1, 2, 3, 4\}} \frac{w_{kj} q_k q_j}{|\mathbf{r}_{kj}|} \right) \quad (\text{B.50})$$

$$\nabla_k E_{corr} = -\nabla_k \sum_{i \neq k \in \{1, 2, 3, 4\}} \frac{w_{ik} q_i q_k}{|\mathbf{r}_{ik}|} \quad (\text{B.51})$$

$$\nabla_k E_{corr} = \sum_{i \neq k \in \{1, 2, 3, 4\}} \frac{w_{ik} q_i q_k}{|\mathbf{r}_{ik}|^3} \mathbf{r}_{ik} \quad (\text{B.52})$$

APPENDIX C

MM Dynamics

Note: $\sigma_{ij} = \sigma_i + \sigma_j$ and $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$. Usually these are true; however, different conventions can be used.

C.1 van der Waal's Energy

C.1.1 vdW Energy

$$E_{vdW} = \frac{1}{2} \sum_{i \neq j} E_{ij}^{vdW} \quad (\text{C.1})$$

$$E_{ij}^{vdW} = \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{|\mathbf{r}_{ij}|} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{|\mathbf{r}_{ij}|} \right)^6 \right] \quad (\text{C.2})$$

$$E_{ij}^{vdW} = \varepsilon_{ij} \left(\frac{\sigma_{ij}}{|\mathbf{r}_{ij}|} \right)^6 \left[\left(\frac{\sigma_{ij}}{|\mathbf{r}_{ij}|} \right)^6 - 2 \right] \quad (\text{C.3})$$

C.1.2 Switching Function

$$f(|\mathbf{r}_{ij}|) = \frac{(|\mathbf{r}_{ij}| - r_c)^2 (3r_c - r_1 - 2|\mathbf{r}_{ij}|)}{(r_1 - r_c)^3} \quad (\text{C.4})$$

Note: r_c is a cutoff radius that can be set to anything. In Q-Chem, if it is not user-defined, then the default is $r_c = 5 \sigma_{ij}$. $r_1 = 0.9 r_c$ for the default $r_1 = 4.5 \sigma_{ij}$.

$$f(|\mathbf{r}_{ij}|) = \frac{(|\mathbf{r}_{ij}| - 5\sigma_{ij})^2 (15\sigma_{ij} - 4.5\sigma_{ij} - 2|\mathbf{r}_{ij}|)}{(4.5\sigma_{ij} - 5\sigma_{ij})^3} \quad (\text{C.5})$$

$$f(|\mathbf{r}_{ij}|) = \frac{(|\mathbf{r}_{ij}| - 5\sigma_{ij})^2 (10.5\sigma_{ij} - 2|\mathbf{r}_{ij}|)}{(-0.5\sigma_{ij})^3} \quad (\text{C.6})$$

$$f(|\mathbf{r}_{ij}|) = \frac{-8}{\sigma_{ij}^3} [(|\mathbf{r}_{ij}|^2 - 10\sigma_{ij}|\mathbf{r}_{ij}| + 25\sigma_{ij}^2) (10.5\sigma_{ij} - 2|\mathbf{r}_{ij}|)] \quad (\text{C.7})$$

$$f(|\mathbf{r}_{ij}|) = \frac{-8}{\sigma_{ij}^3} [10.5\sigma_{ij}|\mathbf{r}_{ij}|^2 - 105\sigma_{ij}^2|\mathbf{r}_{ij}| + 262.5\sigma_{ij}^3 - 2|\mathbf{r}_{ij}|^3 + 20\sigma_{ij}|\mathbf{r}_{ij}|^2 - 50\sigma_{ij}^2|\mathbf{r}_{ij}|] \quad (\text{C.8})$$

$$f(|\mathbf{r}_{ij}|) = \frac{-8}{\sigma_{ij}^3} [-2|\mathbf{r}_{ij}|^3 + 30.5\sigma_{ij}|\mathbf{r}_{ij}|^2 - 155\sigma_{ij}^2|\mathbf{r}_{ij}| + 262.5\sigma_{ij}^3] \quad (\text{C.9})$$

$$f(|\mathbf{r}_{ij}|) = \frac{-8}{\sigma_{ij}^3} [-2|\mathbf{r}_{ij}|^3 + 30.5\sigma_{ij}|\mathbf{r}_{ij}|^2 - 155\sigma_{ij}^2|\mathbf{r}_{ij}| + 262.5\sigma_{ij}^3] \quad (\text{C.10})$$

$$f(|\mathbf{r}_{ij}|) = 16 \left(\frac{|\mathbf{r}_{ij}|}{\sigma_{ij}} \right)^3 - 244 \left(\frac{|\mathbf{r}_{ij}|}{\sigma_{ij}} \right)^2 + 1240 \left(\frac{|\mathbf{r}_{ij}|}{\sigma_{ij}} \right) - 2100 \quad (\text{C.11})$$

C.1.3 Smooth Energy

$$E_{vdW} = \frac{1}{2} \sum_{i \neq j} \begin{cases} E_{ij}^{vdW} & \text{if } |\mathbf{r}_{ij}| \leq r_1 \\ E_{ij}^{vdW} f(|\mathbf{r}_{ij}|) & \text{if } r_1 < |\mathbf{r}_{ij}| < r_c \\ 0 & \text{if } |\mathbf{r}_{ij}| > r_c \end{cases} \quad (\text{C.12})$$

C.2 van der Waal's Gradient

C.2.1 vdW Gradient

$$\nabla_k E_{vdW} = \nabla_k \left[\frac{1}{2} \sum_{\substack{i \neq j \\ i \neq k}} E_{ij}^{vdW} + \frac{1}{2} \sum_{j \neq k} E_{kj}^{vdW} \right] \quad (C.13)$$

$$\nabla_k E_{vdW} = \nabla_k \left[\frac{1}{2} \sum_{\substack{i \neq j \\ i \neq k}} E_{ij}^{vdW} + \frac{1}{2} \sum_{j \neq k} E_{kj}^{vdW} \right] \quad (C.14)$$

$$\nabla_k E_{vdW} = \nabla_k \left[\frac{1}{2} \sum_{\substack{i \neq j \\ ij \neq k}} E_{ij}^{vdW} + \frac{1}{2} \sum_{j \neq k} E_{kj}^{vdW} + \frac{1}{2} \sum_{i \neq k} E_{ik}^{vdW} \right] \quad (C.15)$$

$$\nabla_k E_{vdW} = \nabla_k \sum_{i \neq k} E_{ik}^{vdW} \quad (C.16)$$

$$\nabla_k E_{ik}^{vdW} = \nabla_k \varepsilon_{ik} \left[\left(\frac{\sigma_{ik}}{|\mathbf{r}_{ik}|} \right)^{12} - 2 \left(\frac{\sigma_{ik}}{|\mathbf{r}_{ik}|} \right)^6 \right] \quad (C.17)$$

$$\nabla_k E_{ik}^{vdW} = \varepsilon_{ik} \left[\sigma_{ik}^{12} \nabla_k \left(\frac{1}{|\mathbf{r}_{ik}|^{12}} \right) - 2 \sigma_{ik}^6 \nabla_k \left(\frac{1}{|\mathbf{r}_{ik}|^6} \right) \right] \quad (C.18)$$

$$\nabla_k E_{ik}^{vdW} = \varepsilon_{ik} \left[-12 \sigma_{ik}^{12} \frac{1}{|\mathbf{r}_{ik}|^{13}} \nabla_k (|\mathbf{r}_{ik}|) - 2 \cdot -6 \sigma_{ik}^6 \frac{1}{|\mathbf{r}_{ik}|^7} \nabla_k (|\mathbf{r}_{ik}|) \right] \quad (C.19)$$

$$\nabla_k E_{ik}^{vdW} = \varepsilon_{ik} \left[-12 \frac{\sigma_{ik}^{12}}{|\mathbf{r}_{ik}|^{13}} + 12 \frac{\sigma_{ik}^6}{|\mathbf{r}_{ik}|^7} \right] \nabla_k (|\mathbf{r}_{ik}|) \quad (C.20)$$

$$\nabla_k E_{ik}^{vdW} = 12 \varepsilon_{ik} \frac{\sigma_{ik}^6}{|\mathbf{r}_{ik}|^7} \left[-\frac{\sigma_{ik}^6}{|\mathbf{r}_{ik}|^6} + 1 \right] \frac{\mathbf{r}_{ik}}{|\mathbf{r}_{ik}|} \quad (C.21)$$

$$\nabla_k E_{ik}^{vdW} = 12 \varepsilon_{ik} \frac{\sigma_{ik}^6}{|\mathbf{r}_{ik}|^8} \left[1 - \frac{\sigma_{ik}^6}{|\mathbf{r}_{ik}|^6} \right] \mathbf{r}_{ik} \quad (\text{C.22})$$

C.2.2 Derivative of Switching Function

$$\nabla_k f(|\mathbf{r}_{ik}|) = \nabla_k \frac{(|\mathbf{r}_{ik}| - r_c)^2 (3r_c - r_1 - 2|\mathbf{r}_{ik}|)}{(r_1 - r_c)^3} \quad (\text{C.23})$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \frac{1}{(r_1 - r_c)^3} \nabla_k (|\mathbf{r}_{ik}| - r_c)^2 (3r_c - r_1 - 2|\mathbf{r}_{ik}|) \quad (\text{C.24})$$

$$\begin{aligned} \nabla_k f(|\mathbf{r}_{ik}|) &= \frac{1}{(r_1 - r_c)^3} \left[(3r_c - r_1 - 2|\mathbf{r}_{ik}|) \nabla_k (|\mathbf{r}_{ik}| - r_c)^2 \right. \\ &\quad \left. + (|\mathbf{r}_{ik}| - r_c)^2 \nabla_k (3r_c - r_1 - 2|\mathbf{r}_{ik}|) \right] \quad (\text{C.25}) \end{aligned}$$

$$\begin{aligned} \nabla_k f(|\mathbf{r}_{ik}|) &= \frac{1}{(r_1 - r_c)^3} \left[(3r_c - r_1 - 2|\mathbf{r}_{ik}|) 2 (|\mathbf{r}_{ik}| - r_c) \nabla_k |\mathbf{r}_{ik}| \right. \\ &\quad \left. + (|\mathbf{r}_{ik}| - r_c)^2 \cdot -2 \nabla_k |\mathbf{r}_{ik}| \right] \quad (\text{C.26}) \end{aligned}$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \frac{2}{(r_1 - r_c)^3} \left[(3r_c - r_1 - 2|\mathbf{r}_{ik}|) (|\mathbf{r}_{ik}| - r_c) - (|\mathbf{r}_{ik}| - r_c)^2 \right] \nabla_k |\mathbf{r}_{ik}| \quad (\text{C.27})$$

$$\begin{aligned} \nabla_k f(|\mathbf{r}_{ik}|) &= \frac{2}{(r_1 - r_c)^3} \left[(3r_c |\mathbf{r}_{ik}| - r_1 |\mathbf{r}_{ik}| - 2|\mathbf{r}_{ik}|^2 - 3r_c^2 + r_1 r_c + 2r_c |\mathbf{r}_{ik}|) \right. \\ &\quad \left. - (|\mathbf{r}_{ik}|^2 - 2r_c |\mathbf{r}_{ik}| + r_c^2) \right] \frac{\mathbf{r}_{ik}}{|\mathbf{r}_{ik}|} \quad (\text{C.28}) \end{aligned}$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \frac{2}{(r_1 - r_c)^3} [3r_c |\mathbf{r}_{ik}| - r_1 |\mathbf{r}_{ik}| - 3|\mathbf{r}_{ik}|^2 - 4r_c^2 + r_1 r_c] \frac{\mathbf{r}_{ik}}{|\mathbf{r}_{ik}|} \quad (\text{C.29})$$

Using the substitution in the Note.

$$\nabla_k f(|\mathbf{r}_{ik}|) = \nabla_k \left[16 \left(\frac{|\mathbf{r}_{ik}|}{\sigma_{ik}} \right)^3 - 244 \left(\frac{|\mathbf{r}_{ik}|}{\sigma_{ik}} \right)^2 + 1240 \left(\frac{|\mathbf{r}_{ik}|}{\sigma_{ik}} \right) - 2100 \right] \quad (\text{C.30})$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \frac{16}{\sigma_{ik}^3} \nabla_k |\mathbf{r}_{ik}|^3 - \frac{244}{\sigma_{ik}^2} \nabla_k |\mathbf{r}_{ik}|^2 + \frac{1240}{\sigma_{ik}} \nabla_k |\mathbf{r}_{ik}| \quad (\text{C.31})$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \left[\frac{16}{\sigma_{ik}^3} 3|\mathbf{r}_{ik}|^2 - \frac{244}{\sigma_{ik}^2} 2|\mathbf{r}_{ik}| + \frac{1240}{\sigma_{ik}} \right] \nabla_k |\mathbf{r}_{ik}| \quad (\text{C.32})$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \left[\frac{48}{\sigma_{ik}^3} |\mathbf{r}_{ik}|^2 - \frac{488}{\sigma_{ik}^2} |\mathbf{r}_{ik}| + \frac{1240}{\sigma_{ik}} \right] \frac{\mathbf{r}_{ik}}{|\mathbf{r}_{ik}|} \quad (\text{C.33})$$

$$\nabla_k f(|\mathbf{r}_{ik}|) = \left[\frac{48 |\mathbf{r}_{ik}|}{\sigma_{ik}^3} - \frac{488}{\sigma_{ik}^2} + \frac{1240}{\sigma_{ik} |\mathbf{r}_{ik}|} \right] \mathbf{r}_{ik} \quad (\text{C.34})$$

C.2.3 Smooth vdW Gradient

$$\nabla_k E_{vdW} = \sum_{i \neq k} \begin{cases} \nabla_k E_{ik}^{vdW} & \text{if } |\mathbf{r}_{ik}| \leq r_1 \\ f(|\mathbf{r}_{ij}|) \nabla_k E_{ik}^{vdW} + E_{ik}^{vdW} \nabla_k f(|\mathbf{r}_{ik}|) & \text{if } r_1 < |\mathbf{r}_{ik}| < r_c \\ \mathbf{0} & \text{if } |\mathbf{r}_{ik}| > r_c \end{cases} \quad (\text{C.35})$$

C.3 Minimum Image Convention

This is done in the *ndistance* routine. *w* is a place holder for *x*, *y*, and *z*.

- $\mathbf{r}^{ij} = \mathbf{r}_i - \mathbf{r}_j$
- Take \mathbf{r}^{ij} and break it down into components $r_x^{ij}, r_y^{ij}, r_z^{ij}$.
- If $|r_w^{ij}| < \frac{L_w}{2}$, do nothing.
- If $|r_w^{ij}| > \frac{L_w}{2}$
 - If $r_w^{ij} > 0$ then $r_w^{ij} = r_w^{ij} - L_w$
 - If $r_w^{ij} < 0$ then $r_w^{ij} = r_w^{ij} + L_w$
- Then, if needed, add the appropriate box vector, $[n_x L_x \quad n_y L_y \quad n_z L_z]$.

C.4 Confining Potential

$$V_{confine} = \sum_i V_0 (1 + \tanh[\ell_c (|\mathbf{r}_i| - r_c)]) \quad (\text{C.36})$$

$2V_0$ is the maximum value of the function. Note that this will be in kcal/mol. It will get turned into hartree by the code. r_c is the cutoff radius. ℓ_c is the width of the function. $4/\ell_c$ is the approximate width that it takes for the function to go from 0 to $2V_0$ which is centered around r_c . i is the set of atoms whose original positions, \mathbf{r}_i , are within the cutoff sphere, i.e. $|\mathbf{r}_i(t=0)| < r_c$.

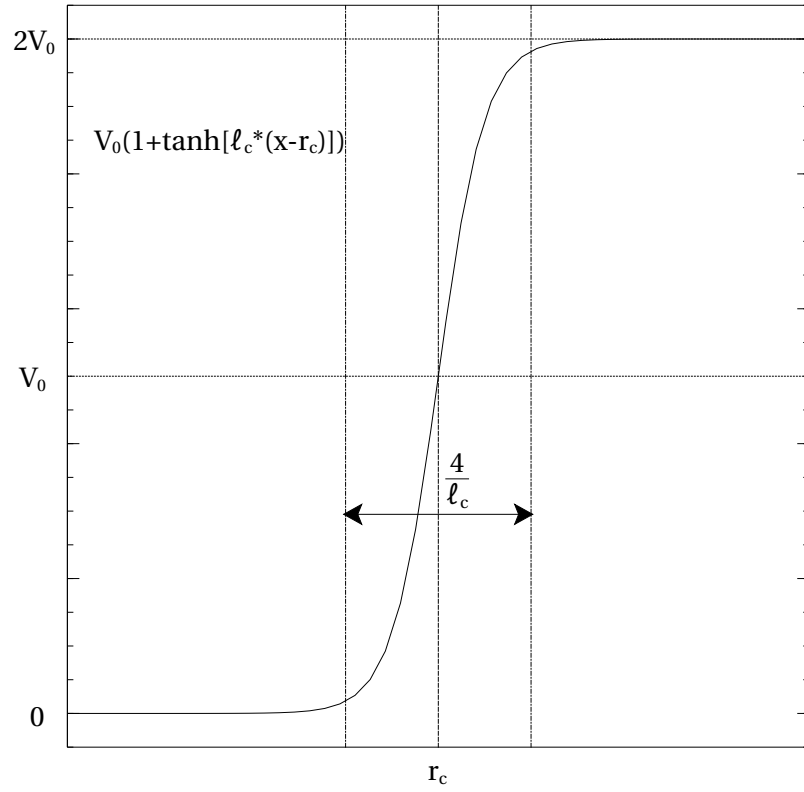


Figure C.1: Graph of the hyperbolic tangent function with parameters of interest shown.

$$\nabla_k V_{confine} = \frac{V_0 \ell_c}{|\mathbf{r}_k|} \text{sech}^2 [\ell_c (|\mathbf{r}_k| - r_c)] \mathbf{r}_k \quad (\text{C.37})$$

Note:

$$\tanh z = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

$$\frac{d \tanh z}{dz} = \frac{d}{dz} (e^z - e^{-z}) (e^z + e^{-z})^{-1}$$

$$\frac{d \tanh z}{dz} = (e^z + e^{-z})^{-1} \frac{d}{dz} (e^z - e^{-z}) + (e^z - e^{-z}) \frac{d}{dz} (e^z + e^{-z})^{-1}$$

$$\frac{d \tanh z}{dz} = (e^z + e^{-z})^{-1} (e^z + e^{-z}) - (e^z - e^{-z})^2 (e^z + e^{-z})^{-2}$$

$$\frac{d \tanh z}{dz} = 1 - \left(\frac{e^z - e^{-z}}{e^z + e^{-z}} \right)^2$$

$$\frac{d \tanh z}{dz} = \frac{(e^z + e^{-z})^2 - (e^z - e^{-z})^2}{(e^z + e^{-z})^2}$$

$$\frac{d \tanh z}{dz} = \frac{e^{2z} + 2 + e^{-2z} - e^{2z} + 2 - e^{-2z}}{(e^z + e^{-z})^2}$$

$$\frac{d \tanh z}{dz} = \frac{4}{(e^z + e^{-z})^2}$$

$$\frac{d \tanh z}{dz} = \left(\frac{2}{e^z + e^{-z}} \right)^2$$

$$\frac{d \tanh z}{dz} = \text{sech}^2 z$$

APPENDIX D

Ewald Energy Correction

Note: SS = supersystem (simulation cell plus images) aka PB (periodic bound); RS = real space (simulation cell only); PI (periodic images only) aka PBC (periodic boundary correction)

$$E_{total}^{SS} = E_{QM/QM}^{SS} + E_{QM/MM}^{SS} + E_{MM/MM}^{SS} \quad (D.1)$$

$$E_{total}^{SS} = E_{QM/QM}^{RS} + \Delta E_{QM/QM}^{PI} + E_{QM/MM}^{RS} + \Delta E_{QM/MM}^{PI} + E_{MM/MM}^{SS} \quad (D.2)$$

$E_{MM/MM}^{SS}$ is the term for normal MM Ewald

$E_{QM/QM}^{RS}$ is some flavor of electronic structure

$E_{QM/MM}^{RS}$ is some flavor of QM/MM scheme such as Janus or ONIOM

D.1 $\Delta E_{QM/QM}^{PI}$

$$\Delta E_{QM/QM}^{PI} = E_{QM/QM}^{SS} - E_{QM/QM}^{RS} \quad (D.3)$$

Assume $\sum_J q_J = 0$

Assume a tin-foil boundary or use enough boxes to reach a desired convergence so

$$E_{dipole} = 0$$

$$\Delta E_{QM/QM}^{PI} = E_{real}^{QM/QM} + E_{recip}^{QM/QM} + E_{self}^{QM/QM} - E_{QM/QM}^{RS} \quad (D.4)$$

$$\begin{aligned} \Delta E_{QM/QM}^{PI} = & \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \left[\sum_{\mathbf{b}_n} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} (1 - \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}) \right] + \right. \\ & \left. \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{e^{-\pi^2 \mathbf{a}_m^2 / \eta^2}}{\pi V \mathbf{a}_m^2} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{\alpha\beta}} \right] - \frac{2\eta}{\sqrt{\pi}} \delta_{\alpha\beta} \right\} - \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \frac{1 - \delta_{\alpha\beta}}{|\mathbf{r}_{\alpha\beta}|} \quad (D.5) \end{aligned}$$

Note: In the reciprocal term, $|S(\mathbf{a}_m)|^2$ or $|S'(\mathbf{a}_m)|^2$ can be used. The primed quantity is defined in the next section. The latter is currently implemented so the same function can be used for QM/QM and QM/MM evaluation.

$$\begin{aligned} \Delta E_{QM/QM}^{PI} = & \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \right. \\ & \left. \left[\sum_{\mathbf{b}_n} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} (1 - \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}) \right] - \frac{2\eta}{\sqrt{\pi}} \delta_{\alpha\beta} - \frac{1 - \delta_{\alpha\beta}}{|\mathbf{r}_{\alpha\beta}|} \right\} \quad (D.6) \end{aligned}$$

$$\begin{aligned} \Delta E_{QM/QM}^{PI} = & \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \right. \\ & \left. \left[\sum_{\mathbf{b}_n} \frac{1 - \text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} (1 - \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}) \right] - \frac{2\eta}{\sqrt{\pi}} \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}} - \frac{1 - \delta_{\alpha\beta}}{|\mathbf{r}_{\alpha\beta}|} \delta_{\mathbf{b}_n \mathbf{0}} \right\} \quad (D.7) \end{aligned}$$

$$\Delta E_{QM/QM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \sum_{\mathbf{b}_n} \right. \quad (\text{D.8})$$

$$\left. \frac{1}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} - \frac{\delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} + \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|) \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right. \quad (\text{D.9})$$

$$\left. - \frac{2\eta}{\sqrt{\pi}} \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}} - \frac{\delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} + \frac{\delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right\} \quad (\text{D.10})$$

$$\Delta E_{QM/QM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \sum_{\mathbf{b}_n} \right. \quad (\text{D.11})$$

$$\left. \frac{1 - \delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} + \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|) \delta_{\alpha\beta} \delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} - \frac{2\eta}{\sqrt{\pi}} \right\}$$

Note: $\lim_{x \rightarrow 0} \frac{\text{erf}(\kappa x)}{x} = \frac{2\kappa}{\sqrt{\pi}}$

$$\Delta E_{QM/QM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \sum_{\mathbf{b}_n} \frac{1 - \delta_{\mathbf{b}_n \mathbf{0}}}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right. \quad (\text{D.12})$$

$$\left. - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right\}$$

$$\Delta E_{QM/QM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{1}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right. \quad (\text{D.13})$$

$$\left. - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta}|)}{|\mathbf{r}_{\alpha\beta}|} \right\}$$

$$\Delta E_{QM/QM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \left\{ -\frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta}|)}{|\mathbf{r}_{\alpha\beta}|} + \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right\} \quad (\text{D.14})$$

$$\Delta E_{QM/QM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S(\mathbf{a}_m)|^2 \right] + \left[\frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right] - \left[\frac{1}{2} \sum_{\alpha \neq \beta}^{N_{QM}} Q_\alpha Q_\beta \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta}|)}{|\mathbf{r}_{\alpha\beta}|} \right] - \left[\frac{\eta}{\sqrt{\pi}} \sum_{\alpha}^{N_{QM}} Q_\alpha^2 \right] \quad (\text{D.15})$$

D.2 $\Delta E_{QM/MM}^{PI}$

$$\Delta E_{QM/MM}^{PI} = E_{QM/MM}^{SS} - E_{QM/MM}^{RS} \quad (\text{D.16})$$

Assume $\sum_{\alpha}^{N_{QM}} Q_\alpha = 0$ and $\sum_j^{N_{MM}} q_j = 0$. $\therefore j \neq \alpha$ there is no self term or primes on sums.

$$\Delta E_{QM/MM}^{PI} = E_{real} + E_{recip} + E_{charge} - E_{QM/MM}^{RS} \quad (\text{D.17})$$

D.2.1 Real Space Term

$$E_{real} - E_{QM/MM}^{RS} = \sum_{\alpha}^{N_{QM}} \sum_j^{N_{MM}} \left[Q_\alpha q_j \sum_{\mathbf{b}_n} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha j} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha j} + \mathbf{b}_n|} \right] - \frac{Q_\alpha q_j}{|\mathbf{r}_{\alpha j}|} \quad (\text{D.18})$$

$$E_{real} - E_{QM/MM}^{RS} = \sum_{\alpha} Q_\alpha \sum_j q_j \left\{ \left[\sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha j} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha j} + \mathbf{b}_n|} \right] + \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha j}|)}{|\mathbf{r}_{\alpha j}|} - \frac{1}{|\mathbf{r}_{\alpha j}|} \right\} \quad (\text{D.19})$$

$$E_{real} - E_{QM/MM}^{RS} = \sum_{\alpha} Q_{\alpha} \sum_j q_j \left\{ \left[\sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\operatorname{erfc}(\eta |\mathbf{r}_{\alpha j} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha j} + \mathbf{b}_n|} \right] - \frac{\operatorname{erf}(\eta |\mathbf{r}_{\alpha j}|)}{|\mathbf{r}_{\alpha j}|} \right\} \quad (\text{D.20})$$

D.2.2 Reciprocal Term

$$E_{recip} = \sum_{\alpha} Q_{\alpha} \sum_j q_j \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi |\mathbf{a}_m|^2} e^{\frac{-\pi^2 \mathbf{a}_m^2}{\eta^2}} e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{j\alpha}} \quad (\text{D.21})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_{j\alpha}} \quad (\text{D.22})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j e^{2\pi i \mathbf{a}_m \cdot \mathbf{r}_j} e^{-2\pi i \mathbf{a}_m \cdot \mathbf{r}_{\alpha}} \quad (\text{D.23})$$

Let: $x = 2\pi \mathbf{a}_m \cdot \mathbf{r}_j$ and $y = 2\pi \mathbf{a}_m \cdot \mathbf{r}_{\alpha}$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j e^{ix} e^{-iy} \quad (\text{D.24})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j \frac{1}{2} [e^{ix} e^{-iy} + e^{-ix} e^{iy}] \quad (\text{D.25})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j \frac{1}{2} [(\cos x + i \sin x)(\cos y - i \sin y) + (\cos x - i \sin x)(\cos y + i \sin y)] \quad (\text{D.26})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j \frac{1}{2} [(\cos x \cos y - i \cos x \sin y + i \sin x \cos y + \sin x \sin y) + (\cos x \cos y + i \cos x \sin y - i \sin x \cos y + \sin x \sin y)] \quad (\text{D.27})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j [(\cos x \cos y + \sin x \sin y)] \quad (\text{D.28})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j [\cos(x - y)] \quad (\text{D.29})$$

$$E_{recip} = \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) Q_{\alpha} q_j [\cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_{j\alpha})] \quad (\text{D.30})$$

$$\text{Let: } |S'(\mathbf{a}_m)|^2 = \sum_{\alpha} \sum_j Q_{\alpha} q_j [\cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_{j\alpha})]$$

$$\begin{aligned} \Delta E_{QM/MM}^{PI} &= \sum_{\alpha} Q_{\alpha} \sum_j q_j \left\{ \left[\sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha j} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha j} + \mathbf{b}_n|} \right] - \frac{\text{erf}(\eta |\mathbf{r}_{\alpha j}|)}{|\mathbf{r}_{\alpha j}|} \right\} \\ &+ \sum_{\alpha} \sum_j \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi |\mathbf{a}_m|^2} e^{-\frac{\pi^2 \mathbf{a}_m^2}{\eta^2}} Q_{\alpha} q_j [\cos(2\pi \mathbf{a}_m \cdot \mathbf{r}_{j\alpha})] \quad (\text{D.31}) \end{aligned}$$

D.3 Forming Pair Potential

$$\begin{aligned} \Delta E_{QM/QM}^{PI} &= \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S'(\mathbf{a}_m)|^2 \right] + \left[\frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_{\alpha} Q_{\beta} \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right] \\ &- \left[\frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_{\alpha} Q_{\beta} \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta}|)}{|\mathbf{r}_{\alpha\beta}|} \right] \quad (\text{D.32}) \end{aligned}$$

Note: The self term can be added back into the erf term to remove the restriction on the sum.

$$\Delta E_{QM/QM}^{PI} = \frac{1}{2} \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S'(\mathbf{a}_m)|^2 \right] + \left[\frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\beta} + \mathbf{b}_n|} \right] - \left[\frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \frac{\text{erf}(\eta |\mathbf{r}_{\alpha\beta}|)}{|\mathbf{r}_{\alpha\beta}|} \right] \quad (\text{D.33})$$

$$\Delta E_{QM/MM}^{PI} = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} w(\mathbf{a}_m) |S'(\mathbf{a}_m)|^2 \right] + \left[\sum_{\alpha}^{N_{QM}} Q_\alpha \sum_j^{N_{MM}} q_j \sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha j} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha j} + \mathbf{b}_n|} \right] - \left[\frac{1}{2} \sum_{\alpha}^{N_{QM}} Q_\alpha \sum_j^{N_{MM}} q_j \frac{\text{erf}(\eta |\mathbf{r}_{\alpha j}|)}{|\mathbf{r}_{\alpha j}|} \right] \quad (\text{D.34})$$

$\therefore \Delta E_{QM/QM}^{PI}$ and $\Delta E_{QM/MM}^{PI}$ have the same forms, if we invoke tin-foil boundary conditions to remove the dipole terms, we can define a pair potential $\omega(\mathbf{r})$.

$$\Delta E^{PI} = \Delta E_{QM/QM}^{PI} + \Delta E_{QM/MM}^{PI} \quad (\text{D.35})$$

$$\Delta E^{PI} = \frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_\alpha Q_\beta \omega(\mathbf{r}_{\alpha\beta}) + \sum_{\alpha}^{N_{QM}} Q_\alpha \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\alpha j}) \quad (\text{D.36})$$

$$\omega(\mathbf{r}_{\alpha\aleph}) = \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2\mathbf{a}_m^2}{\eta^2}} \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\alpha\aleph}) \right] + \left[\sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta |\mathbf{r}_{\alpha\aleph} + \mathbf{b}_n|)}{|\mathbf{r}_{\alpha\aleph} + \mathbf{b}_n|} \right] - \left[\frac{\text{erf}(\eta |\mathbf{r}_{\alpha\aleph}|)}{|\mathbf{r}_{\alpha\aleph}|} \right] \quad (\text{D.37})$$

D.4 Maximum Vectors and Timings

Both the real space sum and the reciprocal space sum converge as Gaussians. A constant is chosen such that e^{-C^2} is small. In the case of an SCF convergence the sum must converge to a value that is equal to or less than the threshold.

$$\therefore C = \sqrt{-\ln(\text{SCF Threshold})} \quad (\text{D.38})$$

Every vector is subject to the constraint that its distance must be less than the distance of the max vectors, $|n| \leq |n_{max}|$ or $|m| \leq |m_{max}|$. Setting the maximum value of each direction to the max value yields a supercube (SC). Adding in the constraint yields a supersphere (SS). The code creates the supercube and makes each vector subject to the constraint so that a supersphere is obtained. The total vectors for both SC and SS are shown; however, SS is non-trivial. The derivation of the equation is shown in appendix I.

D.4.1 Real Space

Note:

$$\mathbf{n}_{max} \equiv \begin{bmatrix} n_{max} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ n_{max} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ n_{max} \end{bmatrix}$$

$$\mathbf{1} \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

n_{total} is the total number of \mathbf{n} 's used.

$$C \leq \eta |\mathbf{r}_{\alpha\beta} + \mathbf{b}_{n_{max}}| \quad (\text{D.39})$$

$$C \leq \eta |\mathbf{r}_{\alpha\mathfrak{N}} + \mathbf{n}_{max}L| \quad (\text{D.40})$$

$$\because \mathbf{r}_{\alpha\mathfrak{N}} \leq L\mathbf{1} \quad (\text{D.41})$$

Redefine \mathbf{n}_{max} so it includes the extra $\mathbf{1}$

$$C \leq \eta |\mathbf{n}_{max}L| \quad (\text{D.42})$$

$$\frac{C}{\eta L} \leq |\mathbf{n}_{max}| \quad (\text{D.43})$$

$$\frac{C}{\eta L} \leq n_{max} \quad (\text{D.44})$$

$$n_{max} = \text{ceiling} \left(\frac{C}{\eta L} \right) \quad (\text{D.45})$$

If $\frac{C}{\eta L} \leq \frac{1}{2}$ then $r_c < \frac{L}{2}$ and $n_{max} = 0$, assuming Minimum Image Convention is used.

$$n_{total}^{SC} = (2n_{max} + 1)^3 \quad (\text{D.46})$$

n_{max}	$n_{total}^{SS} = N(3, n_{max})$
0	1
1	7
2	33
3	123
4	257
5	515
6	925
7	1419
8	2109
9	3071
10	4169

D.4.2 Reciprocal Space

Notes are the same as for previous subsection.

$$C \leq \frac{\pi |\mathbf{m}_{max}|}{\eta L} \quad (\text{D.47})$$

$$\frac{CL\eta}{\pi} \leq |\mathbf{m}_{max}| \quad (\text{D.48})$$

$$\frac{CL\eta}{\pi} \leq \sqrt{m_{max}^2} \quad (\text{D.49})$$

$$\frac{CL\eta}{\pi} \leq m_{max} \quad (\text{D.50})$$

$$m_{max} = \text{ceiling} \left(\frac{CL\eta}{\pi} \right) \quad (\text{D.51})$$

$$m_{total}^{SC} = (2m_{max} + 1)^3 \quad (\text{D.52})$$

The above equation should technically be $(2m_{max} + 1)^3 - 1$ because $\mathbf{0}$ is not included; however, since the program must still go through the *if* loops for the $\mathbf{0}$ vector, the -1 has been dropped. Conversion from SC to SS uses the same chart as above.

D.4.3 Finding η_{min}

$$\frac{\partial}{\partial \eta} \left[(n_{total}^{SC})^3 + (m_{total}^{SC})^3 \right] \Big|_{\eta=\eta_{min}} = 0 \quad (\text{D.53})$$

$$\frac{\partial}{\partial \eta} \left[(2n_{max} + 1)^3 + (2m_{max} + 1)^3 \right] \Big|_{\eta=\eta_{min}} = 0 \quad (\text{D.54})$$

$$\left[6(2n_{max} + 1)^2 \frac{\partial n_{max}}{\partial \eta} + 6(2m_{max} + 1)^2 \frac{\partial m_{max}}{\partial \eta} \right] \Big|_{\eta=\eta_{min}} = 0 \quad (\text{D.55})$$

$$\left(\frac{2C}{\eta_{min}L} + 1 \right)^2 - \frac{C}{\eta_{min}^2 L} + \left(\frac{2CL\eta_{min}}{\pi} + 1 \right)^2 \frac{CL}{\pi} = 0 \quad (\text{D.56})$$

$$\left(\frac{2CL\eta_{min} + \pi}{\pi} \right)^2 \frac{CL}{\pi} = \left(\frac{2C + \eta_{min}L}{\eta_{min}L} \right)^2 \frac{C}{\eta_{min}^2 L} \quad (\text{D.57})$$

$$(2CL\eta_{min} + \pi)^2 \frac{L}{\pi^3} = (2C + \eta_{min}L)^2 \frac{1}{\eta_{min}^4 L^3} \quad (\text{D.58})$$

$$(2CL\eta_{min} + \pi)^2 \frac{\eta_{min}^4 L^4}{\pi^3} = (2C + \eta_{min}L)^2 \quad (\text{D.59})$$

$$(2CL\eta_{min} + \pi) \frac{\eta_{min}^2 L^2}{\pi\sqrt{\pi}} = 2C + \eta_{min}L \quad (\text{D.60})$$

$$\frac{2CL^3\eta_{min}^3}{\pi\sqrt{\pi}} + \frac{\eta_{min}^2L^2}{\sqrt{\pi}} - \eta_{min}L - 2C = 0 \quad (\text{D.61})$$

D.4.4 Orthorhombic Cells

Real and Reciprocal Spaces

L_w will be used to represent L_x , L_y , and L_z

The equations do not change; however, the maxes are not the same in each of the three directions.

$$n_w^{max} = \text{ceiling} \left(\frac{C}{\eta L_w} \right) \quad (\text{D.62})$$

$$m_w^{max} = \text{ceiling} \left(\frac{CL_w\eta}{\pi} \right) \quad (\text{D.63})$$

Similar to cubic cells, the vectors are still subject to the constraint that $|n| \leq |n^{max}|$; however, n^{max} is now defined as the maximum of n_x^{max} , n_y^{max} , or n_z^{max} . Likewise, the same process holds for m^{max} as well.

$$n^{max} = \begin{cases} \begin{bmatrix} n_x^{max} & 0 & 0 \end{bmatrix} & \text{if } n_x^{max} \geq n_y^{max} \cup n_x^{max} \geq n_z^{max} \\ \begin{bmatrix} 0 & n_y^{max} & 0 \end{bmatrix} & \text{if } n_y^{max} \geq n_x^{max} \cup n_y^{max} \geq n_z^{max} \\ \begin{bmatrix} 0 & 0 & n_z^{max} \end{bmatrix} & \text{if } n_z^{max} \geq n_x^{max} \cup n_z^{max} \geq n_y^{max} \end{cases} \quad (\text{D.64})$$

Finding η_{min}

Quick Way: Construct a cubic box using n^{max} and m^{max} . Note: L_{max} and L_{min} will stand for the maximum and minimum components, respectively, of L_x , L_y , and L_z .

$$\frac{\partial}{\partial \eta} [(2n^{max} + 1)^3 + (2m^{max} + 1)^3] \Big|_{\eta=\eta_{min}} = 0 \quad (\text{D.65})$$

$$\left[6(2n_{max} + 1)^2 \frac{\partial n^{max}}{\partial \eta} + 6(2m^{max} + 1)^2 \frac{\partial m^{max}}{\partial \eta} \right] \Big|_{\eta=\eta_{min}} = 0 \quad (\text{D.66})$$

$$\left(\frac{2C}{\eta_{min} L_{min}} + 1 \right)^2 - \frac{C}{\eta_{min}^2 L_{min}} + \left(\frac{2CL_{max}\eta_{min}}{\pi} + 1 \right)^2 \frac{CL_{max}}{\pi} = 0 \quad (\text{D.67})$$

$$\left(\frac{2CL_{max}\eta_{min}}{\pi} + 1 \right)^2 \frac{CL_{max}}{\pi} = \frac{C}{\eta_{min}^2 L_{min}} \left(\frac{2C}{\eta_{min} L_{min}} + 1 \right)^2 \quad (\text{D.68})$$

$$\left(\frac{2CL_{max}\eta_{min} + \pi}{\pi} \right)^2 \frac{L_{max}}{\pi} = \frac{1}{\eta_{min}^2 L_{min}} \left(\frac{2C + \eta_{min} L_{min}}{\eta_{min} L_{min}} \right)^2 \quad (\text{D.69})$$

$$(2CL_{max}\eta_{min} + \pi)^2 \frac{L_{max}}{\pi^3} = \frac{1}{\eta_{min}^4 L_{min}^3} (2C + \eta_{min} L_{min})^2 \quad (\text{D.70})$$

$$(2CL_{max}\eta_{min} + \pi)^2 \frac{\eta_{min}^4 L_{min}^3 L_{max}}{\pi^3} = (2C + \eta_{min} L_{min})^2 \quad (\text{D.71})$$

$$(2CL_{max}\eta_{min} + \pi) \eta_{min}^2 \sqrt{\frac{L_{min}^3 L_{max}}{\pi^3}} = 2C + \eta_{min} L_{min} \quad (\text{D.72})$$

Let $\zeta = \sqrt{\frac{L_{min}^3 L_{max}}{\pi^3}}$

$$2CL_{max}\zeta\eta_{min}^3 + \pi\zeta\eta_{min}^2 - L_{min}\eta_{min} - 2C = 0 \quad (\text{D.73})$$

Long Way:

$$\frac{\partial}{\partial \eta} \left[(2n_x^{max} + 1) (2n_y^{max} + 1) (2n_z^{max} + 1) + (2m_x^{max} + 1) (2m_y^{max} + 1) (2m_z^{max} + 1) \right] \Big|_{\eta=\eta_{min}} = 0 \quad (D.74)$$

$$\frac{\partial}{\partial \eta} \left[(2n_x^{max} + 1) (4n_y^{max} n_z^{max} + 2n_y^{max} + 2n_z^{max} + 1) + (2m_x^{max} + 1) (4m_y^{max} m_z^{max} + 2m_y^{max} + 2m_z^{max} + 1) \right] \Big|_{\eta=\eta_{min}} = 0 \quad (D.75)$$

$$\begin{aligned} \frac{\partial}{\partial \eta} & \left[(8n_x^{max} n_y^{max} n_z^{max} + 4n_x^{max} n_y^{max} + 4n_x^{max} n_z^{max} + 2n_x^{max} + 4n_y^{max} n_z^{max} \right. \\ & \left. + 2n_y^{max} + 2n_z^{max} + 1) + (8m_x^{max} m_y^{max} m_z^{max} + 4m_x^{max} m_y^{max} + 4m_x^{max} m_z^{max} \right. \\ & \left. + 2m_x^{max} + 4m_y^{max} m_z^{max} + 2m_y^{max} + 2m_z^{max} + 1) \right] \Big|_{\eta=\eta_{min}} = 0 \quad (D.76) \end{aligned}$$

$$\begin{aligned} 2 \frac{\partial}{\partial \eta} & \left[(4n_x^{max} n_y^{max} n_z^{max} + 2n_x^{max} n_y^{max} + 2n_x^{max} n_z^{max} + 2n_y^{max} n_z^{max} + n_x^{max} + n_y^{max} \right. \\ & \left. + n_z^{max}) + (4m_x^{max} m_y^{max} m_z^{max} + 2m_x^{max} m_y^{max} + 2m_x^{max} m_z^{max} + 2m_y^{max} m_z^{max} \right. \\ & \left. + m_x^{max} + m_y^{max} + m_z^{max}) \right] \Big|_{\eta=\eta_{min}} = 0 \quad (D.77) \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial \eta} & \left[\frac{4C^3}{\eta^3 L_x L_y L_z} + \frac{2C^2}{\eta^2} \left(\frac{1}{L_x L_y} + \frac{1}{L_x L_z} + \frac{1}{L_y L_z} \right) + \frac{C}{\eta} \left(\frac{1}{L_x} + \frac{1}{L_y} + \frac{1}{L_z} \right) \right. \\ & \left. + \frac{4C^3 L_x L_y L_z \eta^3}{\pi^3} + \frac{2C^2 \eta^2}{\pi^2} (L_x L_y + L_x L_z + L_y L_z) + \frac{C \eta}{\pi} (L_x + L_y + L_z) \right] \Big|_{\eta=\eta_{min}} = 0 \quad (D.78) \end{aligned}$$

$$\left[\frac{-12C^3}{\eta^4 L_x L_y L_z} - \frac{4C^2}{\eta^3} \left(\frac{1}{L_x L_y} + \frac{1}{L_x L_z} + \frac{1}{L_y L_z} \right) - \frac{C}{\eta^2} \left(\frac{1}{L_x} + \frac{1}{L_y} + \frac{1}{L_z} \right) + \frac{12C^3 L_x L_y L_z \eta^2}{\pi^3} + \frac{4C^2 \eta}{\pi^2} (L_x L_y + L_x L_z + L_y L_z) + \frac{C}{\pi} (L_x + L_y + L_z) \right] \Big|_{\eta=\eta_{min}} = 0 \quad (\text{D.79})$$

$$\begin{aligned} & \frac{12C^2 L_x L_y L_z \eta_{min}^2}{\pi^3} + \frac{4C \eta_{min}}{\pi^2} (L_x L_y + L_x L_z + L_y L_z) + \frac{1}{\pi} (L_x + L_y + L_z) \\ &= \frac{12C^2}{\eta_{min}^4 L_x L_y L_z} + \frac{4C}{\eta_{min}^3} \left(\frac{1}{L_x L_y} + \frac{1}{L_x L_z} + \frac{1}{L_y L_z} \right) + \frac{1}{\eta_{min}^2} \left(\frac{1}{L_x} + \frac{1}{L_y} + \frac{1}{L_z} \right) \end{aligned} \quad (\text{D.80})$$

Let: $V = L_x L_y L_z$, $L_s = L_x + L_y + L_z$, $L_p = \frac{1}{L_x} + \frac{1}{L_y} + \frac{1}{L_z}$,
 $A_s = L_x L_y + L_x L_z + L_y L_z$, and $A_p = \frac{1}{L_x L_y} + \frac{1}{L_x L_z} + \frac{1}{L_y L_z}$

$$\frac{12C^2 V \eta_{min}^2}{\pi^3} + \frac{4C \eta_{min} A_s}{\pi^2} + \frac{L_s}{\pi} = \frac{12C^2}{\eta_{min}^4 V} + \frac{4C A_p}{\eta_{min}^3} + \frac{L_p}{\eta_{min}^2} \quad (\text{D.81})$$

$$\frac{12C^2 V^2 \eta_{min}^6}{\pi^3} + \frac{4CV \eta_{min}^5 A_s}{\pi^2} + \frac{L_s V \eta_{min}^4}{\pi} = 12C^2 + 4C A_p V \eta_{min} + L_p V \eta_{min}^2 \quad (\text{D.82})$$

$$12C^2 V^2 \eta_{min}^6 + 4CV \eta_{min}^5 A_s \pi + L_s V \eta_{min}^4 \pi^2 = 12C^2 \pi^3 + 4C A_p V \eta_{min} \pi^3 + L_p V \eta_{min}^2 \pi^3 \quad (\text{D.83})$$

Note: $L_p V = A_s$ and $A_p V = L_s$

$$12C^2 V^2 \eta_{min}^6 + 4CV \eta_{min}^5 A_s \pi + L_s V \eta_{min}^4 \pi^2 = 12C^2 \pi^3 + 4C L_s \eta_{min} \pi^3 + A_s \eta_{min}^2 \pi^3 \quad (\text{D.84})$$

$$12C^2V^2\eta_{min}^6 + 4CV A_s\pi\eta_{min}^5 + L_sV\pi^2\eta_{min}^4 - A_s\pi^3\eta_{min}^2 - 4CL_s\pi^3\eta_{min} - 12C^2\pi^3 = 0$$

(D.85)

APPENDIX E

Ewald Fock Matrix Correction

Note: Greek letters are QM atoms; Roman letters are MM atoms; Hebrew letters
are QM or MM atoms

$$\Delta F_{\mu\nu}^{PI} = \frac{\partial \Delta E^{PI}}{\partial P_{\mu\nu}} \quad (\text{E.1})$$

$$\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial \Delta E^{PI}}{\partial Q_{\gamma}} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \quad (\text{E.2})$$

$$\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \frac{\partial}{\partial Q_{\gamma}} \left[\frac{1}{2} \sum_{\alpha\beta}^{N_{QM}} Q_{\alpha} Q_{\beta} \omega(\mathbf{r}_{\alpha\beta}) + \sum_{\alpha}^{N_{QM}} Q_{\alpha} \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\alpha j}) \right] \quad (\text{E.3})$$

$$\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \frac{\partial}{\partial Q_{\gamma}} \left[\frac{1}{2} \sum_{\substack{\alpha\beta \\ \alpha \neq \gamma}}^{N_{QM}} Q_{\alpha} Q_{\beta} \omega(\mathbf{r}_{\alpha\beta}) + \frac{1}{2} \sum_{\beta}^{N_{QM}} Q_{\gamma} Q_{\beta} \omega(\mathbf{r}_{\gamma\beta}) \right. \\ \left. + \sum_{\alpha \neq \gamma}^{N_{QM}} Q_{\alpha} \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\alpha j}) + Q_{\gamma} \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\gamma j}) \right] \quad (\text{E.4})$$

$$\begin{aligned}
\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \frac{\partial}{\partial Q_{\gamma}} & \left[\frac{1}{2} \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}}^{N_{QM}} Q_{\alpha} Q_{\beta} \omega(\mathbf{r}_{\alpha\beta}) + \frac{1}{2} \sum_{\alpha \neq \gamma}^{N_{QM}} Q_{\alpha} Q_{\gamma} \omega(\mathbf{r}_{\alpha\gamma}) \right. \\
& \left. + \frac{1}{2} \sum_{\beta \neq \gamma}^{N_{QM}} Q_{\gamma} Q_{\beta} \omega(\mathbf{r}_{\gamma\beta}) + \frac{1}{2} Q_{\gamma}^2 \omega(\mathbf{r}_{\gamma\gamma}) + \sum_{\alpha \neq \gamma}^{N_{QM}} Q_{\alpha} \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\alpha j}) + Q_{\gamma} \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\gamma j}) \right] \quad (\text{E.5})
\end{aligned}$$

$$\begin{aligned}
\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} & \left[\frac{1}{2} \sum_{\alpha \neq \gamma}^{N_{QM}} Q_{\alpha} \omega(\mathbf{r}_{\alpha\gamma}) + \frac{1}{2} \sum_{\beta \neq \gamma}^{N_{QM}} Q_{\beta} \omega(\mathbf{r}_{\gamma\beta}) + Q_{\gamma} \omega(\mathbf{r}_{\gamma\gamma}) \right. \\
& \left. + \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\gamma j}) \right] \quad (\text{E.6})
\end{aligned}$$

$$\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \left[\sum_{\alpha \neq \gamma}^{N_{QM}} Q_{\alpha} \omega(\mathbf{r}_{\alpha\gamma}) + Q_{\gamma} \omega(\mathbf{r}_{\gamma\gamma}) + \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\gamma j}) \right] \quad (\text{E.7})$$

$$\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \left[\sum_{\alpha}^{N_{QM}} Q_{\alpha} \omega(\mathbf{r}_{\alpha\gamma}) + \sum_j^{N_{MM}} q_j \omega(\mathbf{r}_{\gamma j}) \right] \quad (\text{E.8})$$

$$\Delta F_{\mu\nu}^{PI} = \sum_{\gamma} \frac{\partial Q_{\gamma}}{\partial P_{\mu\nu}} \left[\sum_{\aleph}^{N_a} Q_{\aleph} \omega(\mathbf{r}_{\aleph\gamma}) \right] \quad (\text{E.9})$$

APPENDIX F

Spatial Derivatives

F.1 Normal HF Derivative

$$\varepsilon_{HF} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + V_{nuc} \quad (\text{F.1})$$

$$\begin{aligned} \frac{d\varepsilon_{HF}}{dx} = & \sum_{\mu\nu} \frac{dP_{\mu\nu}}{dx} H_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{dH_{\mu\nu}}{dx} + \sum_{\mu\nu\lambda\sigma} \frac{dP_{\mu\nu}}{dx} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \\ & + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{d}{dx} \langle \mu\nu || \lambda\sigma \rangle + \frac{dV_{nuc}}{dx} \end{aligned} \quad (\text{F.2})$$

$$\begin{aligned} \frac{d\varepsilon_{HF}}{dx} = & \underbrace{\sum_{\mu\nu} P_{\mu\nu} \frac{dH_{\mu\nu}}{dx} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{d}{dx} \langle \mu\nu || \lambda\sigma \rangle + \frac{dV_{nuc}}{dx}}_{E_D} + \sum_{\mu\nu} \frac{dP_{\mu\nu}}{dx} H_{\mu\nu} \\ & + \sum_{\mu\nu\lambda\sigma} \frac{dP_{\mu\nu}}{dx} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \end{aligned} \quad (\text{F.3})$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D + \sum_{\mu\nu} \frac{dP_{\mu\nu}}{dx} \left(H_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \right) \quad (\text{F.4})$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D + \sum_{\mu\nu} \frac{dP_{\mu\nu}}{dx} F_{\mu\nu} \quad (\text{F.5})$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D + \sum_{\mu\nu} \sum_i \left[\frac{dc_{i\mu}^*}{dx} F_{\mu\nu} c_{\nu i} + c_{i\mu}^* F_{\mu\nu} \frac{dc_{\nu i}}{dx} \right] \quad (\text{F.6})$$

Aside:

$$\begin{aligned} \sum_{\nu} F_{\mu\nu} c_{\nu i} &= \sum_{\nu} \varepsilon_i S_{\mu\nu} c_{\nu i} \\ \sum_{\mu} c_{i\mu}^* F_{\mu\nu} &= \sum_{\mu} c_{i\mu}^* \varepsilon_i S_{\mu\nu} \end{aligned}$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D + \sum_{\mu\nu} \sum_i \left[\frac{dc_{i\mu}^*}{dx} \varepsilon_i S_{\mu\nu} c_{\nu i} + c_{i\mu}^* \varepsilon_i S_{\mu\nu} \frac{dc_{\nu i}}{dx} \right] \quad (\text{F.7})$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D + \sum_i \varepsilon_i \sum_{\mu\nu} \left[\frac{dc_{i\mu}^*}{dx} S_{\mu\nu} c_{\nu i} + c_{i\mu}^* S_{\mu\nu} \frac{dc_{\nu i}}{dx} \right] \quad (\text{F.8})$$

Aside:

$$\begin{aligned} \sum_{\mu\nu} c_{i\mu}^* S_{\mu\nu} c_{\nu i} &= 1 \\ \sum_{\mu\nu} \left[\frac{dc_{i\mu}^*}{dx} S_{\mu\nu} c_{\nu i} + c_{i\mu}^* \frac{dS_{\mu\nu}}{dx} c_{\nu i} + c_{i\mu}^* S_{\mu\nu} \frac{dc_{\nu i}}{dx} \right] &= 0 \\ \sum_{\mu\nu} \left[\frac{dc_{i\mu}^*}{dx} S_{\mu\nu} c_{\nu i} + c_{i\mu}^* S_{\mu\nu} \frac{dc_{\nu i}}{dx} \right] &= - \sum_{\mu\nu} c_{i\mu}^* \frac{dS_{\mu\nu}}{dx} c_{\nu i} \end{aligned}$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D - \sum_i \varepsilon_i \sum_{\mu\nu} c_{i\mu}^* \frac{dS_{\mu\nu}}{dx} c_{\nu i} \quad (\text{F.9})$$

$$\frac{d\varepsilon_{HF}}{dx} = E_D - \sum_{\mu\nu} \underbrace{\sum_i \varepsilon_i c_{i\mu}^* c_{\nu i}}_{W_{\mu\nu}} \frac{dS_{\mu\nu}}{dx} \quad (\text{F.10})$$

$$\frac{d\varepsilon_{HF}}{dx} = \sum_{\mu\nu} P_{\mu\nu} \frac{dH_{\mu\nu}}{dx} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{d}{dx} \langle \mu\nu || \lambda\sigma \rangle + \frac{dV_{nuc}}{dx} - \sum_{\mu\nu} W_{\mu\nu} \frac{dS_{\mu\nu}}{dx} \quad (\text{F.11})$$

F.2 Ewald HF Derivative

Note: α and β are over QM atoms, i is over MM atoms, \aleph is over all (QM+MM) atoms, and ω_{xy} is an abbreviation for $\omega(\mathbf{r}_{xy})$.

F.2.1 w.r.t. QM Atom Position

$$\tilde{\epsilon}_{HF} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + V_{nuc} + \frac{1}{2} \sum_{\alpha\beta} Q_{\alpha} Q_{\beta} \omega_{\alpha\beta} + \sum_i \sum_{\alpha} q_i Q_{\alpha} \omega_{\alpha i} \quad (\text{F.12})$$

$$\begin{aligned} \tilde{\epsilon}_{HF} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + V_{nuc} + \frac{1}{2} \sum_{\substack{\alpha \neq \gamma \\ \beta}} Q_{\alpha} Q_{\beta} \omega_{\alpha\beta} + \frac{1}{2} \sum_{\beta} Q_{\gamma} Q_{\beta} \omega_{\gamma\beta} \\ + \sum_i \sum_{\alpha \neq \gamma} q_i Q_{\alpha} \omega_{\alpha i} + \sum_i q_i Q_{\gamma} \omega_{\gamma i} \quad (\text{F.13}) \end{aligned}$$

$$\begin{aligned} \tilde{\epsilon}_{HF} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + V_{nuc} + \frac{1}{2} \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} Q_{\alpha} Q_{\beta} \omega_{\alpha\beta} + \frac{1}{2} \sum_{\alpha \neq \gamma} Q_{\alpha} Q_{\gamma} \omega_{\alpha\gamma} \\ + \frac{1}{2} \sum_{\beta \neq \gamma} Q_{\gamma} Q_{\beta} \omega_{\gamma\beta} + \frac{1}{2} Q_{\gamma}^2 \omega_{\gamma\gamma} + \sum_i \sum_{\alpha \neq \gamma} q_i Q_{\alpha} \omega_{\alpha i} + \sum_i q_i Q_{\gamma} \omega_{\gamma i} \quad (\text{F.14}) \end{aligned}$$

$$\begin{aligned} \tilde{\epsilon}_{HF} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + V_{nuc} + \frac{1}{2} \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} Q_{\alpha} Q_{\beta} \omega_{\alpha\beta} + \sum_{\alpha \neq \gamma} Q_{\alpha} Q_{\gamma} \omega_{\alpha\gamma} \\ + \frac{1}{2} Q_{\gamma}^2 \omega_{\gamma\gamma} + \sum_i \sum_{\alpha \neq \gamma} q_i Q_{\alpha} \omega_{\alpha i} + \sum_i q_i Q_{\gamma} \omega_{\gamma i} \quad (\text{F.15}) \end{aligned}$$

$$\begin{aligned}
\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} &= \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} H_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_\gamma} + \sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \\
&\quad + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x_\gamma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{nuc}}{\partial x_\gamma} + \frac{1}{2} \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\beta \omega_{\alpha\beta} \\
&\quad + \frac{1}{2} \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} Q_\alpha \frac{\partial Q_\beta}{\partial x_\gamma} \omega_{\alpha\beta} + \sum_{\alpha \neq \gamma} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\gamma \omega_{\alpha\gamma} + \sum_{\alpha \neq \gamma} Q_\alpha \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\alpha\gamma} + \sum_{\alpha \neq \gamma} Q_\alpha Q_\gamma \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma} \\
&\quad + Q_\gamma \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\gamma\gamma} + \frac{1}{2} Q_\gamma^2 \frac{\partial \omega_{\gamma\gamma}}{\partial x_\gamma} + \sum_i \sum_{\alpha \neq \gamma} q_i \frac{\partial Q_\alpha}{\partial x_\gamma} \omega_{\alpha i} + \sum_i q_i \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\gamma i} \\
&\quad + \sum_i q_i Q_\gamma \frac{\partial \omega_{\gamma i}}{\partial x_\gamma}
\end{aligned} \tag{F.16}$$

$$\begin{aligned}
\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} &= \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} H_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_\gamma} + \sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \\
&\quad + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x_\gamma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{nuc}}{\partial x_\gamma} + \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\beta \omega_{\alpha\beta} \\
&\quad + \sum_{\alpha \neq \gamma} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\gamma \omega_{\alpha\gamma} + \sum_{\alpha} Q_\alpha \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\alpha\gamma} + \sum_{\alpha \neq \gamma} Q_\alpha Q_\gamma \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma} + \frac{1}{2} Q_\gamma^2 \frac{\partial \omega_{\gamma\gamma}}{\partial x_\gamma} \\
&\quad + \sum_i \sum_{\alpha} q_i \frac{\partial Q_\alpha}{\partial x_\gamma} \omega_{\alpha i} + \sum_i q_i Q_\gamma \frac{\partial \omega_{\gamma i}}{\partial x_\gamma}
\end{aligned} \tag{F.17}$$

$$\begin{aligned}
\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} &= \underbrace{\sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_\gamma} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x_\gamma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{nuc}}{\partial x_\gamma} + \sum_{\alpha \neq \gamma} Q_\alpha Q_\gamma \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma}}_{E_D} \\
&\quad + \frac{1}{2} Q_\gamma^2 \frac{\partial \omega_{\gamma\gamma}}{\partial x_\gamma} + \sum_i q_i Q_\gamma \frac{\partial \omega_{\gamma i}}{\partial x_\gamma} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} H_{\mu\nu} + \sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \\
&\quad + \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\beta \omega_{\alpha\beta} + \sum_{\alpha \neq \gamma} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\gamma \omega_{\alpha\gamma} + \sum_\alpha Q_\alpha \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\alpha\gamma} + \sum_i \sum_\alpha q_i \frac{\partial Q_\alpha}{\partial x_\gamma} \omega_{\alpha i}
\end{aligned} \tag{F.18}$$

$$\begin{aligned}
\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} &= E_D + \sum_{\alpha \neq \gamma} Q_\alpha Q_\gamma \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma} + \frac{1}{2} Q_\gamma^2 \frac{\partial \omega_{\gamma\gamma}}{\partial x_\gamma} + \sum_i q_i Q_\gamma \frac{\partial \omega_{\gamma i}}{\partial x_\gamma} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} \\
&\quad + \sum_{\substack{\alpha \neq \gamma \\ \beta \neq \gamma}} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\beta \omega_{\alpha\beta} + \sum_{\alpha \neq \gamma} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\gamma \omega_{\alpha\gamma} + \sum_\alpha Q_\alpha \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\alpha\gamma} + \sum_i \sum_\alpha q_i \frac{\partial Q_\alpha}{\partial x_\gamma} \omega_{\alpha i}
\end{aligned} \tag{F.19}$$

$$\begin{aligned}
\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} &= E_D + \sum_\alpha Q_\alpha Q_\gamma \frac{\partial \omega_{\alpha\gamma}}{\partial x_\gamma} \left(1 - \frac{\delta_{\alpha\gamma}}{2}\right) + \sum_i q_i Q_\gamma \frac{\partial \omega_{\gamma i}}{\partial x_\gamma} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} \\
&\quad + \sum_{\substack{\alpha \neq \gamma \\ \beta}} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\beta \omega_{\alpha\beta} + \sum_\alpha Q_\alpha \frac{\partial Q_\gamma}{\partial x_\gamma} \omega_{\alpha\gamma} + \sum_i \sum_\alpha q_i \frac{\partial Q_\alpha}{\partial x_\gamma} \omega_{\alpha i}
\end{aligned} \tag{F.20}$$

$$\begin{aligned}
\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} &= \underbrace{E_D + \sum_{\aleph} Q_\aleph Q_\gamma \frac{\partial \omega_{\aleph\gamma}}{\partial x_\gamma} \left(1 - \frac{\delta_{\aleph\gamma}}{2}\right)}_{\tilde{E}_D} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} + \sum_{\alpha\beta} \frac{\partial Q_\alpha}{\partial x_\gamma} Q_\beta \omega_{\alpha\beta} \\
&\quad + \sum_i \sum_\alpha q_i \frac{\partial Q_\alpha}{\partial x_\gamma} \omega_{\alpha i}
\end{aligned} \tag{F.21}$$

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} + \sum_\alpha \frac{\partial Q_\alpha}{\partial x_\gamma} \sum_{\aleph} Q_\aleph \omega_{\alpha\aleph} \tag{F.22}$$

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} + \sum_\alpha \frac{\partial Q_\alpha}{\partial x_\gamma} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{F.23})$$

Note: $\frac{df(y(x),z(x))}{dx} = \frac{\partial f}{\partial y} \frac{\partial y}{\partial x} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial x}$

$\therefore \frac{df(y(x),z(x),\dots)}{dx} = \sum_M \frac{\partial f}{\partial M} \frac{\partial M}{\partial x}$, where M are all functions of f that depend on x .

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \frac{\partial E^{PI}}{\partial Q_\alpha} + \sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{F.24})$$

Note: In this derivation, M will be all variables (and the appropriate sums) needed to take a complete derivative of Q_α excluding $P_{\mu\nu}$. See section H.5 for what these particular derivatives are.

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} F_{\mu\nu} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \Delta F_{\mu\nu}^{PI} + \sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{F.25})$$

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \tilde{F}_{\mu\nu} + \sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{F.26})$$

Continue from equation F.5 of section F.1 to finish the derivation to achieve the result:

$$\begin{aligned} \frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_\gamma} = & \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_\gamma} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial}{\partial x_\gamma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{nuc}}{\partial x_\gamma} - \sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} \\ & + \sum_{\aleph} Q_\aleph Q_\gamma \frac{\partial \omega_{\aleph\gamma}}{\partial x_\gamma} \left(1 - \frac{\delta_{\aleph\gamma}}{2} \right) + \sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} \end{aligned} \quad (\text{F.27})$$

F.2.2 w.r.t. MM Atom Position

$$\tilde{\varepsilon}_{HF} = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + V_{nuc} + \frac{1}{2} \sum_{\alpha\beta} Q_{\alpha} Q_{\beta} \omega_{\alpha\beta} + \sum_i \sum_{\alpha} q_i Q_{\alpha} \omega_{\alpha i} \quad (\text{F.28})$$

$$\begin{aligned} \frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} &= \sum_{\mu\nu} \left(\frac{\partial P_{\mu\nu}}{\partial x_j} H_{\mu\nu} + P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} \right) \\ &+ \frac{1}{2} \sum_{\mu\nu\lambda\sigma} \left(2 \frac{\partial P_{\mu\nu}}{\partial x_j} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + P_{\mu\nu} P_{\lambda\sigma} \underbrace{\frac{\partial \langle \mu\nu || \lambda\sigma \rangle}{\partial x_j}}_{=0} \right) + \frac{\partial V_{nuc}}{\partial x_j} \\ &+ \frac{1}{2} \sum_{\alpha\beta} \left(2 \frac{\partial Q_{\alpha}}{\partial x_j} Q_{\beta} \omega_{\alpha\beta} + Q_{\alpha} Q_{\beta} \underbrace{\frac{\partial \omega_{\alpha\beta}}{\partial x_j}}_{=0} \right) \\ &+ \sum_i \sum_{\alpha} \left(\underbrace{\frac{\partial q_i}{\partial x_j}}_{=0} Q_{\alpha} \omega_{\alpha i} + q_i \frac{\partial Q_{\alpha}}{\partial x_j} \omega_{\alpha i} + q_i Q_{\alpha} \frac{\partial \omega_{\alpha i}}{\partial x_j} \delta_{ij} \right) \end{aligned} \quad (\text{F.29})$$

$$\begin{aligned} \frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} &= \sum_{\mu\nu} \left(\frac{\partial P_{\mu\nu}}{\partial x_j} H_{\mu\nu} + P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} \right) + \sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x_j} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{nuc}}{\partial x_j} \\ &+ \sum_{\alpha\beta} \frac{\partial Q_{\alpha}}{\partial x_j} Q_{\beta} \omega_{\alpha\beta} + \sum_i \sum_{\alpha} q_i \frac{\partial Q_{\alpha}}{\partial x_j} \omega_{\alpha i} + \sum_{\alpha} q_j Q_{\alpha} \frac{\partial \omega_{\alpha j}}{\partial x_j} \end{aligned} \quad (\text{F.30})$$

$$\begin{aligned} \frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} &= \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} H_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} \sum_{\lambda\sigma} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + \frac{\partial V_{nuc}}{\partial x_j} \\ &+ \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial x_j} \sum_{\beta} Q_{\beta} \omega_{\alpha\beta} + \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial x_j} \sum_i q_i \omega_{\alpha i} + q_j \sum_{\alpha} Q_{\alpha} \frac{\partial \omega_{\alpha j}}{\partial x_j} \end{aligned} \quad (\text{F.31})$$

$$\begin{aligned} \frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = & \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} \left(H_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \right) + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} + \frac{\partial V_{nuc}}{\partial x_j} \\ & + \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial x_j} \sum_{\mathbb{N}} Q_{\beta\omega\alpha\mathbb{N}} + q_j \sum_{\alpha} Q_{\alpha} \frac{\partial \omega_{\alpha j}}{\partial x_j} \quad (\text{F.32}) \end{aligned}$$

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} F_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} + \frac{\partial V_{nuc}}{\partial x_j} + \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial x_j} \frac{\partial E^{PI}}{\partial Q_{\alpha}} + q_j \sum_{\alpha} Q_{\alpha} \frac{\partial \omega_{\alpha j}}{\partial x_j} \quad (\text{F.33})$$

$$\begin{aligned} \frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = & \underbrace{\sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} + \frac{\partial V_{nuc}}{\partial x_j} + q_j \sum_{\alpha} Q_{\alpha} \frac{\partial \omega_{\alpha j}}{\partial x_j}}_{\tilde{E}_D} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} F_{\mu\nu} + \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial x_j} \frac{\partial E^{PI}}{\partial Q_{\alpha}} \\ & + \sum_M \frac{\partial M}{\partial x_j} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial E^{PI}}{\partial Q_{\alpha}} \quad (\text{F.34}) \end{aligned}$$

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} F_{\mu\nu} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \frac{\partial E^{PI}}{\partial Q_{\alpha}} + \sum_M \frac{\partial M}{\partial x_j} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial E^{PI}}{\partial Q_{\alpha}} \quad (\text{F.35})$$

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_j} \tilde{F}_{\mu\nu} + \sum_M \frac{\partial M}{\partial x_j} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial E^{PI}}{\partial Q_{\alpha}} \quad (\text{F.36})$$

Using the second aside in section F.1:

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = \tilde{E}_D + \sum_{\mu\nu} \frac{\partial \tilde{S}_{\mu\nu}}{\partial x_j} \tilde{W}_{\mu\nu} + \sum_M \frac{\partial M}{\partial x_j} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial E^{PI}}{\partial Q_{\alpha}} \quad (\text{F.37})$$

But $\frac{\partial \tilde{S}_{\mu\nu}}{\partial x_j} = 0$ so

$$\frac{\partial \tilde{\varepsilon}_{HF}}{\partial x_j} = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_j} + \frac{\partial V_{nuc}}{\partial x_j} + q_j \sum_{\alpha} Q_{\alpha} \frac{\partial \omega_{\alpha j}}{\partial x_j} + \sum_M \frac{\partial M}{\partial x_j} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial M} \frac{\partial E^{PI}}{\partial Q_{\alpha}} \quad (\text{F.38})$$

F.2.3 Derivative of the Pair Potential

$$\omega(\mathbf{r}_{\gamma\aleph}) = \underbrace{\left[\sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2\mathbf{a}_m^2}{\eta^2}} \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) \right]}_{recip} + \underbrace{\left[\sum_{\mathbf{b}_n \neq \mathbf{0}} \frac{\text{erfc}(\eta|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \right]}_{realrest} - \underbrace{\left[\frac{\text{erf}(\eta|\mathbf{r}_{\gamma\aleph}|)}{|\mathbf{r}_{\gamma\aleph}|} \right]}_{real0} \quad (\text{F.39})$$

$$\nabla_{\gamma} \omega_{recip}(\mathbf{r}_{\gamma\aleph}) = \nabla_{\gamma} \left[\sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2\mathbf{a}_m^2}{\eta^2}} \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) \right] \quad (\text{F.40})$$

$$\nabla_{\gamma} \omega_{recip}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2\mathbf{a}_m^2}{\eta^2}} \nabla_{\gamma} \cos(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) \quad (\text{F.41})$$

$$\nabla_{\gamma} \omega_{recip}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2\mathbf{a}_m^2}{\eta^2}} [-\sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) \nabla_{\gamma} (2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph})] \quad (\text{F.42})$$

$$\nabla_{\gamma} \omega_{recip}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{a}_m \neq \mathbf{0}} \frac{1}{V\pi\mathbf{a}_m^2} e^{-\frac{\pi^2\mathbf{a}_m^2}{\eta^2}} [-\sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) (2\pi\mathbf{a}_m)] \quad (\text{F.43})$$

$$\nabla_{\gamma} \omega_{recip}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{a}_m \neq \mathbf{0}} -2\pi\omega(\mathbf{a}_m) \sin(2\pi\mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) \mathbf{a}_m \quad (\text{F.44})$$

$$\nabla_{\gamma} \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = -\nabla_{\gamma} \left[\frac{\text{erf}(\eta|\mathbf{r}_{\gamma\aleph}|)}{|\mathbf{r}_{\gamma\aleph}|} \right] \quad (\text{F.45})$$

$$\nabla_\gamma \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = - \left[\operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|) \nabla_\gamma \frac{1}{|\mathbf{r}_{\gamma\aleph}|} + \frac{1}{|\mathbf{r}_{\gamma\aleph}|} \nabla_\gamma \operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|) \right] \quad (\text{F.46})$$

$$\nabla_\gamma \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = - \left[\operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|) \frac{-\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^3} + \frac{1}{|\mathbf{r}_{\gamma\aleph}|} \nabla_\gamma \frac{2}{\sqrt{\pi}} \int_0^{\eta |\mathbf{r}_{\gamma\aleph}|} e^{-x^2} dx \right] \quad (\text{F.47})$$

$$\nabla_\gamma \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = \operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|) \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^3} - \frac{1}{|\mathbf{r}_{\gamma\aleph}|} \frac{2}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph}|^2} \nabla_\gamma \eta |\mathbf{r}_{\gamma\aleph}| \quad (\text{F.48})$$

$$\nabla_\gamma \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = \operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|) \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^3} - \frac{1}{|\mathbf{r}_{\gamma\aleph}|} \frac{2}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph}|^2} \eta \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|} \quad (\text{F.49})$$

$$\nabla_\gamma \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = \operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|) \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^3} - \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph}|^2} \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^2} \quad (\text{F.50})$$

$$\nabla_\gamma \omega_{real0}(\mathbf{r}_{\gamma\aleph}) = \left[\frac{\operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|)}{|\mathbf{r}_{\gamma\aleph}|} - \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph}|^2} \right] \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^2} \quad (\text{F.51})$$

$$\nabla_\gamma \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = \nabla_\gamma \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\frac{\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \right] \quad (\text{F.52})$$

$$\begin{aligned} \nabla_\gamma \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|) \nabla_\gamma \frac{1}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \right. \\ \left. + \frac{1}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \nabla_\gamma \operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|) \right] \quad (\text{F.53}) \end{aligned}$$

$$\nabla_{\gamma} \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|) \frac{-(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^3} + \frac{1}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \nabla_{\gamma} \frac{2}{\sqrt{\pi}} \int_{\eta|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|}^{\infty} e^{-x^2} dx \right] \quad (\text{F.54})$$

$$\nabla_{\gamma} \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|) \frac{-(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^3} + \frac{1}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \left[-\frac{2}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \nabla_{\gamma} \eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n| \right] \right] \quad (\text{F.55})$$

$$\nabla_{\gamma} \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = - \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|) \frac{(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^3} + \frac{1}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \frac{2}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \frac{\eta (\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} \right] \quad (\text{F.56})$$

$$\nabla_{\gamma} \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = - \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|) \frac{(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^3} + \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \frac{(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \right] \quad (\text{F.57})$$

$$\nabla_{\gamma} \omega_{realrest}(\mathbf{r}_{\gamma\aleph}) = - \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\left[\frac{\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} + \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \right] \frac{(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \right] \quad (\text{F.58})$$

$$\begin{aligned} \nabla_{\gamma} \omega(\mathbf{r}_{\gamma\aleph}) = & \left[\frac{\operatorname{erf}(\eta |\mathbf{r}_{\gamma\aleph}|)}{|\mathbf{r}_{\gamma\aleph}|} - \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph}|^2} \right] \frac{\mathbf{r}_{\gamma\aleph}}{|\mathbf{r}_{\gamma\aleph}|^2} - \sum_{\mathbf{a}_m \neq \mathbf{0}} 2\pi w(\mathbf{a}_m) \sin(2\pi \mathbf{a}_m \cdot \mathbf{r}_{\gamma\aleph}) \mathbf{a}_m \\ & - \sum_{\mathbf{b}_n \neq \mathbf{0}} \left[\left[\frac{\operatorname{erfc}(\eta |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|} + \frac{2\eta}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \right] \frac{(\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n)}{|\mathbf{r}_{\gamma\aleph} + \mathbf{b}_n|^2} \right] \end{aligned} \quad (\text{F.59})$$

F.3 Logistics in Q-Chem

F.3.1 Outputs and Relationships

Q-Chem Output			
Energy		Derivative	
QM			
SCF	E_{SCF}	SCF	G_{SCF}
Kinetic	E_T	Overlap and Kinetic	G_{TS}
Nuclear Attraction	E_V	Nuclear Attraction	G_V
Coulomb	E_J	Coulomb and Exchange	G_{2e}
Exchange	$E_{K^\alpha}, E_{K^\beta}$		
Nuclear Repulsion	E_{NN}	Nuclear-Nuclear	G_{NN}
QM/MM			
QM/MM Energy (E_qmmm)	E_{tot}	total grad after...	G_{tot}
		External Charge (E)	G_{QE}
		External Charge (N)	G_{QN}
MM Energy (Etot)	E_{MM}	MM energy gradient	G_{vdW}
QM/MM Ewald			
Ewald QM/MM	E_{ewald}	Ewald Potential	G_ω
		Ewald Charge	G_M

$$E_{1e} = E_T + E_V (+E_Z) \quad (\text{F.60})$$

$$E_{2e} = E_J + E_{K^\alpha} + E_{K^\beta} \quad (\text{F.61})$$

$$E_{SCF} = E_{1e} + E_{2e} + E_{NN} (+E_{ewald}) \quad (\text{F.62})$$

$$E_{tot} = E_{SCF} + E_{MM} \quad (\text{F.63})$$

$$G_Q = G_{QE} + G_{QN} \quad (\text{F.64})$$

$$G_{SCF} = G_{TS} + G_{2e} + G_V + G_{NN} (+G_Q + G_\omega + G_M) \quad (\text{F.65})$$

$$G_{tot} = G_{SCF} + G_{vdW} \quad (\text{F.66})$$

F.3.2 QM Derivatives

$$E_{SCF} = \underbrace{\sum_{\mu\nu} P_{\mu\nu} T_{\mu\nu}}_{E_T} + \underbrace{\sum_{\mu\nu} P_{\mu\nu} V_{\mu\nu}}_{E_V} + \underbrace{\frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle}_{E_{2e} = \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} (J_{\mu\nu} + K_{\mu\nu})} + \underbrace{V_{NN}}_{E_{NN}} \quad (\text{F.67})$$

$$\frac{\Delta E_T}{\Delta x} = \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} T_{\mu\nu} + \sum_{\mu\nu} P_{\mu\nu} \frac{\partial T_{\mu\nu}}{\partial x} \quad (\text{F.68})$$

$$\frac{\Delta E_V}{\Delta x} = \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} V_{\mu\nu} + \underbrace{\sum_{\mu\nu} P_{\mu\nu} \frac{\partial V_{\mu\nu}}{\partial x}}_{G_V^x} \quad (\text{F.69})$$

$$\frac{\Delta E_{2e}}{\Delta x} = \sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle + \underbrace{\frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \frac{\partial \langle \mu\nu || \lambda\sigma \rangle}{\partial x}}_{G_{2e}^x} \quad (\text{F.70})$$

$$\frac{\Delta E_{NN}}{\Delta x} = \underbrace{\frac{\partial V_{NN}}{\partial x}}_{G_{NN}^x} \quad (\text{F.71})$$

$$G_{TS}^x = \sum_{\mu\nu} P_{\mu\nu} \frac{\partial T_{\mu\nu}}{\partial x} - \sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x} \quad (\text{F.72})$$

$$-\sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x} = \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} T_{\mu\nu} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} V_{\mu\nu} + \sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle \quad (\text{F.73})$$

$$G_{TS}^x = \underbrace{\sum_{\mu\nu} P_{\mu\nu} \frac{\partial T_{\mu\nu}}{\partial x}}_{\frac{\Delta E_T}{\Delta x}} + \underbrace{\sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} T_{\mu\nu}}_{\frac{\Delta E_V}{\Delta x} - G_V^x} + \underbrace{\sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} V_{\mu\nu}}_{\frac{\Delta E_{2e}}{\Delta x} - G_{2e}^x} + \underbrace{\sum_{\mu\nu\lambda\sigma} \frac{\partial P_{\mu\nu}}{\partial x} P_{\lambda\sigma} \langle \mu\nu || \lambda\sigma \rangle}_{\frac{\Delta E_{2e}}{\Delta x} - G_{2e}^x} \quad (\text{F.74})$$

F.3.3 QM/MM Derivatives

$$E_{SCF} += \underbrace{\sum_{\mu\nu} P_{\mu\nu} Z_{\mu\nu}}_{E_Z} + Z_{NN} \quad (\text{F.75})$$

Note: E_T in the output is calculated by taking $E_{1e} - E_V$. This leads to an incorrect value for E_T in QM/MM jobs because $E_{1e} - E_V = E_T + E_Z$ because $H_{\mu\nu} = T_{\mu\nu} + V_{\mu\nu} + Z_{\mu\nu}$. In order to get the proper E_T use $\text{Tr}(\mathbf{TP})$.

$$\frac{\Delta Z_{NN}}{\Delta x} = G_{QN}^x \quad (\text{F.76})$$

$$\frac{\Delta E_Z}{\Delta x} = \underbrace{\sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} Z_{\mu\nu}}_{\text{in } G_{TS}^x} + \underbrace{\sum_{\mu\nu} P_{\mu\nu} \frac{\partial Z_{\mu\nu}}{\partial x}}_{G_{QE}^x} \quad (\text{F.77})$$

F.3.4 QM/MM Ewald Derivatives

$$E_{SCF} += \underbrace{\sum_{\alpha} \left(\frac{1}{2} \sum_{\beta} Q_{\beta} \omega_{\alpha\beta} + \sum_j q_j \omega_{\alpha j} \right)}_{E_{ewald}} \quad (\text{F.78})$$

$$\begin{aligned} \frac{\Delta E_{ewald}}{\Delta x} = & \underbrace{\sum_{\alpha} Q_{\alpha} \sum_{\aleph} Q_{\aleph} \frac{\partial \omega_{\alpha\aleph}}{\partial x}}_{G_{\omega}^x} + \underbrace{\sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial P_{\mu\nu}} \sum_{\aleph} Q_{\aleph} \omega_{\alpha\aleph}}_{\text{in } G_{TS}^x} \\ & + \underbrace{\sum_M \frac{\partial M}{\partial x} \sum_{\alpha} \frac{\partial Q_{\alpha}}{\partial M} \sum_{\aleph} Q_{\aleph} \omega_{\alpha\aleph}}_{G_M^x} \quad (\text{F.79}) \end{aligned}$$

F.3.5 Scratch Files

Q-Chem Scratch Files		
Variable	File #	Notes
F	58	Stored unvectorized
J	55	Not normally stored
K	56	Not normally stored
S	320	Not normally stored
S^x	23	Stored vectorized
F^x	24	
P	54	Stored unvectorized
H_{1e}	51	Stored vectorized
V	57	Temporary file number
W	111	Temporary file number
T	100	Temporary file number
\vec{c} and $\vec{\varepsilon}$	53	See Note

Note: File 53.0 (FILE_MO_COEFS) is stored in the following way (variable followed by number of entries): $c_{i\mu}^{\alpha}$ ($N_{basis} \times N_{orbs}$), $c_{i\mu}^{\beta}$ ($N_{basis} \times N_{orbs}$), ε_i^{α} (N_{orbs}), ε_i^{β} (N_{orbs})

Entries for FILE_ENERGY (99.0), scfman.C lines 4230-4263		
Entry #	Variable	Name of Energy
0	E_{SCF}	SCF (see note)
1	E_{1e}	One-Electron
2	E_J	Total Coulomb
3	$E_{K\alpha}$	Alpha Exchange
4	$E_{K\beta}$	Beta Exchange
5	E_X	DFT Exchange
6	E_C	DFT Correlation
7	E_T	Kinetic
8	E_{NN}	Nuclear Repu.
9	E_V	Nuclear Attr.
11	E_{tot}	Total (with MM energy, if included)
25	E_{MM}	MM
26	E_{ewald}	Ewald QM/MM

F.3.6 Outline of Code for Computing $\frac{\Delta E}{\Delta P_{\mu\nu}}$ Directly

- Define and zero variables
- Call GetH
- Unvectorize jHv
- Load P from file
- Modify so that you have jPv (vectorized total density), jPA and jPB (unvectorized alpha & beta densities)
- Call MakeJK
- Unvectorize jJv
- for μ

– for ν

- * Load P from file (need unvectorized)
- * Add/Subtract $\delta P_{\mu\nu}$
- * Compute Ewald energy (use new density for charges)
- * Take trace of density matrix with jH, jJ, & jK
- * Add and store energies in the appropriate energy vector

- Take the negative of the negative energy vector
- Add the positive and negative energy vectors
- Scale by $\frac{1}{4 \times \text{stepsize}}$ (4 if you added $\delta P_{\mu\nu}$ to jPA then doubled to get jP)

APPENDIX G

Charge Schemes

G.1 Mulliken

$$Q_\alpha = Z_\alpha - \text{Tr}_\alpha(\mathbf{PS}) \quad (\text{G.1})$$

$$Q_\alpha = Z_\alpha - \sum_{\zeta \in \alpha} (\mathbf{PS})_{\zeta\zeta} \quad (\text{G.2})$$

$$Q_\alpha = Z_\alpha - \sum_{\substack{\zeta \in \alpha \\ \xi}} P_{\zeta\xi} S_{\xi\zeta} \quad (\text{G.3})$$

Note: $S_{\xi\zeta}$ is Hermitian so $S_{\xi\zeta} = S_{\zeta\xi}$

$$Q_\alpha = Z_\alpha - \sum_{\substack{\zeta \in \alpha \\ \xi}}^{N_{basis}} P_{\zeta\xi} S_{\zeta\xi} \quad (\text{G.4})$$

G.2 QM Potential at a Grid Point

$$V(\mathbf{r}_k) = \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_k|} d\mathbf{r} \quad (\text{G.5})$$

$$\rho(\mathbf{r}) = \sum_{\mu\nu} \sum_i c_{\mu i}^* \varphi_\mu^*(\mathbf{r}) c_{\nu i} \varphi_\nu(\mathbf{r}) \quad (\text{G.6})$$

$$\rho(\mathbf{r}) = \sum_{\mu\nu} \varphi_{\mu}^*(\mathbf{r}) \varphi_{\nu}(\mathbf{r}) P_{\mu\nu} \quad (\text{G.7})$$

$$V(\mathbf{r}_k) = \sum_{\mu\nu} \int \frac{\varphi_{\mu}^*(\mathbf{r}) \varphi_{\nu}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_k|} P_{\mu\nu} d\mathbf{r} \quad (\text{G.8})$$

$$V(\mathbf{r}_k) = \sum_{\mu\nu} \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle_{\mathbf{r}} P_{\mu\nu} \quad (\text{G.9})$$

$$V(\mathbf{r}_k) = \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \quad (\text{G.10})$$

G.3 ChEIP-G

Let: N_g = the number of grid points, N_a = the number of atoms; N_b = the number of basis functions; w_k = a weighting function; Q = charge of system

$$\text{Note: } (\mathbb{I}_k)_{\mu\nu} = \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle; \phi_k = \sum_J \frac{q_J}{|\mathbf{r}_k - \mathbf{r}_J|}; \Phi_k = \sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu}$$

G.3.1 Charges

ChEIPG charges are the charges that minimize the difference, in a least squares fashion, between a static potential and the potential constructed from the density matrix subject to the constraint that the charges sum to the total system charge.

$$\mathcal{L} = \sum_k^{N_g} w_k (\Phi_k - \phi_k)^2 - \lambda \left(Q - \sum_J^{N_a} q_J \right) \quad (\text{G.11})$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 = Q - \sum_J^{N_a} q_J \therefore Q = \sum_J^{N_a} q_J \quad (\text{G.12})$$

$$\frac{\partial \mathcal{L}}{\partial q_A} = 0 = \sum_k^{N_g} 2w_k (\Phi_k - \phi_k) \left(-\frac{\partial \phi_k}{\partial q_A} \right) + \lambda \quad (\text{G.13})$$

Note: $\frac{\partial \phi_k}{\partial q_A} = \frac{1}{|\mathbf{r}_k - \mathbf{r}_A|}$

$$\lambda = \sum_k 2w_k (\Phi_k - \phi_k) \frac{1}{|\mathbf{r}_k - \mathbf{r}_A|} \quad (\text{G.14})$$

Let: $e_A = \sum_k \frac{w_k \Phi_k}{|\mathbf{r}_k - \mathbf{r}_A|}$

$$\lambda = 2e_A - 2 \sum_k \frac{w_k \phi_k}{|\mathbf{r}_k - \mathbf{r}_A|} \quad (\text{G.15})$$

$$\lambda = 2e_A - 2 \sum_k \sum_J \frac{w_k q_J}{|\mathbf{r}_k - \mathbf{r}_A| |\mathbf{r}_k - \mathbf{r}_J|} \quad (\text{G.16})$$

$$\lambda = 2e_A - 2 \sum_J q_J \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_A| |\mathbf{r}_k - \mathbf{r}_J|} \quad (\text{G.17})$$

Let: $G_{JA} = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_A| |\mathbf{r}_k - \mathbf{r}_J|}$

$$\lambda = 2e_A - 2 \sum_J q_J G_{JA} \quad (\text{G.18})$$

$$\lambda = 2e_A - 2(\mathbf{Gq})_A \quad (\text{G.19})$$

$$e_A = (\mathbf{Gq})_A + \frac{\lambda}{2} \quad (\text{G.20})$$

Begin Sidenote

Let: $N = N_a$

$$\mathbf{e}' = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \\ Q \end{bmatrix} \quad \mathbf{q}' = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \\ \lambda \end{bmatrix}$$

$$\mathbf{G}' = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1N} & \frac{1}{2} \\ G_{21} & G_{22} & \cdots & G_{2N} & \frac{1}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}$$

$$\therefore \mathbf{e}' = \mathbf{G}'\mathbf{q}'$$

End Sidenote

$$e_A = (\mathbf{Gq})_A + \frac{\lambda}{2} \quad (\text{G.21})$$

Construct N_a eqns Note: $\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ that is an $N_a \times 1$ vector

$$\mathbf{e} = \mathbf{G}\mathbf{q} + \frac{\lambda}{2} \mathbf{1} \quad (\text{G.22})$$

$$\mathbf{G}\mathbf{q} = \mathbf{e} - \frac{\lambda}{2} \mathbf{1} \quad (\text{G.23})$$

$$\mathbf{q} = \mathbf{G}^{-1} \left(\mathbf{e} - \frac{\lambda}{2} \mathbf{1} \right) \quad (\text{G.24})$$

$$q_A = (\mathbf{G}^{-1}\mathbf{e})_A - \frac{\lambda}{2} (\mathbf{G}^{-1}\mathbf{1})_A \quad (\text{G.25})$$

$$q_A = \sum_B^{N_a} e_B (\mathbf{G}^{-1})_{BA} - \frac{\lambda}{2} \sum_B^{N_a} (\mathbf{G}^{-1})_{BA} \quad (\text{G.26})$$

$$Q = \sum_C^{N_a} q_C \quad (\text{G.27})$$

$$Q = \sum_{BC}^{N_a} e_B (\mathbf{G}^{-1})_{BC} - \frac{\lambda}{2} \sum_{CB}^{N_a} (\mathbf{G}^{-1})_{BC} \quad (\text{G.28})$$

$$\frac{\lambda}{2} \sum_{BC}^{N_a} (\mathbf{G}^{-1})_{BC} = \sum_{BC}^{N_a} e_B (\mathbf{G}^{-1})_{BC} - Q \quad (\text{G.29})$$

$$\lambda = \frac{2 \left(\sum_{BC}^{N_a} e_B (\mathbf{G}^{-1})_{BC} - Q \right)}{\sum_{BC}^{N_a} (\mathbf{G}^{-1})_{BC}} \quad (\text{G.30})$$

$$q_A = \sum_B^{N_a} e_B (\mathbf{G}^{-1})_{BA} - \frac{\left(\sum_{BC}^{N_a} e_B (\mathbf{G}^{-1})_{BC} - Q \right)}{\sum_{BC}^{N_a} (\mathbf{G}^{-1})_{BC}} \sum_B^{N_a} (\mathbf{G}^{-1})_{BA} \quad (\text{G.31})$$

APPENDIX H

Charge Derivatives

H.1 Density Derivative for Mulliken Charges

H.1.1 Non-Symmetric Form

$$Q_\alpha = Z_\alpha - \sum_{\substack{\zeta \in \alpha \\ \xi}}^{N_{basis}} P_{\zeta\xi} S_{\zeta\xi} \quad (\text{H.1})$$

$$\frac{\partial Q_\alpha}{\partial P_{\mu\nu}} = \frac{\partial}{\partial P_{\mu\nu}} \left(- \sum_{\substack{\zeta \in \alpha \\ \xi}}^{N_{basis}} P_{\zeta\xi} S_{\zeta\xi} \right) \quad (\text{H.2})$$

$$\frac{\partial Q_\alpha}{\partial P_{\mu\nu}} = -S_{\mu\nu} \delta_{\alpha\in\mu} \quad (\text{H.3})$$

H.1.2 Symmetric Form

$$Q_\alpha = Z_\alpha - \sum_{\substack{\zeta \in \alpha \\ \xi}}^{N_{basis}} P_{\zeta\xi} S_{\zeta\xi} \quad (\text{H.4})$$

Note: $X_{yz} = \frac{1}{2} (X_{yz} + X_{zy})$ iff \mathbf{X} is Hermitian.

$$Q_\alpha = Z_\alpha - \frac{1}{2} \left(\sum_{\substack{\zeta \in \alpha \\ \xi}}^{N_{basis}} P_{\zeta\xi} S_{\zeta\xi} + \sum_{\substack{\xi \in \alpha \\ \zeta}}^{N_{basis}} P_{\xi\zeta} S_{\xi\zeta} \right) \quad (\text{H.5})$$

$$\frac{\partial Q_\alpha}{\partial P_{\mu\nu}} = \frac{\partial}{\partial P_{\mu\nu}} \left[-\frac{1}{2} \left(\sum_{\substack{\zeta \in \alpha \\ \xi}}^{N_{basis}} P_{\zeta\xi} S_{\zeta\xi} + \sum_{\substack{\xi \in \alpha \\ \zeta}}^{N_{basis}} P_{\xi\zeta} S_{\xi\zeta} \right) \right] \quad (\text{H.6})$$

$$\frac{\partial Q_\alpha}{\partial P_{\mu\nu}} = -\frac{1}{2} (S_{\mu\nu} \delta_{\alpha\in\mu} + S_{\nu\mu} \delta_{\alpha\in\nu}) \quad (\text{H.7})$$

The symmetric form must be used in Q-Chem because the diagonalization of the Fock matrix (which occurs in another step) is dependent upon the Fock matrix being Hermitian.

H.2 Density Derivative for ChEIP-G Charges

$$\text{Note: } \frac{\partial(\mathbf{G}^{-1})_{BA}}{\partial P_{\zeta\xi}} = \sum_k \frac{\partial Q}{\partial P_{\zeta\xi}} = 0$$

$$\frac{\partial e_B}{\partial P_{\zeta\xi}} = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_B|} \frac{\partial \Phi_k}{\partial P_{\zeta\xi}} \quad (\text{H.8})$$

$$\frac{\partial e_B}{\partial P_{\zeta\xi}} = -(\Xi_B)_{\zeta\xi} = \sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_B|} (\mathbb{I}_k)_{\zeta\xi} \quad (\text{H.9})$$

$$\frac{\partial q_A}{\partial P_{\zeta\xi}} = (\Lambda_B)_{\zeta\xi} = -\sum_B^{N_a} (\Xi_B)_{\zeta\xi} (\mathbf{G}^{-1})_{BA} + \frac{\sum_{AB}^{N_a} (\Xi_B)_{\zeta\xi} (\mathbf{G}^{-1})_{BA}}{\sum_{AB}^{N_a} (\mathbf{G}^{-1})_{BA}} \sum_B^{N_a} (\mathbf{G}^{-1})_{BA} \quad (\text{H.10})$$

H.2.1 Simplification of Derivative

$$(\Lambda_M)_{\mu\nu} = -\sum_L (\mathbf{G}^{-1})_{ML} (\Xi_L)_{\mu\nu} + \frac{\sum_J (\mathbf{G}^{-1})_{MJ}}{\sum_{\mathbb{I}J} (\mathbf{G}^{-1})_{\mathbb{I}J}} \sum_{LN} (\mathbf{G}^{-1})_{LN} (\Xi_N)_{\mu\nu} \quad (\text{H.11})$$

$$\begin{aligned}
(\Lambda_M)_{\mu\nu} = & - \sum_L (\mathbf{G}^{-1})_{ML} \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_L|} (\mathbb{I}_k)_{\mu\nu} \\
& + \frac{\sum_J (\mathbf{G}^{-1})_{MJ}}{\sum_{\mathbb{I}J} (\mathbf{G}^{-1})_{\mathbb{I}J}} \sum_{LN} (\mathbf{G}^{-1})_{LN} \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_N|} (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.12})
\end{aligned}$$

$$\text{Let } \Gamma = \sum_{\mathbb{I}J} (\mathbf{G}^{-1})_{\mathbb{I}J} \text{ and } a_{kLN} = (\mathbf{G}^{-1})_{LN} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_N|}$$

$$(\Lambda_M)_{\mu\nu} = - \sum_{kL} a_{kML} (\mathbb{I}_k)_{\mu\nu} + \frac{1}{\Gamma} \sum_J (\mathbf{G}^{-1})_{MJ} \sum_{kLN} a_{kLN} (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.13})$$

$$\text{Let } b_{kN} = \sum_L a_{kNL}$$

$$(\Lambda_M)_{\mu\nu} = - \sum_k b_{kM} (\mathbb{I}_k)_{\mu\nu} + \frac{1}{\Gamma} \sum_J (\mathbf{G}^{-1})_{MJ} \sum_{kL} b_{kL} (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.14})$$

$$\text{Let } c_k = \sum_L b_{kL}$$

$$(\Lambda_M)_{\mu\nu} = - \sum_k b_{kM} (\mathbb{I}_k)_{\mu\nu} + \frac{1}{\Gamma} \sum_J (\mathbf{G}^{-1})_{MJ} \sum_k c_k (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.15})$$

$$\text{Let } d_{kM} = \frac{1}{\Gamma} \sum_J (\mathbf{G}^{-1})_{MJ} c_k$$

$$(\Lambda_M)_{\mu\nu} = - \sum_k b_{kM} (\mathbb{I}_k)_{\mu\nu} + \sum_k d_{kM} (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.16})$$

Let $f_{kM} = d_{kM} - b_{kM}$

$$(\Lambda_M)_{\mu\nu} = \sum_k f_{kM} (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.17})$$

$$\text{So } f_{kM} = - \sum_L (\mathbf{G}^{-1})_{ML} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_L|} + \frac{\sum_J (\mathbf{G}^{-1})_{MJ}}{\sum_{IJ} (\mathbf{G}^{-1})_{IJ}} \sum_{LN} (\mathbf{G}^{-1})_{LN} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_N|}$$

H.2.2 Alternate Simplification of Derivative

$$\begin{aligned} (\Lambda_M)_{\mu\nu} = & - \sum_L (\mathbf{G}^{-1})_{ML} \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_L|} (\mathbb{I}_k)_{\mu\nu} \\ & + \frac{\sum_J (\mathbf{G}^{-1})_{MJ}}{\sum_{IJ} (\mathbf{G}^{-1})_{IJ}} \sum_{LN} (\mathbf{G}^{-1})_{LN} \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_N|} (\mathbb{I}_k)_{\mu\nu} \end{aligned} \quad (\text{H.18})$$

$$\text{Let } \gamma_M = \frac{\sum_J (\mathbf{G}^{-1})_{MJ}}{\sum_{IJ} (\mathbf{G}^{-1})_{IJ}} \text{ and } \lambda_{kM} = \sum_L (\mathbf{G}^{-1})_{ML} \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_L|}$$

$$(\Lambda_M)_{\mu\nu} = - \sum_k \lambda_{kM} (\mathbb{I}_k)_{\mu\nu} + \gamma_M \sum_N \sum_k \lambda_{kN} (\mathbb{I}_k)_{\mu\nu} \quad (\text{H.19})$$

$$\text{Let } (\Omega_M)_{\mu\nu} = \sum_k \lambda_{kM} (\mathbb{I}_k)_{\mu\nu}$$

$$(\Lambda_M)_{\mu\nu} = - (\Omega_M)_{\mu\nu} + \gamma_M \sum_N (\Omega_N)_{\mu\nu} \quad (\text{H.20})$$

$$(\Lambda_M)_{\mu\nu} = \sum_N (\Omega_N)_{\mu\nu} (\gamma_M - \delta_{MN}) \quad (\text{H.21})$$

H.2.3 Application of Derivative to Fock Matrix

$$\Delta F_{\mu\nu} = \sum_J \frac{\partial E}{\partial Q_J} (\Lambda_J)_{\mu\nu} \quad (\text{H.22})$$

$$\Delta F_{\mu\nu} = \sum_J \frac{\partial E}{\partial Q_J} \sum_N (\Omega_N)_{\mu\nu} (\gamma_J - \delta_{JN}) \quad (\text{H.23})$$

$$\Delta F_{\mu\nu} = \sum_N (\Omega_N)_{\mu\nu} \sum_J \frac{\partial E}{\partial Q_J} (\gamma_J - \delta_{JN}) \quad (\text{H.24})$$

H.3 Spatial Derivative for ChEIP-G Charges

Note: All sums are over the total number of (QM) atoms. Note: In all cases in this section $\nabla = \nabla_F$, (where F is an arbitrary atom) but the subscript atom will be omitted for clarity.

$$q_A = \sum_B e_B (\mathbf{G}^{-1})_{BA} - \frac{\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \sum_I (\mathbf{G}^{-1})_{IA} \quad (\text{H.25})$$

$$\begin{aligned} q_A \sum_{CD} (\mathbf{G}^{-1})_{CD} &= \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B (\mathbf{G}^{-1})_{BA} \\ &\quad - \left[\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q \right] \sum_I (\mathbf{G}^{-1})_{IA} \end{aligned} \quad (\text{H.26})$$

$$\begin{aligned} \nabla \left[q_A \sum_{CD} (\mathbf{G}^{-1})_{CD} \right] &= \nabla \left[\sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B (\mathbf{G}^{-1})_{BA} \right. \\ &\quad \left. - \sum_{FH} e_F (\mathbf{G}^{-1})_{FH} \sum_I (\mathbf{G}^{-1})_{IA} + Q \sum_I (\mathbf{G}^{-1})_{IA} \right] \end{aligned} \quad (\text{H.27})$$

$$\begin{aligned}
q_A \sum_{CD} (\mathbf{G}^{-1})_{CD} + q_A \sum_{CD} (\mathbf{G}^{-1})_{CD}^\nabla &= \sum_{CD} (\mathbf{G}^{-1})_{CD}^\nabla \sum_B e_B (\mathbf{G}^{-1})_{BA} \\
&+ \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B^\nabla (\mathbf{G}^{-1})_{BA} + \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B (\mathbf{G}^{-1})_{BA}^\nabla \\
&- \sum_{FH} e_F^\nabla (\mathbf{G}^{-1})_{FH} \sum_I (\mathbf{G}^{-1})_{IA} - \sum_{FH} e_F (\mathbf{G}^{-1})_{FH}^\nabla \sum_I (\mathbf{G}^{-1})_{IA} \\
&- \sum_{FH} e_F (\mathbf{G}^{-1})_{FH} \sum_I (\mathbf{G}^{-1})_{IA}^\nabla + Q \sum_I (\mathbf{G}^{-1})_{IA}^\nabla \quad (\text{H.28})
\end{aligned}$$

$$\begin{aligned}
q_A \sum_{CD} (\mathbf{G}^{-1})_{CD} &= \left[\sum_B e_B (\mathbf{G}^{-1})_{BA} - q_A \right] \sum_{CD} (\mathbf{G}^{-1})_{CD}^\nabla \\
&+ \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B^\nabla (\mathbf{G}^{-1})_{BA} + \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B (\mathbf{G}^{-1})_{BA}^\nabla \\
&- \sum_{FH} e_F^\nabla (\mathbf{G}^{-1})_{FH} \sum_I (\mathbf{G}^{-1})_{IA} - \sum_{FH} e_F (\mathbf{G}^{-1})_{FH}^\nabla \sum_I (\mathbf{G}^{-1})_{IA} \\
&- \left[\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q \right] \sum_I (\mathbf{G}^{-1})_{IA}^\nabla \quad (\text{H.29})
\end{aligned}$$

$$\begin{aligned}
q_A \sum_{CD} (\mathbf{G}^{-1})_{CD} &= \underbrace{\left[\sum_B e_B (\mathbf{G}^{-1})_{BA} - q_A \right] \sum_{CD} (\mathbf{G}^{-1})_{CD}^\nabla}_{\text{term 1}} \\
&+ \underbrace{\sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B^\nabla (\mathbf{G}^{-1})_{BA}}_{\text{term 2}} + \underbrace{\sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B (\mathbf{G}^{-1})_{BA}^\nabla}_{\text{term 3}} \\
&- \underbrace{\sum_{FH} e_F^\nabla (\mathbf{G}^{-1})_{FH} \sum_I (\mathbf{G}^{-1})_{IA}}_{\text{term 4}} - \underbrace{\sum_{FH} e_F (\mathbf{G}^{-1})_{FH}^\nabla \sum_I (\mathbf{G}^{-1})_{IA}}_{\text{term 5}} \\
&- \underbrace{\left[\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q \right] \sum_I (\mathbf{G}^{-1})_{IA}^\nabla}_{\text{term 6}} \quad (\text{H.30})
\end{aligned}$$

H.3.1 Simplify Terms 1 and 6

$$T_{16} = \left[\sum_B e_B (\mathbf{G}^{-1})_{BA} - q_A \right] \sum_{CD} (\mathbf{G}^{-1})_{CD}^{\nabla} - \left[\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q \right] \sum_I (\mathbf{G}^{-1})_{IA}^{\nabla} \quad (\text{H.31})$$

$$\begin{aligned} T_{16} = & \left[\sum_B e_B (\mathbf{G}^{-1})_{BA} - \sum_J e_J (\mathbf{G}^{-1})_{JA} \right. \\ & \left. + \frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \sum_I (\mathbf{G}^{-1})_{IA} \right] \sum_{CD} (\mathbf{G}^{-1})_{CD}^{\nabla} \\ & - \left[\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q \right] \sum_I (\mathbf{G}^{-1})_{IA}^{\nabla} \quad (\text{H.32}) \end{aligned}$$

$$\begin{aligned} T_{16} = & \left[\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \sum_I (\mathbf{G}^{-1})_{IA} \right] \sum_{CD} (\mathbf{G}^{-1})_{CD}^{\nabla} \\ & - \left[\sum_{FH} e_F (\mathbf{G}^{-1})_{FH} - Q \right] \sum_I (\mathbf{G}^{-1})_{IA}^{\nabla} \cdot \frac{\sum_{MN} (\mathbf{G}^{-1})_{MN}}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \quad (\text{H.33}) \end{aligned}$$

$$\begin{aligned} T_{16} = & \left[\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \right] \left[\sum_I (\mathbf{G}^{-1})_{IA} \sum_{CD} (\mathbf{G}^{-1})_{CD}^{\nabla} \right. \\ & \left. - \sum_I (\mathbf{G}^{-1})_{IA}^{\nabla} \sum_{MN} (\mathbf{G}^{-1})_{MN} \right] \quad (\text{H.34}) \end{aligned}$$

$$\begin{aligned} T_{16} = & \left[\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \right] \left[\sum_{PQ} (\mathbf{G}^{-1})_{PQ} \delta_{QA} \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} \right. \\ & \left. - \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} \delta_{SA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \right] \quad (\text{H.35}) \end{aligned}$$

$$T_{16} = \left[\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \right] \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} \left[\sum_{PQ} (\mathbf{G}^{-1})_{PQ} \delta_{QA} - \delta_{SA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \right] \quad (\text{H.36})$$

$$T_{16} = \left[\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \right] \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} \left[\sum_P (\mathbf{G}^{-1})_{PA} - \delta_{SA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \right] \quad (\text{H.37})$$

$$T_{16} = \left[\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q \right] \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} \left[\frac{\sum_P (\mathbf{G}^{-1})_{PA}}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} - \delta_{SA} \frac{\sum_{PQ} (\mathbf{G}^{-1})_{PQ}}{\sum_{MN} (\mathbf{G}^{-1})_{MN}} \right] \quad (\text{H.38})$$

$$\text{Let: } \gamma_A = \frac{\sum_P (\mathbf{G}^{-1})_{PA}}{\sum_{MN} (\mathbf{G}^{-1})_{MN}}$$

$$T_{16} = \left[\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q \right] \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} [\gamma_A - \delta_{SA}] \quad (\text{H.39})$$

H.3.2 Terms 2 and 4

$$T_{24} = \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B^{\nabla} (\mathbf{G}^{-1})_{BA} - \sum_{FH} e_F^{\nabla} (\mathbf{G}^{-1})_{FH} \sum_I (\mathbf{G}^{-1})_{IA} \quad (\text{H.40})$$

$$T_{24} = \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \sum_{RS} e_R^{\nabla} (\mathbf{G}^{-1})_{RS} \delta_{SA} - \sum_{RS} e_R^{\nabla} (\mathbf{G}^{-1})_{RS} \sum_P (\mathbf{G}^{-1})_{PA} \quad (\text{H.41})$$

$$T_{24} = \sum_{RS} e_R^{\nabla} (\mathbf{G}^{-1})_{RS} \left[\delta_{SA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} - \sum_P (\mathbf{G}^{-1})_{PA} \right] \quad (\text{H.42})$$

H.3.3 Terms 3 and 5

$$T_{35} = \sum_{CD} (\mathbf{G}^{-1})_{CD} \sum_B e_B (\mathbf{G}^{-1})_{BA}^{\nabla} - \sum_{FH} e_F (\mathbf{G}^{-1})_{FH}^{\nabla} \sum_I (\mathbf{G}^{-1})_{IA} \quad (\text{H.43})$$

$$T_{35} = \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \sum_{RS} e_R (\mathbf{G}^{-1})_{RS}^{\nabla} \delta_{SA} - \sum_{RS} e_R (\mathbf{G}^{-1})_{RS}^{\nabla} \sum_P (\mathbf{G}^{-1})_{PA} \quad (\text{H.44})$$

$$T_{35} = \sum_{RS} e_R (\mathbf{G}^{-1})_{RS}^{\nabla} \left[\delta_{SA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} - \sum_P (\mathbf{G}^{-1})_{PA} \right] \quad (\text{H.45})$$

H.3.4 Recombine

$$q_A^{\nabla} \sum_{CD} (\mathbf{G}^{-1})_{CD} = T_{16} + T_{24} + T_{35} \quad (\text{H.46})$$

$$\begin{aligned} q_A^{\nabla} \sum_{CD} (\mathbf{G}^{-1})_{CD} &= \left[\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q \right] \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} [\gamma_A - \delta_{SA}] \\ &\quad + \sum_{MN} e_M^{\nabla} (\mathbf{G}^{-1})_{MN} \left[\delta_{NA} \sum_{IJ} (\mathbf{G}^{-1})_{IJ} - \sum_I (\mathbf{G}^{-1})_{IA} \right] \\ &\quad + \sum_{WX} e_W (\mathbf{G}^{-1})_{WX}^{\nabla} \left[\delta_{XA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} - \sum_P (\mathbf{G}^{-1})_{PA} \right] \end{aligned} \quad (\text{H.47})$$

$$\begin{aligned} q_A^{\nabla} &= \frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \sum_{RS} (\mathbf{G}^{-1})_{RS}^{\nabla} [\gamma_A - \delta_{SA}] \\ &\quad + \sum_{MN} e_M^{\nabla} (\mathbf{G}^{-1})_{MN} [\delta_{NA} - \gamma_A] + \sum_{WX} e_W (\mathbf{G}^{-1})_{WX}^{\nabla} [\delta_{XA} - \gamma_A] \end{aligned} \quad (\text{H.48})$$

$$q_A^\nabla = \frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \sum_{RS} (\mathbf{G}^{-1})_{RS}^\nabla [\gamma_A - \delta_{SA}] + \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla \right] [\delta_{NA} - \gamma_A] \quad (\text{H.49})$$

$$q_A^\nabla = -\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla [\delta_{NA} - \gamma_A] + \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla \right] [\delta_{NA} - \gamma_A] \quad (\text{H.50})$$

$$\text{Let: } \Xi = -\frac{\sum_{KL} e_K (\mathbf{G}^{-1})_{KL} - Q}{\sum_{CD} (\mathbf{G}^{-1})_{CD}}$$

$$q_A^\nabla = \Xi \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla [\delta_{NA} - \gamma_A] + \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla \right] [\delta_{NA} - \gamma_A] \quad (\text{H.51})$$

$$q_A^\nabla = \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla + \Xi (\mathbf{G}^{-1})_{MN}^\nabla \right] [\delta_{NA} - \gamma_A] \quad (\text{H.52})$$

H.3.5 Alternate Derivation

Let:

$$\alpha = \frac{Q - \sum_P^{N_{atoms}} g_P}{\sum_{MN}^{N_{atoms}} (\mathbf{G}^{-1})_{MN}}$$

$$g_J = \sum_R^{N_{atoms}} e_R (\mathbf{G}^{-1})_{RJ}$$

$$q_A = g_A + \alpha \sum_B (\mathbf{G}^{-1})_{AB} \quad (\text{H.53})$$

$$q_A = g_A + \left(\frac{Q - \sum_C g_C}{\sum_{DE} (\mathbf{G}^{-1})_{DE}} \right) \sum_B (\mathbf{G}^{-1})_{AB} \quad (\text{H.54})$$

$$q_A \sum_{DE} (\mathbf{G}^{-1})_{DE} = g_A \sum_{FH} (\mathbf{G}^{-1})_{FH} + \left(Q - \sum_C g_C \right) \sum_B (\mathbf{G}^{-1})_{AB} \quad (\text{H.55})$$

$$\begin{aligned} q_A^\nabla \sum_{DE} (\mathbf{G}^{-1})_{DE} + q_A \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla &= g_A^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH} + g_A \sum_{KL} (\mathbf{G}^{-1})_{KL}^\nabla \\ &\quad - \sum_C g_C^\nabla \sum_B (\mathbf{G}^{-1})_{AB} + \left(Q - \sum_M g_M \right) \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \end{aligned} \quad (\text{H.56})$$

$$\begin{aligned} q_A^\nabla \sum_{DE} (\mathbf{G}^{-1})_{DE} &= \underbrace{g_A^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH}}_{\text{Terms 2 \& 3}} - \underbrace{\sum_C g_C^\nabla \sum_B (\mathbf{G}^{-1})_{AB}}_{\text{Terms 4 \& 5}} \\ &\quad + \underbrace{g_A \sum_{KL} (\mathbf{G}^{-1})_{KL}^\nabla - q_A \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla}_{\text{Term 1}} + \underbrace{\left(Q - \sum_M g_M \right) \sum_N (\mathbf{G}^{-1})_{AN}^\nabla}_{\text{Term 6}} \end{aligned} \quad (\text{H.57})$$

$$\begin{aligned} q_A^\nabla \sum_{DE} (\mathbf{G}^{-1})_{DE} &= g_A^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH} - \sum_C g_C^\nabla \sum_B (\mathbf{G}^{-1})_{AB} + g_A \sum_{KL} (\mathbf{G}^{-1})_{KL}^\nabla \\ &\quad - \left(g_A + \alpha \sum_P (\mathbf{G}^{-1})_{AP} \right) \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla + \left(Q - \sum_M g_M \right) \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \end{aligned} \quad (\text{H.58})$$

$$\begin{aligned}
q_A^\nabla \sum_{DE} (\mathbf{G}^{-1})_{DE} &= g_A^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH} - \sum_C g_C^\nabla \sum_B (\mathbf{G}^{-1})_{AB} \\
&\quad - \alpha \sum_P (\mathbf{G}^{-1})_{AP} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla + \left(Q - \sum_M g_M \right) \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \quad (\text{H.59})
\end{aligned}$$

$$\begin{aligned}
q_A^\nabla \sum_{DE} (\mathbf{G}^{-1})_{DE} &= g_A^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH} - \sum_C g_C^\nabla \sum_B (\mathbf{G}^{-1})_{AB} \\
&\quad - \frac{Q - \sum_Q g_Q}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \sum_P (\mathbf{G}^{-1})_{AP} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla + \left(Q - \sum_M g_M \right) \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \quad (\text{H.60})
\end{aligned}$$

$$\begin{aligned}
q_A^\nabla &= g_A^\nabla \frac{\sum_{FH} (\mathbf{G}^{-1})_{FH}}{\sum_{DE} (\mathbf{G}^{-1})_{DE}} - \sum_C g_C^\nabla \frac{\sum_B (\mathbf{G}^{-1})_{AB}}{\sum_{TU} (\mathbf{G}^{-1})_{TU}} \\
&\quad - \frac{Q - \sum_Q g_Q}{\left[\sum_{RS} (\mathbf{G}^{-1})_{RS} \right]^2} \sum_P (\mathbf{G}^{-1})_{AP} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla + \frac{\left(Q - \sum_M g_M \right)}{\sum_{VW} (\mathbf{G}^{-1})_{VW}} \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \quad (\text{H.61})
\end{aligned}$$

$$\text{Let: } \gamma_A = \frac{\sum_B (\mathbf{G}^{-1})_{AB}}{\sum_{TU} (\mathbf{G}^{-1})_{TU}} \text{ and } \Xi = \frac{\left(Q - \sum_M g_M \right)}{\sum_{VW} (\mathbf{G}^{-1})_{VW}}$$

$$q_A^\nabla = g_A^\nabla - \sum_C g_C^\nabla \gamma_A - \frac{\Xi}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \sum_P (\mathbf{G}^{-1})_{AP} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla + \Xi \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \quad (\text{H.62})$$

$$\begin{aligned}
q_A^\nabla = g_A^\nabla - \sum_C g_C^\nabla \gamma_A - \frac{\Xi}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \sum_P (\mathbf{G}^{-1})_{AP} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla \\
+ \frac{\Xi}{\sum_{DE} (\mathbf{G}^{-1})_{DE}} \sum_N (\mathbf{G}^{-1})_{AN}^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH} \quad (\text{H.63})
\end{aligned}$$

$$\begin{aligned}
q_A^\nabla = g_A^\nabla - \sum_C g_C^\nabla \gamma_A + \frac{\Xi}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \left(\sum_N (\mathbf{G}^{-1})_{AN}^\nabla \sum_{FH} (\mathbf{G}^{-1})_{FH} \right. \\
\left. - \sum_P (\mathbf{G}^{-1})_{AP} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla \right) \quad (\text{H.64})
\end{aligned}$$

$$\begin{aligned}
q_A^\nabla = \sum_C (g_C^\nabla \delta_{CA} - g_C^\nabla \gamma_A) + \frac{\Xi}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \left(\sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla \delta_{MA} \sum_{FH} (\mathbf{G}^{-1})_{FH} \right. \\
\left. - \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \delta_{QA} \sum_{IJ} (\mathbf{G}^{-1})_{IJ}^\nabla \right) \quad (\text{H.65})
\end{aligned}$$

$$\begin{aligned}
q_A^\nabla = \sum_C g_C^\nabla (\delta_{CA} - \gamma_A) + \frac{\Xi}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \left(\sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla \delta_{MA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \right. \\
\left. - \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \delta_{QA} \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla \right) \quad (\text{H.66})
\end{aligned}$$

$$\begin{aligned}
q_A^\nabla = \sum_C g_C^\nabla (\delta_{CA} - \gamma_A) + \frac{\Xi \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla}{\sum_{RS} (\mathbf{G}^{-1})_{RS}} \left(\delta_{MA} \sum_{PQ} (\mathbf{G}^{-1})_{PQ} - \sum_{PQ} (\mathbf{G}^{-1})_{PQ} \delta_{QA} \right) \\
\quad (\text{H.67})
\end{aligned}$$

$$q_A^\nabla = \sum_C g_C^\nabla (\delta_{CA} - \gamma_A) + \Xi \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla (\delta_{MA} - \gamma_A) \quad (\text{H.68})$$

$$q_A^\nabla = \sum_C \left(\sum_D e_D (\mathbf{G}^{-1})_{DC} \right)^\nabla (\delta_{CA} - \gamma_A) + \Xi \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla (\delta_{MA} - \gamma_A) \quad (\text{H.69})$$

$$q_A^\nabla = \sum_{CD} \left(e_D^\nabla (\mathbf{G}^{-1})_{DC} + e_D (\mathbf{G}^{-1})_{DC}^\nabla \right) (\delta_{CA} - \gamma_A) + \Xi \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla (\delta_{MA} - \gamma_A) \quad (\text{H.70})$$

$$q_A^\nabla = \sum_{MN} \left(e_N^\nabla (\mathbf{G}^{-1})_{NM} + e_N (\mathbf{G}^{-1})_{NM}^\nabla \right) (\delta_{MA} - \gamma_A) + \Xi \sum_{MN} (\mathbf{G}^{-1})_{MN}^\nabla (\delta_{MA} - \gamma_A) \quad (\text{H.71})$$

$$q_A^\nabla = \sum_{MN} \left(e_N^\nabla (\mathbf{G}^{-1})_{NM} + e_N (\mathbf{G}^{-1})_{NM}^\nabla + \Xi (\mathbf{G}^{-1})_{MN}^\nabla \right) (\delta_{MA} - \gamma_A) \quad (\text{H.72})$$

H.3.6 Programmed Versions

$$(q_A^\nabla)_{fast} = \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla + \Xi (\mathbf{G}^{-1})_{MN}^\nabla \right] [\delta_{NA} - \gamma_A] \quad (\text{H.73})$$

$$\begin{aligned} (q_A^\nabla)_{slow} &= \frac{-\Xi}{\sum_{CD} (\mathbf{G}^{-1})_{CD}} \sum_{MN} \left[(\mathbf{G}^{-1})_{MA} (\mathbf{G}^{-1})_{MN}^\nabla - (\mathbf{G}^{-1})_{MN} (\mathbf{G}^{-1})_{MA}^\nabla \right] \\ &\quad + \sum_M \left[e_M^\nabla (\mathbf{G}^{-1})_{MA} + e_M (\mathbf{G}^{-1})_{MA}^\nabla \right] \\ &\quad - \gamma_A \sum_{MN} \left[e_M^\nabla (\mathbf{G}^{-1})_{MN} + e_M (\mathbf{G}^{-1})_{MN}^\nabla \right] \end{aligned} \quad (\text{H.74})$$

H.4 Needed Derivatives for q_A^∇

H.4.1 $(\mathbf{G}^{-1})_{CD}^{\nabla A}$

Begin Sidenote $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

$$\nabla(\mathbf{A}\mathbf{A}^{-1}) = \nabla\mathbf{I}$$

$$(\mathbf{A})^\nabla \mathbf{A}^{-1} + \mathbf{A} (\mathbf{A}^{-1})^\nabla = 0$$

$$\mathbf{A} (\mathbf{A}^{-1})^\nabla = -(\mathbf{A})^\nabla \mathbf{A}^{-1}$$

$$(\mathbf{A}^{-1})^\nabla = -\mathbf{A}^{-1} (\mathbf{A})^\nabla \mathbf{A}^{-1}$$

End Sidenote

$$G_{CD}^{\nabla A} = \nabla_A \left(\sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|} \right) \quad (\text{H.75})$$

$$G_{CD}^{\nabla A} = \sum_k \frac{w_k^{\nabla A}}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|} + \sum_k \frac{-w_k (\mathbf{r}_k - \mathbf{r}_C)}{|\mathbf{r}_k - \mathbf{r}_C|^3 |\mathbf{r}_k - \mathbf{r}_D|} \nabla_A (\mathbf{r}_k - \mathbf{r}_C) \\ + \sum_k \frac{-w_k (\mathbf{r}_k - \mathbf{r}_D)}{|\mathbf{r}_k - \mathbf{r}_C| |\mathbf{r}_k - \mathbf{r}_D|^3} \nabla_A (\mathbf{r}_k - \mathbf{r}_D) \quad (\text{H.76})$$

Fixed Grid Points

Does not require the weight derivative term.

	$\nabla_A (\mathbf{r}_k - \mathbf{r}_X)$
$X = A$	-1
$X \neq A$	0

		$G_{CD}^{\nabla A}$ (Last two terms)
$C = A$	$D = A$	$\sum_k \frac{2w_k}{ \mathbf{r}_k - \mathbf{r}_A ^4} (\mathbf{r}_k - \mathbf{r}_A)$
$C = A$	$D \neq A$	$\sum_k \frac{w_k}{ \mathbf{r}_k - \mathbf{r}_A ^3 \mathbf{r}_k - \mathbf{r}_D } (\mathbf{r}_k - \mathbf{r}_A)$
$C \neq A$	$D = A$	$\sum_k \frac{w_k}{ \mathbf{r}_k - \mathbf{r}_C \mathbf{r}_k - \mathbf{r}_A ^3} (\mathbf{r}_k - \mathbf{r}_A)$
$C \neq A$	$D \neq A$	0

Dynamic Grid Points

Requires the weight derivative term.

		$\nabla_A (\mathbf{r}_k - \mathbf{r}_X)$
$k \in A$	$X = A$	0
$k \notin A$	$X = A$	-1
$k \in A$	$X \neq A$	1
$k \notin A$	$X \neq A$	0

			$G_{CD}^{\nabla A}$ (Last two terms)
$k \in A$	$C = A$	$D = A$	0
$k \notin A$	$C = A$	$D = A$	$\sum_{k \notin A} \frac{2w_k}{ \mathbf{r}_k - \mathbf{r}_A ^4} (\mathbf{r}_k - \mathbf{r}_A)$
$k \in A$	$C = A$	$D \neq A$	$\sum_k \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_A \mathbf{r}_k - \mathbf{r}_D ^3} (\mathbf{r}_k - \mathbf{r}_D)$
$k \notin A$	$C = A$	$D \neq A$	$\sum_{k \notin A} \frac{w_k}{ \mathbf{r}_k - \mathbf{r}_A ^3 \mathbf{r}_k - \mathbf{r}_D } (\mathbf{r}_k - \mathbf{r}_A)$
$k \in A$	$C \neq A$	$D = A$	$\sum_k \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_C ^3 \mathbf{r}_k - \mathbf{r}_A } (\mathbf{r}_k - \mathbf{r}_C)$
$k \notin A$	$C \neq A$	$D = A$	$\sum_{k \notin A} \frac{w_k}{ \mathbf{r}_k - \mathbf{r}_C \mathbf{r}_k - \mathbf{r}_A ^3} (\mathbf{r}_k - \mathbf{r}_A)$
$k \in A$	$C \neq A$	$D \neq A$	$\sum_{k \in A} \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_C \mathbf{r}_k - \mathbf{r}_D } \left[\frac{(\mathbf{r}_k - \mathbf{r}_C)}{ \mathbf{r}_k - \mathbf{r}_C ^2} + \frac{(\mathbf{r}_k - \mathbf{r}_D)}{ \mathbf{r}_k - \mathbf{r}_D ^2} \right]$
$k \notin A$	$C \neq A$	$D \neq A$	0

H.4.2 $e_C^{\nabla A}$

$$e_C^{\nabla A} = \nabla_A \left(\sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right] \right) \quad (\text{H.77})$$

$$\begin{aligned} e_C^{\nabla A} &= \underbrace{\sum_k \nabla_A \left(\frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \right) \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right]}_{\text{term1}} \\ &+ \underbrace{\sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \nabla_A \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} \right)}_{\text{term2}} + \underbrace{\sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \left[-\nabla_A \left(\sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right) \right]}_{\text{term3+term4}} \end{aligned} \quad (\text{H.78})$$

$$\text{term1} = \sum_k \nabla_A \left(\frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \right) \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right] \quad (\text{H.79})$$

$$\begin{aligned} \text{term1} &= \sum_k \frac{w_k^{\nabla A}}{|\mathbf{r}_k - \mathbf{r}_C|} \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right] \\ &+ \sum_k \frac{-w_k (\mathbf{r}_k - \mathbf{r}_C)}{|\mathbf{r}_k - \mathbf{r}_C|^3} \nabla_A (\mathbf{r}_k - \mathbf{r}_C) \left[\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right] \end{aligned} \quad (\text{H.80})$$

$$\text{term2} = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \nabla_A \left(\sum_J \frac{Z_J}{|\mathbf{r}_k - \mathbf{r}_J|} \right) \quad (\text{H.81})$$

$$\text{term2} = \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_J \frac{-Z_J (\mathbf{r}_k - \mathbf{r}_J)}{|\mathbf{r}_k - \mathbf{r}_J|^3} \nabla_A (\mathbf{r}_k - \mathbf{r}_J) \quad (\text{H.82})$$

$$\text{term3} + \text{term4} = \sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \nabla_A \left(\sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right) \quad (\text{H.83})$$

$$\begin{aligned}
term3 + term4 = & \underbrace{\sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} \nabla_A \left((\mathbb{I}_k)_{\mu\nu} \right) P_{\mu\nu}}_{term3} \\
& + \underbrace{\sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} \nabla_A (P_{\mu\nu})}_{term4} \quad (H.84)
\end{aligned}$$

$term4$ is taken care of by \mathbf{WS}^x term.

$$term3 = \sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} \nabla_A \left((\mathbb{I}_k)_{\mu\nu} \right) P_{\mu\nu} \quad (H.85)$$

$$term3 = \sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} \nabla_A \left(\left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle \right) P_{\mu\nu} \quad (H.86)$$

$$\begin{aligned}
term3 = \sum_k \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} & \left(\left\langle \nabla_A \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle + \left\langle \mu \left| \frac{-\nabla_A (-\mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|^3} \right| \nu \right\rangle \right. \\
& \left. + \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|} \right| \nabla_A \nu \right\rangle \right) P_{\mu\nu} \quad (H.87)
\end{aligned}$$

Fixed Grid Points

Does not require weight derivatives

	$\nabla_A (\mathbf{r}_k - \mathbf{r}_X)$
$X = A$	-1
$X \neq A$	0

	$term1$ (Last term only)
$C = A$	$\sum_k \frac{w_k (\mathbf{r}_k - \mathbf{r}_A)}{ \mathbf{r}_k - \mathbf{r}_A ^3} \left[\sum_J \frac{Z_J}{ \mathbf{r}_k - \mathbf{r}_J } - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right]$
$C \neq A$	0

	<i>term2</i>
$J = A$	$\sum_k \frac{w_k}{ \mathbf{r}_k - \mathbf{r}_C } \frac{Z_A(\mathbf{r}_k - \mathbf{r}_A)}{ \mathbf{r}_k - \mathbf{r}_A ^3}$
$J \neq A$	0

		<i>term3</i>
$\mu \in A$	$\nu \in A$	$\sum_k \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_C } \sum_{\mu\nu} \left(\left\langle \mu^{\nabla A} \left \frac{1}{ \mathbf{r} - \mathbf{r}_k } \right \nu \right\rangle + \left\langle \mu \left \frac{1}{ \mathbf{r} - \mathbf{r}_k } \right \nu^{\nabla A} \right\rangle \right) P_{\mu\nu}$
$\mu \in A$	$\nu \notin A$	$\sum_k \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_C } \sum_{\mu\nu} \left\langle \mu^{\nabla A} \left \frac{1}{ \mathbf{r} - \mathbf{r}_k } \right \nu \right\rangle P_{\mu\nu}$
$\mu \notin A$	$\nu \in A$	$\sum_k \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_C } \sum_{\mu\nu} \left\langle \mu \left \frac{1}{ \mathbf{r} - \mathbf{r}_k } \right \nu^{\nabla A} \right\rangle P_{\mu\nu}$
$\mu \notin A$	$\nu \notin A$	0

For *term3*, Integral Job # -111 is

$J^{-111} = \sum_k \sum_{\mu\nu} \left(\left\langle \mu^{\nabla A} \left| \frac{b_k}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu \right\rangle + \left\langle \mu \left| \frac{b_k}{|\mathbf{r} - \mathbf{r}_k|} \right| \nu^{\nabla A} \right\rangle \right) P_{\mu\nu}$. b_k will have a dependence on C because $b_k = \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|}$. Thus, J_C^{-111} is dependent on C , so there are C integral calls.

Dynamic Grid Points

Requires weight derivatives

		$\nabla_A(\mathbf{r}_k - \mathbf{r}_X)$
$k \in A$	$X = A$	0
$k \notin A$	$X = A$	-1
$k \in A$	$X \neq A$	1
$k \notin A$	$X \neq A$	0

		<i>term1</i> (Last term only)
$k \in A$	$C = A$	0
$k \notin A$	$C = A$	$\sum_{k \notin A} \frac{w_k(\mathbf{r}_k - \mathbf{r}_A)}{ \mathbf{r}_k - \mathbf{r}_A ^3} \left[\sum_J \frac{Z_J}{ \mathbf{r}_k - \mathbf{r}_J } - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right]$
$k \in A$	$C \neq A$	$\sum_{k \in A} \frac{-w_k(\mathbf{r}_k - \mathbf{r}_C)}{ \mathbf{r}_k - \mathbf{r}_C ^3} \left[\sum_J \frac{Z_J}{ \mathbf{r}_k - \mathbf{r}_J } - \sum_{\mu\nu} (\mathbb{I}_k)_{\mu\nu} P_{\mu\nu} \right]$
$k \notin A$	$C \neq A$	0

		<i>term2</i>
$k \in A$	$J = A$	0
$k \notin A$	$J = A$	$\sum_{k \notin A} \frac{w_k}{ \mathbf{r}_k - \mathbf{r}_C } \frac{Z_A(\mathbf{r}_k - \mathbf{r}_A)}{ \mathbf{r}_k - \mathbf{r}_A ^3}$
$k \in A$	$J \neq A$	$\sum_{k \in A} \frac{-w_k}{ \mathbf{r}_k - \mathbf{r}_C } \sum_J \frac{Z_J(\mathbf{r}_k - \mathbf{r}_J)}{ \mathbf{r}_k - \mathbf{r}_J ^3}$
$k \notin A$	$J \neq A$	0

$$term3 = J_C^{-111} + \sum_{k \in A} \frac{-w_k}{|\mathbf{r}_k - \mathbf{r}_C|} \sum_{\mu\nu} \left\langle \mu \left| \frac{1}{|\mathbf{r} - \mathbf{r}_k|^3} \right| \nu \right\rangle P_{\mu\nu} \quad (\text{H.88})$$

I'm not sure what this second term is in terms of integral calls. Possibly 112?

H.5 M Derivatives

This section is the derivation of the derivatives $\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha}$ for QM atoms and $\sum_M \frac{\partial M}{\partial x_j} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha}$ for MM atoms.

Remember that in the M sum, $P_{\mu\nu}$ is excluded.

H.5.1 Mulliken Charges

w.r.t. QM Atoms

$$\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial S_{\mu\nu}} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{H.89})$$

$$\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_\alpha \sum_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} \frac{\partial Q_\alpha}{\partial S_{\mu\nu}} \sum_{\aleph} Q_{\aleph} \omega_{\aleph\alpha} \quad (\text{H.90})$$

$$\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_\alpha \sum_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} (-P_{\mu\nu} \delta_{\mu \in \alpha}) \sum_{\aleph} Q_{\aleph} \omega_{\aleph\alpha} \quad (\text{H.91})$$

$$\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = - \sum_\alpha \sum_{\mathfrak{N}} Q_{\mathfrak{N}\omega\mathfrak{N}\alpha} \sum_{\substack{\mu \in \alpha \\ \nu}} P_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_\gamma} \quad (\text{H.92})$$

w.r.t. MM Atoms

$$\sum_M \frac{\partial M}{\partial x_j} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_{\mu\nu} \underbrace{\frac{\partial S_{\mu\nu}}{\partial x_j}}_{=0} \sum_\alpha \frac{\partial Q_\alpha}{\partial S_{\mu\nu}} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{H.93})$$

$$\sum_M \frac{\partial M}{\partial x_j} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = 0 \quad (\text{H.94})$$

H.5.2 ChEIP-G Charges

w.r.t. QM Atoms

$$\sum_\alpha \frac{\partial Q_\alpha}{\partial x_\gamma} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} + \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{H.95})$$

$$\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_\alpha \frac{\partial Q_\alpha}{\partial x_\gamma} \frac{\partial E^{PI}}{\partial Q_\alpha} - \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \frac{\partial E^{PI}}{\partial Q_\alpha} \quad (\text{H.96})$$

$$\sum_M \frac{\partial M}{\partial x_\gamma} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \left(\frac{\partial Q_\alpha}{\partial x_\gamma} - \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\gamma} \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \right) \quad (\text{H.97})$$

Switch indicies from γ to ζ . Remember that $\alpha\beta$ will be QM atoms and $\mu\nu$ will be basis functions.

$$\sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \left(\frac{\partial Q_\alpha}{\partial x_\zeta} - \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} \frac{\partial Q_\alpha}{\partial P_{\mu\nu}} \right) \quad (\text{H.98})$$

$$\sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \left(\frac{\partial Q_\alpha}{\partial x_\zeta} - \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} \sum_\beta (\Omega_\beta)_{\mu\nu} (\gamma_\alpha - \delta_{\alpha\beta}) \right) \quad (\text{H.99})$$

$$\begin{aligned} \sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} &= \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \left(\frac{\partial Q_\alpha}{\partial x_\zeta} \right. \\ &\quad \left. - \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} \sum_{\beta\phi} (\mathbf{G}^{-1})_{\beta\phi} (\gamma_\alpha - \delta_{\alpha\beta}) \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_\phi|} (\mathbb{I}_k)_{\mu\nu} \right) \quad (\text{H.100}) \end{aligned}$$

$$\begin{aligned} \sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} &= \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \left[\sum_{\beta\phi} \left(e_\phi^\nabla (\mathbf{G}^{-1})_{\beta\phi} + e_\phi (\mathbf{G}^{-1})_{\beta\phi}^\nabla + \Xi (\mathbf{G}^{-1})_{\beta\phi}^\nabla \right) \right. \\ &\quad \left. (\delta_{\alpha\beta} - \gamma_\alpha) + \sum_{\beta\phi} (\mathbf{G}^{-1})_{\beta\phi} (\delta_{\alpha\beta} - \gamma_\alpha) \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_\phi|} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} (\mathbb{I}_k)_{\mu\nu} \right] \quad (\text{H.101}) \end{aligned}$$

$$\begin{aligned} \sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} &= \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \sum_{\beta\phi} (\delta_{\alpha\beta} - \gamma_\alpha) \left[e_\phi^\nabla (\mathbf{G}^{-1})_{\beta\phi} + e_\phi (\mathbf{G}^{-1})_{\beta\phi}^\nabla \right. \\ &\quad \left. + \Xi (\mathbf{G}^{-1})_{\beta\phi}^\nabla + (\mathbf{G}^{-1})_{\beta\phi} \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_\phi|} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} (\mathbb{I}_k)_{\mu\nu} \right] \quad (\text{H.102}) \end{aligned}$$

$$\begin{aligned} \sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} &= \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \sum_{\beta\phi} (\delta_{\alpha\beta} - \gamma_\alpha) \\ &\quad \left[\left(e_\phi^\nabla + \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_\phi|} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} (\mathbb{I}_k)_{\mu\nu} \right) (\mathbf{G}^{-1})_{\beta\phi} + e_\phi (\mathbf{G}^{-1})_{\beta\phi}^\nabla + \Xi (\mathbf{G}^{-1})_{\beta\phi}^\nabla \right] \quad (\text{H.103}) \end{aligned}$$

Begin Sidenote:

$$\text{Let: } \tilde{e}_\phi^\nabla = e_\phi^\nabla + \sum_k \frac{w_k}{|\mathbf{r}_k - \mathbf{r}_\phi|} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}}{\partial x_\zeta} (\mathbb{I}_k)_{\mu\nu}$$

Looking at equation H.84, \tilde{e}_ϕ^∇ is exactly the position derivative of e_ϕ^∇ excluding the position derivative of the density matrix term.

End Sidenote

$$\begin{aligned} \sum_M \frac{\partial M}{\partial x_\zeta} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} &= \sum_\alpha \frac{\partial E^{PI}}{\partial Q_\alpha} \sum_{\beta\phi} \\ &(\delta_{\alpha\beta} - \gamma_\alpha) \left[\tilde{e}_\phi^\nabla (\mathbf{G}^{-1})_{\beta\phi} + e_\phi (\mathbf{G}^{-1})_{\beta\phi}^\nabla + \Xi (\mathbf{G}^{-1})_{\beta\phi}^\nabla \right] \end{aligned} \quad (\text{H.104})$$

w.r.t. MM Atoms

$$\sum_M \frac{\partial M}{\partial x_j} \sum_\alpha \frac{\partial Q_\alpha}{\partial M} \frac{\partial E^{PI}}{\partial Q_\alpha} = 0 \quad (\text{H.105})$$

H.6 ChElPG Rectangular Weights

H.6.1 ChElPG Weighting Scheme

Note:

h is the headspace, d is the distance between grid points (dx), R_i^{vdW} is the vdW radius of atom i

y_k is a long-range weighting function, A_k^J is an atomic switching function

subscript c stand for cutoff, superscript s stands for short, superscript ℓ stands for long

$$R_{c,J}^s = R_J^{vdW} - \frac{d}{200} \quad R_J^{on} = R_{c,J}^s + 0.1 \quad (\text{H.106})$$

$$R^{off} = h + \frac{d}{200} \quad R_c^\ell = R^{off} - 0.1 \quad R_k^{min} = \min_J |\mathbf{R}_k - \mathbf{R}_J| - R_J^{vdW} \quad (\text{H.107})$$

$$w_k = y_k \prod_J^{N_{atoms}} A_k^J \quad (\text{H.108})$$

$$A_k^J = \begin{cases} 0 & \text{if } |\mathbf{R}_k - \mathbf{R}_J| < R_{c,J}^s \\ \tau(|\mathbf{R}_k - \mathbf{R}_J|; R_{c,J}^s; R_J^{on}) & \text{if } R_{c,J}^s < |\mathbf{R}_k - \mathbf{R}_J| < R_J^{on} \\ 1 & \text{if } |\mathbf{R}_k - \mathbf{R}_J| > R_J^{on} \end{cases} \quad (\text{H.109})$$

$$y_k = \begin{cases} 1 & \text{if } R_k^{min} < R_c^\ell \\ 1 - \tau(R_k^{min}; R_c^\ell; R^{off}) & \text{if } R_c^\ell < R_k^{min} < R^{off} \\ 0 & \text{if } R_k^{min} > R^{off} \end{cases} \quad (\text{H.110})$$

$$\tau(R; R_c; R_o) = \frac{(R - R_c)^2 (3R_o - R_c - 2R)}{(R_o - R_c)^3} \quad (\text{H.111})$$

H.6.2 Weight Derivatives

$$\nabla_L w_k = (\nabla_L y_k) \prod_J^{N_{atoms}} A_k^J + y_k (\nabla_L A_k^L) \prod_{J \neq L}^{N_{atoms}} A_k^J \quad (\text{H.112})$$

$$\nabla_L A_k^L = \begin{cases} 0 & \text{if } |\mathbf{R}_k - \mathbf{R}_L| < R_{c,L}^s \\ \tau'(|\mathbf{R}_k - \mathbf{R}_L|; R_{c,L}^s; R_L^{on}) & \text{if } R_{c,L}^s < |\mathbf{R}_k - \mathbf{R}_L| < R_L^{on} \\ 0 & \text{if } |\mathbf{R}_k - \mathbf{R}_L| > R_L^{on} \end{cases} \quad (\text{H.113})$$

$$\nabla_L y_k = \begin{cases} 0 & \text{if } R_k^{min} < R_c^\ell \\ -\tau'(R_k^{min}; R_c^\ell; R^{off}) & \text{if } R_c^\ell < R_k^{min} < R^{off} \text{ and } R_k^{min} = |\mathbf{R}_k - \mathbf{R}_L| \\ 0 & \text{if } R_c^\ell < R_k^{min} < R^{off} \text{ and } R_k^{min} \neq |\mathbf{R}_k - \mathbf{R}_L| \\ 0 & \text{if } R_k^{min} > R^{off} \end{cases} \quad (\text{H.114})$$

$$\tau(R; R_c; R_o) = \frac{(R - R_c)^2 (3R_o - R_c - 2R)}{(R_o - R_c)^3} \quad (\text{H.115})$$

$$\tau'(R; R_c; R_o) = \frac{2(R - R_c) R^x (3R_o - R_c - 2R) - 2R^x (R - R_c)^2}{(R_o - R_c)^3} \quad (\text{H.116})$$

$$\tau'(R; R_c; R_o) = 2(R - R_c) R^x \frac{(3R_o - R_c - 2R) - (R - R_c)}{(R_o - R_c)^3} \quad (\text{H.117})$$

$$\tau'(R; R_c; R_o) = \frac{2(R - R_c) (3R_o - 3R)}{(R_o - R_c)^3} R^x \quad (\text{H.118})$$

$$\tau'(|\mathbf{R}_k - \mathbf{R}_L|; R_c; R_o) = \frac{6(|\mathbf{R}_k - \mathbf{R}_L| - R_c) (R_o - |\mathbf{R}_k - \mathbf{R}_L|)}{(R_o - R_c)^3} \frac{\mathbf{R}_k - \mathbf{R}_L}{|\mathbf{R}_k - \mathbf{R}_L|} \quad (\text{H.119})$$

APPENDIX I

Vectors Inside a Supersphere

The number of vectors are computed by use of a recursion formula. F stands for the floor function. This process need not be done with integers. It will work for any real numbers. n is the number of dimensions and r is the radius of interest.

I.1 All Vectors Inside a Supersphere, N

$$N(n, r) = N(n - 1, r) + 2 \sum_{k=1}^{F(r)} N\left(n - 1, \sqrt{r^2 - k^2}\right) \quad (\text{I.1})$$

I.1.1 $n = 0$

$$N(0, r) = 1 \quad (\text{I.2})$$

I.1.2 $n = 1$

$$N(1, r) = 1 + 2 \sum_{k=1}^{F(r)} N\left(0, \sqrt{r^2 - k^2}\right) \quad (\text{I.3})$$

$$N(1, r) = 1 + 2 \sum_{k=1}^{F(r)} 1 \quad (\text{I.4})$$

$$N(1, r) = 1 + 2F(r) \quad (\text{I.5})$$

I.1.3 $n = 2$

$$N(2, r) = 1 + 2F(r) + 2 \sum_{k=1}^{F(r)} N\left(1, \sqrt{r^2 - k^2}\right) \quad (\text{I.6})$$

$$N(2, r) = 1 + 2F(r) + 2 \sum_{k=1}^{F(r)} \left[1 + 2F\left(\sqrt{r^2 - k^2}\right)\right] \quad (\text{I.7})$$

$$N(2, r) = 1 + 4F(r) + 4 \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) \quad (\text{I.8})$$

I.1.4 $n = 3$

$$N(3, r) = 1 + 4F(r) + 4 \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) + 2 \sum_{j=1}^{F(r)} N\left(2, \sqrt{r^2 - j^2}\right) \quad (\text{I.9})$$

$$\begin{aligned} N(3, r) = 1 + 4F(r) + 4 \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) \\ + 2 \sum_{j=1}^{F(r)} \left[1 + 4F\left(\sqrt{r^2 - j^2}\right) + 4 \sum_{k=1}^{F(r)} F\left(\sqrt{\sqrt{r^2 - j^2} - k^2}\right)\right] \end{aligned} \quad (\text{I.10})$$

$$N(3, r) = 1 + 6F(r) + 12 \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) + 8 \sum_{j=1}^{F(r)} \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - j^2 - k^2}\right) \quad (\text{I.11})$$

This is Sloane's series A000605.

r	$N(3, r)$
0	1
1	7
2	33
3	123
4	257
5	515
6	925
7	1419
8	2109
9	3071
10	4169
11	5575
12	7153
13	9171
14	11513
15	14147
16	17077
17	20479
18	24405
19	28671
20	33401

I.2 All Positive Vectors in Sphere (First Octant), M

$$M(n, r) = M(n-1, r) + \sum_{k=1}^{F(r)} M\left(n-1, \sqrt{r^2 - k^2}\right) \quad (\text{I.12})$$

I.2.1 $n = 0$

$$M(0, r) = 1 \quad (\text{I.13})$$

I.2.2 $n = 1$

$$M(1, r) = 1 + \sum_{k=1}^{F(r)} M\left(0, \sqrt{r^2 - k^2}\right) \quad (\text{I.14})$$

$$M(1, r) = 1 + \sum_{k=1}^{F(r)} 1 \quad (\text{I.15})$$

$$M(1, r) = 1 + F(r) \quad (\text{I.16})$$

I.2.3 $n = 2$

$$M(2, r) = 1 + F(r) + \sum_{k=1}^{F(r)} M\left(1, \sqrt{r^2 - k^2}\right) \quad (\text{I.17})$$

$$M(2, r) = 1 + F(r) + \sum_{k=1}^{F(r)} \left[1 + F\left(\sqrt{r^2 - k^2}\right)\right] \quad (\text{I.18})$$

$$M(2, r) = 1 + 2F(r) + \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) \quad (\text{I.19})$$

I.2.4 $n = 3$

$$M(3, r) = 1 + 2F(r) + \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) + \sum_{j=1}^{F(r)} M\left(2, \sqrt{r^2 - j^2}\right) \quad (\text{I.20})$$

$$\begin{aligned} N(3, r) = & 1 + 2F(r) + \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) \\ & + \sum_{j=1}^{F(r)} \left[1 + 2F\left(\sqrt{r^2 - j^2}\right) + \sum_{k=1}^{F(r)} F\left(\sqrt{\sqrt{r^2 - j^2}^2 - k^2}\right)\right] \end{aligned} \quad (\text{I.21})$$

$$N(3, r) = 1 + 3F(r) + 3 \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - k^2}\right) + \sum_{j=1}^{F(r)} \sum_{k=1}^{F(r)} F\left(\sqrt{r^2 - j^2 - k^2}\right) \quad (\text{I.22})$$

This is Sloane's series A000604.

r	$M(3, r)$
0	1
1	4
2	11
3	29
4	54
5	99
6	163
7	239
8	344
9	486
10	648
11	847
12	1069
13	1355
14	1680
15	2046
16	2446
17	2911
18	3443
19	4022
20	4662

Bibliography

- [1] N. Rega, G. Brancato, and V. Barone, *Chem. Phys. Lett.* **422**, 367 (2006).
- [2] K. Nam, J. Gao, and D. M. York, *J. Chem. Theory Comput.* **1**, 2 (2005).
- [3] Z. C. Holden, R. M. Richard, and J. M. Herbert, *J. Chem. Phys.* **139**, 244108 (2013).
- [4] C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.* **11**, 361 (1990).
- [5] D. Frenkel and B. Smit, Understanding Molecular Simulation: From Algorithms to Applications, Academic Press, San Diego, 2002.
- [6] A. Y. Toukmaji and J. A. Board Jr., *Comput. Phys. Commun.* **95**, 73 (1996).
- [7] G. Hummer, L. R. Pratt, and A. E. García, *J. Phys. Chem.* **100**, 1206 (1996).
- [8] S. Sakane, H. S. Ashbaugh, and R. H. Wood, *J. Phys. Chem. B* **102**, 5673 (1998).
- [9] P. H. Hünenberger and J. A. McCammon, *Biophys. Chem.* **78**, 69 (1999).
- [10] S. W. de Leeuw, J. W. Perram, and E. R. Smith, *Proc. R. Soc. Lond. A* **373**, 27 (1980).
- [11] C. J. Fennell and J. D. Gezelter, *J. Chem. Phys.* **124**, 234104 (2006).

- [12] S. Boresch and O. Steinhauser, *Ber. Bunsenges. Phys. Chem.* **101**, 1019 (1997).
- [13] T. Vreven and K. Morokuma, *Annu. Rep. Comput. Chem.* **2**, 35 (2006).
- [14] L. D. Jacobson and J. M. Herbert, *J. Chem. Phys.* **134**, 094118 (2011).
- [15] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem.* **13**, 225 (1979).
- [16] N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, 1973.
- [17] N. J. A. Sloane, A000605 - oeis, <http://oeis.org/a000605>.
- [18] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- [19] U. Essmann et al., *J. Chem. Phys.* **103**, 8577 (1995).
- [20] J. M. Herbert, L. D. Jacobson, K. U. Lao, and M. A. Rohrdanz, *Phys. Chem. Chem. Phys.* **14**, 7679 (2012).
- [21] L. E. Chirlian and M. M. Francl, *J. Comput. Chem.* **8**, 894 (1987).
- [22] M. A. Spackman, *J. Comput. Chem.* **17**, 1 (1996).
- [23] Y. Shao et al., *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).
- [24] A. I. Krylov and P. M. W. Gill, *WIREs Comput. Mol. Sci.* **3**, 317 (2013).
- [25] R. C. Walker, M. F. Crowley, and D. A. Case, *J. Chem. Theory Comput.* **29**, 1019 (2007).

- [26] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Imprey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [27] A. Bondi, *J. Phys. Chem.* **68**, 441 (1964).
- [28] A. W. Lange and J. M. Herbert, *J. Chem. Phys.* **133**, 244111 (2010).
- [29] C.-M. Chang, Y. Shao, and J. Kong, *J. Chem. Phys.* **136**, 114112 (2012).
- [30] A. Alvarez-Ibarra, A. M. Köster, R. Zhang, and D. R. Salahub, *J. Chem. Theory Comput.* **8**, 4232 (2012).