**Введение**

Межмолекулярное взаимодействие является ключевым фактором как при исследовании органических систем на молекулярном уровне, так и при разработке новых материалов с заданными свойствами, которые находят свое применение в медицине, промышленности и сельском хозяйстве. Межмолекулярное взаимодействие определяет существование жидкостей, молекулярных кристаллов и реальных газов. Атомные заряды определяют электростатическое взаимодействие, которое нередко доминирует в межмолекулярном взаимодействии. Это объясняется тем, что кулоновский потенциал относительно медленно убывает как функция расстояния.

Атомные заряды позволяют объяснять важные качественные закономерности реакционной способности молекул и невалентного межмолекулярного взаимодействия. Аппарат частичных атомных зарядов нашел свое применение в различных областях химии, в частности в органической химии для теоритического обоснования свойств химических веществ. Результаты описаний и обоснований хорошо согласуются с экспериментальными данными, что свидетельствует о наличие связи между распределениями зарядовой плотности и физико-химическими свойствами молекулы.

Различные области молекулярного моделирования (докинг, молекулярная динамика), а так же другие области применения атомных зарядов сформировали целый ряд требований к методам расчета зарядов. Высокая роль электростатических взаимодействий определяет важность быстрой и качественной оценки атомных зарядов в молекулярном моделировании. Современные методы должны учитывать различные эффекты, приводящие к изменению атомной зарядовой плотности, а это включает поляризацию и перенос заряда. Предложенные, в настоящие время, схемы расчета атомных зарядов не являются универсальными, а их теоретическое обоснование не всегда согласовано с классической физической теорией. Данная проблема существенно осложняет последующую модификацию этих методов для улучшения точности результатов и скорости вычисления. Изучение потребностей различных областей показывает, что, в настоящий момент, существует острая потребность в новых эмпирических схемах расчета атомных зарядов, применимых как для небольших неорганических структур, так и для многоатомных молекул. Это означает, что задача разработки и анализа таких методов является актуальной.

Приведем простой пример:

(Формула расчета)

Этот пример показывает, что неточность в определении заряда в пределах 10-2е уже существенно влияет на значение энергии взаимодействия этих зарядов.

**Анализ существующих методов**

Для расчета атомных зарядов применяют различные подходы, которые сильно отличаются по трудоемкости и качеству. Ниже представлен краткий обзор основных методов.

**Квантово-химические методы**

Квантово-химические методы являются наиболее физически обоснованными. Суть методов этой группы заключается в подгоне атомных зарядов под молекулярный электростатический потенциал, полученный из решения уравнений квантовой механики. Метод получил большую популярность из-за наилучшей согласованности найденных атомных зарядов с экспериментальными данными. По этой причине, заряды, полученные этим методом, являются эталонными для других эмпирических методов.

Первым широко используемым квантово-химическим методом стал ESP-метод, который, несмотря на отличные результаты, обладает следующими недостатками:

* неэмпирический трудоемкий расчет
* зависимость от конформации молекулы

Данные недостатки могут приводить к артефактам в конформационном анализе и молекулярно-динамическом моделировании.

Дальнейшие работы в области квантово-химических методов смогли решить часть проблем ESP-схем. Следует отметить особым вниманием RESP-метод, который, в настоящее время, считается основным представителем из этой группы методов.

Квантово-химические методы имеют высокую математическую сложность, но нашли широкое применение во многих программных пакетах молекулярного моделирования. В частности, RESP-заряды используются при моделировании в пакете AMBER.

**Топологические методы**

Первые публикации, посвященные топологическим методам, появились в 1991 году. В то время наука имела представление о распределении атомных зарядов для небольшого числа аминокислот, которые достаточно точно соответствовали экспериментальным данным. Следует добавить, что эти заряды были фиксированными для молекулы и не менялись в различных электростатических средах. Это приводило к появлению различных артефактов в молекулярной механике и динамике. Общая схема методов этой группы основана на принципе выравнивания электроотрицательностей атомов при образовании молекулы. Сущность этих моделей в том, что соседние атомы «борются» за электроны и окончательная картина распределения зарядов согласована с «силами влияния» атомов (электроотрицательность) на заряды. Таким образом, из информации о структуре необходима только связность или молекулярная топология, поэтому методы называются топологическими.

**QEq-модель**

Основой метода является функция энергии самого общего вида, включающая члены разложения до 2-го порядка включительно.

Где i и j – это индексы атомов, , и – некоторые параметры. Очевидно, что параметр J должен иметь асимптотическое поведение как (взаимодействие точечных зарядов), а при малых (на расстоянии химической связи и менее) его поведение может быть существенно сложнее.

Одноатомная часть отвечает за притяжение заряда к данному атому и берется из общей зависимости энергии атома от заряда.

Сделав допущение, что рассматриваем энергию около заряда, равного нулю, получаем:

Отсюда следует:

Где *IP* - потенциал ионизации, *EA* - сродство к электрону

Параметр известен как элетроотрицательность, а как жесткость. Оба параметры легко определяются экспериментально.

Коэффициент межатомного взаимодействия определяется как экранированный кулоновский член, вычисляемый как интеграл по ns-орбиталям Слэтеровского(STO) или Гауссова(GTO) типа:

Где и обозначают положение электронов и ядра, а – расстояние между i-м и j-м атомом.

Имея все необходимые параметры, атомные заряды вычисляется минимизацией функции энергии системы с условием сохранения полного заряда системы , что эквивалентно полному выравниванию электроотрицательностей атомов. Данный подход нахождения атомных зарядов получил название QEq(charge equilibrium). Вычислительная сложность данного метода для самого общего случая равна .

Одним из фундаментальных недостатков QEq-модели является проблема диссоциации, то есть получение некорректных значений зарядов атомов в состоянии разрыва связей. Заряд может быть перенесен между любыми атомами, даже теми, которые находятся за сотни ангстрем друг от друга.

Действительно, рассмотрим простейшую систему из двух атомов с полным нулевым зарядом. Тогда имеем:

Сделав минимизацию по q получим:

|  |  |
| --- | --- |
|  | (1) |

Решение этой проблемы было предложено в модели QTPIE.

**QTPIE-модель**

Фундаментальными переменными в QTPIE-модели являются переменные, связанные с переносом заряда (интегральные токи , где i и j – индексы атомов), а не частичные атомные заряды как в QEq схеме. Переменные описывают поток поляризации, то есть склонность к перемещению электронной плотности от одного атома к другому. Интегральные токи описывают количество заряда перенесенного от одного атома к другому и связываются с общим атомным зарядом как:

Стоит отметить антисимметричность интегральных токов .

Используя аппарат интегральных токов, можно переписать уравнение энергии системы как:

|  |  |
| --- | --- |
|  | (2) |

где – жесткость атома.

В данное уравнение вводиться функция ограничения переноса заряда:

|  |  |
| --- | --- |
|  | (3) |

где коэффициент – коэффициент масштабирования. Тогда уравнение энергии принимает следующий вид:

Функция определенна как интеграл перекрытия атомных орбиталей, а значит Другими словами, это функция ограничивает перенос заряда на дальние расстояния. Если мы вернемся к изначальному примеру двухатомной молекулы, тогда уравнение (1) примет вид:

Как мы видим, это уравнение имеет правильное асимптотическое поведение, а, значит, QTPIE-модель дает более точные результат, по сравнению с QEq-моделью.

Интегральные токи находятся минимизацией уравнения полной энергии (2). QTPIE-модель использует идентичные значения параметров для электроотрицательности, жесткости и орбитальной геометрии, что и QEq-модель. Дополнительная оптимизация этих параметров может положительно сказаться на результатах.

Классическая схема QTPIE не является вычислительно эффективной, поскольку общее уравнение энергии содержит неизвестных переменных, тогда как QEq схема использует n неизвестных переменных. Метод QTPIE корректно работают со слабовзаимодействующими атомами, то есть решает проблему диссоциации, а также более точно учитывает молекулярную поляризацию. Авторы отмечают, что вычислительная сложность их метода, при использовании стандартных способов решения линейных уравнений, связанных с обращением матриц, в самом общем случае равна , что на порядок выше сложность QEq-модели, у которой вычислительная сложность равна .

**Динамическая модель флуктуирующих зарядов**

*Динамика движения заряда*

Одним из путей улучшения существующих схем флуктуирующих зарядов является введение динамики течения заряда, то есть использование реальных токов, которые зависят от времени. Следует напомнить, что в QTPIE-модели используется аппарат интегральных токов, который показывает общее количество заряда, перенесенное от одного атома к другому. Динамика движение заряда вводиться добавлением кинетического члена в полное уравнение энергии. Данный кинетический член должен зависеть от элементарных токов, которые возникают между атомами с разными зарядами. Тогда общее уравнение энергии должно принимать следующий вид:

Определим кинетический член как:

где функции являются функциями, ограничивающими перенос заряда, введенные по аналогии с (3).

Потенциальный вклад в общее уравнение энергии определяется членом:

В таком случае, модельный гамильтониан может выглядеть как:

Данное уравнение не может быть использовано для получения уравнения движения, по причине того что переменные и не являются канонически сопряженными. Кроме того, число переменных по порядку величины равно , а число переменных имеет порядок .

Чтобы найти уравнение движения заряда надо уровнять количество переменных и , и добиться их сопряжения.

*Ограничения на модель*

Для уменьшения количества переменных , можно ограничить модель требованием, что токи возможно только между атомами, имеющие химическую связь. Это требование позволяет существенно сократить порядок переменных . В этом случае число элементарных токов будет связано с числом зарядов соотношением:

где – число простых циклов (ароматических колец) в системе.

Дополнительные ограничения берутся из фундаментальных соображений. К элементарным токам можно применить первое правило Кирхгофа, которое гласит, что алгебраическая сумма токов в любом узле любой цепи равно нуля. В нашем случае, этот закон запишется так:

Последнее ограничение вводит требование на сохранение полного заряда системы и записывается следующим способом:

Вышеупомянутые ограничения на модель позволяют полностью сравнять число переменных, связанных с зарядами и элементарными токами.

**Построение уравнения движения на основе QEq-модели**

Общая схема решения состоит из следующих шагов:

* строим лагранжиан системы и находим канонические уравнение
* строим гамильтониан в канонических переменных и выводим уравнение движения

Для вывода уравнения движения необходима информация о топологии молекулы. Согласно ограничениям на модель, элементарные токи возможны только между химически связанными атомами. Для задания топологической матрицы представим молекулярный граф в ориентированном виде, введя направления на ребрах произвольным образом. Формально зададим граф как:

Проведем нумерацию всех ребер, то есть введем биективную нумерующую функцию:

Тогда топология молекулы задается следующим образом:

Тогда для

Матрица A является матрицей инцидентности для молекулярного ориентированного графа и определяет топологические свойства графа молекулы. Зная топологию, запишем изменение заряда на атоме:

Где обозначает ток, текущий по связи . Другими словами, если связи j соответствует пара атомов , то – ток, втекающий в узел со стороны узла . Перейдем к матричной форме описания, тогда изменения заряда на атоме примет вид:

|  |  |
| --- | --- |
|  | (4) |

Кинетический член запишется как:

где – функции, ограничивающие перенос заряда по связи

В матричной форме кинетический член запишется как:

|  |  |
| --- | --- |
|  | (5) |

Выразим из уравнения (4) и подставим в (5):

|  |  |
| --- | --- |
|  | (6) |

где включает в себя топологию(A) и геометрию(K) молекулы

Лагранжиан системы записывается как:

|  |  |
| --- | --- |
|  | (7) |

Позволяет определить каноническую переменную , которая сопряжена с переменной следующим образом:

|  |  |
| --- | --- |
|  | (8) |

Подставляя (7) в (8) получаем выражение:

|  |  |
| --- | --- |
|  | (9) |

Далее выразим из (9) и подставим в (6), тогда получим выражение:

Окончательно, гамильтониан записывается как:

Из гамильтониана следует уравнение движения:

Полученные уравнения движения включают матрицы , которые не требуют обращение. Это означает, что возможна реализация метода с линейной сложностью, поскольку матрица A редкая.

Мы получили уравнение движения заряда, которое позволяет изучать динамику движения заряда в молекулярной системе. Данный метод учитывает перенос заряда на дальние расстояния, а окончательные уравнения не требуют обращения матриц, что, при введении дополнительного порога на кулоновское взаимодействие, позволяет производить подсчет одной итерации динамики с линейной вычислительной сложностью. Следует отметить, что динамика не приводит к определенному конечному состоянию распределения зарядовой плотности. Результатом работы метода являются колебания зарядов около своего положения равновесия, которые и являются оптимальными значениями для зарядов.

**Построение уравнения движения на основе QTPIE-модели**

Построение уравнения движения на основе QTPIE-модели осуществляется аналогичным способом. Основное отличие заключается в использовании потенциального члена модели QTPIE, который оперирует интегральными токами для выражения зарядовой плотности. Модельный гамильтониан в данном случае записывается в виде:

где - матрица реальных токов, зависящих от времени, – матрица интегральных токов, показывающих количество заряда, перенесенное по связи , - коэффициенты масштабирования, влияющие на амплитуду полученных колебаний. Следует заметить, что потенциальный член QTPIE-модели содержит переменных, что позволяет не вводить дополнительные ограничения для уравнивания количества переменных токов. Это означает, что перенос зарядовой плотности возможен между химически несвязанными атомами.

*Ограничения на модель*

Модель ограничивается фундаментальными требованиями о равенстве суммы входящих и выходящих токов из узла молекулярного графа. Формально эти условия записываются как:

Легко показать, что данных условий достаточно для обеспечения сохранности заряда в любом несвязанном подграфе.

*Построение уравнения движения*

Лагранжиан системы записывается как:

позволяет определить обобщенный импульс системы , сопряженный с интегральными токами следующим образом:

Модельный гамильтониан системы в канонических переменных выглядит как:

А уравнения движения принимают вид:

Метод на основе QTPIE-модели уступает по производительности методу на основе QEq-модели, так как оперирует с переменными, поэтому вычислительная сложность одной итерации динамики равно . Использование потенциального члена схемы QTPIE обеспечивает более точное соответствие положения равновесия колебаний атомного заряда со значением, определенными химико-физическим способом. Следует добавить, что предложенный метод является расширением QTPIE-модели, а, следовательно, наследует свойства подхода QTPIE, то есть метод дает правильные значения зарядов при диссоциации молекулы.

**Реализация**

В данной работе была реализована классическая схема QEq и метод уравнения движения на базе QEq-модели (далее динамический QEq или dQEq). Выбор обусловлен линейной вычислительной сложностью динамического QEq-метода, и способностью блокировать перенос заряда на дальние расстояния. Алгоритмы были реализованы в составе программного комплекса BMMKERN, позволяющего обеспечить эффективное выполнение базовых задач молекулярного моделирования. Данный программный комплекс находится под непрерывным развитием в НИИ ЦТиГ СО РАН. BMMKERN реализован на языке C++ с использованием библиотек STL, boost, СBLAS, FFTW и др. Все ее объекты построены на технологии статического полиморфизма. Базовый код программы MOLKERN векторизован (использует SSE и SSE2 расширения), использует многопоточность и выполняется в операционной системе Linux. Кроме того, он включает ряд программных компонент для параллельных расчётов с использованием технологий OpenMP и MPI. На данное время пакет оптимизирован для выполнения на многоядерных векторных процессорах с общей памятью.

Реализация алгоритмов выполнена отдельными модулями к программному комплексу BMMKERN. Все эмпирические параметры задавались в виде словарей, реализованных с использованием порождающего шаблона проектирования «одиночка». Были взяты канонические значения параметров жесткости, электроотрицательности и коэффициенты Кулона, которые были использованы в классическом QEq-методе. Задача минимизации уравнения энергии и нахождения оптимальных значений зарядов решалась с использованием квази-ньютоновсвого метода BFGS (Broyden–Fletcher–Goldfarb–Shanno).

Использовалась оптимизация для умножения редких матриц, позволяющая приравнять число операций к количеству ненулевых элементов.

**Результаты**

В этой главе представлены основные результаты разработанного метода. В нашей работе мы использовали минимально возможную типизацию атомов – атомный тип отвечает химическому элементу. Данных подход соответствует классической QEq-схеме. Это означает, что для каждого атома мы используем один набор параметров. Положение равновесия определяется потенциальным членом, следовательно, более детальная типизация параметров потенциального члена уравнения энергии, учитывающая гибритизацию атома и его окружение, должна повысить точность полученных результатов. Это означает, что динамический подход применим к различным модификациям статических QEq и QTPIE моделей, и позволяет получать динамику молекулярной системы с более точными результатами для узкопрофильных входных данных.

На графиках приведены результаты расчетов для простого тестового случая – молекулы воды, показывающие основные отличия разработанного динамического метода от QEq и QTPIE моделей.

В отличие от QEq-метода, предложенный метод дает правильные значения зарядов атомов при диссоциации молекулы, в частности, нулевое значение атома водорода на бесконечности (рис. 1). Это объясняется тем, что кинетический член блокирует перенос заряда при определенном расстоянии между атомами.

В отличие от статических QEq и QTPIE методов, расширенный метод демонстрирует динамику зарядовой плотности (колебания около положения равновесия) (рис. 2) и итоговый атомный заряд может определяться средним значением заряда за определенный временной интервал (рис. 3).

Амплитуда и период колебаний зависит от начального значения заряда, обобщенного импульса, топологии и геометрии молекулы. Графики колебаний с различными начальными параметрами показаны на рис. 4. На рис. 5 показано изменение периода, при увеличении межатомного расстояния между атомами кислорода и водорода в молекуле воды. В предельном случае () кинетический член блокирует перенос заряда, что соответствует бесконечному периоду. Заряд на отдельно взятом атоме зависит от зарядов всех атомов, входящих в молекулу. Это приводит к несимметричности колебаний, что хорошо показано на рис. 4.

Следует отметить, что амплитуда и период колебаний не зависят от химических параметров атомов, таких как электроотрицательность и жесткость. Это означает, что характер колебаний, полученный для молекулы воды, справедлив и для любых других химических соединений.

**Заключение**

В данной работе предложен новый метод для нахождения атомных зарядов. Данный подход является естественным расширением статической модели флуктуирующих зарядов и может быть использован для получения динамической системы из существующих статических методов. Теоретически показано получение уравнений движения для классических QEq и QTPIE моделей. Динамическая модель на базе схемы QEq (dQEq) была успешно реализована в составе программного комплекса BMMKERN. Результаты показывают, что dQEq обладает следующими преимуществами, по сравнению с QEq-схемой:

* вводит динамику движения зарядовой плотности
* вводит неэквивалентность путей переноса заряда, не увеличивая вычислительную сложность алгоритма.

Результаты работы dQEq-метода разобраны и проанализированы.

Возможным вариантом дальнейшего расширения данной работы является получение уравнений движения, которые включают в себя не только движения заряда, но и самих атомов.