THE THEORY OF EWALD SUMMATION
APPLIED TO QM/MM SYSTEMS

THESIS

Presented in Partial Fulfillment of the Requirements for
the Degree Master of Science in the Graduate
School of The Ohio State University

By

Zachary Conner Holden, B.S.
Graduate Program in Chemistry

* * * * *

The Ohio State University

2012

Thesis Committee:

John M. Herbert, Adviser
Sherwin J. Singer
Abstract

Theoretical models of condensed phases can be quite large and computationally expensive because of the number of atoms used. The computational expense has been resolved due to QM-MM methods; however, the number of atoms can still be quite large. Periodic boundary conditions can be used to propagate a system in space without explicitly representing atoms. Periodic boundary conditions can be used with QM-MM methods by making the appropriate changes to the Fock matrix. The change is a result of performing a modified Ewald summation using the QM atoms. This combined Ewald QM-MM method is a way to theoretically construct condensed phase systems without explicitly specifying an overabundance of atoms. The method can then be applied to various types of systems that are sensitive to their environments.
In memory of Dr. Peter E. Siska who first got me interested in electronic structure.
I am truly indebted and thankful to JMH for his patience and understanding throughout my graduate school career. I owe my position in graduate school to my parents who have helped me, both emotionally and financially, get through these many years of school. I would also like to thank OSU for giving me a chance to show my potential.
VITA

2005 ...........................................Chambersburg Senior High School

2009 ...........................................B.S. Chemistry, B.S. Physics, University of Pittsburgh

2009 to present ..............................Graduate Teaching Associate, The Ohio State University

FIELDS OF STUDY

Major Field: Chemistry
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>The Ewald Summation</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Energy Partitioning and Embedding Schemes</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>Derivation of the Fock Matrix Correction</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>Future Work</td>
<td>17</td>
</tr>
</tbody>
</table>

BIBLIOGRAPHY ........................................ 20
LIST OF FIGURES

FIGURE  PAGE

1  A comparison of the terms in the Ewald summation. ($\eta = 0.5$ Å) . . 4
2  A depiction of a macrocell inside a dielectric material. . . . . . . . 8
3  A macrocell showing replicas of the QM/MM simulation cell. . . . . 10
CHAPTER 1

Introduction

One of the largest problems in theoretical chemistry today is how to quickly and accurately model condensed phase systems. Gas systems are easy to model because their motion is completely random. When calculating the electronics of gas molecules, the approximation that the molecule is isolated is an acceptable approximation; therefore, it is possible to ignore long range effects. Solids are completely ordered. The long range effects can be accurately captured using the Madelung constant for simple systems. For liquids none of these approximations are possible. Liquids are quasi-random. The motion can be modeled by either Molecular Dynamics or Monte Carlo simulations, but the electronics are difficult to model. The long range interactions cannot be ignored because the motion of the molecules is neither completely random nor is it ordered; therefore, it cannot be assumed that the interactions will average to 0. The most accurate way to obtain the electronics of a liquid is to do a quantum mechanical calculation over a large number of molecules; however, this is impractical because all QM calculations scale non-linearly in time and CPU usage with the number of basis functions and hence the number of atoms. The scaling is particular to the method used. One method used to attempt to reduce the computational cost of the
calculation is to use a QM/MM method. This method is done by taking a small QM system inside of a much larger MM system. The cost of the QM calculation greatly outweighs the MM calculation, so it is possible to make the MM system extremely large without increasing the cost of the calculation by a large amount. This enables the calculation to include more long range effects. It is difficult to make the QM-MM system large enough to capture all long range effects with explicit atoms because of the $\frac{1}{r}$ dependence of the Coulomb potential. Periodic boundary conditions can eliminate the need for explicit atoms. The periodic boundaries are evaluated through the use of the Ewald summation. Historically Ewald summation has been used for MM systems. It is possible to apply the Ewald method to the QM system in order to generate the Ewald QM/MM method by appropriate modification the Fock matrix.
CHAPTER 2

The Ewald Summation

The Coulomb sum is a slowly convergent sum because of the $\frac{1}{r}$ dependence. It is possible to use periodic boundary conditions instead of specifying all the atoms explicitly. Periodic boundary conditions require a user defined simulation cell and replicating it an infinite number of times in every direction. The Ewald summation\textsuperscript{1,2} entails taking a slowly convergent sum and transforming it into two quickly convergent sums by use of the error function.

$$E_{cc} = \frac{1}{2} \sum_n \sum_{ij} q_i q_j \left[ \frac{\text{erf} (\eta |\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} + \frac{\text{erfc} (\eta |\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} \right]$$

A user chosen parameter is represented by $\eta$ which controls how much of the sum is evaluated by the erf and erfc terms. Often $\alpha$ is used in Ewald summation; however, in this paper $\eta$ will be used in order to avoid confusion later with $\alpha$ being a QM atom.

There is a quickly convergent sum (erfc) and slowly convergent sum (erf) shown in figure 1. This formulation makes it easy to see the convergence. In practice the goal is to modify the equation so one sum is performed in real space and one in reciprocal space. The easiest formulation is to transform the second term in equation (2.1) by
using the identity

$$|r_{ij} + n|^{-2s} = \frac{1}{\Gamma(s)} \int_0^\infty dt \ t^{s-1} e^{-t|r_{ij} + n|^2}$$  \hspace{1cm} (2.2)$$

where $n$ is the vector that represents how many box lengths the two atoms are from one another.

$$E_{cc} = \frac{1}{2} \lim_{\sigma \to 0} \sum_n \sum_{ij} \frac{q_i q_j}{\sqrt{\pi}} \left[ \int_0^{\eta^2} dt \ t^{-\frac{1}{2}} e^{-t|r_{ij} + n|^2} e^{-\sigma n^2} + \int_{\eta^2}^{\infty} dt \ t^{-\frac{1}{2}} e^{-t|r_{ij} + n|^2} e^{-\sigma n^2} \right]$$ \hspace{1cm} (2.3)$$

is obtained by plugging equation (2.2) into equation (2.1). $\sigma$ is introduced because of a singularity that exists at 0 and the prime on the $ij$ sum denotes that $i \neq j$ when $n = 0$. The first term in (2.3) converges quickly. This term is named the "real space

Figure 1: A comparison of the terms in the Ewald summation. ($\eta = 0.5 \text{ Å}$)
term” because it is summed in real space and reduces to equation (2.4).

\[ E_{\text{real}} = \frac{1}{2} \sum_n \sum_{ij} q_i q_j \frac{\text{erfc} \left( \frac{\eta |r_{ij} + n|}{|r_{ij} + n|} \right)}{|r_{ij} + n|} \]  

(2.4)

There is a small implementation issue with this term so that interactions are not overcounted. Often in force fields 1-2, 1-3, 1-4, and sometimes 1-5 interactions are treated separately from the Coulombic contribution to the MM energy. If \( i \) and \( j \) run over all atoms in the system, then a correction term (2.5) is introduced.

\[ E_{\text{corr}} = \sum_{ij \in 1-2} \sum_{ij \in 1-3} \sum_{ij \in 1-4} \sum_{ij \in 1-5} -\kappa_{ij} \frac{q_i q_j}{|r_{ij}|} \]  

(2.5)

\( \kappa_{ij} \) is a damping parameter that is dependent on the force field used. Often \( \kappa_{ij} \) will take on the values in (2.6). It is possible to avoid calculating this term if more restrictions are introduced onto the \( ij \) sum in (2.4).

\[ \kappa_{ij} = \begin{cases} 1 & \text{if 1-2 or 1-3} \\ 1 \leq \kappa_{ij} \leq 0.5 & \text{if 1-4} \\ 0.5 \leq \kappa_{ij} \leq 0 & \text{if 1-5} \end{cases} \]  

(2.6)

The second term in equation (2.3) does not converge quickly; therefore, to make the sum converge quickly it is performed in reciprocal space. To accomplish this a few modifications must be made to the term. When performing spatial transformations, it is easier if there are no restrictions on the sums. This has been obtained by adding and subtracting the \( i = j \) and \( n = 0 \) term as shown in (2.7).

\[ E_{\text{nonreal}} = \frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_n \sum_{ij} q_i q_j \frac{\eta^2}{\sqrt{\pi}} \int_0^{\eta} dt \, t^{-\frac{1}{2}} e^{-\frac{t}{2}|r_{ij} + n|^2} e^{-\sigma n^2} \]

\[ -\frac{1}{2} \lim_{\sigma \rightarrow 0} \sum_n \sum_{ij} q_i q_j \frac{\eta^2}{\sqrt{\pi}} \int_0^{\eta} dt \, t^{-\frac{1}{2}} e^{-\frac{t}{2}|r_{ij} + n|^2} e^{-\sigma n^2} \delta_{ij} \delta_{n0} \]  

(2.7)
The first term in (2.7) can now be transformed into reciprocal space by utilizing the Jacobi Imaginary Transformation for $\theta$-functions:

$$
\sum_{n} e^{-t|\mathbf{r}+n|^2} = \frac{1}{V} \left( \frac{\pi}{t} \right)^{\frac{3}{2}} \sum_{\mathbf{k}} e^{-\frac{\mathbf{k}^2}{4t} + i\mathbf{k} \cdot \mathbf{r}}
$$

(2.8)

where $V$ is the volume and $\mathbf{k}$ is the reciprocal space vector. There is still a singularity at $t = 0$ when $\mathbf{k} = \mathbf{0}$ so that term can be treated separately. The rest of the term simplifies down to become $E_{\text{recip}}$.

$$
E_{\text{recip}} = \frac{1}{2} \frac{4\pi}{V} \sum_{k \neq 0} e^{-\frac{k^2}{4\eta^2}} \sum_{ij} q_i q_j e^{i\mathbf{r}_{ij} \cdot \mathbf{k}}
$$

(2.9)

It is worth mentioning that the reciprocal term can be found using equation (2.9); however, it is not the most practical way to find this term. There is a double sum over indices $i$ and $j$. By using Euler’s relation for imaginary numbers it is possible to simplify this term down to the addition of single sums which can be performed simultaneously.

$$
E_{\text{recip}} = \frac{1}{2} \frac{4\pi}{V} \sum_{k \neq 0} e^{-\frac{k^2}{4\eta^2}} \left\{ \left[ \sum_{i} q_i \cos (\mathbf{r}_i \cdot \mathbf{k}) \right]^2 + \left[ \sum_{i} q_i \sin (\mathbf{r}_i \cdot \mathbf{k}) \right]^2 \right\}
$$

(2.10)

Even though this equation may not look simpler it is easier to program and faster to evaluate. The previous sum had $N^2$ terms and required the use of complex numbers. It has now been transformed into a sum with $2N$ terms without complex variables.

The whole reciprocal term, including the sum over $\mathbf{k}$, scales as $2NN_{\text{recip}}$ where $N_{\text{recip}}$ is the number of reciprocal vectors. This can still be an expensive process.

It is possible to evaluate the sum over $i$ and $j$ using a discrete grid (which scales as $\log N$). This can be done by two different methods. The particles can either lie on the
mesh points and this process is known as Particle Mesh Ewald (PME). Alternatively the atom can lie in between the mesh points which is called Smooth Particle Mesh Ewald (SPME). The sum is then evaluated using Lagrange interpolation or cardinal B-Splines. At the current time, PME has not been implemented in the program QChem.

Returning to the second term in (2.7), this is the term that is coined the self term. It reduces to equation (2.11).

$$E_{self} = -\frac{1}{2} \sum_i q_i^2 \frac{\eta^2}{\sqrt{\pi}} \int_0^\eta dt \ t^{-\frac{1}{2}} = -\frac{\eta}{\sqrt{\pi}} \sum_i q_i^2$$  \hspace{1cm} (2.11)

This term arises solely because of the restriction in (2.3). If $i$ and $j$ are over a different group of particles, then the self term does not exist, as will be observed in a later result.

The $k = 0$ term that resulted from (2.8) is evaluated and two terms result.

$$E_{surf} = -\frac{\pi}{3V} \sum_{ij} q_i q_j r_{ij}^2$$  \hspace{1cm} (2.12)

This is the surface term because it is dependent on the surface outside of the replicated cells. If the system is charge neutral, it is possible to simplify $E_{surf}$.

$$E_{surf} = -\frac{2\pi}{3V} \left( \sum_i q_i r_i \right)^2$$  \hspace{1cm} (2.13)

The meaning of the surface term can be interpreted by imagining that the simulation cell is replicated a finite number of times. This macrocell is placed within a spherical cavity inside some dielectric material as shown in figure 2.

This term is the energy that arises from the macrocell interacting with the surface of the material. When alternate derivation is performed using a dielectric material
outside of the macrocell instead of vacuum the denominator of 3 is replaced with the quantity $2\varepsilon + 1$. There are two dielectrics that are commonly utilized as the material outside the macrocell. One is to use vacuum, as the material ($\varepsilon = 1$). The more common convention is to use tin-foil boundary conditions ($\varepsilon = \infty$). The benefit of using tin-foil is that it results in $E_{\text{surf}} = 0$. For large systems this term can be sizeable and should not be neglected since a system is rarely placed inside a metal to be analyzed.

When a system is not charge neutral it must be placed inside a uniform charge compensating background. If this is not done, the Coulomb sum becomes infinite and Ewald summation is neither accurate nor appropriate. A charge term arises out of the singularity that occurs in the $\boldsymbol{k} = \boldsymbol{0}$ from the transformation in (2.8).
\[ E_{\text{charge}} = -\frac{Q^2 \pi}{2V \eta^2} \quad (2.14) \]

Q is the charge of the simulation cell. The charge term is 0 if the system is neutral (2.14). Once all the terms have been found equations (2.4), (2.9), (2.11), (2.12), and (2.14) can be combined so that the complete cell energy can be calculated.
CHAPTER 3

Energy Partitioning and Embedding Schemes

It is possible to capture long range effects by using periodic boundary conditions instead of using an extremely large system as shown in figure 3. It is ideal to formulate a theory which can be used with various QM and QM/MM features. When the energy is partitioned, it must be calculated through a procedure where the terms are independent, so that the calculation of one term does not affect the calculation of other

Figure 3: A macrocell showing replicas of the QM/MM simulation cell.
terms. Using a QM-MM system it is possible to treat each region separately so there
are three types of interactions (QM-QM, QM-MM, and MM-MM). The treatment will
be different whether computing the interaction between two atoms in the simulation
cell (real space (RS)) or an atom in the simulation cell and another atom in an image
cell (periodic boundary conditions (PBC)). This will yield six different energy terms
to compute

\[ E^{PB} = E_{QM-QM}^{RS} + \Delta E_{QM-QM}^{PBC} + E_{QM-MM}^{RS} + \Delta E_{QM-MM}^{PBC} + E_{MM-MM}^{RS} + \Delta E_{MM-MM}^{PBC} \] (3.1)

where \( E^{PB} \) is the energy of total system which is over all space. All MM-MM terms
are treated with a standard MM Ewald formulation. \( E_{QM-QM}^{RS} \) is treated by the user’s
choice self-consistent field (SCF) procedure. \( E_{QM-MM}^{RS} \) can be treated by the user’s
choice link atom technique. This leaves only two terms to be evaluated.

\[ \Delta E^{PBC} = \Delta E_{QM-QM}^{PBC} + \Delta E_{QM-MM}^{PBC} \] (3.2)

An embedding scheme must chosen to treat these interactions. The atoms can
be treated by using mechanical embedding, which means they are treated classically
and are integrated into the MM portion of the procedure. Conversely they can also
be treated by electrical embedding, which means they are treated as a correction to
the Fock matrix. Mechanical embedding is typically done by use of

\[ E_{total} = E_{MM}(QM + MM) + E_{QM}(QM) - E_{MM}(QM) \] (3.3)

where the subscript is the method used and the subsystem is in parenthesis. This
approach completely ignores the effect that the MM subsystem atoms can have on the
electronic structure of the QM subsystem. This is obviously very detrimental when examining solvent effects, charge transfer systems, or electronic states lying close in energy. As a further drawback, it also requires MM parameters for the QM system. Trying to determine the MM parameters through an algorithm can increase amount of computational time greatly.\textsuperscript{7}

Electrical embedding is treated by means of a Fock matrix correction. While this method is preferred over mechanical embedding, it is more expensive. The correction has to be updated for every SCF cycle. Calculating the correction is generally a much faster process than doing the QM calculation itself; however, if the QM region is small and the MM region is large, the correction itself may become the longest process in the calculation. This embedding scheme includes the effect that MM atoms have on the electronics of the QM subsystem. The MM parameters required are usually readily available from standard force fields. This is the embedding scheme chosen for the QM/MM Ewald formulation.\textsuperscript{7}
CHAPTER 4

Derivation of the Fock Matrix Correction

The precursor to this method was developed by Nam, Gao, and York. Up to the current time this method has always been employed using Neglect of Diatomic Differential Overlap (NDDO) and other semi-empirical methods. The goal of this work is to create a generalized method that does not use NDDO and can be combined with different SCF methods.

The Fock matrix is defined as the derivative of energy with respect to density matrix. The total energy can be defined as the real space energy plus a correction term due to the periodic boundary conditions. When the derivative of energy is taken with respect to density matrix this leads to the corrected Fock matrix that is generated by using an SCF procedure plus the correction introduced by this method. The correction term is the total energy of all the periodic replicas minus the energy that is required in real space. The total energy is known and is the result of the Ewald summation. The Ewald sum is in terms of charges and not densities. In order to make use of the sum, it is assumed that the atoms are far enough away to approximate their densities by point charges so Mulliken charges will be shown as an example in
Mulliken charges are defined by

\[ Q_\alpha = Z_\alpha - \sum_{\mu \nu \in \alpha} \rho_{\mu \nu} S_{\mu \nu} \]  \hspace{1cm} (4.1)

The derivative of energy with respect to density matrix may be transformed into a derivative of energy with respect to the charge.

\[ \frac{\partial \Delta E^{PBC}}{\partial \rho_{\mu \nu}} = \sum_{\alpha \in \mu \nu} \frac{\partial Q_\alpha}{\partial \rho_{\mu \nu}} \frac{\partial \Delta E^{PBC}}{\partial Q_\alpha} = \sum_{\alpha \in \mu \nu} -S_{\mu \nu} \frac{\partial \Delta E^{PBC}}{\partial Q_\alpha} \]  \hspace{1cm} (4.2)

A different set of charges could be used instead of Mulliken, such as Löwdin or ChElP-G. The only variable that would change in this scheme is the derivative of charge with respect to the density matrix. Analysis of (4.2) reveals that the correction to the Fock matrix will be block diagonal due to the fact that \( \mu \) and \( \nu \) must be centered on the atom \( \alpha \). The energy can be further broken down into a QM-QM term and a QM-MM term. A potential can be defined as a function of the distance between the two atoms being analyzed.

\[ \Delta E^{PBC} = \frac{1}{2} \sum_{\alpha \beta} Q_\alpha Q_\beta \psi (R_{\alpha \beta}) + \sum_{\alpha} \sum_{i} Q_\alpha q_i \psi (R_{\alpha i}) \]  \hspace{1cm} (4.3)

The derivative of the energy with respect to the charge is given by

\[ \frac{\partial \Delta E^{PBC}}{\partial Q_\alpha} = \sum_{\beta \in \mu \nu} Q_\beta \psi (R_{\alpha \beta}) + \sum_{i} q_i \psi (R_{\alpha i}) \]  \hspace{1cm} (4.4)

The same potential can be used for both QM-QM interactions and QM-MM interactions. It can be broken up into terms just like the energy in Ewald.

\[ \psi (R_{\alpha \beta}) = \psi_{\text{real}} (R_{\alpha \beta}) + \psi_{\text{recip}} (R_{\alpha \beta}) \]  \hspace{1cm} (4.5)
There are many similarities between normal MM/MM Ewald and QM/MM Ewald. An SCF calculation is performed and because the interaction is not limited to pairs of atoms within the same molecule in the QM-QM calculation the correction term, \((-\frac{1}{r})\), arises between all pairs of atoms. The same correction applies to the QM-MM term because all QM-MM pairs within the simulation cell have been included by some QM/MM scheme. This correction only affects the pairs that are within the simulation cell \((n = 0)\).

\[
\psi_{\text{real}}(R_{\alpha\beta}) = -\frac{\text{erf}(\eta |R_{\alpha\beta}|)}{|R_{\alpha\beta}|} + \sum_{n \neq 0} \frac{\text{erfc}(\eta |R_{\alpha\beta} + n|)}{|R_{\alpha\beta}|} \quad (4.6)
\]

There is no self term in the QM-MM term because an atom cannot be both QM and MM. The QM-QM self term is included within the real space sum. The sum over \(\beta\) is unrestricted so upon taking the limit as the distance goes to 0 a term analogous to the classical self term arises.

\[
\lim_{r \to 0} \frac{\text{erf}(\eta r)}{r} = \frac{2\eta}{\sqrt{\pi}} \quad (4.7)
\]

The reciprocal term is the same except for how the imaginary exponential is treated. The previous treatment (2.10) is not appropriate for this situation because that transformation assumes that both atoms are within the same subsystem (\(i\) and \(j\) are both MM atoms). Both atoms being in the same subsystem cannot be assumed in this case. The k-vectors used contain only positive numbers when a different transformation are used.

\[
\psi_{\text{recip}}(R_{\alpha\beta}) = \frac{4\pi}{V} \sum_{k > 0} \frac{\exp(-k^2/4\eta^2)}{k^2} \cos(k \cdot R_{\alpha\beta}) \quad (4.8)
\]
The gradients for the potential are easily determined analytically. The reciprocal term in the gradient of the potential is not much more complicated than the reciprocal term in the potential.

\[ \nabla \psi_{\text{recip}}(\mathbf{R}_{\alpha\beta}) = -\frac{4\pi}{V} \sum_{k>0} \frac{\exp(-k^2/4\eta^2)}{k^2} \sin(k \cdot \mathbf{R}_{\alpha\beta}) \mathbf{k} \] (4.9)

The gradient of the real term is more complicated.

\[ \nabla \psi_{\text{real}}(\mathbf{R}_{\alpha\beta}) = \left[ \frac{\text{erf}(\eta |\mathbf{R}_{\alpha\beta}|)}{|\mathbf{R}_{\alpha\beta}|} - \frac{2}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{R}_{\alpha\beta}|^2} \right] \frac{\mathbf{R}_{\alpha\beta}}{|\mathbf{R}_{\alpha\beta}|^2} \]

\[ - \sum_{n \neq 0} \left( \frac{\text{erfc}(\eta |\mathbf{R}_{\alpha\beta} + \mathbf{n}|)}{|\mathbf{R}_{\alpha\beta} + \mathbf{n}|} + \frac{2}{\sqrt{\pi}} e^{-\eta^2 |\mathbf{R}_{\alpha\beta} + \mathbf{n}|^2} \right) \frac{\mathbf{R}_{\alpha\beta} + \mathbf{n}}{|\mathbf{R}_{\alpha\beta} + \mathbf{n}|^2} \] (4.10)

In order to speed up the calculation values can be stored in memory or the \( \eta \) value can be optimized. There is no need to calculate these terms twice because there are parts of the gradient that are similar to the potential so they can be stored in memory.

An easier solution to this problem is to choose an appropriate \( \eta \) value that will make the real space gradient go to 0 very quickly. If this is done, a cutoff radius can be used for the real term once the convergence criteria has been chosen. It must be ensured that there are enough reciprocal space vectors when performing the sum. The less interactions that are performed in real space means there are more interactions to perform in reciprocal space. It is not apparent what the optimal \( \eta \) value in the Ewald QM/MM method.
CHAPTER 5

Future Work

There are many ways that this theory can be utilized. At the current time there have been no benchmarking studies for this theory. Even with the NDDO approximation, no benchmarking studies have been performed. Most of the current studies have centered on implementation and comparison of this method with other methods. At the current time no studies have been conducted on choosing an $\eta$ value. The $\eta$ value determines the number of reciprocal vectors needed in the calculation. A difference will be made in the time it takes to construct the Fock matrix correction if an optimal $\eta$ value is used. How a cutoff radius effects the accuracy of the calculation should be investigated as well. The limitations of the theory have not be discovered because previous research focuses on comparing this method to other methods. The limitations of the theory must be known so that the method can be applied to appropriate systems.

One type of system that could benefit from the use of the Ewald QM/MM method is bioinorganic chemistry. It has been a rapidly growing field in recent times; however, it is important to know how the metal centers in the molecules function before they are put in vivo. Most metals are poisonous in one way or another, especially when the
metal complex includes organic substituents such as dimethyl mercury (H₃CHgCH₃); therefore, it is essential to know how an inorganic molecule will react or cleave *in vivo*.

An important factor in knowing how the molecule will react is the HOMO-LUMO energy gap. It is a well known fact that the solvent can affect the energy gap between these orbitals.⁸ In order to represent these energy levels the solvent must be included in the calculation. With this method it would be possible to set up a calculation where there are many solvent molecules. Metals are computationally expensive to compute because of the number of basis functions needed to represent the atomic orbitals. It is not feasible to include many, if any, solvation shells in a QM calculation because all SCF and post-SCF methods scale as the number of basis functions to some power. Mechanical embedding cannot be used to determine the energetics of the metal. With the electrical embedding in this scheme it is possible to include many solvent molecules without making the QM part of the calculation significantly more expensive. In an organometallic reaction often charged species are formed and this could result in the need to implement a polarizable solvent. The benefit of performing these calculations with periodic replications rather than a cluster calculation is that it helps to ensure that there are enough solvent-metal ion interactions in the calculation.

Often density functional theory (DFT) is used for inorganic calculations because it is a monodeterminant method and is computationally cheap compared to other methods. It has been demonstrated that photoelectron binding energies in the gas
phase calculated with Møller-Plesset Perturbation Theory (MP2) are closer to Coupled Cluster (CCSD(T)) energies than are DFT energies. This is because occupancy of the d atomic orbitals calculated by DFT is greater than that calculated by MP2. This leads to overstabilization of the complex which results in the binding energies being lower.\textsuperscript{9} Overstabilization helps to explain the controversy that exists when comparing properties of the molecule to experiment. Some researchers claim that DFT fails because of the overstabilization.\textsuperscript{10} Others claim that an accurate environment is not being modeled in the simulations. Experimental values, such as bond distances and angles, are achieved by observing the solid phase. The claim is that boundary effects are non-negligible in these systems; therefore, it is not possible to compare gas phase system calculations with condensed phase system data.\textsuperscript{11} It may become possible to finally resolve this debate using the Ewald QM/MM method with high levels of theory.
Bibliography


