



И. Майер

# Избранные главы квантовой химии

Доказательства  
теорем  
и вывод формул



Лаборатория  
ЗНАНИИ



---

# Simple Theorems, Proofs, and Derivations in Quantum Chemistry

István Mayer

*Chemical Research Center  
Hungarian Academy of Sciences  
Budapest, Hungary*



Kluwer Academic/Plenum Publishers  
New York • Boston • Dordrecht • London • Moscow

---

И. Майер

---

# Избранные главы квантовой химии



## Доказательства теорем и вывод формул



Перевод с английского  
канд. физ. мат. наук М. Б. Дарховского и  
канд. физ. мат. наук А. М. Токмачева

под редакцией  
д-ра физ. мат. наук А. Л. Чугреева

4-е издание, электронное



Москва  
Лаборатория знаний  
2021

УДК 530.145+541.1  
ББК 17.8я73  
М14

**Майер И.**

М14 Избранные главы квантовой химии: доказательства теорем и вывод формул / И. Майер ; пер. с англ. — 4-е изд., электрон. — М. : Лаборатория знаний, 2021. — 387 с. — Систем. требования: Adobe Reader XI ; экран 10". — Загл. с титул. экрана. — Текст : электронный.

ISBN 978-5-93208-516-5

В учебном издании, написанном специалистом из Венгрии, рассмотрены основные результаты и точные утверждения квантовой химии с выводами и доказательствами. Приведены примеры использования квантово-химических утверждений при анализе конкретных систем.

Для студентов, аспирантов, преподавателей и научных работников в области теоретической химии, молекулярной физики и квантовой механики.

УДК 530.145+541.1  
ББК 17.8я73

**Деривативное издание на основе печатного аналога:** Избранные главы квантовой химии: доказательства теорем и вывод формул / И. Майер ; пер. с англ. — М. : БИНОМ. Лаборатория знаний, 2006. — 384 с. : ил. — ISBN 5-94774-499-6.

В соответствии со ст. 1299 и 1301 ГК РФ при устранении ограничений, установленных техническими средствами защиты авторских прав, правообладатель вправе требовать от нарушителя возмещения убытков или выплаты компенсации

Copyright © Kluwer Academic  
Publishers/Springer Science +  
Business Media, 2003  
© Перевод на русский язык,  
Лаборатория знаний, 2015

ISBN 978-5-93208-516-5

# Предисловие редактора перевода

«А разве в квантовой химии бывают теоремы?» — может спросить любознательный студент, а иногда и не менее любознательный преподаватель. «Да и зачем они, — спросит продвинутый аспирант, а нередко и его руководитель, — если за 200 руб. можно купить компакт-диск, на котором записаны не только квантовохимические программы, заботливо «кракнутые» неведомыми умельцами, но и инструкции и даже программы графического ввода данных и анализа результатов?». Любителям подобных рассуждений даю справку: в магазине медтехники любой из вас может приобрести скальпель всего за 50 руб. Пользуясь этим инструментом, покупатель легко мог бы разрешить все хирургические проблемы, как свои, так и своих близких (конечно, если бы прилагалась понятная ему инструкция).

Так вот, любознательный читатель, квантовая химия — не «бродилка», и не «стрелялка», и даже не «стратегия», а один из разделов теоретической физики и прикладной математики. В ней есть формулы, и надо знать, как они выводятся, теоремы, и надо знать, как их доказывать. Надо знать, какова область применимости того или иного приближения (и почему она такая) и основанного на этом приближении квантовохимического метода. *Все это надо знать*, иначе пользы от компьютера, на котором вы проводите свои расчеты, будет ровно столько, сколько от обогревателя равной мощности.

Книга профессора Иштвана Майера задает нужное направление — она развивает у читателя представление о квантовой химии, как о части математики и теоретической физики. Это самодостаточный учебник по квантовой химии, но учебник довольно высокого уровня, где приведен последовательный вывод основных результатов квантовой химии с использованием методов квантовой механики. В книге изложены классические результаты хорошо известные профессионалам, но не включенные в учебники, которые обычно ориентированы либо на самый начальный уровень изучения квантовой химии, либо слишком озабочены «практическими» приложениями. Замечательно, что все результаты и утверждения выводятся или доказываются явно, а важнейшие из них доказаны двумя или более способами, что способствует более многостороннему их восприятию.

От читателя, однако, требуется определенный опыт обращения с понятиями квантовой механики: операторами, волновыми функциями, уравнением Шрёдингера, математическими ожиданиями физических величин, т. е. предполагается, что данная книга, скорее всего, не первая в его жизни книга по квантовой химии, даже если речь идет о студенте или аспиранте. Для понимания

математического аппарата книги необходимо, конечно же, свободное владение основными фактами и приемами математического анализа и матричной (линейной) алгебры. При этом весь материал, который может отсутствовать в «багаже» знаний типичного аспиранта и молодого ученого, специализирующихся в области физической и теоретической химии, а также спектроскопии, материаловедения, молекулярной физики, и т. д. (на такую молодежь собственно и рассчитана книга), включен либо в основной текст либо в приложения. Но самое главное для тех, кто открывает эту книгу — это их готовность следовать за автором и осознание того, что нет смысла заучивать формулировки теорем без изучения их доказательств. Особенно полезна книга будет тем, кто, имея начальные представления о квантовой химии, хотел бы получить более глубокое понимание ее фундаментальных концепций.

Предлагаемая книга вышла при непосредственном участии автора — профессора Майера, который неоднократно прочитал выполненный нами перевод и предложил несколько важных изменений, в том числе к моим примечаниям и дополнению. Я хочу выразить в связи с этим ему свою искреннюю благодарность. В русском издании были исправлены также замеченные ошибки.

Перевод выполнен канд. физ.-мат. наук А. М. Токмачевым (предисловие и гл. 1) и канд. физ.-мат. наук М. Б. Дарховским (гл. 2–9 и приложения).

*А. Л. Чугреев*



# Предисловие автора к русскому изданию

*Посвящаю русское издание книги моим  
преподавателям в Харьковском  
политехническом институте (1961-1963) и  
Харьковском государственном университете  
(1963-1967)*

«Navigare necesse est, vivere non est necesse» (плавать по морю необходимо, жить не обязательно), — сказал Помпей и поднялся на палубу корабля, когда непогода поставила под угрозу снабжение Рима зерном. Это выражение обычно цитируют тогда, когда какую-то задачу нужно выполнить независимо от затрат (сил, денег, времени и т. д.), которых она может стоить. Так и я отношусь к этой книге: считал нужным написать ее, так как чувствовал некоторый зазор между тем материалом, который рассматривается в стандартной учебной литературе, и тем, что (и как) следует знать молодому специалисту, занимающемуся квантовой химией. Я уверен в том, что нельзя по-настоящему ориентироваться в своей отрасли науки, основываясь на пустых декларациях типа «можно показать, что», часто встречающихся в учебниках. Для настоящего понимания полученных результатов нужно самому проработать их выводы от начала до конца.

У меня имеется такое же чувство необходимости и относительно русского издания моей книги, которое даже усиливается тем, что в русскоязычной научной культуре развитие математического формализма традиционно пользуется высоким уважением и мало ценят «рецептурные» подходы типа «делай так и так, а что получится, то и будет твой результат». Как выпускник Харьковского государственного университета 1967 г. (о, боже, как давно это было!) я уделял большое внимание русской версии своей книги и изо всех сил старался улучшить ее. Я надеюсь, переводчики и редактор простят мне то, что я придирался чуть ли не к каждому слову по ходу их работы.

Русское издание книги строго придерживается английского оригинала — сохранена и нумерация уравнений. Однако, следуя совету редактора перевода А. Л. Чугреева, я добавил небольшой вывод теоремы вириала в приближении Борна—Оппенгеймера для приближенных волновых функций, полученных вариационным методом. (В оригинале рассматривался только случай точной волновой функции.) Кроме этого, в русском издании сделано несколько мелких замечаний (либо моих, либо редактора) — они добавлены в виде сносок; наряду с этим редактор перевода написал небольшую самостоятельную главу, в которой на примере простой полуэмпирической модели детально разъяснены некоторые проблемы, кратко затронутые мною в гл. 6, разд. 4.3. Конечно, эти формулы нельзя непосредственно применить к *ab initio* расчетам, но они хорошо передают суть рассматриваемых явлений.

В заключение я хочу поблагодарить всех, кто работал над русским переводом моей книги. По существу, хорошее качество русского издания является

их заслугой. В то же время я считаю, что ответственность за все ошибки, оставшиеся в книге, я должен взять на себя; ведь я — автор, и не раз проверял как английский, так и русский вариант. Так что я могу только просить читателя быть снисходительным, если он встретит в книге какую-то ошибку или неточность, и надеяться на то, что он найдет здесь не только ошибки, но и достаточно полезного материала для того, чтобы считать результирующее сальдо положительным.

*Будапешт, декабрь 2005 г.*





# Предисловие

*Посвящаю моей жене Марти, а также  
памяти моей матери и теток, которые меня  
спасали и вырастили в трудные времена.*

Я читаю лекции в Будапештском университете с 1983 г. На мои лекции приходят студенты-химики, уже знакомые с основами квантовой химии в объеме университетской программы и желающие приобрести более глубокие знания, возможно, для подготовки дипломной или кандидатской (PhD) работы в области теоретической химии. В этой аудитории я посчитал правильным ограничить число тем для подробного обсуждения, выбрав те, которые, с моей точки зрения, являются наиболее важными. Я изменяю содержание своих лекций из года в год, но при этом я обязательно даю основные квантовохимические принципы и теоремы, а также теоретические результаты, которые, как я полагаю, могут сохранить свое значение для быстро развивающихся квантовохимических исследований. Любая тема, обсуждаемая мною на лекции, рассматривается с «азов» без ссылок на предварительную подготовку по квантовой химии; предполагается только самое общее знакомство с ее целями, подходами и, в меньшей степени, техническими приемами. Я акцентирую внимание на формулах и их доказательствах, полагая, что аудитория в основном понимает, для чего, собственно, нужны получаемые нами результаты.

Эта книга во многом основана на материале моих лекций. Особенность, отличающая ее от большинства других учебников, состоит в том, что все результаты явно доказаны или выведены и выкладки представлены полностью шаг за шагом. Действительное понимание теоретического результата может быть достигнуто только путем его вывода. В связи с такой особенностью эта книга может рассматриваться как учебник «с начала, но не для начинающих». Для полного понимания от читателя требуется только знакомство с простейшими понятиями квантовой механики, а именно: волновые функции, операторы и их матричные элементы, зависящее и не зависящее от времени уравнение Шрёдингера; также необходимо иметь представление об основах математического анализа, линейной и матричной алгебры. Я полагал, что в багаже знаний химика может отсутствовать кое-что из необходимого для нашего общения, поэтому я включил этот материал либо в основной текст, либо в приложения. Следовательно, у каждого, кого эта книга может заинтересовать даже только по заглавию, скорее всего, будет достаточно знаний. Однако по этому учебнику вовсе нельзя начинать знакомство с квантовой химией (возможно, это может позволить себе только физик-профессионал). Здесь много внимания уделяется выводу формул, а не их практическому использованию<sup>12</sup>. Таким образом, несмотря на то что изначально предполагалось, что читателю достаточно

<sup>1</sup> Для изучения практических аспектов я могу порекомендовать книгу: Veszprémi T., Fehér M. *Quantum Chemistry*, Kluwer Academic/Plenum Publishers, 1999

<sup>2</sup> Из русских учебников по квантовой химии можно рекомендовать: Абаренков И. В., Братцев В. Ф., Тулуб А. В. Начала квантовой химии. — М.: Высшая школа, 1989; Степанов Н. Ф., Пупышев В. И. Квантовая механика молекул и квантовая химия. — М.: Изд-во МГУ, 1991; Минкин В. И., Симкин Б. Я., Миняев Р. М.

иметь весьма ограниченный объем знаний, эта книга не для легкого чтения: требуется определенная собранность и напряжение, чтобы отслеживать математические выкладки (иногда, прямо скажем, не очень короткие) и подробные доказательства. Я думаю, что уже само заглавие книги<sup>1</sup> отражает все эти аспекты — относительную простоту и упор на формализм.

Многие из результатов, изложенных в этой книге, надо полагать, должны быть известны профессионалу, работающему в данной области. Тем не менее, этот материал не входит в элементарные учебники и найти такие источники, которые содержат детальные выводы или доказательства этих вроде бы «общеизвестных фактов», особенно те, которые не используют более утонченный формализм, например вторичное квантование, довольно-таки трудно. В этой книге такие методы, как вторичное квантование, матрицы плотности, диаграммы и теория групп, не используются, так что она может удовлетворять потребностям тех читателей, которые имеют химическое, а не (теоретико-)физическое образование. Я адресовал свою книгу в основном будущим профессионалам (студентам и молодым специалистам) в области теоретической химии и смежных областях (спектроскопия, материаловедение, молекулярная физика и т. д.). Тем не менее я смею надеяться, что мои коллеги — специалисты по квантовой химии и молекулярной физике — также обратятся к моему труду как полезному источнику простых доказательств и выводов при подготовке своих лекций.

Я старался сделать изложение настолько простым, насколько это было возможно, оставив рассмотрение на так называемом «физическом уровне строгости». Я не следовал принятому в литературе способу представления материала, если сам нашел нечто лучшее. В частности, я рассматривал все вариационные задачи на основе одной и той же формулировки вариационного принципа (2.13), которая позволяет, например, очень просто вывести уравнения Хартри—Фока из теоремы Бриллюэна. (Однако в некоторых случаях альтернативные подходы, использующие метод множителей Лагранжа, также представлены для полноты картины.) С моей точки зрения, книга самодостаточна. Я постарался избежать вопросов приоритета. Поэтому здесь нет ссылок на обширные литературные обзоры, можно встретить только короткие библиографические заметки, цитирующие некоторые основные работы и указывающие те источники, из которых я почерпнул какой-то нестандартный материал.

Наконец, я добрался до того места в своем предисловии, где я с удовольствием могу выразить глубочайшую благодарность моей коллеге д-ру Андреа Хамза, которая прочитала всю рукопись и внесла значительный вклад в ее улучшение. Я также благодарю профессора Петера Р. Шурьяна, профессора Далю Шатковскене, д-ра Имре Папай и д-ра Габора Шуберта за полезные комментарии к некоторым разделам. Подготовка книги к печати была в основном проделана Жужей Кертес, и я чувствую себя ее должником за ту тщательность, которую она проявила при наборе объемных формул в  $\text{LaTeX}^e$ .

<sup>1</sup> Имеется в виду английский оригинал.

# Гамильтониан Борна—Оппенгеймера

## 1. Отделение движения центра масс в квантовой механике

### 1.1. Переход от задачи двух тел к двум задачам одного тела

Как известно, гамильтониан двух взаимодействующих частиц (точечных тел) имеет вид

$$\hat{H} = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + V(\vec{r}_1, \vec{r}_2), \quad (1.1)$$

где

$$\Delta_1 \equiv \nabla_1^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2} \quad (1.2)$$

и  $\Delta_2$  определяется аналогично. В замкнутой системе потенциальная энергия  $V$  зависит только от относительного положения частиц, так что

$$\hat{H} = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + V(\vec{r}_1 - \vec{r}_2). \quad (1.3)$$

Точно так же, как это делается в классической механике (см. приложение П2), определим полную массу системы

$$M = m_1 + m_2 \quad (1.4)$$

и новые координаты  $\vec{R}$  и  $\vec{r}$ , описывающие положение центра масс системы и относительное положение частиц, а также декартовы компоненты этих векторов:

$$\vec{r} = \vec{r}_1 - \vec{r}_2 = \vec{i}x + \vec{j}y + \vec{k}z; \quad (1.5)$$

$$\vec{R} = \frac{m_1}{M}\vec{r}_1 + \frac{m_2}{M}\vec{r}_2 = \vec{i}X + \vec{j}Y + \vec{k}Z; \quad (1.6)$$

т. е.

$$x = x_1 - x_2 \quad (1.7)$$

( $y$  и  $z$  определяются аналогично); и

$$X = \frac{m_1}{M}x_1 + \frac{m_2}{M}x_2 \quad (1.8)$$

(и аналогично для  $Y$  и  $Z$ ).

Исследуем сумму

$$\frac{1}{m_1} \frac{\partial^2 \Psi}{\partial x_1^2} + \frac{1}{m_2} \frac{\partial^2 \Psi}{\partial x_2^2}, \quad (1.9)$$

рассматривая  $\Psi$  как функцию координат  $\vec{r}$  и  $\vec{R}$ ; тогда она также является сложной функцией  $x_1$  и  $x_2$  (и других координат частиц). Применяя правила дифференцирования сложных функций и соотношения  $\frac{\partial y}{\partial x_1} = \frac{\partial Y}{\partial x_1} = 0$  и т. д., получаем

$$\frac{\partial \Psi}{\partial x_1} = \frac{\partial \Psi}{\partial x} \frac{\partial x}{\partial x_1} + \frac{\partial \Psi}{\partial X} \frac{\partial X}{\partial x_1} = \frac{\partial \Psi}{\partial x} \cdot 1 + \frac{\partial \Psi}{\partial X} \frac{m_1}{M} = \left( \frac{\partial}{\partial x} + \frac{m_1}{M} \frac{\partial}{\partial X} \right) \Psi. \quad (1.10)$$

Действуем затем подобным образом:

$$\frac{\partial^2 \Psi}{\partial x_1^2} = \left( \frac{\partial}{\partial x} + \frac{m_1}{M} \frac{\partial}{\partial X} \right)^2 \Psi = \left[ \frac{\partial^2}{\partial x^2} + 2 \frac{m_1}{M} \frac{\partial^2}{\partial x \partial X} + \left( \frac{m_1}{M} \right)^2 \frac{\partial^2}{\partial X^2} \right] \Psi. \quad (1.11)$$

Аналогично получается и  $\frac{\partial^2 \Psi}{\partial x_2^2}$ , но в этом случае  $\frac{\partial x}{\partial x_2} = -1$  (а не  $+1$ , как было в предыдущем случае):

$$\begin{aligned} \frac{\partial \Psi}{\partial x_2} &= \left( -\frac{\partial}{\partial x} + \frac{m_2}{M} \frac{\partial}{\partial X} \right) \Psi \\ \frac{\partial^2 \Psi}{\partial x_2^2} &= \left[ \frac{\partial^2}{\partial x^2} - 2 \frac{m_2}{M} \frac{\partial^2}{\partial x \partial X} + \left( \frac{m_2}{M} \right)^2 \frac{\partial^2}{\partial X^2} \right] \Psi. \end{aligned} \quad (1.12)$$

При суммировании слагаемые, содержащие  $\frac{\partial^2}{\partial x \partial X}$ , уничтожаются:

$$\frac{1}{m_1} \frac{\partial^2 \Psi}{\partial x_1^2} + \frac{1}{m_2} \frac{\partial^2 \Psi}{\partial x_2^2} = \left[ \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \frac{\partial^2}{\partial x^2} + \frac{m_1 + m_2}{M^2} \frac{\partial^2}{\partial X^2} \right] \Psi. \quad (1.13)$$

Определим приведенную массу  $\mu$  как

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}. \quad (1.14)$$

Тогда, принимая во внимание определение (1.4), получаем

$$\frac{1}{m_1} \frac{\partial^2 \Psi}{\partial x_1^2} + \frac{1}{m_2} \frac{\partial^2 \Psi}{\partial x_2^2} = \left( \frac{1}{\mu} \frac{\partial^2}{\partial x^2} + \frac{1}{M} \frac{\partial^2}{\partial X^2} \right) \Psi. \quad (1.15)$$

Аналогичные преобразования могут быть проделаны и для компонент  $y$  и  $z$ .

Следовательно, имеем:

$$\hat{H} = -\frac{\hbar^2}{2\mu}\Delta_{rel} - \frac{\hbar^2}{2M}\Delta_{ЦМ} + V(\vec{r}_{rel}) \quad (\vec{r}_{rel} \equiv \vec{r}) . \quad (1.16)$$

Первое и третье слагаемые этого выражения образуют гамильтониан относительного движения двух частиц

$$\hat{H}_{rel} = -\frac{\hbar^2}{2\mu}\Delta_{rel} + V(\vec{r}_{rel}) , \quad (1.17)$$

который представляет собой эффективный гамильтониан задачи о движении одной частицы с массой, равной приведенной массе  $\mu$ , в потенциале, совпадающем с исходным потенциалом межчастичного взаимодействия  $V(\vec{r})$ . (Мы предположили, что система замкнута, т. е. внешние поля отсутствуют). Второе слагаемое является гамильтонианом, отвечающим свободному движению центра масс (ЦМ) — это также эффективный гамильтониан одной частицы (с массой  $M$ ):

$$\hat{H}_{ЦМ} = -\frac{\hbar^2}{2M}\Delta_{ЦМ} . \quad (1.18)$$

Собственные функции такого гамильтониана, как хорошо известно, — плоские волны.

Если гамильтониан какой-либо задачи представляет собой сумму двух независимых гамильтонианов, которые не содержат общих переменных, то волновая функция является простым произведением собственных функций двух гамильтонианов, а энергия — суммой двух энергий. Это значит, что в нашем случае волновая функция  $\Psi$ , зависящая от всех переменных, характеризующих систему, может быть записана как

$$\Psi(\vec{r}_1, \vec{r}_2) = \Psi(\vec{R}, \vec{r}) = \Phi(\vec{R})\varphi(\vec{r}) , \quad (1.19)$$

где  $\Phi(\vec{R})$  и  $\varphi(\vec{r})$  являются решениями не зависящих от времени уравнений Шрёдингера, записанных для движения центра масс и для относительного движения соответственно:

$$\hat{H}_{ЦМ}\Phi(\vec{R}) = E_{ЦМ}\Phi(\vec{R}) \quad (1.20)$$

$$\hat{H}_{rel}\varphi(\vec{r}) = E_{rel}\varphi(\vec{r}) , \quad (1.21)$$

а энергия  $E$  является суммой  $E_{\text{ЦМ}} + E_{\text{rel}}$ . В самом деле, благодаря отсутствию общих переменных, можно переставлять гамильтониан  $\hat{H}_{\text{ЦМ}}$  и функцию  $\varphi(\vec{r})$ , а также  $\hat{H}_{\text{rel}}$  и  $\Phi(\vec{R})$ ; в соответствии с этим имеем:

$$\begin{aligned} (\hat{H}_{\text{ЦМ}} + \hat{H}_{\text{rel}}) \Phi(\vec{R})\varphi(\vec{r}) &= \hat{H}_{\text{ЦМ}}\Phi(\vec{R})\varphi(\vec{r}) + \hat{H}_{\text{rel}}\Phi(\vec{R})\varphi(\vec{r}) \\ &= (\hat{H}_{\text{ЦМ}}\Phi(\vec{R}))\varphi(\vec{r}) + \Phi(\vec{R})(\hat{H}_{\text{rel}}\varphi(\vec{r})) \\ &= (E_{\text{ЦМ}} + E_{\text{rel}})\Phi(\vec{R})\varphi(\vec{r}). \quad \text{Ч. т. д.}^1 \end{aligned} \quad (1.22)$$

Существует другой способ вывода приведенного выше гамильтониана относительного движения. Мы можем предположить, что используем координатную систему, движущуюся с той же скоростью, что и центр масс двух частиц. Это значит, что суммарный импульс системы по отношению к этой системе координат равен нулю. (Это эквивалентно требованию неподвижности центра масс, но только мы не делаем теперь никакого предположения о его местоположении.)

Известно, что оператор полного импульса системы равен

$$\hat{\vec{P}} = \frac{\hbar}{i} \sum_i \nabla_i \equiv \frac{\hbar}{i} \sum_i \frac{\partial}{\partial \vec{r}_i}, \quad (1.23)$$

где мы использовали обозначение

$$\frac{\partial}{\partial \vec{a}} = \vec{i} \frac{\partial}{\partial a_x} + \vec{j} \frac{\partial}{\partial a_y} + \vec{k} \frac{\partial}{\partial a_z} \quad (1.24)$$

для обозначения дифференцирования по вектору. Известно также, что оператор  $\hat{\vec{P}}$  коммутирует с гамильтонианом, и, следовательно, полный импульс может быть измерен одновременно с энергией и представляет собой интеграл движения (см. гл. 2, разд. 3.1).

Теперь выразим сумму  $\frac{1}{m_1}\Delta_1 + \frac{1}{m_2}\Delta_2$  в новых координатах:

$$\begin{aligned} \vec{r} &= \vec{r}_1 - \vec{r}_2 \\ \vec{\rho} &= \vec{r}_1 + \vec{r}_2 \end{aligned} \quad (1.25)$$

Первая координата  $\vec{r}$  в (1.25) это та же координата относительного положения частиц, что использовалась ранее, тогда как  $\vec{\rho}$  (вторая координата) — это просто сумма  $\vec{r}_1$  и  $\vec{r}_2$ , т. е. она не является масс-взвешенной,

<sup>1</sup> Что и требовалось доказать.

как радиус-вектор центра масс  $\vec{R}$ . Используя правила дифференцирования сложных функций, получаем

$$\frac{\partial}{\partial \vec{r}_1} = \frac{\partial}{\partial \vec{\rho}} \frac{\partial \vec{\rho}}{\partial \vec{r}_1} + \frac{\partial}{\partial \vec{r}} \frac{\partial \vec{r}}{\partial \vec{r}_1} = \frac{\partial}{\partial \vec{\rho}} + \frac{\partial}{\partial \vec{r}}, \quad (1.26)$$

$$\frac{\partial}{\partial \vec{r}_2} = \frac{\partial}{\partial \vec{\rho}} \frac{\partial \vec{\rho}}{\partial \vec{r}_2} + \frac{\partial}{\partial \vec{r}} \frac{\partial \vec{r}}{\partial \vec{r}_2} = \frac{\partial}{\partial \vec{\rho}} - \frac{\partial}{\partial \vec{r}}$$

Теперь мы потребуем, чтобы волновая функция  $\Psi$  была собственной функцией оператора  $\hat{P} = \frac{\hbar}{i} \left( \frac{\partial}{\partial \vec{r}_1} + \frac{\partial}{\partial \vec{r}_2} \right)$  с собственным значением 0. Из этого требования и равенств (1.26) следует

$$\left( \frac{\partial}{\partial \vec{r}_1} + \frac{\partial}{\partial \vec{r}_2} \right) \Psi = 2 \frac{\partial}{\partial \vec{\rho}} \Psi = 0, \quad (1.27)$$

т. е.  $\Psi$  не зависит от  $\vec{\rho}$ . Следовательно, мы можем пренебречь всеми дифференцированиями по  $\vec{\rho}$  при вычислении  $\Delta_1 \Psi$  и  $\Delta_2 \Psi$  в этом частном случае, и останутся только дифференцирования по  $\vec{r}$ . Так мы получим

$$\left( \frac{1}{m_1} \Delta_1 + \frac{1}{m_2} \Delta_2 \right) \Psi \Big|_{P=0} = \left( \frac{1}{m_1} + \frac{1}{m_2} \right) \Delta_{rel} \Psi, \quad (1.28)$$

что в свою очередь означает, что если суммарный импульс системы равен нулю, то его кинетическая энергия совпадает с кинетической энергией относительного движения частицы с массой, равной приведенной массе.

Из квантовой механики известно, что энергия электрона, движущегося в кулоновском поле неподвижного единичного положительного заряда, равна  $\frac{m_e e_0^4}{2\hbar^2}$ ; это дает величину постоянной Ридберга в спектроскопических единицах:

$$R = 109737.31 \text{ см}^{-1}.$$

Из спектроскопии мы знаем, что для легкого изотопа водорода H

$$R_H = 109677.581 \text{ см}^{-1}.$$

Поскольку масса протона составляет приблизительно 1836.15 масс электрона, приведенная масса атома H

$$\mu = \frac{m_e}{1 + \frac{m_e}{M_p}} \cong \frac{m_e}{1 + \frac{1}{1836.15}}.$$

Из экспериментального значения  $R_H$  и этого соотношения мы можем оценить постоянную Ридберга, которая отвечала бы бесконечной массе ядра (неподвижному ядру):

$$R = R_\infty = R_H \left( 1 + \frac{1}{1836.15} \right) \cong 109737.31 \text{ см}^{-1},$$

что находится в замечательном согласии с теоретическим значением  $R$ .

Поскольку протон является самым легким ядром, столь малое расхождение между  $R_H$  и  $R_\infty$  может рассматриваться как указание на целесообразность введения приближения Борна—Оппенгеймера.

## 1.2. Центр масс в квантовой механике

Если система состоит из более чем двух частиц, то исследование ее движения можно провести следующим образом. Сначала координаты первых двух частиц преобразуются в координаты относительного движения и координаты их центра масс, как описано ранее. Затем необходимо повторить эту процедуру для центра масс первых двух частиц (как будто бы это была частица с массой  $m_1 + m_2$ ) и третьей частицы, и т. д.: в каждом случае первый набор трех новых координат описывает относительное движение новой частицы по отношению к центру масс предыдущих, тогда как второй набор описывает центр масс всех частиц, рассмотренных до этого момента. Последние три из этих «координат Якоби» представляют собой декартовы координаты центра масс всей системы. Эта схема преобразования координат может использоваться как в классической, так и в квантовой механике. Однако она имеет серьезный недостаток, так как она не симметрична относительно отдельных частиц, что означает, что в квантовомеханическом случае неразличимость электронов не «встроена» в расчетную схему. (Очевидно, что наличие симметрии по отношению к перестановкам идентичных частиц является весьма желательным свойством теории.)

Поскольку процедура, рассмотренная здесь, представляет собой обобщение схемы, описанной в предыдущем разделе, снова получается, что движение центра масс описывается гамильтонианом, характерным для свободной частицы, в которой сосредоточена вся масса системы. Очень важно то, что центр масс замкнутой системы ведет себя как свободная частица, и в квантовой механике он должен описываться плоской волной. Одно из следствий этого заключается в том, что корпускулярно-волновой дуализм не связан с внутренней структурой отдельных частиц, а является общим законом природы в атомном и субатомном масштабе.



Другой вывод состоит в том, что нельзя просто предположить, что мы используем систему координат, в которой центр масс покоится и одновременно находится в начале координат — это противоречило бы принципам квантовой механики. (Это возможно в классической механике, нужно только соответствующим образом выбрать начальные условия. См. приложение П1). В квантовой механике, в соответствии с принципом неопределенности, мы можем сделать утверждение либо о положении центра масс, либо о его скорости (импульсе), но не об обоих одновременно.

Формально проблема состоит в том, что в квантовой механике координаты, которые мы рассматриваем, не являются действительными координатами частиц, движущихся вдоль хорошо определенных траекторий, как это обстоит в классической механике, а представляют собой *независимые переменные* волновой функции, описывающей вероятностное поведение системы. Волновая функция определена для *всех возможных* значений координат индивидуальных частиц, а не только для тех, для которых центр масс закреплен в начале координат, т. е. не только для таких, для которых выполняется соотношение

$$\vec{R} = \sum_{i=1}^N \frac{m_i}{M} \vec{r}_i = 0. \quad (1.29)$$

Здесь и далее мы обозначаем полную массу системы как  $M$

$$M = \sum_{i=1}^N m_i. \quad (1.30)$$

В этой связи заметим, что необходимо различать понятие центра масс в смысле функции независимых координат (1.29) от квантовомеханического среднего радиус-вектора, которое может быть получено усреднением оператора координаты центра масс по волновой функции всей системы.

Очевидно, что когда центр масс уже отделен, мы можем использовать любой набор *относительных* (или внутренних) координат, а не только набор координат Якоби, обсуждавшийся ранее. Вероятно, наиболее простая возможность состоит в том, чтобы рассмотреть относительные координаты первых  $N - 1$  частиц относительно центра масс всей системы:

$$\vec{r}'_i = \vec{r}_i - \vec{R} = \vec{r}_i - \sum_{j=1}^N \frac{m_j}{M} \vec{r}_j. \quad (1.31)$$

В этом случае координаты  $N$ -й частицы уже не являются независимыми переменными, а могут быть определены из условия, что в системе ЦМ центр масс находится в начале координат, откуда

$$\vec{r}'_N = - \sum_{j=1}^{N-1} \frac{m_j}{m_N} \vec{r}'_j. \quad (1.32)$$

Это значит, что вместо исходного набора координат  $\vec{r}_i$ ,  $i = 1, 2, \dots, N$ , мы описываем систему координатами  $\vec{R}$  и  $\vec{r}'_i$ ,  $i = 1, 2, \dots, N-1$ .

При вычислении частных производных необходимо принять во внимание, что координаты центра масс зависят от всех исходных координат, так что  $\frac{\partial x'_j}{\partial x_i} = \delta_{ij} - \frac{m_i}{M}$ . Затем, используя правила дифференцирования сложных функций, получаем

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial X} \frac{\partial X}{\partial x_i} + \sum_{j=1}^{N-1} \frac{\partial}{\partial x'_j} \frac{\partial x'_j}{\partial x_i} = \frac{m_i}{M} \frac{\partial}{\partial X} + \frac{\partial}{\partial x'_i} - \frac{m_i}{M} \sum_{j=1}^{N-1} \frac{\partial}{\partial x'_j} \quad (1.33)$$

для  $i = 1, 2, \dots, N-1$  и

$$\frac{\partial}{\partial x_N} = \frac{\partial}{\partial X} \frac{\partial X}{\partial x_N} + \sum_{j=1}^{N-1} \frac{\partial}{\partial x'_j} \frac{\partial x'_j}{\partial x_N} = \frac{m_N}{M} \left( \frac{\partial}{\partial X} - \sum_{j=1}^{N-1} \frac{\partial}{\partial x'_j} \right) \quad (1.34)$$

для  $N$ -й частицы.

Следовательно, вторые производные имеют вид

$$\frac{\partial^2}{\partial x_i^2} = \frac{m_i^2}{M^2} \left( \frac{\partial^2}{\partial X^2} + \sum_{j,k=1}^{N-1} \frac{\partial^2}{\partial x'_j \partial x'_k} - 2 \sum_{j=1}^{N-1} \frac{\partial^2}{\partial X \partial x'_j} \right) \quad (1.35)$$

$$+ \frac{\partial^2}{\partial x_i'^2} + 2 \frac{m_i}{M} \left( \frac{\partial^2}{\partial X \partial x'_i} - \sum_{j=1}^{N-1} \frac{\partial^2}{\partial x'_j \partial x'_i} \right)$$

и

$$\frac{\partial^2}{\partial x_N^2} = \frac{m_N^2}{M^2} \left( \frac{\partial^2}{\partial X^2} + \sum_{j,k=1}^{N-1} \frac{\partial^2}{\partial x'_j \partial x'_k} - 2 \sum_{j=1}^{N-1} \frac{\partial^2}{\partial X \partial x'_j} \right). \quad (1.36)$$

Поделив каждое из этих уравнений на соответствующую массу  $m_i$  ( $m_N$ ) и суммируя, после тривиальных алгебраических преобразований и применения равенства (1.30) получаем

$$\sum_{i=1}^N \frac{1}{m_i} \frac{\partial^2}{\partial x_i^2} = \frac{1}{M} \frac{\partial^2}{\partial X^2} + \sum_{i=1}^{N-1} \frac{1}{m_i} \frac{\partial^2}{\partial x_i'^2} - \frac{1}{M} \sum_{i,j=1}^{N-1} \frac{\partial^2}{\partial x'_i \partial x'_j}. \quad (1.37)$$

Этот результат означает, что слагаемые со смешанными частными производными, связывающие движение центра масс и внутреннее движение, сокращаются, как и в случае, обсуждавшемся в предыдущем разделе, но остаются смешанные частные производные, содержащие пары различных внутренних координат.

Выполняя аналогичные вычисления для оставшихся двух направлений и используя оператор  $\nabla'_i = \vec{i} \frac{\partial}{\partial x_{i'}} + \vec{j} \frac{\partial}{\partial y_{i'}} + \vec{k} \frac{\partial}{\partial z_{i'}}$ , который ведет себя как вектор, получаем гамильтониан в виде

$$\hat{H} = -\frac{\hbar^2}{2} \left( \frac{1}{M} \Delta_R + \sum_{i=1}^{N-1} \frac{1}{m_i} \Delta'_i - \frac{1}{M} \sum_{i,j=1}^{N-1} \nabla'_i \nabla'_j \right) + V. \quad (1.38)$$

Часть этого гамильтониана, отвечающая кинетической энергии, является суммой кинетических энергий движения центра масс и внутреннего движения; однако последняя не является простой суммой операторов кинетической энергии для  $N - 1$  индивидуальных частиц, относительные координаты которых рассматриваются как независимые переменные, а содержит также перекрестные слагаемые типа  $\nabla'_i \nabla'_j$ . Появление подобных слагаемых неизбежно, если используется любой набор внутренних координат, отличный от ортогональных координат Якоби. Можно ожидать, что сумма, содержащая эти слагаемые, относительно мала, если суммарная масса системы  $M$  много больше, чем каждая из  $N - 1$  масс  $m_i$ ,  $i = 1, 2, \dots, N - 1$ . Практически это имеет место только тогда, когда мы рассматриваем случай свободного атома и  $N$ -я частица (координаты которой не считаются независимыми переменными) есть ядро атома. В таком случае, однако, подход, описанный в следующем разделе, намного более пригоден. Заметим также, что в случае свободного атома и при использовании ядра в качестве  $N$ -й частицы, пренебрежение этими относительно малыми перекрестными вкладами приводит часть гамильтониана (1.38), ответственную за внутренние движения составляющих систему частиц, к виду гамильтониана Борна—Оппенгеймера (см. разд. 2).

Другой недостаток гамильтониана (1.38) состоит в том, что он содержит потенциальную энергию, которая зависит от относительных координат всех частиц, а не только  $N - 1$  выбранных. Это значит, что в рамках такого рассмотрения относительные координаты  $N$ -й частицы всегда имеют вид (1.32), что делает потенциальную энергию более сложной функцией независимых координат. В случае свободных атомов этого усложнения можно избежать при помощи схемы, обсуждаемой в следующем разделе.

### 1.3. Свободные атомы и атомоподобные системы<sup>1</sup>

Как обсуждалось в предыдущем разделе, для системы  $N$  тел мы должны выделить движение центра масс и выбрать подходящий набор  $3N - 3$  относительных (внутренних) координат. В случае двух тел адекватными относительными координатами были три декартовы компоненты радиус-вектора одной частицы в системе координат, связанной с другой частицей. Такой подход может быть обобщен на случай большего числа частиц: выберем одну частицу как главную и введем декартовы компоненты радиус-векторов других частиц относительно главной как набор внутренних координат. Эти внутренние координаты будут играть роль независимых переменных, от которых зависит волновая функция, описывающая внутреннее движение. (Волновая функция движения центра масс снова есть плоская волна, которая для нас не интересна.)

Для многоэлектронного атома вполне естественно выбрать *ядро* в качестве главной частицы и рассматривать координаты электронов относительно ядра. Такие координаты, в отличие от координат Якоби, но как и координаты, рассмотренные в предыдущем разделе, имеют тот недостаток, что часть гамильтониана, отвечающая кинетической энергии, не может быть выражена строго в виде суммы слагаемых, каждое из которых содержит лапласиан (оператор  $\Delta \equiv \nabla^2$ ) относительных координат какой-либо одной частицы. Важным преимуществом указанного набора координат, однако, является то, что координаты электронов входят в него на равном основании. Кроме того, этот способ позволяет рассмотреть различные «экзотические атомы», например «мюонный атом», в котором один из электронов заменен на мюон ( $\mu$ -мезон), и атомоподобную систему, обычно атом He, в котором один электрон заменен на антипротон. (На уровне нерелятивистского гамильтониана эти изменения проявляются только в большей массе одной из отрицательно заряженных частиц).

Описанная система координат подобна той, которая обсуждалась в предыдущем разделе и в которой центр масс был использован как начало координат, но не совпадает с ней: масса ядра велика по сравнению с массой электронов, но отношение масс не бесконечно; следовательно, положения ядра и центра масс не совпадают точно. Поскольку потенциальная энергия системы с кулоновским взаимодействием есть простая функция расстояний между частицами, такой выбор координат более выгоден для исследования свободных атомов.

<sup>1</sup> Автор в английском оригинале использует термин *atomcule*, который при калькировании режет слух носителя языка. — *Прим. ред.*

В соответствии с данным обсуждением, мы будем описывать атом (атомоподобную частицу), используя следующий набор координат:

а) координаты центра масс (включая ядро)

$$\vec{R} = \frac{m_N}{M} \vec{R}_N + \sum_{i=1}^{N_e} \frac{m_i}{M} \vec{r}_i, \quad (1.39)$$

где  $m_N$  и  $\vec{R}_N$  — масса и радиус-вектор ядра,  $m_i$  и  $\vec{r}_i$  — те же величины для  $i$ -го электрона (отрицательно заряженной частицы),  $N_e$  — количество электронов и  $M$  — суммарная масса системы

$$M = m_N + \sum_{i=1}^{N_e} m_i. \quad (1.40)$$

б) координаты  $N_e$  отрицательно заряженных частиц (электроны, возможно, мюон или антипротон) по отношению к положению ядра, рассматриваемого как начало координат,

$$\vec{r}'_i = \vec{r}_i - \vec{R}_N \quad (i = 1, 2, \dots, N_e). \quad (1.41)$$

(Заметьте, что здесь определение  $\vec{r}'_i$  отличается от использованного в предыдущем разделе.)

В этом случае  $\frac{\partial X}{\partial x_i} = \frac{m_i}{M}$ ,  $\frac{\partial X}{\partial X_N} = \frac{m_N}{M}$ ,  $\frac{\partial x'_j}{\partial x_i} = \delta_{ij}$ , и  $\frac{\partial x'_j}{\partial X_N} = -1$ , так что использование правил дифференцирования сложных функций дает для первых производных:

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial X} \frac{\partial X}{\partial x_i} + \sum_{j=1}^{N_e} \frac{\partial}{\partial x'_j} \frac{\partial x'_j}{\partial x_i} = \frac{m_i}{M} \frac{\partial}{\partial X} + \frac{\partial}{\partial x'_i} \quad (1.42)$$

для  $i = 1, 2, \dots, N_e$ ; и

$$\frac{\partial}{\partial X_N} = \frac{\partial}{\partial X} \frac{\partial X}{\partial X_N} + \sum_{j=1}^{N_e} \frac{\partial}{\partial x'_j} \frac{\partial x'_j}{\partial X_N} = \frac{m_N}{M} \frac{\partial}{\partial X} - \sum_{j=1}^{N_e} \frac{\partial}{\partial x'_j}. \quad (1.43)$$

Для вторых производных:

$$\begin{aligned} \frac{\partial^2}{\partial x_i^2} &= \left( \frac{m_i}{M} \frac{\partial}{\partial X} + \frac{\partial}{\partial x'_i} \right)^2 = \frac{m_i^2}{M^2} \frac{\partial^2}{\partial X^2} + 2 \frac{m_i}{M} \frac{\partial^2}{\partial X \partial x'_i} + \frac{\partial^2}{\partial x_i'^2} \\ \frac{\partial^2}{\partial X_N^2} &= \left( \frac{m_N}{M} \frac{\partial}{\partial X} - \sum_{i=1}^{N_e} \frac{\partial}{\partial x'_i} \right) \left( \frac{m_N}{M} \frac{\partial}{\partial X} - \sum_{j=1}^{N_e} \frac{\partial}{\partial x'_j} \right) \\ &= \frac{m_N^2}{M^2} \frac{\partial^2}{\partial X^2} - 2 \frac{m_N}{M} \sum_{i=1}^{N_e} \frac{\partial^2}{\partial X \partial x'_i} + \sum_{i,j=1}^{N_e} \frac{\partial^2}{\partial x'_i \partial x'_j}. \end{aligned} \quad (1.44)$$

Поделив каждую вторую производную на соответствующую массу и суммируя по всем частицам, легко получаем

$$\frac{1}{m_N} \frac{\partial^2}{\partial X_N^2} + \sum_{i=1}^{N_e} \frac{1}{m_i} \frac{\partial^2}{\partial x_i^2} = \frac{1}{M} \frac{\partial^2}{\partial X^2} + \sum_{i=1}^{N_e} \frac{1}{m_i} \frac{\partial^2}{\partial x_i'^2} + \frac{1}{m_N} \sum_{i,j=1}^{N_e} \frac{\partial^2}{\partial x_i' \partial x_j'} . \quad (1.45)$$

Этот результат показывает, что слагаемые, перекрестные между движением центра масс и внутренним движением, опять исчезают, но вновь появляются слагаемые, относящиеся к внутренним движениям, которые содержат смешанные частные производные по внутренним координатам различных частиц.

Выполняя аналогичные дифференцирования для направлений  $y$  и  $z$ , приходим к преобразованному гамильтониану:

$$\hat{H} = -\frac{\hbar^2}{2} \left( \frac{1}{M} \Delta_R + \sum_{i=1}^{N_e} \frac{1}{m_i} \Delta_i' + \frac{1}{m_N} \sum_{i,j=1}^{N_e} \nabla_i' \nabla_j' \right) + V . \quad (1.46)$$

Отбрасывая слагаемые, содержащие  $\nabla_i' \nabla_j'$ , снова приходим к сумме гамильтониана свободного движения центра масс и электронного гамильтониана Борна—Оппенгеймера для внутреннего движения (см. разд. 2); это эквивалентно пределу  $m_N \rightarrow \infty$ , который отвечает картине «закрепленного» ядра.

Используя гамильтониан (1.46), легко частично учесть и движение ядра. Это может быть особенно важно в случае «атомоподобной системы», для которой одна из отрицательно заряженных частиц имеет массу  $m_i$ , сравнимую с массой ядра  $m_N$ . Необходимо заметить, что индексы суммирования  $i, j$  в последней сумме (1.46) пробегают номера частиц независимо, так что сумма содержит слагаемые  $\nabla_i' \nabla_j'$  с  $i = j$ . Поскольку  $(\nabla_i')^2 \equiv \Delta_i'$ , можно тривиально преобразовать (1.46) в

$$\hat{H} = -\frac{\hbar^2}{2} \left( \frac{1}{M} \Delta_R + \sum_{i=1}^{N_e} \left( \frac{1}{m_i} + \frac{1}{m_N} \right) \Delta_i' + \frac{2}{m_N} \sum_{i < j}^{N_e} \nabla_i' \nabla_j' \right) + V . \quad (1.47)$$

Данный результат означает, что мы должны заменить обратную массу каждой отрицательной частицы  $\frac{1}{m_i}$  на обратную приведенную массу, отвечающую ее взаимодействию с ядром, т. е.  $\frac{1}{\mu_i} = \frac{1}{m_i} + \frac{1}{m_N}$ . Последняя сумма уравнения (1.47) содержит только нетривиальные (т. е., у которых  $i \neq j$ ) смешанные частные производные. Эти остающиеся перекрестные слагаемые могут быть учтены с использованием, например, методов теории возмущений.

Хотя для электрона приведенная масса отличается от массы свободного электрона лишь ненамного, такое изменение очень важно для антипротона в описанной выше атомоподобной системе, где в атоме гелия один из электронов заменен на антипротон: в предположении, что положительно заряженная частица — ядро  ${}^4\text{He}$ , приведенная масса составляет только около 80% массы свободного антипротона. Кроме того, в этом случае необходимо явно рассматривать вклад от произведения  $\nabla'_i \nabla'_j$ , поскольку коэффициент  $\frac{1}{m_N}$  только лишь в четыре раза меньше, чем коэффициент  $\frac{1}{m_i}$ , отвечающий антипротону. В качестве альтернативы можно предложить рассмотреть эту атомоподобную систему в рамках приближения Борна—Оппенгеймера, рассматривая антипротон как второе (хотя и отрицательно заряженное) ядро.



## 2. Приближение Борна—Оппенгеймера

### 2.1. Вводные замечания

Большинство исследований в квантовой химии и теоретической молекулярной физике имеют дело с (приближенными) решениями не зависящего от времени нерелятивистского уравнения Шрёдингера, записанного в приближении Борна—Оппенгеймера.

Можно попытаться наглядно обосновать схему Борна—Оппенгеймера, приняв, что движение легких электронов следует за движением тяжелых ядер, т. е. ядра обычно движутся достаточно медленно, так, что электронное облако может мгновенно перестроиться вслед за изменениями положений ядер.

По-другому ту же идею можно сформулировать так: поскольку электроны движутся быстрее, они «видят» мгновенные положения ядер, тогда как ядра «чувствуют» только усредненный потенциал электронного облака (электростатическое поле, вызванное пространственно распределенным электронным зарядом). Следовательно, распределение электронов можно изучать, рассматривая ядра как фиксированные (так называемое приближение «закрепленных ядер»). Получившаяся электронная энергия будет зависеть от ядерной конфигурации; эта электронная энергия вместе с энергией взаимного отталкивания ядер определяет (гипер-)поверхность потенциальной энергии (ППЭ), управляющую движением ядер. Это называется «разделением электронного и ядерного движения по Борну и Оппенгеймеру», или просто *приближением Борна—Оппенгеймера*.

В действительности существует слабая («вибронная», т. е. электронно-колебательная) связь между уравнениями, описывающими движения электронов и ядер; этими вкладами, однако, можно пренебречь в большинстве реальных расчетов.

Не зависящее от времени уравнение Шрёдингера имеет вид:

$$\hat{H}(R, r)\Psi(R, r) = \varepsilon\Psi(R, r). \quad (1.48)$$

Здесь  $\{R\}$  и  $\{r\}$  — наборы всех ядерных и электронных координат, соответственно, а  $\hat{H}(R, r)$  является полным (включающим и электроны, и ядра) молекулярным гамильтонианом:

$$\hat{H}(R, r) = \underbrace{\sum_{\alpha=1}^{N_N} -\frac{\hbar^2}{2M_{\alpha}}\Delta_{\alpha}}_{\hat{T}_N} + \underbrace{\sum_{i=1}^{N_e} -\frac{\hbar^2}{2m_e}\Delta_i}_{\hat{T}_e} + \underbrace{V(R, r)}_V, \quad (1.49)$$

ядерная                      электронная                      потенциальная энергия  
кинетическая энергия

где

$$V \equiv V(R, r) = - \underbrace{\sum_{\alpha=1}^{N_N} \sum_{i=1}^{N_e} e_0^2 \frac{Z_{\alpha}}{r_{\alpha i}}}_{V_{eN}} + \underbrace{\sum_{\alpha < \beta} e_0^2 \frac{Z_{\alpha} Z_{\beta}}{R_{\alpha\beta}}}_{V_{NN}} + \underbrace{\sum_{i < j} e_0^2 \frac{1}{r_{ij}}}_{V_{ee}} \quad (1.50)$$

Поскольку мы используем нерелятивистское приближение, гамильтониан  $\hat{H}$  содержит только кинетическую энергию и электростатические взаимодействия; более слабыми, например магнитными (спин-спиновым, спин-орбитальным), взаимодействиями мы пренебрегаем. Эти слабые взаимодействия должны учитываться либо в *очень* точных расчетах либо при описании собственно магнитных явлений. Важность этих (в сущности релятивистских) эффектов увеличивается с увеличением атомного номера, так что их часто приходится учитывать при рассмотрении тяжелых атомов.

## 2.2. Разделение электронных и ядерных переменных по Борну—Оппенгеймеру

Рассмотрим  $\Psi(R, r)$  просто как произведение:

$$\Psi(R, r) \approx \Psi_e(r; R)\Psi_N(R), \quad (1.51)$$

где  $\Psi_e(r; R)$  есть электронная волновая функция в поле *фиксированных* ядер ( $\{R\} = \text{const}$ ), а  $\Psi_N(R)$  — ядерная волновая функция. Это позволяет разделить решение задачи на два отдельных «шага».



# 1. Решение электронного уравнения Шрёдингера:

$$\hat{H}_e \Psi_e(r; R) = E(R) \Psi_e(r; R), \quad (1.52)$$

где

$$\hat{H}_e = \hat{T}_e + V(R, r) = -\frac{\hbar^2}{2m_e} \sum_{i=1}^{N_e} \Delta_i + V(R, r) \quad (1.53)$$

представляет собой электронный гамильтониан (включающий энергию межъядерного отталкивания).

Вычисление  $E(R)$  и  $\Psi_e(r; R)$  происходит при *фиксированных* значениях координат  $\{R\}$ ; соответственно, они зависят от ядерных координат как от параметров. Уравнение (1.52) обычно имеет много различных решений  $\Psi_e(r; R)$ ; каждое из них определяет ППЭ  $E(R)$  для данного (основного или возбужденного) электронного состояния.  $E(R)$ , определенная таким образом, затем используется как потенциальная энергия для исследования движения ядер.

# 2. Решение ядерного уравнения Шрёдингера:

$$\hat{H}_N \Psi_N(R) = \varepsilon \Psi_N(R), \quad (1.54)$$

где

$$\hat{H}_N = -\sum_{\alpha=1}^{N_N} \frac{\hbar^2}{2M_\alpha} \Delta_\alpha + E(R) \quad (1.55)$$

представляет собой гамильтониан движения ядер. Заметим, что  $\Psi_N$  (в отличие от  $\Psi_e$ ) является функцией *только* ядерных координат<sup>1</sup>.

<sup>1</sup> Выражение «зависит от координат ядер, как от параметров» превратилось в заклинание, смысл которого часто ускользает от читателя, а особенно от слушателя на лекциях. Речь идет вот о чем. На самом деле уравнение (1.52) представляет собой не одно, а целое семейство разных уравнений: для каждого значения  $R$  ( $R$  — *параметр*, позволяющий различить уравнения в этом семействе). Каждое уравнение этого семейства есть дифференциальное уравнение (второго порядка в частных производных), служащее для определения функций, зависящих от *переменных*  $r$ . Понятно, что решения *разных* уравнений сами тоже разные и для них указание на то, какое уравнение, собственно, решалось, обозначается как зависимость собственных функций  $\Psi_e(r; R)$  от переменных  $r$  и еще от параметров  $R$ . От  $R$  зависят также собственные значения  $E(R)$ . Для каждого  $R$  уравнение (1.52) имеет *много* решений, отвечающих различным состояниям  $\Psi_e(r; R)$ . Эти решения (состояния) нумеруются какими-то квантовыми числами и соответствующие им ППЭ  $E(R)$ , конечно же, являются разными функциями. Уравнение (1.54) написано для какой-то одной функции  $E(R)$ , получившейся решением всего параметризованного при помощи  $R$  семейства уравнений (1.52). Это уже *одно* дифференциальное уравнение (тоже второго порядка в частных производных), служащее для определения функций, зависящих от *переменных*  $R$ . — *Прим. ред.*

### 2.3. Почему разделение переменных по Борну—Оппенгеймеру не является точным

Подставим волновую функцию, записанную в виде произведения (1.51), в левую часть полного уравнения Шрёдингера (1.48) для электронов и ядер:

$$\hat{H}(R, r)\Psi(R, r) = \hat{H}(R, r)\Psi_e(r; R)\Psi_N(R) . \quad (1.56)$$

При исследовании этого уравнения критической величиной оказывается кинетическая энергия ядер:



$$\hat{T}_N = \sum_{\alpha} -\frac{\hbar^2}{2M_{\alpha}}\Delta_{\alpha} . \quad (1.57)$$

(Далее мы будем опускать пределы суммирования, когда они очевидны.)

$$\Delta_{\alpha} = \frac{\partial^2}{\partial X_{\alpha}^2} + \frac{\partial^2}{\partial Y_{\alpha}^2} + \frac{\partial^2}{\partial Z_{\alpha}^2} \quad (1.58)$$

Вычисляя вторую производную произведения

$$\frac{\partial^2}{\partial X_{\alpha}^2} [\Psi_e(r; R)\Psi_N(R)] = \Psi_N \frac{\partial^2 \Psi_e}{\partial X_{\alpha}^2} + 2 \frac{\partial \Psi_e}{\partial X_{\alpha}} \frac{\partial \Psi_N}{\partial X_{\alpha}} + \Psi_e \frac{\partial^2 \Psi_N}{\partial X_{\alpha}^2} \quad (1.59)$$

и выполняя аналогично дифференцирование по  $Y_{\alpha}$  и  $Z_{\alpha}$ , мы получаем

$$\begin{aligned} \sum_{\alpha} -\frac{\hbar^2}{2M_{\alpha}}\Delta_{\alpha} [\Psi_e(r; R)\Psi_N(R)] \\ = \sum_{\alpha} -\frac{\hbar^2}{2M_{\alpha}} [\Psi_N \Delta_{\alpha} \Psi_e + 2(\nabla_{\alpha} \Psi_e \nabla_{\alpha} \Psi_N) + \Psi_e \Delta_{\alpha} \Psi_N] . \end{aligned} \quad (1.60)$$



Первые два слагаемых  $\Psi_N \Delta_{\alpha} \Psi_e$  и  $2(\nabla_{\alpha} \Psi_e \nabla_{\alpha} \Psi_N)$  содержат производные электронной волновой функции по ядерным координатам, а также первые производные волновой функции ядер по их координатам — никакие такие производные не входят ни в электронное, ни в ядерное уравнение Шрёдингера (1.52) и (1.54), что свидетельствует о том, что решая сначала (1.52), а потом (1.54), нельзя получить точное решение полного молекулярного уравнения Шрёдингера (1.48). Указанные вклады приводят к некоторому (обычно слабому, но не всегда пренебрежимо малому) зацеплению между различными электронными состояниями приближения Борна—Оппенгеймера, возникающему за счет движения ядер.

Подставляя волновую функцию (1.51) в общее уравнение Шрёдингера (для электронов и ядер) и обозначая сумму всех таких зацепляющих вкладов

$$B = \sum_{\alpha} -\frac{\hbar^2}{2M_{\alpha}} [\Psi_N \Delta_{\alpha} \Psi_e + 2(\nabla_{\alpha} \Psi_e \nabla_{\alpha} \Psi_N)], \quad (1.61)$$

мы можем записать

$$\begin{aligned} \hat{H} \Psi_e \Psi_N &= B + \underbrace{\Psi_e \sum_{\alpha} -\frac{\hbar^2}{2M_{\alpha}} \Delta_{\alpha} \Psi_N}_{\hat{T}_N \Psi_N} + \underbrace{\sum_i -\frac{\hbar^2}{2m_e} \Delta_i \Psi_e \Psi_N + V \Psi_e \Psi_N}_{(\hat{T}_e \Psi_e) \cdot \Psi_N = \Psi_N \cdot (\hat{T}_e \Psi_e)} \\ &= B + \Psi_e \hat{T}_N \Psi_N + \underbrace{\Psi_N (\hat{T}_e \Psi_e + V \Psi_e)}_{\substack{\text{левая часть электронного} \\ \text{уравнения Шрёдингера} = E(R) \Psi_e}} \quad (1.62) \\ &= B + \Psi_e \hat{T}_N \Psi_N + \Psi_N E(R) \Psi_e = B + \underbrace{\Psi_e [\hat{T}_N + E(R)] \Psi_N}_{\substack{\text{левая часть ядерного} \\ \text{уравнения Шрёдингера} = \varepsilon \Psi_N}} \\ &= B + \varepsilon \Psi_e \Psi_N, \end{aligned}$$

так что  $\Psi_e(r; R) \Psi_N(R)$  не может быть (точной) собственной функцией из-за присутствия вкладов  $B$ .

## 2.4. Приближенное разделение уравнений

### а) Приближение Борна—Оппенгеймера

Полное пренебрежение зацепляющими вкладами — самый обычный подход. Они относительно малы, поскольку представляют собой производные электронной волновой функции, деленные на массы ядер, которые много больше массы электрона. (Электронная волновая функция сильно зависит от электрон-ядерных расстояний, так что можно ожидать, что производные по  $\{r\}$  и  $\{R\}$  — величины одного порядка.) Обычно можно пренебречь зацепляющими вкладами вблизи равновесных конфигураций молекул, поскольку они обычно (много) меньше, чем ошибки, которые мы делаем при решении самих уравнений, получающихся в приближении Борна—Оппенгеймера.

### б) Адиабатическое приближение

(Этот термин не связан с адиабатическими процессами в термодинамике. Не вполне корректно также называть приближение Борна—Оппенгеймера адиабатическим.)

В этом случае вычисляется среднее значение зацепляющих вкладов по рассматриваемому электронному состоянию  $\Psi_e$  (их диагональные матричные элементы):

$$\sum_{\alpha} -\frac{\hbar^2}{2M_{\alpha}} (\langle \Psi_e | \Delta_{\alpha} | \Psi_e \rangle + 2 \langle \Psi_e | \nabla_{\alpha} | \Psi_e \rangle \nabla_{\alpha}), \quad (1.63)$$

которые добавляют к  $\hat{H}_N$ , используемому в ядерном уравнении Шрёдингера. Заметьте, что второе слагаемое в скобках все еще является дифференциальным оператором, зависящим от  $R$ , который действует на  $\Psi_N(R)$ <sup>1</sup>.

Интересно сравнить результаты, полученные на различных уровнях теории на примере простейшей молекулы, для которой доступны очень точные расчеты: молекулы  $\text{H}_2$ . Ниже в таблице приведены энергии диссоциации (в  $\text{см}^{-1}$ ) молекул различного изотопного состава (нулевая энергия колебаний включена).

Приближение	$\text{H}_2$	HD	$\text{D}_2$
Борна—Оппенгеймера	36112.2	36401.5	36745.6
Адиабатическое	36118.0	36405.7	36748.3
Неадиабатическое	36114.7	36402.9	36746.2
Эксперимент	36113.6	36400.5	36744.2

Приближение Борна—Оппенгеймера дает вполне приемлемые результаты и, более того, здесь имеет место (по меньшей мере в этом случае) случайная компенсация ошибок.

Можно заметить, что большинство (термических) химических реакций являются «адиабатическими» в том смысле, что система остается на ППЭ, отвечающей данному (обычно основному) электронному состоянию. Однако существуют и «неадиабатические» реакции с переходом с одной ППЭ на другую, происходящим обычно при тех конфигурациях ядер, когда две ППЭ, отвечающие различным электронным состояниям, подходят близко друг к другу.

<sup>1</sup> В выражении (1.63) угловые скобки (определение обозначений «бра» и «кет», см. приложение П7) подразумевают интегрирование по координатам электронов. — *Прим. ред.*

## 2.5. Замечания по поводу разделения переменных по Борну—Оппенгеймеру

Необходимо подчеркнуть, что запись волновой функции  $\Psi(R, r)$  в виде произведения  $\Psi_e(r; R)\Psi_N(R)$  сама по себе не является приближением: приближение состоит в том, что предполагается, что эти функции являются решениями отдельных электронного и ядерного уравнений Шрёдингера (1.52) и (1.54). Фактически можно разделить и умножить  $\Psi(R, r)$  на любую  $\Psi_N(R)$  и формально написать  $\Psi_e(r; R) = \Psi(R, r)/\Psi_N(R)$ . Более того, если  $\Psi(R, r)$  нормирована, то она может быть представлена как произведение вида  $\Psi_e(r; R)\Psi_N(R)$ , в котором как электронная волновая функция  $\Psi_e$ , так и ядерная волновая функция  $\Psi_N(R)$  нормированы по отдельности. (При вычислении нормы  $\Psi_e$  необходимо интегрировать только по электронным координатам<sup>1</sup>.)

Чтобы получить представление такого типа, мы интегрируем  $|\Psi(R, r)|^2$  по электронным координатам (включая суммирование по спинам  $d\tau_i = dv_i d\sigma_i$ ) и определяем функцию  $q(R)$  как результат этого интегрирования:

$$q(R) = \int |\Psi(R, r)|^2 d\tau_1 d\tau_2 \dots d\tau_{N_e} . \quad (1.64)$$

Тогда нормированная электронная волновая функция для каждой ядерной конфигурации  $\{R\}$  может быть записана, как

$$\Psi_e(r; R) = \frac{\Psi(R, r)}{\sqrt{q(R)}} , \quad (1.65)$$

и ядерная функция есть просто

$$\Psi_N(R) = \sqrt{q(R)} . \quad (1.66)$$

Ее нормированность обусловлена предположением, что полная волновая функция  $\Psi(R, r)$  нормирована. К сожалению, по-видимому, не существует прямой связи между функциями (1.65), (1.66) и функциями, полученными путем решения отдельных электронного и ядерного уравнений Шрёдингера (1.52) и (1.54). (Если среди ядер есть идентичные фермионы, то волновая функция обращается в нуль для геометрических конфигураций, в которых два из них находятся в одной и той же точке пространства и имеют один и тот же спин. Для таких систем предыдущее определение должно быть расширено с помощью дополнительного условия: в этих точках  $\Psi_N(R) = 0$ ; тогда в этих точках  $\Psi_e(r; R)$  может считаться произвольной при условии, что она остается конечной.)

<sup>1</sup> Ведь от  $R$  она зависит, как от параметров. — *Прим. ред.*

## Библиографические заметки

### Раздел 1.1.

Большинство авторов выполняют это преобразование без обсуждения каких-либо деталей его вывода (например, [1-3]) или просто отделяют движение центра масс, принимая это как нечто заведомо известное (из классической механики; см. приложение П2). Вывод, основанный на требовании равенства нулю суммарного импульса, возможно, оригинален (но претензии на это отсутствуют).

### Раздел 1.2.

Представленный вывод отличается от вывода в [4] только нумерацией частиц.

### Раздел 1.3.

Вывод подобного типа кратко обсуждается, например, в [5, 6] в связи с отделением движения центра масс. Этот подход обсуждается в деталях в [7] для специального случая атома He.

### Разделы 2.1.-2.4.

По существу, стандартное изложение (например, [8]). Концепция приближения Борна—Оппенгеймера сначала появилась в [9] на основе качественных аргументов: если движение электронов и ядер рассматривается как колебательное, то частоты, приписываемые им, соотносятся как обратные квадратные корни масс, т. е. электроны движутся с частотой, в  $\sqrt{\frac{m_N}{m_e}} \sim 100$  раз большей, чем частота движения ядер. Разделение переменных по Борну—Оппенгеймеру было впервые введено в [10]. В литературе повторяется ошибочное утверждение, что в этой статье точность разделения электронного и ядерного движения была установлена вплоть до порядка по параметру  $\kappa = \sqrt{\frac{m_e}{m_N}} \sim \frac{1}{10}$ . Однако это не так; параметр  $\kappa$  был использован в обсуждаемой работе в некоторых разложениях в ряды, необходимых для установления относительной величины электронной, колебательной и вращательной энергий; точность же разделения электронного и ядерного движений не исследовалась. При ссылках на современную форму приближения Борна—Оппенгеймера обычно упоминают работу [11].

### Раздел 2.5.

Авторское рассмотрение.

## Для дальнейшего изучения

В [4] можно найти детальное обсуждение использования правил дифференцирования сложной функции для преобразования координат, а также «трюка Подольского», используемого для перехода от классической функции Гамильтона к квантовомеханическому гамильтониану в криволинейных координатах.

В связи с задачами, возникающими «за пределами приближения Борна—Оппенгеймера» можно рекомендовать две обзорные статьи [12, 13] и упомянуть интересную попытку [14] предложить полностью независимую теорию, в которой «кластерное разложение» волновой функции используется для выра-

жения сильной взаимной зависимости между электронами и ядрами без фиксации положений последних.

## Литература



1. Fermi E. *Notes on Quantum Mechanics*, University of Chicago Press, 1960. (Имеется русский перевод: Ферми Э. *Лекции по квантовой механике*. — М.: РИХД, 2000.)
2. Mott N., Sneddon I. *Wave Mechanics and its Applications*, Clarendon Press, Oxford 1948. (Имеется русский перевод: Мотт Н., Снеддон И. *Волновая механика и ее применения*. — М.: Наука, 1966.)
3. Ландау Л. Д., Лифшиц И. М. *Квантовая механика*. — М.: Наука, 1974.
4. Bunker P. R., Jensen P. *Molecular Symmetry and Spectroscopy*, NRC Research Press, Ottawa 1998. (Имеется русский перевод: Банкер Ф., Йенсен П. *Молекулярная симметрия и спектроскопия*. — М.: Мир, 2004.)
5. Zülücke L. *Quantenchemie*, Deutcher Verlag der Wissenschaften, Berlin 1973 (Имеется русский перевод: Цюлике Л. *Квантовая химия*. — М.: Мир, 1976).
6. Sutcliffe B. T. *Fundamentals in Computational Quantum Chemistry*, in *Computational Techniques in Quantum Chemistry and Molecular Physics*, (ed. G.H.F. Dierksen, B.T. Sutcliffe and A. Veillard) Reidel, Dordrecht 1974.
7. Bethe H. A., *Intermediate Quantum Mechanics*, Benjamin, New York 1964. (Имеется русский перевод: Бете Г. *Квантовая механика*. — М.: Мир, 1965.)
8. Slater J. C. *Quantum Theory of Molecules and Solids vol. 1. Electronic Structure of Molecules*, McGraw-Hill, New York 1963. (Имеется русский перевод: Слэтер Дж. *Электронная структура молекул*. — М.: Мир, 1965.)
9. Born M., Heisenberg W. *Ann. Phys.* **84**, 1 (1924).
10. Born M., Oppenheimer J. R. *Ann. Phys.* **84**, 457 (1927).
11. Born M., Huang K. *Dynamical Theory of Crystal Lattices*, Oxford University Press, London 1954, Appendix VIII. (Имеется русский перевод: Борн М., Хуан Кунь. *Динамическая теория кристаллических решеток*. — М.: ИЛ, 1958.)
12. Azumi T., Matsuzaki K. *Photochem. Photobiol.*, **25**, 315 (1977).
13. Baer M. *Physics Reports*, **358**, 75 (2002).
14. Monkhorst H. J. *Phys. Rev.* **A36**, 1544 (1987).

# Общие теоремы и принципы

## 1. Вариационный принцип



### 1.1. Среднее значение

Квантовомеханическое среднее значение энергии для решений уравнения Шрёдингера равно собственному значению гамильтониана  $\hat{H}$ . В самом деле, точные решения суть «собственные состояния оператора энергии»:

$$\hat{H}\Psi = E\Psi . \quad (2.1)$$

Умножим (2.1) слева на  $\Psi^*$  и проинтегрируем:

$$\langle \Psi | \hat{H} | \Psi \rangle = E \langle \Psi | \Psi \rangle . \quad (2.2)$$

Разделив обе части на  $\langle \Psi | \Psi \rangle$ , мы получим энергию как *среднее значение* («коэффициент Рэлея»)

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} . \quad (2.3)$$

Для приближенной волновой функции можно использовать только среднее значение, так как  $\frac{\hat{H}\Psi}{\Psi} \neq const$ .

### 1.2. Вариационный принцип для основного состояния

В классе волновых функций, удовлетворяющих граничным условиям рассматриваемой задачи, *наименьшее возможное среднее значение  $E_0$  энергии  $E$  принадлежит точной волновой функции основного состояния*. Поэтому для любого  $\Psi$

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0 . \quad (2.4)$$

(Равенство справедливо только для точной волновой функции основного состояния.) Поэтому энергия, вычисленная с использованием приближенной волновой функции, дает верхнюю границу (оценку сверху) энергии точного основного состояния.



### Доказательство

Гамильтониан  $\hat{H}$  эрмитов; следовательно, его собственные векторы  $\Psi_i$  образуют полный ортонормированный набор функций, и любую волновую функцию  $\Psi$  можно разложить в ряд

$$\Psi = \sum_i c_i \Psi_i. \quad (2.5)$$

(Если  $\hat{H}$  имеет непрерывный спектр, тогда сумма должна быть заменена — или дополнена — интегрированием, при этом возникают и другие трудности; мы не будем рассматривать здесь эту возможность, так как она по существу ни на что не влияет.)

Вычислим

$$\langle \Psi | \Psi \rangle = \left\langle \sum_i c_i \Psi_i \left| \sum_j c_j \Psi_j \right. \right\rangle = \sum_{i,j} c_i^* c_j \langle \Psi_i | \Psi_j \rangle = \sum_{i,j} c_i^* c_j \delta_{ij} = \sum_i |c_i|^2 \quad (2.6)$$

и

$$\begin{aligned} \langle \Psi | \hat{H} | \Psi \rangle &= \left\langle \sum_i c_i \Psi_i \left| \hat{H} \left| \sum_j c_j \Psi_j \right. \right. \right\rangle = \sum_{i,j} c_i^* c_j \langle \Psi_i | \hat{H} | \Psi_j \rangle \\ &= \sum_{i,j} c_i^* c_j E_j \langle \Psi_i | \Psi_j \rangle = \sum_{i,j} c_i^* c_j E_j \delta_{ij} = \sum_i |c_i|^2 E_i. \end{aligned} \quad (2.7)$$

Поэтому

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_i |c_i|^2 E_i}{\sum_i |c_i|^2}. \quad (2.8)$$

Так как  $E_i \geq E_0$ , получаем

$$E \geq \frac{E_0 \sum_i |c_i|^2}{\sum_i |c_i|^2} = E_0. \quad \text{Ч. т. д.} \quad (2.9)$$

Тот факт, что среднее значение энергии для точной волновой функции представляет собой минимум, указывает на то, что точность энергии, вычисленной как среднее значение, обычно намного больше, чем точность волновой функции, для которой эта энергия определяется. В самом деле, энергия как функционал волновой функции в своем минимуме должна быть стационарной, т. е. ее первая вариация должна обращаться в нуль. Поэтому ошибка в энергии квадратична относительно ошибки волновой функции: если отклонение волновой функции от точной пропорционально некоему (малому) параметру  $\varepsilon$ , то среднее значение энергии будет отличаться от точного на величину, пропорциональную  $\varepsilon^2$ . На практике это означает, что если параметры разложения волновой функции определены до  $k$ -го знака, то следует ожидать значение энергии до  $2k$ -го знака (на самом деле даже больше).

### 1.3. Вариационный принцип как эквивалент уравнения Шрёдингера. Полезная формулировка вариационного принципа

Уравнение Шрёдингера можно эквивалентно представить и как решение некоторой вариационной задачи: требуется найти волновую функцию, для которой энергия (функционал энергии, среднее значение гамильтониана) стационарна, т. е.

$$\delta E = \delta \left[ \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \right] = 0 \quad (2.10)$$

Условие обращения в нуль первой вариации энергии и уравнение Шрёдингера эквивалентны не только для основного состояния, но и для возбужденных (что недостаточно осознается). Если использовать процедуру минимизации энергии для поиска возбужденных состояний, то необходимо дополнительно потребовать ортогональности искомого решения к состояниям с более низкой энергией. Исключениями являются наименьшие состояния для каждого типа симметрии, для которых ортогональность удовлетворяется автоматически, — они являются основными состояниями для своего типа симметрии.

В выражении (2.10) можно проварьировать дробь аналогично дифференцированию (об аналогии между вариациями и дифференциалами рекомендуется прочитать в приложении ПЗ):

$$\delta E = \frac{\langle \delta \Psi | \hat{H} | \Psi \rangle \langle \Psi | \Psi \rangle - \langle \delta \Psi | \Psi \rangle \langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle^2} + \text{с.с.} = 0 \quad (2.11)$$

(с.с. = комплексно сопряженное).

Так как  $\delta \Psi$  произвольна, ее фаза также произвольна (т. е.  $\delta \Psi$  содержит произвольный фазовый множитель), поэтому написанное явно выражение и комплексно сопряженное к нему слагаемое должны быть равны нулю каждое по отдельности. После умножения на  $\langle \Psi | \Psi \rangle \neq 0$  (мы рассматриваем только нормируемые волновые функции) получаем

$$\underbrace{\langle \delta \Psi | \hat{H} | \Psi \rangle \frac{\langle \Psi | \Psi \rangle}{\langle \Psi | \Psi \rangle}}_1 - \underbrace{\langle \delta \Psi | \Psi \rangle \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}}_E = 0. \quad (2.12)$$

Значит,

$$\langle \delta \Psi | \hat{H} - E | \Psi \rangle = 0, \quad (2.13)$$

что является очень полезной формой вариационного принципа, которую мы будем часто использовать. Вариация  $\delta \Psi$  произвольна, если никакие

ограничения не наложены на  $\Psi$ , поэтому условие (2.13) полностью эквивалентно<sup>1</sup> уравнению Шрёдингера:

$$(\hat{H} - E)\Psi = 0. \quad (2.14)$$

Полученная форма вариационного принципа (2.13) пригодна и в тех случаях, когда мы используем не произвольную пробную волновую функцию  $\Psi$  в процедуре варьирования, а ищем приближение к точной волновой функции, применяя вариационную волновую функцию, относящуюся к некоторому заданному типу («Ansatz»). В данном случае вариация  $\delta\Psi$  больше не является абсолютно произвольной, а соответствует вариации  $\Psi$ , не выходящей за пределы выбранного типа волновых функций. Это означает, что волновая функция  $\Psi + \delta\Psi$ , полученная после вариации, также принадлежит к данному типу, что ограничивает возможные вариации  $\delta\Psi$ . В таких случаях (2.13) неэквивалентно уравнению Шрёдингера, но позволяет получить вариационные уравнения для выбранного типа пробных волновых функций.

### Использование множителей Лагранжа

Эквивалентность вариационного принципа и уравнения Шрёдингера можно доказать, используя также и метод множителей Лагранжа. Задача тогда формулируется следующим образом: потребуем стационарности функционала (интеграла)  $F = \langle \Psi | \hat{H} | \Psi \rangle$  при дополнительном условии  $\langle \Psi | \Psi \rangle = 1$ . Введем вспомогательный функционал  $F'$ , добавив к  $F$  левую часть дополнительного условия, переписанного в виде

$$\langle \Psi | \Psi \rangle - 1 = 0 \quad (2.15)$$

и умноженного на неопределенный множитель Лагранжа  $\lambda$

$$F' = \langle \Psi | \hat{H} | \Psi \rangle + \lambda(\langle \Psi | \Psi \rangle - 1), \quad (2.16)$$

и потребуем обращения в нуль вариации  $\delta F'$ . К получающемуся уравнению необходимо добавить еще условие (2.15), или, что то же самое, дополнительно потребовать стационарности  $F'$ , как функции скаляра  $\lambda$ . [Из условия  $\frac{\partial F'}{\partial \lambda} = 0$  немедленно восстанавливается условие (2.15).] Получаем

$$\langle \delta\Psi | \hat{H} | \Psi \rangle + \lambda \langle \delta\Psi | \Psi \rangle + c. c. = 0. \quad (2.17)$$

<sup>1</sup> Согласно «основной лемме вариационного исчисления», функция, ортогональная к любой произвольно взятой функции, равна нулю по крайней мере «почти всюду», т. е. может иметь ненулевые значения только «на множестве меры нуль».

Так как  $\delta\Psi$  произвольна и содержит произвольный фазовый множитель, из условия (2.17) следует, что

$$\hat{H}\Psi + \lambda\Psi = 0 . \quad (2.18)$$

Теперь умножим (2.18) на  $\Psi^*$ , проинтегрируем, и, учитывая соотношение (2.15), получим

$$\langle\Psi|\hat{H}|\Psi\rangle + \lambda = 0 , \quad (2.19)$$

т. е.  $\lambda = -E$ . Подставив этот результат в уравнение (2.18), получим

$$\hat{H}\Psi - E\Psi = 0 . \text{ Ч. т. д. } \quad (2.20)$$

#### 1.4. Неравенство Эккарта

Вариационный принцип говорит нам, что в точном основном состоянии имеется абсолютный минимум энергии, а неравенство Эккарта показывает, что, как только мы достаточно близки к решению, уменьшение энергии, получающейся с данной пробной функцией, указывает и на улучшение самой волновой функции; другими словами, когда энергия последовательно приближается к своему точному значению, волновая функция также приближается к точной.

Рассмотрим пробную волновую функцию  $\Psi$ . Она является линейной комбинацией точных собственных функций гамильтониана (каждая из которых нормирована на единицу) с коэффициентами  $c_i$ :

$$\Psi = \sum_{i=0} c_i \Psi_i . \quad (2.21)$$

Положим, что

$$\langle\Psi|\Psi\rangle = \sum_{i=0} |c_i|^2 = 1 . \quad (2.22)$$

Рассмотрим перекрывание (скалярное произведение) пробной волновой функции  $\Psi$  и точного основного состояния  $\Psi_0$

$$S = \langle\Psi_0|\Psi\rangle = c_0 , \quad (2.23)$$

где последнее равенство следует из ортонормировки точных решений  $\Psi_i$ . Можно написать

$$\begin{aligned} E &= \sum_{i=0} |c_i|^2 E_i = |S|^2 E_0 + \sum_{i=1} |c_i|^2 E_i \quad (\text{так как } E_i \geq E_1 \text{ если } i \geq 1) \\ &\geq |S|^2 E_0 + \sum_{i=1} |c_i|^2 E_1 = |S|^2 E_0 + E_1 \sum_{i=1} |c_i|^2 . \end{aligned} \quad (2.24)$$

Так как  $\sum_{i=0} |c_i|^2 = 1$  и  $|c_0|^2 = |S|^2$ , очевидно имеем  $\sum_{i=1} |c_i|^2 = 1 - |S|^2$ ,  
так что

$$E \geq |S|^2 E_0 + (1 - |S|^2) E_1, \quad (2.25)$$

откуда

$$|S|^2 \geq \frac{E_1 - E}{E_1 - E_0}. \quad (2.26)$$

Поэтому если  $E \rightarrow E_0$ , то  $|S| \rightarrow 1$ . ( $|S| \leq 1$  в соответствии с определением.) Ч. т. д.

## 1.5. Возбужденные состояния

Гамильтониан является эрмитовым оператором; его собственные векторы, принадлежащие разным собственным значениям, автоматически ортогональны; векторы, принадлежащие одному и тому же (вырожденному, т. е. кратному) собственному значению, также можно выбрать ортогональными. Однако, когда нужно найти собственный вектор возбужденного состояния, можно столкнуться с проблемой «вариационного коллапса»: основное состояние «захватывает» численную процедуру, которая сходится к нему же, вместо того, чтобы сойтись к искомому возбужденному состоянию. В таких ситуациях, как уже было отмечено в разд. 1.3, нужно ввести явное требование ортогональности искомого состояния к уже найденным состояниям, лежащим ниже по энергии. (Мы будем рассматривать явно только одно такое состояние; обобщение на случай нескольких таких состояний тривиально.) Следующие рассуждения применимы также в случае, когда волновая функция рассматриваемого основного состояния не точная, а лишь представляет какое-то приближение к ней.

Мы должны решить следующую задачу. Нужно найти волновую функцию  $|\Psi\rangle$ , имеющую стационарную энергию в подпространстве функций, ортогональных к какой-либо (точной или приближенной) волновой функции основного состояния  $|\Psi_0\rangle$ . (Здесь удобно использовать обозначения «бра-кет», см. приложение П7.) Это означает, что мы должны дополнительно удовлетворить условию:

$$(1 - \hat{P})|\Psi\rangle = |\Psi\rangle \quad (2.27)$$

где

$$\hat{P} = |\Psi_0\rangle\langle\Psi_0| \quad (2.28)$$

есть проекционный оператор на нормированную волновую функцию основного состояния  $|\Psi_0\rangle$ .

Так как мы ищем волновую функцию в подпространстве, ортогональном к  $|\Psi_0\rangle$ , вариация не должна выводить  $|\Psi\rangle$  из этого подпространства. Следовательно, необходимо рассматривать только те вариации  $|\delta\Psi\rangle$ , которые также находятся в том же подпространстве; этого можно добиться, записав допустимую вариацию в виде  $|\delta\Psi\rangle = (1 - \hat{P})|\delta\Phi\rangle$ , где  $|\delta\Phi\rangle$  уже полностью произвольна. Подставляя эту вариацию в общую формулировку вариационного принципа (2.13), мы получим

$$\langle\delta\Phi|(1 - \hat{P})(\hat{H} - E)|\Psi\rangle = 0. \quad (2.29)$$

Используя тот факт, что  $|\delta\Phi\rangle$  совершенно произвольна, получаем

$$(1 - \hat{P})\hat{H}|\Psi\rangle = E(1 - \hat{P})|\Psi\rangle. \quad (2.30)$$

Это уравнение останется верным, если подставить условие  $(1 - \hat{P})|\Psi\rangle = |\Psi\rangle$  в правую часть:

$$(1 - \hat{P})\hat{H}|\Psi\rangle = E|\Psi\rangle. \quad (2.31)$$

Действительно, умножая уравнение (2.31) на  $(1 - \hat{P})$  и используя идемпотентность этого оператора, мы видим, что все решения уравнения (2.31), принадлежащие к ненулевым собственным значениям  $E$ , автоматически ортогональны<sup>1</sup> к  $|\Psi_0\rangle$ :  $|\Psi\rangle = (1 - \hat{P})|\Psi\rangle$ .

Далее, можно использовать это равенство еще раз, вставить дополнительный проектор  $(1 - \hat{P})$  в левую часть уравнения (2.31) и получить уравнение на собственные значения эрмитова оператора:

$$(1 - \hat{P})\hat{H}(1 - \hat{P})|\Psi\rangle = E|\Psi\rangle, \quad (2.32)$$

которое, по существу, является уравнением на собственные значения для гамильтониана, спроектированного на подпространство функций, ортогональных к основному состоянию  $|\Psi_0\rangle$ . Этот оператор обладает тем свойством, что волновая функция основного состояния  $|\Psi_0\rangle$  также является его собственным вектором, причем с нулевым собственным значением. (Следовательно, требуемая ортогональность также гарантируется ортогональностью собственных векторов эрмитовых операторов.)

### Применение множителей Лагранжа

Предыдущую задачу можно решить также с помощью метода множителей Лагранжа. Задача формулируется следующим образом: потребуем

<sup>1</sup> В самом деле, таким образом получаем  $E(1 - \hat{P})|\Psi\rangle = E|\Psi\rangle$ , из чего следует, что  $|\Psi\rangle = (1 - \hat{P})|\Psi\rangle$ , если только  $E \neq 0$ . Это и означает, что  $|\Psi\rangle$  целиком лежит в подпространстве, ортогональном к  $|\Psi_0\rangle$ .

стационарности функционала  $F = \langle \Psi | \hat{H} | \Psi \rangle$  при дополнительных условиях  $\langle \Psi | \Psi \rangle = 1$  и  $\langle \Psi | \Psi_0 \rangle = 0$ .

Как  $\langle \Psi | \Psi \rangle$ , так и  $\langle \Psi | \hat{H} | \Psi \rangle$  необходимо вещественны, но условие  $\langle \Psi | \Psi_0 \rangle = 0$  в общем, комплексном случае эквивалентно двум вещественным условиям  $Re\langle \Psi | \Psi_0 \rangle = 0$  и  $Im\langle \Psi | \Psi_0 \rangle = 0$ . Соответственно, вспомогательный функционал  $F'$  образуется с помощью трех множителей Лагранжа ( $\lambda$ ,  $\mu$ , и  $\nu$ ):

$$F' = \langle \Psi | \hat{H} | \Psi \rangle + \lambda(\langle \Psi | \Psi \rangle - 1) + \mu Re\langle \Psi | \Psi_0 \rangle + \nu Im\langle \Psi | \Psi_0 \rangle . \quad (2.33)$$

Его можно преобразовать следующим образом. Выпишем:

$$Re\langle \Psi | \Psi_0 \rangle = \frac{1}{2}(\langle \Psi | \Psi_0 \rangle + \langle \Psi_0 | \Psi \rangle) \quad (2.34)$$

и

$$Im\langle \Psi | \Psi_0 \rangle = \frac{1}{2i}(\langle \Psi | \Psi_0 \rangle - \langle \Psi_0 | \Psi \rangle) . \quad (2.35)$$

Подставим эти выражения в определение (2.33), соберем слагаемые, содержащие  $\langle \Psi | \Psi_0 \rangle$  и  $\langle \Psi_0 | \Psi \rangle$ , введем обозначения  $\tau = \frac{1}{2}(\mu - i\nu)$  и  $\tau^* = \frac{1}{2}(\mu + i\nu)$  и получим

$$F' = \langle \Psi | \hat{H} | \Psi \rangle + \lambda(\langle \Psi | \Psi \rangle - 1) + \tau\langle \Psi | \Psi_0 \rangle + \tau^*\langle \Psi_0 | \Psi \rangle . \quad (2.36)$$

Потребуем, чтобы  $\delta F' = 0$ , т. е.

$$\delta F' = \langle \delta \Psi | \hat{H} | \Psi \rangle + \lambda\langle \delta \Psi | \Psi \rangle + \tau\langle \delta \Psi | \Psi_0 \rangle + c.c. = 0 . \quad (2.37)$$

Из этого условия (ввиду того, что  $\delta \Psi$  произвольна и содержит произвольный фазовый множитель) получаем систему уравнений

$$\begin{aligned} \hat{H}|\Psi\rangle + \lambda|\Psi\rangle + \tau|\Psi_0\rangle &= 0 \\ \langle \Psi | \Psi \rangle &= 1 \\ \langle \Psi | \Psi_0 \rangle &= 0. \end{aligned} \quad (2.38)$$

Умножая первое уравнение системы (2.38) на  $\langle \Psi |$  и  $\langle \Psi_0 |$  соответственно и используя вспомогательные условия (второе и третье уравнение системы), получим

$$\lambda = -E \quad (2.39)$$

и

$$\tau = -\langle \Psi_0 | \hat{H} | \Psi \rangle . \quad (2.40)$$

( $|\Psi_0\rangle$  предполагается нормированной.) Подставим эти значения в первое из уравнений (2.38):

$$\hat{H}|\Psi\rangle - |\Psi_0\rangle\langle \Psi_0 | \hat{H} | \Psi \rangle = E|\Psi\rangle \quad (2.41)$$

т. е.

$$(1 - \hat{P})\hat{H}|\Psi\rangle = E|\Psi\rangle. \quad (2.42)$$

Мы снова получили уравнение (2.31); его можно привести к эрмитову виду, как описано выше.

## 2. Теорема Гельмана—Фейнмана

При выводе вариационного принципа мы рассматривали произвольные «математические» вариации волновой функции данной системы (данного гамильтониана) с тем, чтобы исследовать поведение функционала энергии при малых изменениях волновой функции в окрестности точного решения. Напротив, теорема Гельмана—Фейнмана рассматривает изменения, вызванные вариациями физических параметров системы (параметров, определяющих ее гамильтониан). Тем не менее, эти результаты имеют много общего.

### 2.1. Дифференциальная теорема Гельмана—Фейнмана

Рассмотрим случай, когда некоторый параметр  $\alpha$ , характеризующий систему, изменяется:

$$\alpha \rightarrow \alpha + d\alpha. \quad (2.43)$$

Тогда изменятся гамильтониан  $\hat{H}$  и, следовательно,  $\Psi$  и  $E$ .

$$\begin{aligned} \hat{H} &\rightarrow \hat{H} + \frac{\partial \hat{H}}{\partial \alpha} d\alpha \\ \Psi &\rightarrow \Psi + \frac{\partial \Psi}{\partial \alpha} d\alpha \\ E &\rightarrow E + \frac{\partial E}{\partial \alpha} d\alpha. \end{aligned} \quad (2.44)$$

(Мы везде опускаем слагаемые выше первого порядка по  $d\alpha$ .) Запишем производную энергии:

$$\begin{aligned} \frac{\partial E}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\ &= \frac{\left( \langle \frac{\partial \Psi}{\partial \alpha} | \hat{H} | \Psi \rangle + \langle \Psi | \frac{\partial \hat{H}}{\partial \alpha} | \Psi \rangle + \langle \Psi | \hat{H} | \frac{\partial \Psi}{\partial \alpha} \rangle \right) \langle \Psi | \Psi \rangle}{\langle \Psi | \Psi \rangle^2} \\ &\quad - \frac{\left( \langle \frac{\partial \Psi}{\partial \alpha} | \Psi \rangle + \langle \Psi | \frac{\partial \Psi}{\partial \alpha} \rangle \right) \langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle^2} \end{aligned}$$



$$= \frac{1}{\langle \Psi | \Psi \rangle} \left\{ \left\langle \frac{\partial \Psi}{\partial \alpha} \right| \hat{H} | \Psi \right\rangle - E \left\langle \frac{\partial \Psi}{\partial \alpha} \right| \Psi \right\rangle + \langle \Psi | \hat{H} \left| \frac{\partial \Psi}{\partial \alpha} \right\rangle - E \langle \Psi | \left| \frac{\partial \Psi}{\partial \alpha} \right\rangle + \langle \Psi | \left| \frac{\partial \hat{H}}{\partial \alpha} \right| \Psi \rangle \right\}. \quad (2.45)$$

Первые четыре слагаемых выражения в фигурных скобках взаимно уничтожаются, если  $\Psi$  есть решение уравнения Шрёдингера, так как в этом случае  $\langle \frac{\partial \Psi}{\partial \alpha} | \hat{H} - E | \Psi \rangle + \text{с.с.} = 0$ . (Можно также сформулировать это так, что вариация  $\delta \Psi = \frac{\partial \Psi}{\partial \alpha} d\alpha$  является некоторой разрешенной «математической» вариацией, для которой выполняется вариационный принцип.) Поэтому получаем

$$\frac{\partial E}{\partial \alpha} = \frac{\langle \Psi | \frac{\partial \hat{H}}{\partial \alpha} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (2.46)$$

Это «дифференциальная теорема Гельмана—Фейнмана» (часто называемая просто теоремой Гельмана—Фейнмана). Она показывает, что для вычисления производной энергии нам не нужна производная волновой функции: производную энергии можно вычислить просто как среднее значение производной гамильтониана, т. е., как это получилось бы и по теории возмущений первого порядка.

Если параметр  $\alpha$  совпадает с одной из ядерных координат  $X_A$  молекулы, описываемой в рамках приближения Борна—Оппенгеймера, то теорема Гельмана—Фейнмана утверждает, что сила, действующая в направлении этой координаты (сопряженная этой координате), определена выражением:

$$F_{X_A} = -\frac{\partial E}{\partial X_A} = -\frac{\langle \Psi | \frac{\partial \hat{H}}{\partial X_A} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \quad (2.47)$$

где  $\Psi$  — электронная волновая функция, а  $\hat{H}$  — полный электронный гамильтониан (учитывающий межъядерное отталкивание).

В электронном гамильтониане от ядерных координат зависят только слагаемые, описывающие электрон-ядерное и межъядерное кулоновские взаимодействия. Поэтому в рамках приближения Борна—Оппенгеймера силы, действующие на данное ядро, определяются электростатическим полем других ядер и зарядового распределения электронного облака, но не содержат никаких двухэлектронных вкладов. Соответственно, чтобы найти эти силы, в принципе не нужно знать электронную волновую функцию, а достаточно знать лишь пространственное распределение электронной плотности. На практике, однако, ситуация не

столь благоприятна, так как мы равно не знаем ни точных волновых функций, ни соответствующих точных зарядовых распределений. (Заметим, что предыдущее заключение, касающееся первостепенного значения электронной плотности, не имеет ничего общего с теоремами, на которых основывается так называемая «теория функционала плотности» — «density functional theory», DFT.)

Полезно показать и явным образом то, что силы, действующие на ядро в молекуле, являются *точно* классическими электростатическими силами, возникающими от распределения зарядов в молекуле, создаваемого ядрами и электронным облаком<sup>1</sup>.

Выписывая выражение (2.47) для каждой из трех координат и предполагая, что электронная волновая функция  $\Psi$  нормирована ( $\langle \Psi | \Psi \rangle = 1$ ), мы можем объединить полученные выражения в одно векторное:

$$\vec{F}_A = -\langle \Psi | \nabla_A \hat{H} | \Psi \rangle. \quad (2.48)$$

Здесь  $\vec{F}_A$  — полная сила, действующая на ядро А. Она является суммой двух сил:  $\vec{F}_{A,NN}$ , которая происходит из потенциальной энергии межъядерного отталкивания  $V_{NN}$ , включенной в гамильтониан, и  $\vec{F}_{A,eN}$ , происходящей из электрон-ядерного притяжения  $V_{eN}$ .

Энергия межъядерного отталкивания  $V_{NN}$  не зависит от электронной волновой функции, так что  $\nabla_A V_{NN}$  можно вынести из-под знака интеграла и последний сводится к единичному множителю. Ядерная составляющая силы определяется просто:

$$\vec{F}_{A,NN} = -\nabla_A \sum_{\alpha < \beta}^{N_N} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} = -\nabla_A \sum_{\substack{B=1 \\ (B \neq A)}}^{N_N} \frac{Z_A Z_B}{R_{AB}}. \quad (2.49)$$

(Мы используем здесь гамильтониан Борна—Оппенгеймера, записанный в атомных единицах; см. приложение Пб; во втором равенстве в формуле (2.49) мы учли, что как  $\alpha$ , так и  $\beta$  могут быть равны А.)

Вычислим стандартным образом

$$\begin{aligned} \frac{\partial}{\partial X_A} \frac{1}{R_{AB}} &= -\frac{1}{R_{AB}^2} \frac{\partial R_{AB}}{\partial X_A} \\ &= -\frac{1}{R_{AB}^2} \frac{\partial}{\partial X_A} \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2 + (Z_B - Z_A)^2} \end{aligned}$$

<sup>1</sup> В русской литературе этот результат обычно называют «электростатической теоремой». — Прим. ред.

$$\begin{aligned}
&= \frac{1}{R_{AB}^2} \frac{X_B - X_A}{\sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2 + (Z_B - Z_A)^2}} \\
&= \frac{1}{R_{AB}^3} (X_B - X_A)
\end{aligned} \tag{2.50}$$



Аналогично вычисляя производные вдоль двух других направлений, получим в векторной форме:

$$\vec{F}_{A,NN} = - \sum_{\substack{B=1 \\ (B \neq A)}}^{N_N} \frac{Z_A Z_B}{R_{AB}^3} \vec{R}_{AB} = - \sum_{\substack{B=1 \\ (B \neq A)}}^{N_N} \frac{Z_A Z_B}{R_{AB}^2} \vec{R}_{AB}^0, \tag{2.51}$$

где  $\vec{R}_{AB}^0$  есть единичный вектор, направленный от ядра А к ядру В.

Для электронно-ядерной составляющей силы вывод совершенно аналогичен, но суммирование идет по электронам, несущим заряд  $-1$ , и все рассматриваемые величины входят в подинтегральное выражение:

$$\vec{F}_{A,eN} = - \langle \Psi | \nabla_A \sum_{\alpha=1}^{N_N} \sum_{i=1}^{N_e} \frac{-Z_\alpha}{r_{\alpha i}} | \Psi \rangle = \langle \Psi | \sum_{i=1}^{N_e} \frac{Z_A}{r_{Ai}^2} \vec{r}_{Ai}^0 | \Psi \rangle. \tag{2.52}$$

Электроны идентичны и предполагается, что волновая функция  $\Psi$  должным образом антисимметризована, так что вклад каждого электрона в выражение (2.52) одинаков; поэтому мы имеем:

$$\vec{F}_{A,eN} = N_e \langle \Psi | \frac{Z_A}{r_{A1}^2} \vec{r}_{A1}^0 | \Psi \rangle \tag{2.53}$$

$$= N_e \int dv_1 \left[ \sum_{\sigma_1} \int \int \int \dots \int d\tau_2 d\tau_3 \dots d\tau_{N_e} |\Psi(1, 2, \dots, N_e)|^2 \right] \frac{Z_A}{r_{A1}^2} \vec{r}_{A1}^0.$$

Здесь  $\int d\tau_i$  обозначает интегрирование по координатам электрона  $i$ , включая суммирование по проекциям его спина, тогда как обозначение  $\int dv_1$  выражает интегрирование по пространственным координатам электрона 1, а  $\sum_{\sigma_1}$  обозначает суммирование по проекциям его спина.

Известно, что функция  $|\Psi(1, 2, \dots, N_e)|^2$  определяет плотность вероятности того, что система частиц имеет координаты  $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_{N_e}$  и проекции спинов  $\sigma_1, \sigma_2, \dots, \sigma_{N_e}$ ; поэтому величина в квадратных скобках в (2.53), которая интегрируется по координатам электронов  $2, 3, \dots, N_e$  и суммируется по проекциям спинов всех электронов, дает *пространственную* плотность вероятности нахождения электрона 1 вблизи точки  $\vec{r}_1$  с произвольной проекцией спина; при этом пространственные и спиновые координаты всех других электронов могут принимать произвольные

значения. Так как электроны идентичны, другие электроны имеют то же самое пространственное распределение вероятности. Поэтому, умножая функцию в квадратных скобках на число электронов  $N_e$ , мы получаем пространственную плотность  $\rho(\vec{r}_1)$  электронного распределения («электронную плотность»). Таким образом, (2.53) можно переписать как

$$\vec{F}_{A,eN} = \int dv_1 \rho(\vec{r}_1) \frac{Z_A}{r_{A1}^2} \vec{r}_{A1}^0, \quad (2.54)$$

что и в самом деле есть результирующая электростатическая сила, действующая на ядро А и со стороны пространственного распределения  $-\rho(\vec{r}_1)$  отрицательного заряда. [Элементарный отрицательный заряд  $-\rho(\vec{r}_1)dv_1$  вызывает элементарную силу  $\frac{dv_1 \rho(\vec{r}_1) Z_A}{r_{A1}^2} \vec{r}_{A1}^0$ ; отметим противоположный знак по сравнению с (2.51).]

Дифференциальная теорема Гельмана—Фейнмана выполняется не только для точных волновых функций, но также для приближенных, полученных с использованием вариационного принципа. Действительно, результат (2.45) можно переписать в виде

$$\frac{\partial E}{\partial \alpha} = \frac{1}{\langle \Psi | \Psi \rangle} \left\{ \langle \frac{\partial \Psi}{\partial \alpha} | \hat{H} - E | \Psi \rangle + \langle \Psi | \hat{H} - E | \frac{\partial \Psi}{\partial \alpha} \rangle + \langle \Psi | \frac{\partial \hat{H}}{\partial \alpha} | \Psi \rangle \right\}, \quad (2.55)$$

т. е. для выполнения теоремы Гельмана—Фейнмана достаточно, чтобы формулировка вариационного принципа (2.13) удовлетворялась для вариации волновой функции вида  $\delta \Psi = \frac{\partial \Psi}{\partial \alpha} d\alpha$ . Таким образом, например, теорема Гельмана—Фейнмана верна для волновой функции Хартри—Фока, полученной решением интегро-дифференциальных уравнений Хартри—Фока, но не для обычно применяемого на практике метода Хартри—Фока—Рутана, где используется ограниченный набор базисных функций. В последнем случае базисные функции «следуют» за ядрами, и изменение геометрического расположения (координат) ядер вызывает также смещение базисных орбиталей. В рамках расчетов по методу Хартри—Фока—Рутана вариационная задача решается для заданных (фиксированных) геометрии и базисного набора. В то же время вариация какой-либо базисной орбитали, вызванная смещением ее центра, в общем случае не может быть разложена по исходному ограниченному базису, который используется в данном расчете. Поэтому бесконечно малое изменение геометрии порождает вариации результирующей волновой функции, не принадлежащие к классу допустимых вариаций, для которых выведены уравнения Хартри—Фока—Рутана. (Сходные рассуждения верны для зависимости базисных орбиталей от геометрии в рамках других вариационных методов, а также методов теории



возмущений и т. д., где используются ограниченные базисные наборы.) Следовательно, в случае ограниченного базиса теорема Гельмана—Фейнмана верна, только если базисные орбитали не фиксированы на ядрах, а положения их центров оптимизируются как независимые вариационные параметры. Это единственный случай, когда вариации многоэлектронной волновой функции, вызванные производными базисных орбиталей, также находятся среди допустимых «математических» вариаций<sup>1</sup>.

Присутствие в вариации энергии вкладов, вызванных тем, что базисные орбитали «следуют за ядрами» (их иногда называют «силами волновой функции»), значительно затрудняет расчеты градиентов энергии, и, следовательно, оптимизацию молекулярных геометрий, представляющую собой очень важную область приложений квантовой химии. Они приводят к тому, что для вычисления сил необходимо знать градиенты двухэлектронных интегралов; подробности см. в гл. 6, разд. 7. (Для невариационных методов, например методов теории возмущений, необходимо определять и градиенты волновой функции.)

## 2.2. Интегральная теорема Гельмана—Фейнмана

Влияние *конечных* приращений параметра  $\alpha$  может быть, вообще говоря, получено интегрированием (например, вычисление потенциала интегрированием силы). Предполагая, что условие нормировки

$$\langle \Psi | \Psi \rangle = 1 \quad (2.56)$$

---

<sup>1</sup> Все это выглядит крайне пессимистично. Теорема (!) Гельмана—Фейнмана не верна ни для одного расчета, выполняемого при помощи общепринятых квантово-химических программ, поскольку все они проводятся в базисе АО, привязанных к ядрам. Спрашивается, как же тогда вычислять градиенты энергии при оптимизации геометрии? На самом деле все не так плохо. В общей формулировке теоремы параметр  $\alpha$  — любой, от которого зависит гамильтониан. В квантовохимических расчетах используется вовсе не сам гамильтониан приближения Борна—Оппенгеймера (1.52), а его проекция (8.22) на подпространство, растянутое детерминантами Слэтера, которые можно построить в подпространстве одноэлектронных состояний, растянутом АО, центрированными на ядрах. Это совсем другой гамильтониан, определенный своими матричными элементами. Его зависимость от ядерных координат гораздо более сложная, чем (1.52), поскольку все матричные элементы, включая двухэлектронные, зависят от базиса, привязанного к ядрам. Производная энергии в этом случае, конечно же, равна среднему от производной этого гамильтониана по любой выбранной координате. Не выполняются формулы (2.51)–(2.53), выражающие градиент энергии исключительно как среднее от производной одноэлектронного потенциала, вычисленное по электронной плотности — так называемая «электростатическая теорема». См. также замечание автора к разд. 2.1 по этому вопросу на с. 60 — *Прим. ред.*

не зависит от  $\alpha$ , и применяя дифференциальную теорему Гельмана—Фейнмана, мы получаем ее «проинтегрированную форму»:

$$\Delta E = E(\alpha_1) - E(\alpha_0) = \int_{\alpha_0}^{\alpha_1} \left( \frac{\partial E}{\partial \alpha} \right) d\alpha = \int_{\alpha_0}^{\alpha_1} \langle \Psi(\alpha) | \frac{\partial \hat{H}}{\partial \alpha} | \Psi(\alpha) \rangle d\alpha. \quad (2.57)$$

$\Delta E$  также можно найти, используя лишь величины, соответствующие граничным значениям ( $\alpha_0$  и  $\alpha_1$ ) для  $\alpha$ . Обозначим разные величины, соответствующие  $\alpha_0$ , как  $\Psi_0$ ,  $\hat{H}_0$ , и  $E_0$ , а те, которые соответствуют  $\alpha_1$ , как  $\Psi_1$ ,  $\hat{H}_1$ , и  $E_1$ , и пусть  $\Delta E$  снова обозначает разность  $E(\alpha_1) - E(\alpha_0)$ . Уравнения Шрёдингера, соответствующие этим двум точкам, суть

$$\begin{aligned} \hat{H}_1 \Psi_1 &= E_1 \Psi_1 \\ \hat{H}_0 \Psi_0 &= E_0 \Psi_0 \end{aligned} \quad (2.58)$$

Умножим первое уравнение (2.58) на  $\Psi_0^*$ , а второе — на  $\Psi_1^*$ , и проинтегрируем их:

$$\begin{aligned} \langle \Psi_0 | \hat{H}_1 | \Psi_1 \rangle &= E_1 \langle \Psi_0 | \Psi_1 \rangle \\ \langle \Psi_1 | \hat{H}_0 | \Psi_0 \rangle &= E_0 \langle \Psi_1 | \Psi_0 \rangle. \end{aligned} \quad (2.59)$$

Вычитая комплексно сопряженное второе уравнение из первого, получаем

$$\langle \Psi_0 | \hat{H}_1 | \Psi_1 \rangle - \langle \Psi_0 | \hat{H}_0 | \Psi_1 \rangle = E_1 \langle \Psi_0 | \Psi_1 \rangle - E_0 \langle \Psi_0 | \Psi_1 \rangle \quad (2.60)$$

т. е.

$$\Delta E = \frac{\langle \Psi_0 | \hat{H}_1 - \hat{H}_0 | \Psi_1 \rangle}{\langle \Psi_0 | \Psi_1 \rangle}. \quad (2.61)$$

Это интегральная теорема Гельмана—Фейнмана. (Она выведена для точных волновых функций, поэтому ее применение к приближенным функциям должно проводиться с определенной осторожностью.)

Дифференциальную теорему Гельмана—Фейнмана также можно вывести из интегральной, рассматривая  $\hat{H}_1 = \hat{H}_0 + \delta \hat{H} = \hat{H}_0 + \frac{\partial \hat{H}}{\partial \alpha} d\alpha$ ,  $\Psi_0 = \Psi$ ,  $\Psi_1 = \Psi + \delta \Psi = \Psi + \frac{\partial \Psi}{\partial \alpha} d\alpha$ ; подставляя их в выражение для интегральной теоремы Гельмана—Фейнмана и оставляя слагаемые только первого порядка. (Можно предположить без ограничения общности, что  $\langle \Psi | \delta \Psi \rangle = 0$ .)

### 3. Теорема вириала в квантовой механике

#### 3.1. Зависимость физической величины от времени

Рассмотрим квантово-механическое среднее значение физической величины  $A$ :

$$\langle A \rangle = \langle \Psi | \hat{A} | \Psi \rangle, \quad (2.62)$$

при условии, что  $\langle \Psi | \Psi \rangle = 1$ . Производная  $\langle A \rangle$  по времени равна

$$\frac{d}{dt} \langle A \rangle = \langle \frac{\partial \Psi}{\partial t} | \hat{A} | \Psi \rangle + \langle \Psi | \frac{\partial \hat{A}}{\partial t} | \Psi \rangle + \langle \Psi | \hat{A} | \frac{\partial \Psi}{\partial t} \rangle. \quad (2.63)$$

Зависящее от времени уравнение Шрёдингера имеет вид

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi \quad (2.64)$$

или

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar} \hat{H} \Psi. \quad (2.65)$$

Подставляя в (2.63), имеем:

$$\frac{d}{dt} \langle A \rangle = \langle \Psi | \frac{\partial \hat{A}}{\partial t} | \Psi \rangle + \frac{i}{\hbar} \langle \hat{H} \Psi | \hat{A} | \Psi \rangle - \frac{i}{\hbar} \langle \Psi | \hat{A} | \hat{H} \Psi \rangle. \quad (2.66)$$

Во втором слагаемом появляется знак «+», так как коэффициент находился в «бра»-части и мы произвели комплексное сопряжение, чтобы вынести его за скобки. Можно преобразовать это слагаемое дальше, пользуясь следующим рассуждением. Оператор  $\hat{L}^\dagger$ , эрмитово сопряженный к  $\hat{L}$ , определяется как оператор, для которого выполняется равенство  $\langle \hat{L}^\dagger \psi | \varphi \rangle = \langle \psi | \hat{L} \varphi \rangle$ , для любых функций  $\psi$  и  $\varphi$ . Известно, что оператор Гамильтона  $\hat{H}$  является эрмитовым (самосопряженным), т. е.  $\hat{H} = \hat{H}^\dagger$ . Следовательно, мы можем переносить его из «бра»- в «кет»-часть выражения для матричных элементов (отметим, что  $\langle \psi | \hat{L} | \varphi \rangle \equiv \langle \psi | \hat{L} \varphi \rangle$ ):

$$\frac{d}{dt} \langle A \rangle = \langle \Psi | \frac{\partial \hat{A}}{\partial t} | \Psi \rangle + \frac{i}{\hbar} \langle \Psi | \hat{H} \hat{A} - \hat{A} \hat{H} | \Psi \rangle \quad (2.67)$$

т. е.

$$\frac{d}{dt} \langle A \rangle = \langle \frac{\partial \hat{A}}{\partial t} \rangle + \frac{i}{\hbar} \langle [\hat{H}; \hat{A}] \rangle \quad (2.68)$$

где

$$[\hat{H}; \hat{A}] = \hat{H} \hat{A} - \hat{A} \hat{H} \quad (2.69)$$

есть коммутатор операторов  $\hat{H}$  и  $\hat{A}$ .

Если  $\hat{A}$  не зависит явно от времени и коммутирует с гамильтонианом (т. е.  $[\hat{H}; \hat{A}] = \hat{H}\hat{A} - \hat{A}\hat{H} = 0$ ), то  $\langle A \rangle = \text{const}$ , т. е.  $A$  сохраняется (является интегралом движения).

(Операция взятия коммутатора обладает свойством дистрибутивности, т. е.  $[\hat{A}; (\hat{B} + \hat{C})] = \hat{A}(\hat{B} + \hat{C}) - (\hat{B} + \hat{C})\hat{A} = \hat{A}\hat{B} + \hat{A}\hat{C} - \hat{B}\hat{A} - \hat{C}\hat{A} = [\hat{A}; \hat{B}] + [\hat{A}; \hat{C}]$ .)

### 3.2. Теорема вириала

Будем рассматривать только стационарные состояния и используем тот факт, что среднее значение любой физической величины в стационарном состоянии должно быть постоянно. (Теорема вириала в классической механике — см. приложение П5, — верна для ограниченных (финитных) движений, а в квантовой механике — для стационарных состояний; эти два случая, в некотором смысле, являются аналогами друг друга в классической и квантовой теориях, соответственно.) Рассмотрим

$$\frac{d}{dt} \langle \sum_i \vec{r}_i \hat{p}_i \rangle = 0. \quad (2.70)$$

Так как  $\frac{\partial}{\partial t} \langle \vec{r}_i \hat{p}_i \rangle = 0$ , это значит, что

$$\langle \Psi | [\hat{H}; \sum_i \vec{r}_i \hat{p}_i] | \Psi \rangle = 0. \quad (2.71)$$

Далее,  $\hat{p}_i = \frac{\hbar}{i} \nabla_i$ , и подставляя выражение для  $\hat{H}$  в предположении, что система является замкнутой, получаем (исключая общий множитель  $\frac{\hbar}{i}$ )

$$\langle \Psi | [ \sum_j \frac{-\hbar^2}{2m_j} \Delta_j + V; \sum_i \vec{r}_i \nabla_i ] | \Psi \rangle = 0. \quad (2.72)$$

Так как  $[\Delta_j; \vec{r}_i \nabla_i] = 0$  для  $j \neq i$ , остается только одна сумма:

$$\sum_i \left\{ \langle \Psi | [ \frac{-\hbar^2}{2m_i} \Delta_i; \vec{r}_i \nabla_i ] | \Psi \rangle + \langle \Psi | [ V; \vec{r}_i \nabla_i ] | \Psi \rangle \right\} = 0. \quad (2.73)$$

Необходимо учесть, что  $\frac{-\hbar^2}{2m_i} \Delta_i = \frac{1}{2m_i} \hat{p}_i^2$  и что разные компоненты импульса коммутируют друг с другом, так же как и с координатами, соответствующим «другим» направлениям. Поэтому коммутатор в первом слагаемом сводится к выражению

$$[ \frac{-\hbar^2}{2m_i} \Delta_i; \vec{r}_i \nabla_i ] = \frac{-\hbar^2}{2m_i} \left( [ \frac{\partial^2}{\partial x_i^2}; x_i \frac{\partial}{\partial x_i} ] + [ \frac{\partial^2}{\partial y_i^2}; y_i \frac{\partial}{\partial y_i} ] + [ \frac{\partial^2}{\partial z_i^2}; z_i \frac{\partial}{\partial z_i} ] \right). \quad (2.74)$$



Используя формулу для вторых производных произведения функций, имеем

$$\begin{aligned} \left[ \frac{\partial^2}{\partial x_i^2}; x_i \frac{\partial}{\partial x_i} \right] \Psi &= \frac{\partial^2}{\partial x_i^2} x_i \frac{\partial \Psi}{\partial x_i} - x_i \frac{\partial}{\partial x_i} \frac{\partial^2 \Psi}{\partial x_i^2} \\ &= 2 \frac{\partial^2 \Psi}{\partial x_i^2} + x_i \frac{\partial^3 \Psi}{\partial x_i^3} - x_i \frac{\partial^3 \Psi}{\partial x_i^3} = 2 \frac{\partial^2 \Psi}{\partial x_i^2} \end{aligned} \quad (2.75)$$

так как  $\frac{\partial^2}{\partial x_i^2} x_i = 0$ , и  $\frac{\partial}{\partial x_i} x_i = 1$ . После аналогичных преобразований для координат  $y$  и  $z$ , записав второй коммутатор, как

$$\begin{aligned} [V; \vec{r}_i \nabla_i] \Psi &= V \vec{r}_i \nabla_i \Psi - \vec{r}_i \nabla_i (V \Psi) = V \vec{r}_i \nabla_i \Psi - \vec{r}_i (\nabla_i V) \Psi - \vec{r}_i V \nabla_i \Psi \\ &= - \vec{r}_i (\nabla_i V) \Psi, \end{aligned} \quad (2.76)$$

получаем

$$\sum_i \left\{ \frac{-\hbar^2}{2m_i} \langle \Psi | 2\Delta_i | \Psi \rangle - \langle \Psi | \vec{r}_i (\nabla_i V) | \Psi \rangle \right\} = 0. \quad (2.77)$$

Сумма в первом слагаемом равна удвоенной кинетической энергии системы, тогда как — если потенциал  $V$  есть однородная функция порядка  $k$ , — второе слагаемое переписывается в соответствии с теоремой Эйлера (приложение П4) как

$$\sum_i \langle \Psi | \vec{r}_i (\nabla_i V) | \Psi \rangle = \langle \Psi | \sum_i \vec{r}_i \frac{\partial V}{\partial \vec{r}_i} | \Psi \rangle = \langle \Psi | kV | \Psi \rangle. \quad (2.78)$$

Поэтому, как и в случае классической механики (приложение П5), мы имеем

$$2\langle T \rangle - k\langle V \rangle = 0. \quad (2.79)$$

Для атомов и молекул (кулоновское взаимодействие,  $k = -1$ ) теорема вириала

$$\langle V \rangle = -2\langle T \rangle; \quad E = \langle T \rangle = \frac{\langle V \rangle}{2} < 0 \quad (2.80)$$

справедлива в случае внутренних движений частиц, а движение центра масс описывается плоской волной, на которую вывод этой формулы не распространяется (впрочем, оно и не представляет интереса). При выводе мы предположили, что волновая функция нормирована; это означает, что теорема вириала справедлива для связанных состояний, т. е. для состояний с отрицательной полной энергией. (Отмеченный факт согласуется с тем, что среднее значение кинетической энергии строго положительное, так как кинетическую энергию можно выразить через квадраты импульсов.)

### 3.3. Масштабирование — связь с вариационным принципом

Рассмотрим какую-либо приближенную волновую функцию системы

$$\Psi = \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N), \quad (2.81)$$

и заменим все  $\vec{r}_i$  на  $\vec{\rho}_i = \eta \vec{r}_i$ . (Предполагается, что мы рассматриваем волновую функцию всей системы, зависящую как от электронных, так и от ядерных координат, и что указанная замена производится для обоих типов координат. В качестве альтернативы можно рассмотреть электронную волновую функцию в приближении БО для случая одного атома с ядром в начале координат.)

Нормированную волновую функцию в новых координатах можно записать как

$$\Psi_\eta = \eta^{\frac{3N}{2}} \Psi(\eta \vec{r}_1, \eta \vec{r}_2, \dots, \eta \vec{r}_N) = \eta^{\frac{3N}{2}} \Psi(\vec{\rho}_1, \vec{\rho}_2, \dots, \vec{\rho}_N), \quad (2.82)$$

предполагая, что исходная функция  $\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$  также была нормирована. В самом деле, так как  $\vec{r}_i = \frac{1}{\eta} \vec{\rho}_i$ , то

$$dv_i = dx_i dy_i dz_i = \frac{1}{\eta^3} d\rho_{x_i} d\rho_{y_i} d\rho_{z_i} = \frac{1}{\eta^3} dv_{\rho_i}, \quad (2.83)$$

и тогда

$$\begin{aligned} & \int \dots \int |\Psi_\eta|^2 dv_1 dv_2 \dots dv_N \\ &= \eta^{3N} \int \dots \int |\Psi(\vec{\rho}_1, \vec{\rho}_2, \dots, \vec{\rho}_N)|^2 dv_1 dv_2 \dots dv_N \\ &= \int \dots \int |\Psi(\vec{\rho}_1, \vec{\rho}_2, \dots, \vec{\rho}_N)|^2 dv_{\rho_1} dv_{\rho_2} \dots dv_{\rho_N} \\ &= \int \dots \int |\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2 dv_1 dv_2 \dots dv_N = 1, \end{aligned} \quad (2.84)$$

где мы учли тот факт, что значение интеграла не зависит от обозначения переменной интегрирования.

Поскольку  $\frac{1}{r} = \frac{\eta}{\rho}$ , получаем

$$\langle \Psi_\eta | \frac{1}{r} | \Psi_\eta \rangle = \eta \langle \Psi_\eta | \frac{1}{\rho} | \Psi_\eta \rangle. \quad (2.85)$$

При вычислении интеграла  $\langle \Psi_\eta | \frac{1}{\rho} | \Psi_\eta \rangle$ , а также  $\langle \Psi_\eta | \frac{\partial^2}{\partial \rho_x^2} | \Psi_\eta \rangle$  (см. ниже), множитель  $\eta^{\frac{3N}{2}}$  уничтожается при вычислении элемента объема, так же, как и в случае интеграла нормировки, и тем же самым способом мы

можем вернуться к исходным обозначениям  $\vec{r}_i$ , что соответствует случаю  $\eta = 1$ . Поэтому (2.85) можно переписать как<sup>1</sup>

$$V(\eta) = \eta V \Big|_{\eta=1} = \eta V(1) . \quad (2.86)$$

Аналогично, так как

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \rho_x} \frac{d\rho_x}{dx} = \eta \frac{\partial}{\partial \rho_x} ; \quad \frac{\partial^2}{\partial x^2} = \eta^2 \frac{\partial^2}{\partial \rho_x^2} \quad (2.87)$$

получаем

$$\langle \Psi_\eta | \frac{\partial^2}{\partial x^2} | \Psi_\eta \rangle = \eta^2 \langle \Psi_\eta | \frac{\partial^2}{\partial \rho_x^2} | \Psi_\eta \rangle , \quad (2.88)$$

т. е.

$$T(\eta) = \eta^2 T \Big|_{\eta=1} = \eta^2 T(1) . \quad (2.89)$$

Получаем энергию как функцию  $\eta$ :

$$E(\eta) = T(\eta) + V(\eta) = \eta^2 T(1) + \eta V(1) . \quad (2.90)$$

Теперь прооптимизируем энергию (2.90) по  $\eta$ , рассматривая параметр  $\eta$  как вариационный:

$$\frac{dE}{d\eta} = 2\eta T(1) + V(1) = 0 , \quad (2.91)$$

откуда находим

$$\eta_{opt} = -\frac{V(1)}{2T(1)} . \quad (2.92)$$

Поэтому

$$T(\eta_{opt}) = \left[ -\frac{V(1)}{2T(1)} \right]^2 T(1) = \frac{V(1)^2}{4T(1)} \quad (2.93)$$

$$V(\eta_{opt}) = -\frac{V(1)}{2T(1)} V(1) = -\frac{V(1)^2}{2T(1)} \quad (2.94)$$

и, соответственно,

$$E_{opt} = -\frac{V(1)^2}{4T(1)} . \quad (2.95)$$

После масштабирования теорема вириала выполняется<sup>2</sup>:

$$\frac{V(\eta_{opt})}{T(\eta_{opt})} = -2 . \quad (2.96)$$

<sup>1</sup> Далее для простоты опускаем обозначение средних величин.

<sup>2</sup> Так как точная волновая функция может считаться оптимизированной по любым возможным параметрам, в том числе и по параметру масштабирования  $\eta$ , этот результат можно рассматривать как независимое доказательство теоремы вириала, причем не только для точной волновой функции, но и для любой достаточно гибкой (чтобы допускать масштабирование) приближенной *вариационной* волновой функции.

### 3.4. Теорема вириала в приближении Борна—Оппенгеймера

В случае фиксированных ядер (приближение Борна—Оппенгеймера), теорема вириала выполняется только для равновесной конфигурации ядер и других стационарных точек ППЭ. (Межъядерное отталкивание включено в  $V$ .) Это можно качественно (классически) объяснить следующим образом. Сила, действующая на ядро  $\alpha$ , в рамках приближения БО равна  $\vec{F}_\alpha = -\frac{\partial E(R)}{\partial \vec{R}_\alpha}$ . Ядра останутся фиксированными, если существуют внешние силы такой же величины, но противоположного направления, которые удерживают ядра в заданных положениях; однако, тогда необходимо учесть и вириал этих внешних сил (ср. приложение П5), т. е. в приближении Борна—Оппенгеймера должно выполняться равенство

$$2\langle T \rangle + \langle V \rangle + \sum_{\alpha} \vec{R}_\alpha \frac{\partial E(R)}{\partial \vec{R}_\alpha} = 0 \quad (2.97)$$

Соответственно, теорема вириала (2.80) выполняется, когда все  $\frac{\partial E(R)}{\partial \vec{R}_\alpha} = 0$ , т. е. в стационарных точках потенциальной поверхности и в случае бесконечного расстояния между атомами.

Перейдем к квантовомеханическому выводу этого результата. Подобно общему случаю, стационарный характер волновой функции приводит снова к уравнению

$$\frac{d}{dt} \langle \sum_i \vec{r}_i \hat{p}_i \rangle = 0, \quad (2.98)$$

где суммирование, однако, ведется только по электронам. Точно так же, как описано выше, можно преобразовать уравнение (2.98) к виду

$$2\langle \Psi | \sum_i -\frac{\hbar^2}{2m_e} \Delta_i | \Psi \rangle - \langle \Psi | \sum_i \vec{r}_i \frac{\partial V}{\partial \vec{r}_i} | \Psi \rangle = 0, \quad (2.99)$$

где мы учли, что все  $m_i$  равны массе электрона  $m_e$ .

Аналогично общему случаю, первое слагаемое в левой части (2.99) равно удвоенной кинетической энергии электронов:

$$2\langle \Psi | \sum_i -\frac{\hbar^2}{2m_e} \Delta_i | \Psi \rangle = 2\langle T_e \rangle. \quad (2.100)$$

Однако, так как в приближении Борна—Оппенгеймера ядра предполагаются *фиксированными*, общий вывод неприменим ко второму слагаемому. Мы обсудим это слагаемое подробно.

Потенциальная энергия молекулярной системы обязана трем составляющим: электрон-электронному отталкиванию, притяжению электронов к ядрам и межъядерному отталкиванию:

$$V = V_{ee} + V_{eN} + V_{NN} . \quad (2.101)$$

Тогда среднее значение  $\langle \Psi | \sum_i \vec{r}_i \frac{\partial V}{\partial \vec{r}_i} | \Psi \rangle$  также сумма трех слагаемых; рассмотрим их по отдельности.

Потенциальная энергия  $V_{ee}$  межэлектронного отталкивания является однородной функцией электронных координат, поэтому предыдущий вывод применим к ней:

$$-\langle \Psi | \sum_i \vec{r}_i \frac{\partial V_{ee}}{\partial \vec{r}_i} | \Psi \rangle = -k \langle V_{ee} \rangle = \langle V_{ee} \rangle \quad (2.102)$$

( $k = -1$  в случае кулоновского потенциала). Энергия межъядерного отталкивания  $V_{NN}$  никак не зависит от электронных координат, так что производные  $\frac{\partial V_{NN}}{\partial \vec{r}_i}$  исчезают и не дают вклада в среднее значение.

Притяжение электронов к ядрам  $V_{eN}$  зависит от расстояний  $r_{i\alpha} = |\vec{r}_i - \vec{R}_\alpha|$  между отдельными электронами и ядрами, и не является однородной функцией электронных координат в случае фиксированных ядер. (В общем случае координаты всех частиц рассматривались на равных, поэтому мы могли считать потенциальную энергию однородной функцией *всех* координат одновременно. Это невозможно в рамках приближения Борна—Оппенгеймера, в котором координаты ядер фиксированы.)

Чтобы пояснить этот момент, рассмотрим энергию взаимодействия между  $i$ -м электроном и ядром  $\alpha$ :

$$V_{i\alpha} = -\frac{Z_\alpha}{r_{i\alpha}} = \frac{-Z_\alpha}{\sqrt{(x_i - X_\alpha)^2 + (y_i - Y_\alpha)^2 + (z_i - Z_\alpha)^2}} . \quad (2.103)$$

Дифференцируя (2.103), получим

$$\frac{\partial V_{i\alpha}}{\partial x_i} = -Z_\alpha \frac{-(x_i - X_\alpha)}{[(x_i - X_\alpha)^2 + (y_i - Y_\alpha)^2 + (z_i - Z_\alpha)^2]^{3/2}} \quad (2.104)$$

и аналогичные выражения можно получить для  $\frac{\partial V_{i\alpha}}{\partial y_i}$  и  $\frac{\partial V_{i\alpha}}{\partial z_i}$ .

Можно установить непосредственной проверкой, что сумму  $x_i \frac{\partial V_{i\alpha}}{\partial x_i} + y_i \frac{\partial V_{i\alpha}}{\partial y_i} + z_i \frac{\partial V_{i\alpha}}{\partial z_i}$  нельзя привести к виду  $kV_{i\alpha}$ . Такое преобразование,

однако, можно провести, если рассмотреть выражение

$$(x_i - X_\alpha) \frac{\partial V_{i\alpha}}{\partial x_i} + (y_i - Y_\alpha) \frac{\partial V_{i\alpha}}{\partial y_i} + (z_i - Z_\alpha) \frac{\partial V_{i\alpha}}{\partial z_i} . \quad (2.105)$$

В самом деле, подставляя (2.104) и его аналог для  $\frac{\partial V_{i\alpha}}{\partial y_i}$  и  $\frac{\partial V_{i\alpha}}{\partial z_i}$  в (2.105) и разделив числитель и знаменатель на

$$r_{i\alpha}^2 = (x_i - X_\alpha)^2 + (y_i - Y_\alpha)^2 + (z_i - Z_\alpha)^2 , \quad (2.106)$$

мы получим  $-V_{i\alpha} = k V_{i\alpha}$ . Этот результат подсказывает, как двигаться дальше в наших выкладках.

Рассмотрим выражение  $-\langle \Psi | \sum_i \vec{r}_i \frac{\partial V_{eN}}{\partial \vec{r}_i} | \Psi \rangle$ , соответствующее притяжению электронов к ядрам в левой части (2.99), подставим в него явное выражение для  $V_{eN}$ ,

$$V_{eN} = \sum_{j,\alpha} V_{j\alpha} = \sum_{j,\alpha} \frac{-Z_\alpha}{r_{j\alpha}} , \quad (2.107)$$

и примем во внимание то, что  $V_{j\alpha}$  зависит от  $\vec{r}_i$  только если  $i = j$ :  $\frac{\partial V_{j\alpha}}{\partial \vec{r}_i} = \delta_{ij} \frac{\partial V_{i\alpha}}{\partial \vec{r}_i}$ , и поэтому  $\frac{\partial V_{eN}}{\partial \vec{r}_i} = \sum_\alpha \frac{\partial V_{i\alpha}}{\partial \vec{r}_i}$ . Получаем

$$-\langle \Psi | \sum_i \vec{r}_i \frac{\partial V_{eN}}{\partial \vec{r}_i} | \Psi \rangle = - \sum_{i,\alpha} \langle \Psi | \vec{r}_i \frac{\partial V_{i\alpha}}{\partial \vec{r}_i} | \Psi \rangle . \quad (2.108)$$

Добавим и вычтем  $\vec{R}_\alpha$  в каждом слагаемом:

$$-\langle \Psi | \sum_i \vec{r}_i \frac{\partial V_{eN}}{\partial \vec{r}_i} | \Psi \rangle = - \sum_{i,\alpha} \langle \Psi | (\vec{r}_i - \vec{R}_\alpha) \frac{\partial V_{i\alpha}}{\partial \vec{r}_i} | \Psi \rangle - \sum_{i,\alpha} \langle \Psi | \vec{R}_\alpha \frac{\partial V_{i\alpha}}{\partial \vec{r}_i} | \Psi \rangle . \quad (2.109)$$

$V_{i\alpha}$  есть функция  $\vec{r}_{i\alpha} = \vec{r}_i - \vec{R}_\alpha = \vec{i}x_{i\alpha} + \vec{j}y_{i\alpha} + \vec{k}z_{i\alpha}$ . Перейдем к переменным  $x_{i\alpha}, y_{i\alpha}$ , и  $z_{i\alpha}$  в первом слагаемом в правой части. Используя правило дифференцирования сложных функций, получаем

$$\frac{\partial V_{i\alpha}}{\partial x_i} = \frac{\partial V_{i\alpha}}{\partial x_{i\alpha}} \frac{\partial x_{i\alpha}}{\partial x_i} = \frac{\partial V_{i\alpha}}{\partial x_{i\alpha}} , \quad (2.110)$$

так как другие переменные не зависят от  $x_i$  и  $\frac{\partial x_{i\alpha}}{\partial x_i} = 1$ . Тогда

$$-\langle \Psi | \sum_i \vec{r}_i \frac{\partial V_{eN}}{\partial \vec{r}_i} | \Psi \rangle = - \sum_{i,\alpha} \langle \Psi | \vec{r}_{i\alpha} \frac{\partial V_{i\alpha}}{\partial \vec{r}_{i\alpha}} | \Psi \rangle - \sum_{i,\alpha} \vec{R}_\alpha \langle \Psi | \nabla_i V_{i\alpha} | \Psi \rangle . \quad (2.111)$$

В последнем слагаемом мы использовали определение  $\frac{\partial V_{i\alpha}}{\partial \vec{r}_i} \equiv \nabla_i V_{i\alpha}$  и вынесли из-под интеграла вектор  $\vec{R}_\alpha$ , не зависящий от переменных интегрирования (электронных координат). Теперь  $V_{i\alpha}$  является однородной функцией  $\vec{r}_{i\alpha}$  ( $k = -1$ ), так что, согласно теореме Эйлера, первое слагаемое в правой части можно записать как  $-k\langle V_{eN} \rangle = \langle V_{eN} \rangle$ . При исследовании второго слагаемого используем тот факт, что  $V_{i\alpha}$  является функцией переменной  $|\vec{r}_{i\alpha}| = |\vec{r}_i - \vec{R}_\alpha|$ , и в этом случае перестановка двух концов вектора меняет знак производной:  $\nabla_i V_{i\alpha} \equiv -\nabla_\alpha V_{i\alpha}$ . Итак, получаем

$$-\langle \Psi | \sum_i \vec{r}_i \frac{\partial V_{eN}}{\partial \vec{r}_i} | \Psi \rangle = -k\langle V_{eN} \rangle + \sum_{i,\alpha} \vec{R}_\alpha \langle \Psi | \nabla_\alpha V_{i\alpha} | \Psi \rangle. \quad (2.112)$$

Подставляя (2.100), (2.102), и (2.112) в (2.99), имеем

$$0 = 2\langle T_e \rangle - k\langle V_{ee} \rangle - k\langle V_{eN} \rangle + \sum_{i,\alpha} \vec{R}_\alpha \langle \Psi | \nabla_\alpha V_{i\alpha} | \Psi \rangle. \quad (2.113)$$

(Коэффициент  $-k = 1$  оставлен только для того, чтобы было ясно происхождение соответствующего слагаемого.)

Вследствие очевидного равенства  $\nabla_\alpha V_{i\beta} = \delta_{\alpha\beta} \nabla_\alpha V_{i\alpha}$ , можно представить  $\nabla_\alpha V_{i\alpha}$  как  $\nabla_\alpha V_{i\alpha} = \nabla_\alpha \sum_\beta V_{i\beta}$ ; таким образом,  $\sum_i \nabla_\alpha V_{i\alpha} = \nabla_\alpha V_{eN}$ , и поэтому последнее слагаемое в правой части (2.113) можно переписать как

$$0 = 2\langle T_e \rangle - k\langle V_{ee} \rangle - k\langle V_{eN} \rangle + \sum_\alpha \vec{R}_\alpha \langle \Psi | \nabla_\alpha V_{eN} | \Psi \rangle. \quad (2.114)$$

Чтобы получить окончательное выражение, воспользуемся тем, что потенциальная энергия межъядерного отталкивания является однородной функцией ядерных координат ( $k = -1$ ), так что, в соответствии с теоремой Эйлера,

$$\sum_\alpha \vec{R}_\alpha \frac{\partial V_{NN}}{\partial \vec{R}_\alpha} = kV_{NN}. \quad (2.115)$$

Собирая все члены (2.115) в одной части, используя тождество  $\frac{\partial V_{NN}}{\partial \vec{R}_\alpha} \equiv \nabla_\alpha V_{NN}$ , и учитывая то, что межъядерное отталкивание не зависит от электронной волновой функции (а значит, его среднее по электронной

волновой функции равно просто его значению, рассчитанному в приближении Борна—Оппенгеймера), получим

$$0 = -k\langle V_{NN} \rangle + \langle \Psi | \sum_{\alpha} \vec{R}_{\alpha} \nabla_{\alpha} V_{NN} | \Psi \rangle . \quad (2.116)$$

Добавляя (2.116) к (2.114) и учитывая, что  $V = V_{ee} + V_{eN} + V_{NN}$ :

$$0 = 2\langle T_e \rangle - k\langle V \rangle + \sum_{\alpha} \vec{R}_{\alpha} \langle \Psi | \nabla_{\alpha} (V_{eN} + V_{NN}) | \Psi \rangle . \quad (2.117)$$

По теореме Гельмана—Фейнмана (разд. 2), если гамильтониан зависит от параметра  $p$ , то для производной энергии по  $p$  имеем

$$\frac{\partial E}{\partial p} = \langle \Psi | \frac{\partial \hat{H}}{\partial p} | \Psi \rangle . \quad (2.118)$$

В электронном гамильтониане, записанном в приближении Борна—Оппенгеймера, имеются только два слагаемых, которые зависят от ядерных координат:  $V_{eN}$  и  $V_{NN}$ . Поэтому  $\left( \nabla_{\alpha} \equiv \frac{\partial}{\partial \vec{R}_{\alpha}} \right)$

$$\frac{\partial E}{\partial \vec{R}_{\alpha}} = \langle \Psi | \nabla_{\alpha} \hat{H} | \Psi \rangle = \langle \Psi | \nabla_{\alpha} (V_{eN} + V_{NN}) | \Psi \rangle , \quad (2.119)$$

где  $E$  — полная энергия, полученная при решении электронного уравнения Шрёдингера. Подставляя производную (2.119) в равенство (2.117), мы получим выражение теоремы вириала для случая приближения Борна—Оппенгеймера:

$$2\langle T_e \rangle + \langle V \rangle + \sum_{\alpha} \vec{R}_{\alpha} \frac{\partial E}{\partial \vec{R}_{\alpha}} = 0 . \quad (2.120)$$

Здесь  $\frac{\partial E}{\partial \vec{R}_{\alpha}} = \nabla_{\alpha} E = -\vec{F}_{\alpha}$ , а  $\vec{F}_{\alpha}$  — сила, действующая на ядро  $\alpha$ .

Поэтому «идеальное» вириальное отношение  $\frac{\langle V \rangle}{\langle T \rangle} = -2$  в рамках приближения Борна—Оппенгеймера получается только для стационарных точек (минимумов, максимумов и седловых точек) ППЭ, где все силы  $\vec{F}_{\alpha}$  равны нулю. (Кроме того, система, в которой ядра бесконечно удалены, тоже будет удовлетворять этому соотношению.)

В случае равновесной конфигурации теорема вириала верна как для точной, так и для некоторых приближенных волновых функций, например, для волновой функции Хартри—Фока. Однако в общем случае



при использовании конечных базисных наборов (метод Хартри—Фока—Рутана; см. гл. 6, разд. 5.1) она не выполняется. Проверка близости «вириального отношения»  $-\frac{\langle V \rangle}{\langle T \rangle}$  к точному значению 2 дает полезную меру качества волновой функции (т. е. применяемого базиса). Однако значение  $-\frac{\langle V \rangle}{\langle T \rangle} \cong 2$  является только необходимым, но не достаточным, условием «пригодности» волновой функции, так как точное значение 2 можно получить и для волновых функций весьма низкого качества, если их подвергнуть процедуре «масштабирования» (см. разд. 3.3).

### Дополнение к русскому изданию

Равенство (2.120) выполняется не только для точных решений уравнения Шредингера, но и для приближенных волновых функций, полученных вариационным путем.

Снова рассмотрим электронную волновую функцию  $\Psi(\vec{r}; \vec{R})$  приближения Борна-Оппенгеймера, в которой радиус-векторы  $\vec{r}_i$  всех электронов заменяются на «растянутые»  $\vec{\rho}_i = \eta \vec{r}_i$ . Поскольку ни кинетическая энергия<sup>1</sup>  $T$ , ни энергия межэлектронного взаимодействия  $V_{ee}$  явно от ядерных координат не зависят, для них остаются в силе результаты разд. 3.3, т. е.

$$T_e(\eta) = \eta^2 T_e(1) \quad , \quad (2.121)$$

и

$$V_{ee}(\eta) = \eta V_{ee}(1) \quad . \quad (2.122)$$

Для того чтобы определить, как ведет себя энергия электронно-ядерного притяжения  $V_{eN}$  при масштабном преобразовании, нам нужно исследовать интегралы

$$\langle \Psi_\eta | \frac{1}{|\vec{r}_i - \vec{R}_\alpha|} | \Psi_\eta \rangle \quad , \quad (2.123)$$

где  $\Psi_\eta = \Psi_\eta(\vec{\rho}; \vec{R})$  — нормированная электронная волновая функция типа (2.82), в которой электронные координаты (и только электронные координаты) подвергнуты растяжению заменой всех  $\vec{r}_i$  на  $\vec{\rho}_i = \eta \vec{r}_i$ . Умножая числитель и знаменатель на  $\eta$ , этот интеграл можно переписать как

$$\langle \Psi_\eta | \frac{\eta}{|\eta \vec{r}_i - \eta \vec{R}_\alpha|} | \Psi_\eta \rangle = \eta \langle \Psi_\eta | \frac{1}{|\vec{\rho}_i - \eta \vec{R}_\alpha|} | \Psi_\eta \rangle \quad . \quad (2.124)$$

<sup>1</sup> См. сноску на с. 51.

Значение интеграла снова не зависит от обозначения переменных интегрирования, так что получаем следующую зависимость для энергии электронно-ядерного притяжения:

$$V_{eN}(\eta; \vec{R}) = \eta V_{eN}(1; \eta \vec{R}) \quad , \quad (2.125)$$

где кроме зависимости от  $\eta$  явно указана и зависимость от параметров  $\vec{R}$ .

Таким образом, полная электронная энергия в приближении Борна-Оппенгеймера, как функция параметра  $\eta$ , равна

$$E(\eta) = \eta^2 T_e(1) + \eta V_{ee}(1) + \eta V_{eN}(1; \eta \vec{R}) + V_{NN} \quad . \quad (2.126)$$

Для оптимизации по  $\eta$  нужно потребовать

$$\frac{dE}{d\eta} = 2\eta T_e(1) + V_{ee}(1) + V_{eN}(1; \eta \vec{R}) + \eta \sum_{\alpha} \vec{R}_{\alpha} \frac{\partial V_{eN}(1; \eta \vec{R})}{\partial (\eta \vec{R}_{\alpha})} = 0 \quad . \quad (2.127)$$

Если волновая функция была оптимизирована вариационным путем, то (ср. сноску на с. 51) это равенство должно выполняться при  $\eta = 1$ , и мы получим, опуская аргументы

$$2T_e + V_{ee} + V_{eN} + \sum_{\alpha} \vec{R}_{\alpha} \frac{\partial V_{eN}}{\partial \vec{R}_{\alpha}} = 0 \quad . \quad (2.128)$$

Добавим к этому равенству уравнение (2.115) и перенесем все члены налево (напомним, что  $k = -1$ ):

$$2T_e + V_{ee} + V_{eN} + V_{NN} + \sum_{\alpha} \vec{R}_{\alpha} \frac{\partial (V_{eN} + V_{NN})}{\partial \vec{R}_{\alpha}} = 0 \quad . \quad (2.129)$$

Теперь замечаем, что  $V_{ee} + V_{eN} + V_{NN} = V$ , используем равенство (2.119), вытекающее из теоремы Гельманна-Фейнмана (она тоже выполняется для вариационных волновых функций) и можем написать

$$2T_e + V + \sum_{\alpha} \vec{R}_{\alpha} \frac{\partial E}{\partial \vec{R}_{\alpha}} = 0. \quad \text{Ч. т. д.} \quad (2.130)$$

### 3.5. Теорема вириала и химическая связь

Важным следствием теоремы вириала является то, что *образование химической связи является следствием понижения потенциальной энергии* (чисто электростатического эффекта, связанного с тем, что в области связывания появляется заселенность перекрывания, а также с тем,

что в среднем электроны несколько приближаются к ядрам), тогда как кинетическая энергия при образовании связи увеличивается (но это увеличение — в силу теоремы вириала — составляет лишь половину от понижения потенциальной энергии). В этом отношении в химической литературе встречаются повторяющиеся ошибки, восходящие к Гельману, который считал, что причиной связывания является уменьшение кинетической энергии, которое становится возможным, так как электроны делокализованы в большей области пространства, по сравнению со свободными атомами, и это уменьшает их импульсы и, следовательно, кинетическую энергию, в соответствии с принципом неопределенности. Но ситуация намного более сложна: теорема вириала указывает на то, что кинетическая энергия в процессе образования связи в итоге *увеличивается*, а не уменьшается.

Проблема была детально исследована Рюденбергом на простых примерах, таких, как  $H_2^+$ . Он рассматривал образование связи как процесс, формально состоящий из последовательных этапов: первым этапом является «*промотирование*» (promotion), представляющее собой сжатие валентных орбиталей. Оно приводит к понижению потенциальной энергии (электроны в среднем подходят ближе к ядрам), но кинетическая энергия увеличивается в *большей* степени, чем уменьшается потенциальная энергия. (Такое увеличение кинетической энергии является следствием соотношения неопределенностей.) Возможность делокализации позволяет кинетической энергии уменьшиться *по отношению к «промотированному» состоянию со сжатыми орбиталями*; это уменьшение кинетической энергии необходимо для того, чтобы можно было сохранить тот выигрыш в потенциальной энергии, который получился в результате сжатия. Однако, в конечном итоге образование связи сопровождается уменьшением потенциальной энергии и увеличением кинетической энергии, в соответствии с теоремой вириала. (Пример: в расчетах молекулы  $H_2$  по Гайтлеру—Лондону с использованием минимального базиса 1s-орбиталей с экспоненциальной радиальной частью оптимальный показатель экспоненты равен  $\cong 1.19$ , что заметно больше единицы — значения показателя, характерного для свободного атома H, что указывает на заметное сжатие орбиталей при образовании связи.)

## Библиографические заметки

### Раздел 1.2.

Стандартное изложение, например см. [1-3]. В некоторых случаях (например, атом водорода) собственные функции гамильтониана образуют полный набор, только если туда включается также *континуум* положительных собственных значений. В таком случае суммирование в уравнении (2.5) должно быть до-

полнено интегрированием по континууму (см., например, [4, 5]); это, однако, не влияет на существо вывода.

### Раздел 1.3.

Автор настоящей книги обычно рассматривает уравнение (2.13) как основное для описания любой вариационной задачи. Вывод, данный здесь, был впервые опубликован в [6]; вывод с использованием множителей Лагранжа является стандартным, см. например [7, 8].

### Раздел 1.4.

Первоначально выведено в [9]. Имеет значительно большее концептуальное, нежели практическое значение.

### Раздел 1.5.

Этот результат недавно опубликовал Шурьян [10], используя несколько другой вывод. За обсуждаемую здесь схему приведения уравнения к эрмитовому виду автор признателен В.А. Куприевичу [11]. Другие авторы (например, [12]) добиваются эрмитовости уравнения (2.31), полагая

$$[\hat{H} - (\hat{P}\hat{H} + \hat{H}\hat{P})]\Psi = E\Psi$$

Этот оператор, однако, не является спроектированным гамильтонианом.

### Раздел 2.1.

Теорема впервые была выведена Гельманом [13] и несколько позже Фейнманом [14]. Приведенное доказательство [15] является простым обобщением вывода вариационного принципа, данного в разделе 1.3; по существу, такой же подход применен и в [5], где также подробно обсуждается и то, что теорема Гельмана—Фейнмана в виде (2.51)–(2.53), т. е. электростатическая теорема, применима к конечным базисным наборам, только если используются «плавающие» функции. Существует другое, очень простое доказательство (например, [2]), которое явно использует предположение о том, что волновая функция удовлетворяет уравнению Шрёдингера и что ее нормировка не зависит от параметра  $\alpha$ . В самом деле, если  $\langle\Psi|\Psi\rangle = 1 = \text{const}$  и  $\hat{H}\Psi = E\Psi$ , то  $\langle\frac{\partial\Psi}{\partial\alpha}|\hat{H}|\Psi\rangle = E\langle\frac{\partial\Psi}{\partial\alpha}|\Psi\rangle$  и

$$\frac{\partial}{\partial\alpha}\langle\Psi|\hat{H}|\Psi\rangle = E\frac{\partial}{\partial\alpha}\langle\Psi|\Psi\rangle + \langle\Psi|\frac{\partial\hat{H}}{\partial\alpha}|\Psi\rangle = \langle\Psi|\frac{\partial\hat{H}}{\partial\alpha}|\Psi\rangle.$$

Можно заметить, что если воспользоваться формализмом вторичного квантования, надлежащим образом обобщенным [16] на случай перекрывающихся базисных наборов, то гамильтониан определяется *интегралами* по базисным орбиталям и теорема Гельмана—Фейнмана формально снова становится применима даже к базисным орбиталям, центрированным на атомах [17].

### Раздел 2.2.

Было выведено в [18] и применено в [19, 20] для объяснения происхождения барьеров внутреннего вращения в молекулах, подобных этану. См. также [2].

### Раздел 3.1.

Стандартные результаты, приведенные в учебниках (см., например, [7]).

### Раздел 3.2.

Описание аналогично приведенному в [2]. Простой альтернативный вывод можно найти в приложении 3 в [21].

### Раздел 3.3.

Вывод во многом следует [22].

### Раздел 3.4.

Авторский вывод (не опубликовано). Обычные выводы или сочетают квантово-механический вывод с классическим вириалом внешних сил [21], или представляются несколько сомнительными [22].

### Раздел 3.5.

Как замечено Лёвдином (см. сноску в [22]), Г. Холл первым обратил внимание на противоречие между простой картиной Гельмана, основанной на принципе неопределенности, и теоремой вириала. Проблема затем была разъяснена Рюденбергом [23].

## Для дальнейшего изучения

Классическая книга Слэтера [21] дает исключительно полезные примеры того, как приложение теории к простым задачам может выделить наиболее важные аспекты химического связывания. Монография Эпштейна [8] содержит разработанный математический аппарат для решения большинства проблем, обсуждаемых в этой главе.

## Литература

1. Bethe H. A. *Intermediate Quantum Mechanics*, Benjamin, New York 1964. (Имеется русский перевод: Бете Г. *Квантовая механика*. — М.: Мир, 1965).
2. Züllicke L. *Quantenchemie*, Deutcher Verlag der Wissenschaften, Berlin 1973. (Имеется русский перевод: Цюлик Л. *Квантовая химия*. — М.: Мир, 1976).
3. Eyring H., Walter J., Kimball G. E. *Quantum Chemistry*, John Wiley & Sons, New York 1944. (Имеется русский перевод: Эйринг Г., Уолтер Д., Кимбалл Г. *Квантовая химия*. — М.: Издательство, 1948).
4. Kapuy E., Török F. *Az Atomok és Molekulák Kvantumelmélete* Akadémiai Kiadó, Budapest 1975.
5. Pilar F. L. *Elementary Quantum Chemistry*, McGraw Hill, New York 1968.
6. Mayer I., Ladik J., Biczó G. *Int. J. Quantum Chem.* **7**, 583 (1973).
7. Ландау Л. Д., Лифшиц И. М. *Квантовая механика*. — М.: Наука, 1974.
8. Epstein S. T. *The Variation Method in Quantum Chemistry*, Academic Press, New York 1974. (Имеется русский перевод: Эпштейн С. *Вариационный метод в квантовой химии*. — М.: Мир, 1976).
9. Eckart C. E. *Phys. Rev.* **36**, 878 (1930).
10. Surján P. R. *Chem. Phys. Lett.* **325**, 120 (2000).
11. Куприевич В. А., частное сообщение, Киев, 1971.
12. Фудзинага С. *Метод молекулярных орбиталей*. — М.: Мир, 1983.
13. Гельман Г. *Квантовая химия*. — М.: Гостехиздат, 1937. Hellmann H. *Einführung in die Quantenchemie*, Deuticke, Leipzig 1937.
14. Feynman R. P. *Phys. Rev.* **56**, 340 (1939).
15. Mayer I. *Fejezetek a Kvantumkémiából*, BME Mérnök-továbbképző Int., Budapest 1987.
16. Mayer I. *Int. J. Quantum Chem.* **23**, 341 (1983).

17. Surján P. R., Mayer I., Poirier R. J. *Mol. Structure (Theochem)* **170**, 1 (1988).
18. Kim H. J., Parr R. G. *J. Chem. Phys.* **41**, 2892 (1964).
19. Wyatt R. E., Parr R. G. *J. Chem. Phys.* **41**, 3262 (1964).
20. Parr R. G. in *Modern Quantum Chemistry (Istanbul Lectures)*, ed. O. Sinanoğlu, Academic Press, New York 1965. (Имеется русский перевод: Парр Р. в *Современная квантовая химия* (ред. О. Синаноглу). — М.: Мир, 1968).
21. Slater J. C. *Quantum Theory of Molecules and Solids vol. 1. Electronic Structure of Molecules*, McGraw-Hill, New York 1963. (Имеется русский перевод: Слэтер Дж. *Электронная структура молекул*. — М.: Мир, 1965).
22. Löwdin P.-O. *J. Mol. Spectroscopy* **3**, 46 (1959).
23. Ruedenberg K. *Rev. Mod. Phys.* **34**, 326 (1962).



# Линейный вариационный метод и схемы ортогонализации по Лёвдину



## 1. Линейный вариационный метод (метод Ритца)

Выберем конечное число  $m$  определенных заранее функций  $\varphi_i$ , каждая из которых удовлетворяет граничным условиям задачи. Функции  $\varphi_i$  зависят от тех же независимых переменных, что и искомая волновая функция, но не содержат других переменных или подбираемых параметров. (Если  $m \rightarrow \infty$  и набор функций  $\varphi_i$  стремится к полному, то использование метода Ритца эквивалентно точному решению уравнения Шрёдингера.)

Функции  $\varphi_i$  называются «базисными функциями» задачи; мы предполагаем, что они линейно независимы, и таким образом растягивают  $m$ -мерное подпространство гильбертова пространства.

Аппроксимируем волновую функцию  $\Psi$  линейной комбинацией базисных функций  $\varphi_i$ :

$$\Psi = \sum_{i=1}^m c_i \varphi_i, \quad (3.1)$$

и определим коэффициенты  $c_i$ , доставляющие минимум функционалу энергии:

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \min. \quad (3.2)$$

Так как функции  $\varphi_i$  фиксированы, варьирование волновой функции вида (3.1) может быть проделано только за счет варьирования коэффициентов  $c_i$ :

$$c_i \rightarrow c_i + \delta c_i \quad (3.3)$$

( $\delta c_i = \Delta c_i = dc_i$  — малое изменение  $c_i$ ; в случае независимой скалярной переменной нет нужды различать ее дифференциал и вариацию.)

Если все  $c_i$  изменяются в соответствии с (3.3),  $\Psi$  изменится как

$$\Psi \rightarrow \Psi + \delta \Psi = \sum_{i=1}^m (c_i + \delta c_i) \varphi_i = \underbrace{\sum_{i=1}^m c_i \varphi_i}_{\Psi} + \underbrace{\sum_{i=1}^m \delta c_i \varphi_i}_{\delta \Psi}. \quad (3.4)$$

Это значит, что *наиболее общей* вариацией волновой функции  $\Psi$ , используемой в линейном вариационном методе, является

$$\delta\Psi = \sum_{i=1}^m \delta c_i \varphi_i, \quad (3.5)$$

и именно это есть та  $\delta\Psi$ , которую мы должны подставить в выражение (2.13) для вариационного принципа  $\langle \delta\Psi | \hat{H} - E | \Psi \rangle = 0$ , установленное в гл. 2, разд. 1.3. Получаем

$$\left\langle \sum_{i=1}^m \delta c_i \varphi_i | \hat{H} - E | \Psi \right\rangle = 0. \quad (3.6)$$

Вариации  $\delta c_i$  произвольны и полностью независимы друг от друга; мы можем, например, рассмотреть случай, когда только одна  $\delta c_i \neq 0$ . Это означает, что равенство

$$\langle \varphi_i | \hat{H} - E | \Psi \rangle = 0 \quad (i = 1, 2, \dots, m) \quad (3.7)$$

должно выполняться для каждого  $i$ .

Подставим  $\Psi = \sum_{j=1}^m c_j \varphi_j$  в условие (3.7) и преобразуем левую часть:

$$\begin{aligned} \langle \varphi_i | \hat{H} - E | \sum_{j=1}^m c_j \varphi_j \rangle &= \sum_{j=1}^m \langle \varphi_i | \hat{H} - E | \varphi_j \rangle c_j \\ &= \sum_{j=1}^m \left( \langle \varphi_i | \hat{H} | \varphi_j \rangle - E \langle \varphi_i | \varphi_j \rangle \right) c_j = \sum_{j=1}^m (H_{ij} - ES_{ij}) c_j = 0, \end{aligned} \quad (3.8)$$

где введены обозначения

$$H_{ij} = \langle \varphi_i | \hat{H} | \varphi_j \rangle \quad \text{и} \quad S_{ij} = \langle \varphi_i | \varphi_j \rangle. \quad (3.9)$$

Равенство

$$\sum_{j=1}^m (H_{ij} - ES_{ij}) c_j = 0 \quad (3.10)$$

должно выполняться для каждого  $i$ . Таким образом, мы получили систему  $m$  линейных уравнений. Число неизвестных равно  $m + 1$ :  $m$  коэффициентов  $c_i$  и энергия  $E$ . Однако требование нормировки волновой функции  $\Psi = \sum_i c_i \varphi_i$  дает одно дополнительное условие; другими словами, для решения задачи достаточно определить  $m - 1$  отношение  $c_i/c_1$  ( $i = 2, 3, \dots, m$ ), вместо  $m$  коэффициентов  $c_i$  ( $i = 1, 2, \dots, m$ ).



Как известно из линейной алгебры, система однородных уравнений (3.10) имеет нетривиальное решение (когда не все  $c_i = 0$ ), тогда и только тогда, когда ее определитель обращается в нуль:

$$\det\{H_{ij} - ES_{ij}\} = 0. \quad (3.11)$$

Уравнения такого типа (в частности, когда  $S_{ij} = \delta_{ij}$ ) часто называют «секулярными», или «вековыми», уравнениями — название, заимствованное из астрономии.

Уравнение (3.11) является уравнением  $m$ -той степени относительно  $E$ . Подставляя его решение в исходную систему уравнений, можно, в принципе, определить  $m - 1$  отношение  $c_i/c_1$ . На практике, однако, так не поступают, так как существуют намного более эффективные методы, обходящиеся без явного решения алгебраического уравнения степени  $m$ .

Введем матрицы  $\mathbf{H}$  и  $\mathbf{S}$  размера  $m \times m$  с элементами  $H_{ij}$  и  $S_{ij}$ , соответственно, и соберем коэффициенты  $c_i$  в вектор-столбец  $\mathbf{c}$ :

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_i \\ \vdots \\ c_m \end{pmatrix}. \quad (3.12)$$

Используя матричные обозначения, запишем задачу (3.10) как

$$(\mathbf{H} - E\mathbf{S})\mathbf{c} = 0, \quad (3.13)$$

т. е.

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c}, \quad (3.14)$$

что представляет собой «обобщенное уравнение на собственные значения матрицы». Если базис  $\{\varphi_i\}$  ортонормирован,  $S_{ij} = \langle \varphi_i | \varphi_j \rangle = \delta_{ij}$ , т. е.  $\mathbf{S} = \mathbf{1}$  (единичная матрица), и тогда мы получаем «стандартное» уравнение на собственные значения матрицы

$$\mathbf{H}\mathbf{c} = E\mathbf{c}. \quad (3.15)$$

Из определений (3.9) следует, что матрицы  $\mathbf{H}$  и  $\mathbf{S}$  эрмитовы:  $H_{ij} = H_{ji}^*$ ;  $S_{ij} = S_{ji}^*$ ; в большинстве практических случаев они также и вещественны (вещественная эрмитова матрица = симметрическая матрица).

Из линейной алгебры известно, что каждая эрмитова матрица размера  $m \times m$  имеет  $m$  собственных векторов. Собственные значения эрмитовой матрицы вещественны. Собственные векторы, соответствующие разным собственным числам, — ортогональны; те же собственные векторы, которые соответствуют одному и тому же (вырожденному) собственному значению, можно также выбрать ортогональными. Это означает, что собственные векторы можно использовать для построения полного ортонормированного базиса в рассматриваемом (под)пространстве. Образуя столбцы матрицы  $\mathbf{U}$  из собственных векторов эрмитовой матрицы, мы убеждаемся, что таким образом получается унитарная матрица:  $(\mathbf{U}^\dagger \mathbf{U})_{ij} = \sum_k (\mathbf{U}^\dagger)_{ik} U_{kj} = \sum_k U_{ki}^* U_{kj} = \delta_{ij}$ , как скалярное произведение  $i$ -го и  $j$ -го собственных векторов.

В соответствии с определением, преобразование подобия матрицы  $\mathbf{A}$  с помощью матрицы  $\mathbf{T}$  дает матрицу  $\mathbf{T}^{-1} \mathbf{A} \mathbf{T}$ . В частном случае, когда  $\mathbf{T}$  является унитарной,  $(\mathbf{T}^{-1} = \mathbf{T}^\dagger)$ , эта матрица сводится к  $\mathbf{T}^\dagger \mathbf{A} \mathbf{T}$ . Поиск всех собственных значений и векторов эрмитовой матрицы является задачей, эквивалентной ее *диагонализации* с помощью преобразования подобия: проводя преобразование подобия матрицы  $\mathbf{H}$  с помощью унитарной матрицы  $\mathbf{U}$ , построенной из собственных векторов  $\mathbf{H}$ , можно получить диагональную матрицу собственных значений. Обратное также верно: столбцы матрицы  $\mathbf{U}$ , диагонализующей эрмитову матрицу  $\mathbf{H}$ , являются собственными векторами последней. В самом деле, пусть

$$\mathbf{U}^\dagger \mathbf{H} \mathbf{U} = \mathbf{D} \quad \text{и} \quad \mathbf{D} = \text{diag}(E_i), \quad (3.16)$$

т. е.  $D_{ij} = E_i \delta_{ij}$ . Умножая (3.16) на  $\mathbf{U}$  слева и используя ее унитарность  $(\mathbf{U} \mathbf{U}^\dagger = \mathbf{1})$ , получим

$$\mathbf{H} \mathbf{U} = \mathbf{U} \mathbf{D}, \quad (3.17)$$

что является матричной формой уравнения на собственные значения (3.15) записанной одновременно для всех собственных векторов. В самом деле, выписывая равенство (3.17) покомпонентно, имеем

$$\sum_k H_{ik} U_{kj} = \sum_k U_{ik} D_{kj} = \sum_k U_{ik} E_j \delta_{kj} = E_j U_{ij}, \quad (3.18)$$

т. е. уравнение на собственные значения для  $j$ -го столбца матрицы  $\mathbf{U}$  с собственным значением  $E_j$ . Можно рассмотреть  $j$ -й столбец матрицы  $\mathbf{U}$  как вектор-столбец  $\mathbf{c}^j$  с элементами  $c_i^j = U_{ij}$  и верхним индексом  $j$ , нумерующим столбец. Подставляя его в соотношение (3.18), имеем

$$\sum_k H_{ik} c_k^j = E_j c_i^j, \quad (3.19)$$

т. е. развернутую форму уравнения на собственные значения (3.15)  $\mathbf{c} = \mathbf{c}^j$  и  $E = E_j$ .

(Обратите внимание на порядок расположения матриц в правой части равенства (3.17): диагональная матрица собственных значений стоит *после* матрицы собственных векторов. Аналогично, матричная форма обобщенного уравнения на собственные значения (3.14) есть  $\mathbf{HC} = \mathbf{SCD}$ .)

Существуют стандартные численные алгоритмы и программы, предназначенные для решения матричной задачи на собственные значения. Их подробное обсуждение выходит за рамки этой книги; сделаем только несколько кратких замечаний. Во-первых, на практике обходятся без прямого вычисления векового определителя (3.11) и решения алгебраических уравнений высоких степеней для собственных значений  $E$ . Вместо этого для «компактных» матриц, полностью сохраняемых в оперативной компьютерной памяти, используют методы, основанные на преобразованиях подобия, выполняющих диагонализацию матрицы. В классическом методе Якоби (1848) собственные векторы находят последовательными двухмерными вращениями базисных векторов. На каждом таком шаге исключается один ненулевой недиагональный элемент. Однако, этот матричный элемент может стать снова ненулевым (обычно так и происходит), когда делается какое-то другое (последующее) вращение, поэтому эта процедура должна циклически повторяться для всех недиагональных матричных элементов, пока все они не станут достаточно малыми. Недостатком метода Якоби является то, что он весьма неэкономичен с точки зрения машинного времени. В настоящее время используются главным образом современные (середины двадцатого столетия) методы (алгоритмы Гивенса, Хаусхолдера, QR или QL), в которых проводится преобразование подобия, приводящее матрицу к «трехдиагональному» виду, в котором только элементы главной диагонали и их ближайшие соседи отличны от нуля. Задача на собственные значения для трехдиагональной матрицы затем решается специальным алгебраическим методом, развитым для этой частной задачи, а результаты преобразуются обратно в исходный базис.

Для решения задачи на собственные значения очень больших, но «разреженных» матриц (т. е. матриц, имеющих относительно мало ненулевых элементов), используются методы, которые не требуют явного вычисления и хранения всей матрицы, что часто было бы невозможно даже для самых мощных компьютеров. (Работы с полной матрицей  $\mathbf{A}$  можно избежать, если иметь в руках процедуру, позволяющую как-то вычислять результат действия  $\mathbf{Ax}$  матрицы  $\mathbf{A}$  на произвольный вектор  $\mathbf{x}$ , и, если необходимо, определить только один или несколько собственных

векторов, принадлежащих к собственным значениям с наибольшими абсолютными величинами. В этом случае требования к памяти могут снизиться до нескольких векторов длины  $m$ , где  $m$  — размерность задачи, тогда как хранение матрицы  $\mathbf{A}$  потребовало бы  $m^2$  ячеек памяти, или по крайней мере  $\frac{m(m+1)}{2}$  для симметрической матрицы  $\mathbf{A}$ . В квантовой химии по такой методике проводятся расчеты методом конфигурационного взаимодействия (КВ) высоких порядков: результат действия  $\mathbf{H}$  матрицы гамильтониана  $\mathbf{H}$  на вектор состояния  $\mathbf{x}$  вычисляется непосредственной обработкой списков одно- и двухэлектронных *интегралов*, входящих в матричные элементы (гл. 5, разд. 6), но без явного вычисления последних.

*Альтернативный вывод с использованием множителей Лагранжа*  
 Существует несколько вариантов вывода вековых уравнений (3.10) — или их матричного представления (3.14), — с использованием метода множителей Лагранжа. Поучительно рассмотреть подробно два из них.

1. Для наиболее общего вывода запишем интегралы, входящие в коэффициент Рэлея в (3.2), в виде

$$\begin{aligned}\langle \Psi | \hat{H} | \Psi \rangle &= \sum_{i,j=1}^m c_i^* H_{ij} c_j \\ \langle \Psi | \Psi \rangle &= \sum_{i,j=1}^m c_i^* S_{ij} c_j\end{aligned}\tag{3.20}$$

и рассмотрим их как функции скалярных переменных  $c_i$ . Тогда вариационная задача, которую предстоит решить, имеет вид

$$\sum_{i,j=1}^m c_i^* H_{ij} c_j = \min\tag{3.21}$$

при условии

$$\sum_{i,j=1}^m c_i^* S_{ij} c_j = 1\tag{3.22}$$

Определим вспомогательную функцию с добавлением множителя Лагранжа  $\lambda$  формулой

$$F' = \sum_{i,j=1}^m c_i^* H_{ij} c_j + \lambda \left( \sum_{i,j=1}^m c_i^* S_{ij} c_j - 1 \right)\tag{3.23}$$

и потребуем  $\frac{\partial F'}{\partial c_i} = 0$  и  $\frac{\partial F'}{\partial c_i^*} = 0$  для всех  $i = 1, 2, \dots, m$ . В этом выводе  $c_i$  и  $c_i^*$  рассматриваются как две независимые переменные; условия  $\frac{\partial F'}{\partial c_i} = 0$  и  $\frac{\partial F'}{\partial c_i^*} = 0$  приводят к комплексно сопряженным уравнениям и их не надо рассматривать по отдельности (см. нижеследующее замечание). Необходимо еще добавить условие (3.22) к полученным уравнениям, или, что то же самое, потребовать  $\frac{\partial F'}{\partial \lambda} = 0$ . Тогда, выписав  $\frac{\partial F'}{\partial c_i^*}$ , получим систему уравнений

$$\begin{aligned} \sum_{j=1}^m H_{ij} c_j + \lambda \sum_{j=1}^m S_{ij} c_j &= 0 \quad (i = 1, 2, \dots, m) \\ \sum_{i,j=1}^m c_i^* S_{ij} c_j &= 1. \end{aligned} \quad (3.24)$$

Умножая первое уравнение на  $c_i^*$ , суммируя по  $i$  и учитывая второе уравнение, мы получаем  $\lambda = - \sum_{i,j=1}^m c_i^* H_{ij} c_j = -E$ , откуда следует, что

$$\sum_{j=1}^m H_{ij} c_j - E \sum_{j=1}^m S_{ij} c_j = 0 \quad \text{Ч. т. д.} \quad (3.25)$$

### Замечание

Очевидно, величины  $c_i$  и  $c_i^*$  не являются строго независимыми параметрами — они комплексно сопряжены друг к другу. Однако за ними скрываются две истинно независимые вещественные переменные — их вещественная и мнимая части  $a_i = \operatorname{Re} c_i = \frac{1}{2}(c_i + c_i^*)$  и  $b_i = \operatorname{Im} c_i = \frac{1}{2i}(c_i - c_i^*)$ . Можно было бы исследовать линейную вариационную задачу с использованием вещественных параметров  $a_i$  и  $b_i$ , но это довольно обременительно. Вместо этого можно положить  $\frac{\partial E}{\partial a_i} = \frac{\partial E}{\partial b_i} = 0$  и, используя правило дифференцирования сложных функций, получить  $(c_i = a_i + ib_i, c_i^* = a_i - ib_i)$

$$\begin{aligned} \frac{\partial E}{\partial a_i} &= \frac{\partial E}{\partial c_i} \frac{\partial c_i}{\partial a_i} + \frac{\partial E}{\partial c_i^*} \frac{\partial c_i^*}{\partial a_i} = \frac{\partial E}{\partial c_i} + \frac{\partial E}{\partial c_i^*} = 0 \\ \frac{\partial E}{\partial b_i} &= \frac{\partial E}{\partial c_i} \frac{\partial c_i}{\partial b_i} + \frac{\partial E}{\partial c_i^*} \frac{\partial c_i^*}{\partial b_i} = i \frac{\partial E}{\partial c_i} - i \frac{\partial E}{\partial c_i^*} = 0, \end{aligned} \quad (3.26)$$

а эта система эквивалентна условиям  $\frac{\partial E}{\partial c_i} = \frac{\partial E}{\partial c_i^*} = 0$ , использованным выше.

2. Можно также непосредственно вывести матричную форму (3.14) векового уравнения, видоизменив вывод уравнения Шрёдингера из вариационного принципа, обсуждаемый в гл. 2, разд. 1.3 для случая, когда волновая функция  $\Psi$  имеет вид разложения (3.1).

Предположим, что интеграл  $\langle \Psi | \hat{H} | \Psi \rangle$  стационарен при дополнительном условии  $\langle \Psi | \Psi \rangle = 1$ . Выпишем вспомогательный функционал

$$F' = \langle \Psi | \hat{H} | \Psi \rangle + \lambda (\langle \Psi | \Psi \rangle - 1) , \quad (3.27)$$

потребуем, чтобы было выполнено условие  $\delta F' = 0$ , и объединим его с условием  $\langle \Psi | \Psi \rangle = 1$  (или, что то же самое, потребуем, чтобы  $\frac{\partial F'}{\partial \lambda} = 0$ ). Получим

$$\begin{aligned} \langle \delta \Psi | \hat{H} | \Psi \rangle + \lambda \langle \delta \Psi | \Psi \rangle + c.c. &= 0 \\ \langle \Psi | \Psi \rangle &= 1 . \end{aligned} \quad (3.28)$$

Так как  $\Psi$  имеет вид (3.1) с определенными заранее функциями  $\varphi_i$ , то  $\delta \Psi$  принимает вид

$$\delta \Psi = \sum_{i=1}^m \delta c_i \varphi_i \quad (3.29)$$

с произвольными коэффициентами  $\delta c_i \rightarrow 0$ . Они могут быть объединены в вектор  $\delta \mathbf{c}$ . Подставляя разложения (3.1) и (3.29) в уравнения (3.28) и используя определения (3.9) матричных элементов  $H_{ij}$  и  $S_{ij}$ , мы получаем в матричных обозначениях

$$\begin{aligned} \delta \mathbf{c}^\dagger \mathbf{H} \mathbf{c} + \lambda \delta \mathbf{c}^\dagger \mathbf{S} \mathbf{c} + c.c. &= 0 \\ \mathbf{c}^\dagger \mathbf{S} \mathbf{c} &= 1 . \end{aligned} \quad (3.30)$$

Так как  $\delta \mathbf{c}$  имеет произвольный фазовый множитель, слагаемое, выписанное явно, и его комплексно сопряженное (с.с.) должны по отдельности обращаться в нуль:

$$\begin{aligned} \delta \mathbf{c}^\dagger \mathbf{H} \mathbf{c} + \lambda \delta \mathbf{c}^\dagger \mathbf{S} \mathbf{c} &= 0 \\ \mathbf{c}^\dagger \mathbf{S} \mathbf{c} &= 1 . \end{aligned} \quad (3.31)$$

Здесь  $\delta \mathbf{c}^\dagger$  — произвольный вектор (например, можно рассмотреть последовательно случаи, когда только один  $\delta c_i \neq 0$ ;  $i = 1, 2, \dots, m$ ), поэтому первое уравнение системы (3.31) может быть верно только если

$$\mathbf{H} \mathbf{c} + \lambda \mathbf{S} \mathbf{c} = 0 . \quad (3.32)$$

Умножая это уравнение на  $\mathbf{c}^\dagger$  слева и учитывая второе уравнение системы (3.31), мы получаем  $\lambda = -\mathbf{c}^\dagger \mathbf{H} \mathbf{c} = -\langle \Psi | \hat{H} | \Psi \rangle = -E$ , т. е.

$$\mathbf{H} \mathbf{c} = E \mathbf{S} \mathbf{c} . \quad \text{Ч. т. д.} \quad (3.33)$$

## 2. Симметричная ортогонализация по Лёвдину

### 2.1. Матрица $\mathbf{S}^{-1/2}$

Как мы выяснили в предыдущем разделе, линейная вариационная задача приводит к *обобщенному* уравнению на собственные значения  $\mathbf{H}\mathbf{c} = \mathbf{E}\mathbf{S}\mathbf{c}$ , если базисные функции  $\varphi_i$  не ортонормированы. Схема симметричной ортогонализации по Лёвдину представляет собой один из многих методов, позволяющих свести обобщенное уравнение на собственные значения к «стандартному» уравнению на собственные значения  $\mathbf{H}'\mathbf{d} = \mathbf{E}\mathbf{d}$ . Однако, симметричная ортогонализация (или «ортогонализация по Лёвдину») является не только одним из алгоритмов решения обобщенной задачи на собственные значения, но также имеет важное концептуальное значение и часто используется в разных областях квантовохимической теории (заметим, что Лёвдин предложил и другие методы ортогонализации, наиболее важным из них с точки зрения вычислений является «каноническая ортогонализация»; см. ниже разд. 3.2).

Матрица интегралов перекрывания (или «метрика»)  $\mathbf{S}$ , имеющая элементы  $S_{ij} = \langle \varphi_i | \varphi_j \rangle$ , является, очевидно, эрмитовой и, если базисный набор  $\{\varphi_i\}$  линейно независим, она также положительно определена (см. ниже разд. 3). Поэтому матрицу  $\mathbf{S}$  можно диагонализировать преобразованием подобия с унитарной матрицей  $\mathbf{V}$ , и при этом все ее собственные значения  $\lambda_i > 0$ :

$$\mathbf{V}^\dagger \mathbf{S} \mathbf{V} = \mathbf{\Lambda}; \quad \mathbf{\Lambda} = \text{diag}(\lambda_i) = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ 0 & 0 & \lambda_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (3.34)$$

Введем диагональную матрицу  $\mathbf{K} = \text{diag}\left(\frac{1}{\sqrt{\lambda_i}}\right)$ . Очевидно, имеем  $\mathbf{\Lambda}\mathbf{K}^2 = \mathbf{1}$ . Подвергнем матрицы  $\mathbf{\Lambda}$  и  $\mathbf{K}$  преобразованию подобия с матрицей, *обратной* к матрице  $\mathbf{V}$ , диагонализующей матрицу  $\mathbf{S}$ . (Так как  $\mathbf{V}$  унитарна,  $\mathbf{V}^{-1} = \mathbf{V}^\dagger$ ,  $\mathbf{V}^{\dagger-1} = \mathbf{V}$ .) Мы получаем матрицы

$$\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger = \mathbf{S} \quad (3.35)$$

и



$$\mathbf{V}\mathbf{K}\mathbf{V}^\dagger = \mathbf{V}\mathbf{K}\mathbf{V}^\dagger. \quad (3.36)$$

Для матрицы  $\mathbf{Z}$ , определенной таким образом, имеет место соотношение  $\mathbf{S}\mathbf{Z}^2 = \mathbf{1}$ . В самом деле ( $\mathbf{V}$  унитарна,  $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger$ ),

$$\mathbf{S}\mathbf{Z}^2 = \underbrace{\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger}_1 \underbrace{\mathbf{V}\mathbf{K}\mathbf{V}^\dagger}_1 \mathbf{V}\mathbf{K}\mathbf{V}^\dagger = \underbrace{\mathbf{V}\mathbf{\Lambda}\mathbf{K}^2}_1 \mathbf{V}^\dagger = \mathbf{V}\mathbf{V}^\dagger = \mathbf{1}. \quad (3.37)$$

Аналогично можно проверить, что  $\mathbf{Z}^2\mathbf{S} = \mathbf{Z}\mathbf{S}\mathbf{Z} = \mathbf{1}$ . Поэтому матрицу  $\mathbf{Z}$  можно рассматривать как  $-\frac{1}{2}$ -ю степень матрицы  $\mathbf{S}$ , и обычно она так и обозначается:

$$\mathbf{S}^{-1/2} = \mathbf{Z} = \mathbf{V} \operatorname{diag} \left\{ \frac{1}{\sqrt{\lambda_i}} \right\} \mathbf{V}^\dagger. \quad (3.38)$$

Матрица  $\mathbf{S}^{-1/2}$ , очевидно, эрмитова:  $(\mathbf{S}^{-1/2})^\dagger = \mathbf{S}^{-1/2}$ . В самом деле,  $\mathbf{V}$  унитарная, а  $\mathbf{S}$  эрмитова положительно определенная матрицы, поэтому в формуле (3.38) все  $\lambda_i > 0$ .

## 2.2. Преобразование $\mathbf{S}^{-1/2}$

Рассмотрим снова обобщенное уравнение на собственные значения:

$$\mathbf{H}\mathbf{c} = E\mathbf{S}\mathbf{c}. \quad (3.39)$$

Умножим уравнение (3.39) на  $\mathbf{S}^{-1/2}$  слева и вставим единичную матрицу  $\mathbf{1} = \mathbf{S}^{-1/2}\mathbf{S}^{1/2}$  справа от матрицы  $\mathbf{H}$  (матрица  $\mathbf{S}^{1/2} = (\mathbf{S}^{-1/2})^{-1} = \mathbf{Z}^{-1}$ ; как будет видно, для решения обобщенной задачи на собственные значения нет необходимости вычислять ее явно). Мы получаем

$$\underbrace{\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}}_{\mathbf{H}'} \underbrace{\mathbf{S}^{1/2}\mathbf{c}}_{\mathbf{d}} = E \underbrace{\mathbf{S}^{1/2}\mathbf{c}}_{\mathbf{d}}. \quad (3.40)$$

Обозначая матрицу  $\mathbf{H}' = \mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$  (она эрмитова) и вектор  $\mathbf{d} = \mathbf{S}^{1/2}\mathbf{c}$ , мы получаем обычную задачу на собственные значения:

$$\mathbf{H}'\mathbf{d} = E\mathbf{d}. \quad (3.41)$$

Ее собственные значения  $E$  совпадают с собственными значениями исходной задачи (3.39), тогда как собственные векторы последней можно найти из очевидного соотношения

$$\mathbf{c} = \mathbf{S}^{-1/2}\mathbf{d}. \quad (3.42)$$

Альтернативное решение обобщенных уравнений на собственные значения достигается при помощи так называемого разложения Холецкого. В этом случае матрица  $\mathbf{S}$  рассматривается не как произведение  $\mathbf{S}^{1/2}\mathbf{S}^{1/2}$ , а как  $\mathbf{S} = \mathbf{L}\mathbf{L}^\dagger$ , где  $\mathbf{L}$  — нижняя треугольная матрица (т. е. имеющая ненулевые элементы только на главной диагонали и ниже ее); преобразование аналогично описанному ранее с тем отличием, что  $\mathbf{S}^{-1/2}$  заменяется на  $\mathbf{L}^{-1}$  или  $\mathbf{L}^{-1\dagger}$ . Однако разложение Холецкого интересно только с точки зрения численных расчетов: оно не играет какой-либо роли в квантовохимической теории.



### 2.3. Лёвдиновский базис

Преобразование  $\mathbf{S}^{-1/2}$ , описанное ранее, эквивалентно изменению базисного набора при решении линейной вариационной задачи. Вместо неортогонального набора  $\{\varphi_i\}$  базисных функций мы вводим «ортогонализированный по Лёвдину» (или «симметрично ортогонализированный») базис состояний  $\{\psi_j\}$ , определенный как

$$\psi_j = \sum_{k=1}^m S_{kj}^{-1/2} \varphi_k, \quad j = 1, 2, \dots, m. \quad (3.43)$$

Здесь мы ввели упрощенное обозначение  $S_{kj}^{-1/2}$  вместо детального, но громоздкого  $(\mathbf{S}^{-1/2})_{kj}$ .

Если базис  $\{\varphi_i\}$  линейно независим, то матрица  $\mathbf{S}^{-1/2}$  не сингулярна. Тогда набор функций  $\{\psi_j\}$  растягивает *то же самое подпространство*, что и исходный набор  $\{\varphi_i\}$ , но при этом функции  $\psi_j$  являются ортонормированными:

$$\begin{aligned} \langle \psi_i | \psi_j \rangle &= \left\langle \sum_k S_{ki}^{-1/2} \varphi_k \middle| \sum_l S_{lj}^{-1/2} \varphi_l \right\rangle \\ &= \sum_{k,l} S_{ki}^{-1/2*} \langle \varphi_k | \varphi_l \rangle S_{lj}^{-1/2} = \sum_{k,l} S_{ik}^{-1/2} S_{kl}^{-1/2} = (\mathbf{S}^0)_{ij} = \delta_{ij}, \end{aligned} \quad (3.44)$$

где мы учли, что к коэффициентам, находящимся в «бра»-части, необходимо применить комплексное сопряжение и что  $S_{ki}^{-1/2*} = S_{ik}^{-1/2}$ , так как  $\mathbf{S}^{-1/2}$  эрмитова.

Матричные элементы гамильтониана  $\hat{H}$  в базисе  $\{\psi_j\}$  имеют вид

$$\langle \psi_i | \hat{H} | \psi_j \rangle = \left\langle \sum_k S_{ki}^{-1/2} \varphi_k \middle| \hat{H} \middle| \sum_l S_{lj}^{-1/2} \varphi_l \right\rangle = \sum_{k,l} S_{ik}^{-1/2} H_{kl} S_{lj}^{-1/2} = H'_{ij}. \quad (3.45)$$

Так как наборы базисных функций  $\{\varphi_i\}$  и  $\{\psi_j\}$  растягивают одно и то же подпространство, любую функцию  $\sum_k c_k \varphi_k$  можно разложить также и по функциям  $\psi_j$ . Из определения (3.43) следует, что функции  $\varphi_k$  могут быть выражены, в свою очередь, как

$$\varphi_k = \sum_l S_{lk}^{+1/2} \psi_l \quad (k = 1, 2, \dots, m). \quad (3.46)$$

Поэтому имеем

$$\sum_k c_k \varphi_k = \sum_k c_k \sum_l S_{lk}^{+1/2} \psi_l = \sum_{k,l} (S_{lk}^{+1/2} c_k) \psi_l = \sum_l d_l \psi_l, \quad (3.47)$$

где вектор  $\mathbf{d} = \mathbf{S}^{1/2} \mathbf{c}$  тот же, что был определен выше.

Предыдущий результат означает, что при переходе к базису  $\{\psi_j\}$ , различные матрицы и векторы преобразуются как

$$\begin{aligned}\mathbf{H} &\rightarrow \mathbf{H}' = \mathbf{S}^{-1/2} \mathbf{H} \mathbf{S}^{-1/2} \\ \mathbf{S} &\rightarrow \mathbf{1} \\ \mathbf{c} &\rightarrow \mathbf{d}\end{aligned}\tag{3.48}$$

Подставляя эти выражения в уравнение (3.15), мы заключаем, что решение линейной вариационной задачи в ортонормированном базисе  $\{\psi_j\}$  сводится к «стандартному» уравнению на собственные значения

$$\mathbf{H}'\mathbf{d} = E\mathbf{d}.\tag{3.49}$$

Данное уравнение — то же самое, что мы получили ранее «преобразованием  $\mathbf{S}^{-1/2}$ » обобщенного уравнения на собственные значения (3.14). Это находится в полном согласии с очевидным требованием, что функция, реализующая стационарное значение функционала энергии в данном подпространстве, не должна зависеть от выбора базиса в этом подпространстве; равным образом и значение функционала должно быть инвариантным.

## 2.4. Свойство экстремальности симметричной ортогонализации по Лёвдину

В 1950 г. Лёвдин ввел свой метод симметричной ортогонализации, преимущественно как алгоритм, который может использоваться в численных расчетах. Однако очень скоро стало ясно, что этот метод имеет значительно более широкое значение. В 1957 г. Карлсон и Келлер доказали, что симметрично ортогонализированные функции имеют следующее замечательное свойство: среди всех возможных функций, полученных ортогонализацией данного набора неортогональных функций, симметрично ортогонализированные функции являются ближайшими в среднеквадратичном смысле (см. ниже) к исходным неортогональным функциям.

То обстоятельство, что орбитали, ортогонализированные по Лёвдину, близки к исходным орбиталям настолько, насколько это только возможно, придает особую важность симметричной ортогонализации. В 1965 г. Фишер-Яльмарс показала, что двухэлектронные интегралы, которыми пренебрегают в теориях, использующих так называемое приближение «нулевого дифференциального перекрытия» (НДП, ZDO) (по существу, она обсуждала метод ССП для  $\pi$ -электронов Паризера—Парра—Поппа), и в самом деле малы, если предположить, что базис получен из исходного базиса АО ортогонализацией по Лёвдину. Это на-

блюдение может служить обоснованием правомерности многих полуэмпирических квантовохимических методов (CNDO, MNDO и т. д.) и упрощенных моделей в физике твердого тела, в которых пренебрегают перекрыванием базисных орбиталей. Можно считать, что в этих моделях неявно используется «базис Лёвдина» и все используемые в этих методах формулы относятся именно к нему (Эти методы обычно также используют и другие приближения).

Рассмотрим упомянутое условие экстремальности и покажем, что оно удовлетворяется функциями, ортогонализированными по Лёвдину<sup>1</sup>.

Пусть  $\{\varphi_i\}$  будет набором  $N$  неортогональных, но линейно независимых базисных функций, и  $\mathbf{S}$  есть соответствующая матрица перекрывания с элементами  $S_{ij} = \langle \varphi_i | \varphi_j \rangle$ , и пусть  $\{\psi_i\}$  есть набор искомым ортонормированных функций. Наборы  $\{\varphi_i\}$  и  $\{\psi_i\}$  растягивают одно и то же подпространство, поэтому ортогональные функции можно разложить по исходным как

$$\psi_i = \sum_k C_{ki} \varphi_k. \quad (3.50)$$

Ортонормировка функций  $\psi_i$  означает, что

$$\langle \psi_i | \psi_j \rangle = \left\langle \sum_k C_{ki} \varphi_k \left| \sum_l C_{lj} \varphi_l \right. \right\rangle = \sum_{k,l} C_{ki}^* S_{kl} C_{lj} = \delta_{ij} \quad (3.51)$$

или, в матричном виде, вводя матрицу  $\mathbf{C}$  коэффициентов разложения и ее эрмитово сопряженную  $\mathbf{C}^\dagger$ , имеем

$$\mathbf{C}^\dagger \mathbf{S} \mathbf{C} = \mathbf{1}. \quad (3.52)$$

Мы допускаем возможность комплексных функций и комплексных коэффициентов  $C_{ij}$ . Так как набор  $\{\varphi_i\}$  линейно независим, матрицы  $\mathbf{S}$  и  $\mathbf{C}$  не сингулярны. Можно поэтому умножить равенство (3.52) справа на  $\mathbf{C}^{-1} \mathbf{S}^{-1}$  и слева на  $\mathbf{C}$ , получая условие ортогональности (3.51) в виде

$$\mathbf{C} \mathbf{C}^\dagger = \mathbf{S}^{-1}. \quad (3.53)$$

Перейдем к условию экстремальности. Мы требуем, чтобы два набора были так близки друг к другу в среднеквадратичном смысле, как только возможно, т. е.

$$\sum_i \int |\psi_i - \varphi_i|^2 d\tau = \sum_i \langle \psi_i - \varphi_i | \psi_i - \varphi_i \rangle = \min, \quad (3.54)$$

<sup>1</sup> Последующее доказательство воспроизводится по моей статье «On Löwdin's Method of Symmetric Orthogonalization,» I. Mayer, Int. J. Quantum Chem. **90**, 63 (2002); Copyright 2002 Wiley Periodicals.

где  $\int d\tau$  обозначает интегрирование по всем переменным, от которых зависят функции  $\varphi_i$ ,  $\psi_i$ . Разлагая интеграл  $\langle \psi_i - \varphi_i | \psi_i - \varphi_i \rangle$  в (3.54) и учитывая, что  $\langle \psi_i | \psi_i \rangle = 1$  по условию ортонормировки, а также, что  $\langle \varphi_i | \varphi_i \rangle = \text{const}$ , мы можем сформулировать наше условие как следующую вариационную задачу: необходимо найти матрицу разложения  $\mathbf{C}$ , такую, что при дополнительном условии ортонормировки  $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ :

$$\sum_i \left\{ \langle \varphi_i | \psi_i \rangle + \langle \psi_i | \varphi_i \rangle \right\} = \max. \quad (3.55)$$

Для решения этой вариационной задачи с дополнительными условиями можно было бы, в принципе, использовать стандартный метод множителей Лагранжа, но представляется более простым, и, в особенности, гораздо более прозрачным, явно рассмотреть разные *специальные вариации* функций  $\psi_i$ , и потребовать обращения в нуль вариации функционала в левой части выражения (3.55) при всех *допустимых* вариациях, т. е. вариациях, не нарушающих условие нормировки:

$$\sum_i \left\{ \langle \varphi_i | \delta \psi_i \rangle + \langle \delta \psi_i | \varphi_i \rangle \right\} = 0 \quad (3.56)$$

для всех наборов  $\delta \psi_i$ , которые не нарушают условия  $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ . Как обычно, это требование позволяет получить необходимые условия экстремума.

### Вариации фазы

Простейшей вариацией функции является вариация ее фазы, когда функция  $\psi_i$  заменяется на

$$\psi_i + \delta \psi_i = e^{i\eta} \psi_i \quad (3.57)$$

с  $\eta \rightarrow 0$ , являющейся произвольным вещественным вариационным параметром. Разлагая экспоненту в ряд ( $e^{i\eta} = 1 + i\eta + \dots$ ), легко увидеть, что вариация первого порядка для  $\psi_i$  равна

$$\delta \psi_i = i\eta \psi_i. \quad (3.58)$$

Вариация фазы (3.57) никак не нарушает условий ортонормировки, тогда как вариация (3.58) сохраняет их в первом порядке по  $\eta$ . Подставляя вариацию (3.58) в уравнение (3.56), получаем

$$\langle \varphi_i | i\eta \psi_i \rangle + \langle i\eta \psi_i | \varphi_i \rangle = 0, \quad (3.59)$$

что (так как  $\eta \rightarrow 0$ , но  $\eta \neq 0$ ) сводится к

$$\langle \varphi_i | \psi_i \rangle = \langle \psi_i | \varphi_i \rangle . \quad (3.60)$$

Это означает, что интеграл  $\langle \psi_i | \varphi_i \rangle$  вещественен. В покомпонентной записи имеем

$$\langle \varphi_i | \sum_l C_{li} \varphi_l \rangle = \langle \sum_l C_{li} \varphi_l | \varphi_i \rangle , \quad (3.61)$$

т. е.

$$\sum_l S_{il} C_{li} = \sum_l C_{li}^* S_{li} . \quad (3.62)$$

В матричных обозначениях это равенство записывается как

$$(\mathbf{S}\mathbf{C})_{ii} = (\mathbf{C}^\dagger \mathbf{S})_{ii} . \quad (3.63)$$

### Вариации типа вращения

Вариация типа  $\delta\psi_i = \eta\psi_j$  ( $i \neq j$ ) сама по себе не допустима, так как нарушает условия ортонормировки уже в первом порядке. Однако, *совместная вариация* (здесь  $\eta \rightarrow 0$  может быть и комплексной)

$$\begin{aligned} \delta\psi_i &= \eta\psi_j \\ \delta\psi_j &= -\eta^* \psi_i \end{aligned} \quad (3.64)$$

является допустимой, так как в этом случае условие ортонормировки сохраняется в первом порядке по параметру  $\eta$ . В случае совместной вариации двух функций  $\psi_i$  и  $\psi_j$ , условие (3.56) сводится к

$$\langle \varphi_i | \delta\psi_i \rangle + \langle \delta\psi_i | \varphi_i \rangle + \langle \varphi_j | \delta\psi_j \rangle + \langle \delta\psi_j | \varphi_j \rangle = 0 . \quad (3.65)$$

Подставляя вариации (3.64) в выражение (3.65), мы получаем

$$\eta(\langle \varphi_i | \psi_j \rangle - \langle \psi_i | \varphi_j \rangle) + c.c. = 0 . \quad (3.66)$$

Так как  $\eta$  содержит произвольный фазовый множитель, условие (3.66) выполняется для произвольных  $\eta$  тогда и только тогда, когда как величина, выписанная явно, так и ее комплексно сопряженная (*c.c.*) равны нулю по отдельности:

$$\langle \varphi_i | \psi_j \rangle - \langle \psi_i | \varphi_j \rangle = 0 . \quad (3.67)$$

Используя разложение (3.50) для функций, имеем

$$\langle \varphi_i | \sum_l C_{lj} \varphi_l \rangle - \langle \sum_l C_{li} \varphi_l | \varphi_j \rangle = 0 , \quad (3.68)$$

т. е.

$$\sum_l \mathbf{S}_{il} \mathbf{C}_{lj}^* = \sum_l C_{li}^* S_{lj}, \quad (3.69)$$

или в матричных обозначениях

$$(\mathbf{S}\mathbf{C})_{ij} = (\mathbf{C}^\dagger \mathbf{S})_{ij}. \quad (3.70)$$

Условие (3.63) должно выполняться для всех функций  $\psi_i$ , а условие (3.70) должно выполняться для любой пары функций  $\psi_i, \psi_j$ ; значит, мы показали, что все элементы матриц  $\mathbf{S}\mathbf{C}$  и  $\mathbf{C}^\dagger \mathbf{S}$  должны быть равны, т. е.

$$\mathbf{S}\mathbf{C} = \mathbf{C}^\dagger \mathbf{S}. \quad (3.71)$$

Умножая это уравнение на  $\mathbf{C}$  справа и используя соотношение (3.52), получаем

$$\mathbf{S}\mathbf{C}^2 = \mathbf{1} \quad (3.72)$$

или

$$\mathbf{C}^2 = \mathbf{S}^{-1}. \quad (3.73)$$

Сравнивая выражения (3.73) и (3.53), мы видим, что матрица  $\mathbf{C}$  должна быть эрмитовой ( $\mathbf{C} = \mathbf{C}^\dagger$ ), значит, она должна представлять *эрмитов квадратный корень* матрицы  $\mathbf{S}^{-1}$ :

$$\mathbf{C} = \mathbf{S}^{-1/2} \quad \text{Ч. т. д.} \quad (3.74)$$

## 2.5. Ортогонализация по Лёвдину. Двумерный пример

Полезно проделать явно выкладки, обсуждаемые в разд. 2.1 и 2.3, для простейшего случая ортогонализации по Лёвдину двух перекрывающихся функций.

Рассмотрим две нормированные функции  $\varphi_1$  и  $\varphi_2$  с перекрыванием  $\langle \varphi_1 | \varphi_2 \rangle = S$ . Тогда матрица перекрывания равна

$$\mathbf{S} = \begin{pmatrix} 1 & S \\ S & 1 \end{pmatrix}. \quad (3.75)$$

Соображения симметрии показывают (это также легко проверить и напрямую), что нормированными собственными векторами матрицы  $\mathbf{S}$  являются векторы

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{и} \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (3.76)$$

с собственными значениями  $\lambda_1^{\text{®}} = 1 + S$  и  $\lambda_2 = 1 - S$  соответственно<sup>1</sup>. Это означает, что унитарная матрица, диагонализующая **S**, равна

$$\mathbf{V} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (3.77)$$

Эта матрица интересна тем, что она является не только унитарной, но и эрмитовой.

Матрица **Λ**, соответственно, равна

$$\mathbf{\Lambda} = \begin{pmatrix} 1 + S & 0 \\ 0 & 1 - S \end{pmatrix}, \quad (3.78)$$

а матрица **K**

$$\mathbf{K} = \begin{pmatrix} \frac{1}{\sqrt{1+S}} & 0 \\ 0 & \frac{1}{\sqrt{1-S}} \end{pmatrix}, \quad (3.79)$$

и мы сразу получаем матрицу **S**<sup>-1/2</sup> как

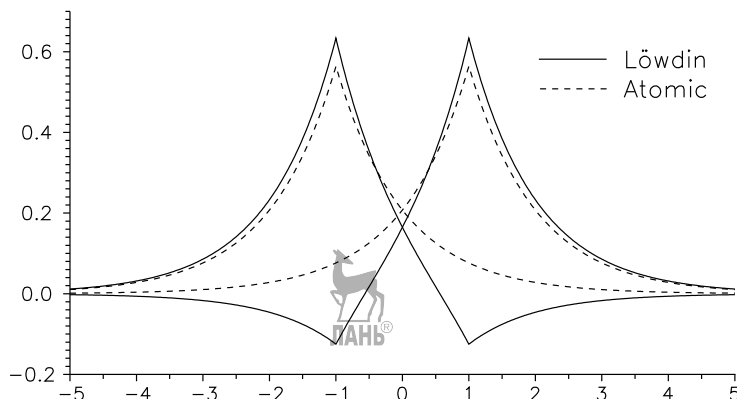
$$\begin{aligned} \mathbf{S}^{-1/2} &= \mathbf{V} \mathbf{K} \mathbf{V}^\dagger \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1+S}} & 0 \\ 0 & \frac{1}{\sqrt{1-S}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \frac{1}{\sqrt{1+S}} + \frac{1}{\sqrt{1-S}} & \frac{1}{\sqrt{1+S}} - \frac{1}{\sqrt{1-S}} \\ \frac{1}{\sqrt{1+S}} - \frac{1}{\sqrt{1-S}} & \frac{1}{\sqrt{1+S}} + \frac{1}{\sqrt{1-S}} \end{pmatrix}. \end{aligned} \quad (3.80)$$

Соответственно, ортогонализированными по Лёвдину функциями  $\psi_1$  и  $\psi_2$  являются

$$\begin{aligned} \psi_1 &= \frac{1}{2} \left\{ \left( \frac{1}{\sqrt{1+S}} + \frac{1}{\sqrt{1-S}} \right) \varphi_1 + \left( \frac{1}{\sqrt{1+S}} - \frac{1}{\sqrt{1-S}} \right) \varphi_2 \right\} \\ \psi_2 &= \frac{1}{2} \left\{ \left( \frac{1}{\sqrt{1+S}} - \frac{1}{\sqrt{1-S}} \right) \varphi_1 + \left( \frac{1}{\sqrt{1+S}} + \frac{1}{\sqrt{1-S}} \right) \varphi_2 \right\}. \end{aligned} \quad (3.81)$$

Можно видеть, что  $\psi_1 \rightarrow \varphi_1$  и  $\psi_2 \rightarrow \varphi_2$ , если  $S \rightarrow 0$ .

<sup>1</sup> Это и есть (с точностью до нормировки) векторы коэффициентов разложения канонически ортогональных функций — т. е. собственные векторы матрицы **S** (см. ниже). — Прим. ред.



На рисунке показаны точные атомные  $1s$ -орбитали двух атомов водорода, расположенных на расстоянии 2 а.е. (большем, чем равновесное расстояние  $\sim 1.4$  а.е. в молекуле  $\text{H}_2$ ), и соответствующие им функции, ортогонализированные по Лёвдину (величина  $S$  составляет 0.5865). Можно видеть, что лёвдиновские орбитали не сильно отличаются от их перекрывающихся аналогов в области, соответствующей собственно атому, и имеют отрицательный пик на другом атоме, что обеспечивает итоговую ортогональность лёвдиновских орбиталей.

Интересно исследовать зависимость коэффициентов орбиталей, ортогонализированных по Лёвдину, от интеграла перекрывания  $S$ . Орбитали в (3.81) можно записать как

$$\begin{aligned}\psi_1 &= \alpha\varphi_1 + \beta\varphi_2 \\ \psi_2 &= \beta\varphi_1 + \alpha\varphi_2\end{aligned}\tag{3.82}$$

где

$$\begin{aligned}\alpha &= \frac{1}{2} \left( \frac{1}{\sqrt{1+S}} + \frac{1}{\sqrt{1-S}} \right) \\ \beta &= \frac{1}{2} \left( \frac{1}{\sqrt{1+S}} - \frac{1}{\sqrt{1-S}} \right)\end{aligned}\tag{3.83}$$

В нижеследующей таблице приведены значения  $\alpha$  и  $\beta$  для нескольких характерных значений межатомного расстояния (значений  $S$ ).

$R$ (а.е.)	$S$	$\alpha$	$\beta$
1.5	0.7252	1.3345	-0.5731
2	0.5865	1.1745	-0.3806
3	0.3485	1.0500	-0.1889

В случае двух функций можно также взглянуть на ортогонализацию по Лёвдину с более специальной точки зрения. Сначала рассмотрим две



ортонормированные функции  $\psi_1$  и  $\psi_2$ , и их нормированные сумму и разность:

$$\begin{aligned} g &= \frac{1}{\sqrt{2}}(\psi_1 + \psi_2) \\ u &= \frac{1}{\sqrt{2}}(\psi_1 - \psi_2). \end{aligned} \quad (3.84)$$

Функции  $g$  и  $u$  ортогональны друг другу.

Исходные ортогональные функции  $\psi_1$  и  $\psi_2$  могут быть, в свою очередь, выражены как нормированные сумма и разность функций  $g$  и  $u$ :

$$\begin{aligned} \psi_1 &= \frac{1}{\sqrt{2}}(g + u) \\ \psi_2 &= \frac{1}{\sqrt{2}}(g - u). \end{aligned} \quad (3.85)$$

В случае, когда мы имеем две нормированные но неортогональные функции  $\varphi_1$  и  $\varphi_2$ , с интегралом перекрывания  $S$ , мы снова можем составить их нормированную сумму и разность, чьи нормировочные множители будут зависеть от  $S$ :

$$\begin{aligned} g &= \frac{1}{\sqrt{2(1+S)}}(\varphi_1 + \varphi_2) \\ u &= \frac{1}{\sqrt{2(1-S)}}(\varphi_1 - \varphi_2). \end{aligned} \quad (3.86)$$

Функции  $u$  и  $g$  снова ортогональны.

Теперь мы можем снова составить нормированные сумму и разность  $g$  и  $u$ , но в этом случае мы получаем не исходные перекрывающиеся функции  $\varphi_1$   $\varphi_2$ , а их аналоги  $\psi_1$  и  $\psi_2$ , ортогонализованные по Лёвдину:

$$\begin{aligned} \psi_1 &= \frac{1}{\sqrt{2}}(g + u) = \alpha\varphi_1 + \beta\varphi_2 \\ \psi_2 &= \frac{1}{\sqrt{2}}(g - u) = \beta\varphi_1 + \alpha\varphi_2 \end{aligned} \quad (3.87)$$

где значения  $\alpha$  и  $\beta$  определены выражениями (3.83).

### 3. Линейная независимость базиса и каноническая ортогонализация по Лёвдину

#### 3.1. Собственные значения матрицы интегралов перекрывания. Мера линейной независимости базиса

Рассмотрим  $m$  функций  $\varphi_i$  с матрицей перекрывания  $\mathbf{S}$ , имеющей элементы  $S_{kl} = \langle \varphi_k | \varphi_l \rangle$ , и пусть  $\mathbf{c}^i$  — собственные векторы  $\mathbf{S}$ , соответствующие собственным значениям  $\lambda_i$ :

$$\mathbf{S} \mathbf{c}^i = \lambda_i \mathbf{c}^i. \quad (3.88)$$

Пусть все  $\mathbf{c}^i$  нормированы на единицу:

$$\mathbf{c}^{i\dagger} \mathbf{c}^i = \sum_k |c_k^i|^2 = 1 \quad (3.89)$$

(в этом разделе все суммирование проводится от 1 до  $m$ ). Далее, зададим функцию

$$\psi_i = \sum_k c_k^i \varphi_k \quad (3.90)$$

и определим квадрат ее нормы (интеграл перекрывания самой с собой):

$$\begin{aligned} \langle \psi_i | \psi_i \rangle &= \left\langle \sum_j c_j^i \varphi_j \left| \sum_k c_k^i \varphi_k \right. \right\rangle = \sum_{j,k} c_j^{i*} S_{jk} c_k^i \\ &= \sum_j c_j^{i*} \left( \sum_k S_{jk} c_k^i \right) = \sum_j c_j^{i*} \lambda_i c_j^i = \lambda_i \sum_j |c_j^i|^2 = \lambda_i, \end{aligned} \quad (3.91)$$

где мы использовали тот факт, что  $\mathbf{c}^i$  является нормированным собственным вектором  $\mathbf{S}$ , и поэтому  $\sum_k S_{jk} c_k^i = \lambda_i c_j^i$  и  $\sum_j |c_j^i|^2 = 1$ .

Из формулы (3.91) следует, что

а) Собственные значения матрицы перекрывания  $\mathbf{S}$  неотрицательны, так как каждое из них представляет собой квадрат нормы  $\langle \psi_i | \psi_i \rangle$  функции  $\psi_i$ . Если базис линейно независим, то все  $\lambda_i > 0$  (для линейно независимого базиса равенство  $\langle \psi_i | \psi_i \rangle = 0$  возможно, только если все  $c_j^i = 0$ ). Другими словами, матрица перекрывания  $\mathbf{S}$  является «положительно определенной», если базисный набор линейно независим и «положительно полуопределенной», если он линейно зависим.

б) Можно рассматривать наименьшее собственное значение матрицы  $\mathbf{S}$ , как меру линейной независимости используемого базиса. Чтобы убедиться в этом, мы докажем следующую теорему.

### Теорема

Пусть  $\psi = \sum_j d_j \varphi_j$  — линейная комбинация базисных функций  $\varphi_j$  с произвольным вектором коэффициентов  $\mathbf{d}$ , нормированным на единицу ( $\sum_j |d_j|^2 = 1$ ). Тогда значение интеграла  $\langle \psi | \psi \rangle$  ограничено наименьшим и наибольшим собственными значениями матрицы перекрывания  $\mathbf{S}$ , т. е. оно удовлетворяет неравенствам

$$\lambda_{\min} \leq \langle \psi | \psi \rangle \leq \lambda_{\max} . \quad (3.92)$$

### Доказательство

Матрица  $\mathbf{S}$  эрмитова и ее собственные векторы образуют полный ортонормированный набор в линейном пространстве  $m$ -мерных векторов. Поэтому всегда можно записать

$$\mathbf{d} = \sum_i \eta_i \mathbf{c}^i \quad (3.93)$$

где  $\sum_i |\eta_i|^2 = 1$ . Подставляя разложение (3.93) для  $\mathbf{d}$  в интеграл  $\langle \psi | \psi \rangle$ , мы получаем

$$\begin{aligned} \langle \psi | \psi \rangle &= \left\langle \sum_j d_j \varphi_j \left| \sum_k d_k \varphi_k \right. \right\rangle = \sum_{j,k} d_j^* S_{jk} d_k \\ &= \mathbf{d}^\dagger \mathbf{S} \mathbf{d} = \sum_{i,j} \eta_i^* \mathbf{c}^{i\dagger} \mathbf{S} \mathbf{c}^j \eta_j = \sum_{i,j} \eta_i^* \eta_j \lambda_j \mathbf{c}^{i\dagger} \mathbf{c}^j \\ &= \sum_{i,j} \eta_i^* \eta_j \lambda_j \delta_{ij} = \sum_i |\eta_i|^2 \lambda_i , \end{aligned} \quad (3.94)$$

где учтено, что  $\mathbf{S} \mathbf{c}^j = \lambda_j \mathbf{c}^j$  и  $\mathbf{c}^{i\dagger} \mathbf{c}^j = \delta_{ij}$ .

Так как  $\lambda_i \leq \lambda_{\max}$ ,  $\sum_i |\eta_i|^2 \lambda_i \leq \lambda_{\max} \sum_i |\eta_i|^2 = \lambda_{\max}$  и, аналогично,  $\lambda_i \geq \lambda_{\min}$ , то, следовательно, имеем  $\sum_i |\eta_i|^2 \lambda_i \geq \lambda_{\min}$ . Ч. т. д.

Если базис почти линейно зависим, то наименьшее из собственных значений  $\lambda_i$  становится очень малым. С учетом ошибок округления, при достижении некоторого достаточно малого значения (скажем,  $10^{-8} - 10^{-10}$ , в зависимости от используемого компьютера) его можно рассматривать как практически равное нулю.

Существование такой приближенной линейной зависимости может затруднить вычисления, например, в результате деления на число, очень близкое к 0, или вычислении малых разностей очень больших чисел. В

таких ситуациях необходимо исключить из базиса те функции  $\psi_i$ , которые приблизительно являются линейными комбинациями других. Такowymi являются функции, построенные согласно рецепту (3.90) с использованием тех собственных векторов  $\mathbf{S}$ , которые соответствуют наименьшим собственным значениям (меньше некоторого заданного предела). В таких случаях решения обобщенных уравнений на собственные значения (разд. 1) лучше всего находить с помощью «канонической ортогонализации» Лёвдина. (Эту процедуру следует отличать от симметричной ортогонализации, которая обсуждалась в разд. 2.)

### 3.2. Каноническая ортогонализация по Лёвдину

Перенормируем в разложениях (3.90) функции, образованные с помощью тех собственных векторов матрицы перекрывания  $\mathbf{S}$ , которые соответствуют ненулевым собственным значениям (на практике тем, которые больше, чем выбранная граница). Из выражения (3.91) для норм векторов получаем функции

$$\psi_i' = \frac{1}{\sqrt{\lambda_i}} \psi_i = \frac{1}{\sqrt{\lambda_i}} \sum_j c_j^i \varphi_j, \quad (3.95)$$

которые теперь нормированы на единицу и образуют ортонормированный базис в пространстве, растянутом исходными функциями  $\{\varphi_j\}$ :

$$\begin{aligned} \langle \psi_i' | \psi_j' \rangle &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \langle \sum_k c_k^j \varphi_k | \sum_l c_l^i \varphi_l \rangle \\ &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \sum_{k,l} c_k^{j*} S_{kl} c_l^i = \frac{1}{\sqrt{\lambda_i \lambda_j}} \sum_k c_k^{j*} \sum_l S_{kl} c_l^i \\ &= \frac{\lambda_i}{\sqrt{\lambda_i \lambda_j}} \sum_k c_k^{j*} c_k^i = \delta_{ij}, \end{aligned} \quad (3.96)$$

где мы использовали равенство  $\sum_l S_{kl} c_l^i = \lambda_i c_k^i$ , и тот факт, что собственные векторы  $\mathbf{c}^i$  эрмитовой матрицы  $\mathbf{S}$  ортонормированы. (В этом разделе суммирование снова идет от 1 до  $m$ , за исключением явно оговоренных случаев.)

Процедура составления ортонормированных векторов  $\{\psi_i'\}$  называется «канонической ортогонализацией» по Лёвдину функций  $\{\varphi_j\}$ .

Определим матрицу  $\mathbf{Q}$ , составленную из коэффициентов разложения функций  $\psi_i'$  по исходному базису  $\{\varphi_j\}$ , образуя ее  $i$ -й столбец из коэффициентов  $i$ -й функции  $\psi_i'$ :

$$Q_{ji} = \frac{1}{\sqrt{\lambda_i}} c_j^i. \quad (3.97)$$

Если базис  $\{\varphi_j\}$  линейно зависим (строго или приближенно), и мы вследствие этого отбрасываем одну или более функций  $\psi_i$ , соответствующих собственным векторам  $\mathbf{S}$  с нулевыми (малыми) собственными значениями, тогда матрица  $\mathbf{Q}$  — прямоугольная и имеет столько столбцов, каково число линейно независимых функций  $\psi_i'$ . Если число исходных базисных функций  $\varphi_j$  равно  $m$ , а  $n$  из них линейно независимы, то матрица  $\mathbf{Q}$  является матрицей  $m \times n$  (если нет линейной зависимости, то, разумеется,  $n = m$ ).

Если имеется линейная зависимость, то разложение разных функций по базису  $\{\varphi_j\}$  неоднозначно. В самом деле, если базис линейно зависимый, то можно составить нетривиальную обращающуюся в нуль комбинацию  $\sum_j p_j \varphi_j = 0$  базисных функций  $\varphi_j$  (не все  $p_j = 0$ ). Если

какую-то функцию  $\Psi$  можно представить как  $\Psi = \sum_j q_j \varphi_j$ , то добавляя

к ней такую исчезающую комбинацию, мы получим, что *ту же*  $\Psi$  можно записать и как  $\Psi = \sum_j (q_j + p_j) \varphi_j = \sum_j q_j' \varphi_j$ , где  $q_j' \neq q_j$ . В то же время,

эта же функция имеет однозначное разложение в ортонормированном (и, следовательно, линейно независимом) базисе  $\{\psi_i'\}$  :  $\Psi = \sum_{j=1}^n r_j \psi_j'$ .

Подставляя сюда разложение (3.95) вектора  $\psi_j'$ , получаем с учетом определения (3.97)

$$\begin{aligned} \Psi &= \sum_{j=1}^n r_j \psi_j' = \sum_{j=1}^n r_j \frac{1}{\sqrt{\lambda_j}} \sum_{k=1}^m c_k^j \varphi_k = \sum_{j=1}^n r_j \sum_{k=1}^m Q_{kj} \varphi_k = \\ &= \sum_{k=1}^m \left( \sum_{j=1}^n Q_{kj} r_j \right) \varphi_k = \sum_{k=1}^m (\mathbf{Qr})_k \varphi_k = \sum_{k=1}^m q_k \varphi_k. \end{aligned} \quad (3.98)$$

Следовательно, одним из возможных векторов  $\mathbf{q}$  является

$$\mathbf{q} = \mathbf{Qr}. \quad (3.99)$$

Матрица  $\mathbf{Q}^\dagger \mathbf{S} \mathbf{Q}$  размера  $n \times n$  является единичной матрицей. В самом деле

$$\begin{aligned} (\mathbf{Q}^\dagger \mathbf{S} \mathbf{Q})_{ij} &= \sum_{k,l} (\mathbf{Q}^\dagger)_{ik} S_{kl} Q_{lj} = \sum_{k,l} Q_{ki}^* S_{kl} Q_{lj} \\ &= \sum_{k,l} \frac{1}{\sqrt{\lambda_i}} c_k^{i*} S_{kl} \frac{1}{\sqrt{\lambda_j}} c_l^j = \frac{1}{\sqrt{\lambda_i \lambda_j}} \sum_k c_k^{i*} \sum_l S_{kl} c_l^j \\ &= \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} \sum_k c_k^{i*} c_k^j = \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} \delta_{ij} = \delta_{ij}, \end{aligned} \quad (3.100)$$

где мы воспользовались тем фактом, что  $\mathbf{c}^i$  и  $\mathbf{c}^j$  ортонормированные собственные векторы матрицы  $\mathbf{S}$ . (Результат (3.100) также следует из того, что функции  $\psi_i'$  ортонормированы.)

Умножим обобщенное уравнение на собственные значения

$$\mathbf{H} \mathbf{c} = \varepsilon \mathbf{S} \mathbf{c} \quad (3.101)$$

на  $\mathbf{Q}^\dagger$  слева:

$$\mathbf{Q}^\dagger \mathbf{H} \mathbf{c} = \varepsilon \mathbf{Q}^\dagger \mathbf{S} \mathbf{c}. \quad (3.102)$$

Если базис линейно зависимый, то вектор  $\mathbf{c}$ , выражающий иско-  
мое решение уравнений (3.101) и (3.102), определен неоднозначно; од-  
нако, все возможные функции  $\sum_i c_i \varphi_i$  имеют одно и то же разложение

$\sum_j d_j \psi_j'$  по ортонормированному базису  $\{\psi_i'\}$ . Как мы видели ранее, в такой ситуации мы можем выбрать вектор  $\mathbf{c}$  как  $\mathbf{c} = \mathbf{Q} \mathbf{d}$ . (Очевидно, мы можем записать вектор  $\mathbf{c}$  в виде  $\mathbf{c} = \mathbf{Q} \mathbf{d}$  также и в случае, когда базис  $\{\varphi_i\}$  линейно независим. В этом случае преобразование между наборами  $\{\varphi_i\}$  и  $\{\psi_i'\}$  не сингулярно, и можно также вычислить обратное к преобразованию (3.99), т. е. обратную к матрице  $\mathbf{Q}$ : имеем  $\mathbf{Q}^\dagger \mathbf{S} \mathbf{Q} = \mathbf{1}$ , согласно (3.100), откуда  $\mathbf{Q}^{-1} = \mathbf{Q}^\dagger \mathbf{S}$ . Это значит, что векторы  $\mathbf{c}$  и  $\mathbf{d}$  определяют друг друга взаимнооднозначно, если базис линейно независим.)

Подставляя  $\mathbf{c} = \mathbf{Q} \mathbf{d}$  в уравнение (3.102), получаем

$$\mathbf{Q}^\dagger \mathbf{H} \mathbf{Q} \mathbf{d} = \varepsilon \mathbf{Q}^\dagger \mathbf{S} \mathbf{Q} \mathbf{d}, \quad (3.103)$$

т. е. обозначая эрмитову матрицу  $\mathbf{Q}^\dagger \mathbf{H} \mathbf{Q}$  как  $\mathbf{H}'$ , и учитывая, что  $\mathbf{Q}^\dagger \mathbf{S} \mathbf{Q} = \mathbf{1}$ , согласно (3.100):

$$\mathbf{H}' \mathbf{d} = \varepsilon \mathbf{d}, \quad (3.104)$$

что есть обычная задача на собственные значения.

Легко видеть, что матрица  $\mathbf{H}'$  является матрицей гамильтониана  $\hat{H}$  в базисе ортонормированных функций  $\psi_i'$ , из которого исключены линейно зависимые векторы исходного базиса  $\{\varphi_i\}$  (если таковые есть). После отыскания собственных векторов (3.104), они могут быть преобразованы обратно к исходному базису  $\{\varphi_i\}$  по соотношению  $\mathbf{c} = \mathbf{Qd}$ ; разумеется, можно найти не больше собственных векторов, чем число  $n$  линейно независимых базисных функций.

## Библиографические заметки



### Раздел 1.

Стандартная задача, включаемая во все книги по квантовой химии, от [1] до [2] (в более ранних книгах рассматривался только случай ортонормированных базисных функций). Вывод уравнений (3.7)-(3.10) находится в тесной связи со схемой, называемой в математике методом Галёркина [3], а в квантовой химии — методом моментов [4, 5]. В методе Галёркина ищут приближенное решение функционального (например, дифференциального) уравнения  $\hat{L}f = 0$ , требуя, чтобы его проекции на разные «весовые функции»  $\varphi_i$  обращались в нуль:  $\int \varphi_i^* \hat{L}f dx = 0$ . Если приблизить функцию  $f$  линейной комбинацией функций набора  $\{\varphi_i\}$  и использовать *те же* функции как весовые, то метод Галёркина эквивалентен линейной вариационной задаче и приводит к вековому уравнению (3.10). (Метод «связанных кластеров» в проблеме электронной корреляции (см. гл. 8, разд. 3.3) является типичным методом моментов.)

### Раздел 2.

Симметричная ортогонализация по Лёвдину была введена в [6]. Ее экстремальность была впервые доказана в [7]; доказательство в разд. 2.4 принадлежит автору [8]. На важность ортогонализации по Лёвдину для понимания успеха методов ПДП было указано в [9].

Симметричная ортогонализация по Лёвдину используется в самых разных областях квантовой химии — от *a posteriori* анализа «натуральных связевых орбиталей» в [10-12] до теории межмолекулярных взаимодействий [13-15] и развития формализма вторичного квантования для неортогональных базисных наборов [16,17] и т. д.

### Раздел 3.

Лёвдин предложил свою «каноническую ортогонализацию» в [18] (см. также [19]). Исключение приблизительных линейных зависимостей является общей практикой в любой стандартной квантовохимической программе (например, [20, 21]). Если используется большой базисный набор, то изменение числа сохраняемых функций может вызвать некоторые «труднопонятные» мелкие разрывы на рассчитанных потенциальных поверхностях.

## Для дальнейшего изучения

Монография Уилкинсона [22] является классической работой, посвященной алгоритмам, используемым в практическом численном решении задач на собственные значения (см. также, например, [23]). Дэвидсон [24] обсуждает задачу со специальной квантовохимической точки зрения (включая его собственные алгоритмы для больших матриц). В большинстве практически важных случаев можно с успехом использовать стандартные пакеты программ по линейной алгебре, например, «Lараск» [25].

Для краткого обзора проблемы неортогональности и разных схем ортогонализации в духе Лёвдина мы отсылаем читателя к [26].

## Литература

1. Eyring H., Walter J., Kimball G. E. *Quantum Chemistry*, John Wiley & Sons, New York 1944. (Имеется русский перевод: Эйринг Г., Уолтер Д., Кимбалл Г. *Квантовая химия*. — М.: Издательство, 1948).
2. Veszprémi T., Fehér M. *Quantum Chemistry: Fundamentals to applications*. Kluwer Academic/Plenum New York, 1999.
3. Korn G. A., Korn T. M. *Mathematical Handbook for Scientists and Engineers*. MacGraw-Hill, New York 1961. (Имеется русский перевод: Корн Г., Корн Т. *Справочник по математике для ученых и инженеров*. — М.: Наука, 1974).
4. Szondy E., Szondy T. *Acta Phys. Hung.* **20**, 253 (1966).
5. Boys S. F. *Proc. Roy. Soc.* **A309**, 195 (1969).
6. Löwdin P.-O. *J. Chem. Phys.* **18**, 365 (1950).
7. Carlson B. C., Keller J. M. *Phys. Rev.* **105**, 102 (1957).
8. Mayer I. *Int. J. Quantum Chem.* **90**, 63 (2002).
9. Fischer-Hjalmars I. in *Modern Quantum Chemistry (Istanbul Lectures)*, ed. O. Sinanoğlu. Academic Press, New York 1965 (Имеется русский перевод: Фишер-Яльмарс И., в *Современная квантовая химия* (ред. О. Синаноглу). — М.: Мир, 1968).
10. Foster J. P., Weinhold F. *J. Am. Chem. Soc.* **102**, 7211 (1980).
11. Reed A. E., Weinstock R. B., Weinhold F. *J. Chem. Phys.* **83**, 735 (1985).
12. Reed A. E., Curtiss L. A., Weinhold F. *Chem. Rev.* **88**, 899 (1988).
13. Kvasnička V., Laurinc V., Hubač I. *Phys. Rev. A* **10**, 2016 (1974).
14. Laurinc V., Lukeš V., Biskupič S. *Theor. Chem. Acc.* **99**, 53 (1989).
15. Lukeš V., Laurinc V., Biskupič S. *Int. J. Quantum Chem.* **75**, 81 (1999).
16. Kvasnička V. *Chem. Phys. Lett.* **51**, 165 (1977).
17. Mayer I. *Int. J. Quantum Chem.* **23**, 341 (1983).
18. Löwdin P.-O. *Adv. Phys.* **5**, 1 (1966).
19. Löwdin P.-O. *J. appl. Phys. Suppl.* **33**, 251 (1962).
20. Программа «HONDO 2000», M. Dupuis, A. Marques and E.R. Davidson. (Основана на программе «HONDO 95.3» тех же авторов, Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405.)
21. Program «Gaussian 98», M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi,



- R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1998.
22. Wilkinson J. H. *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford 1965 (Имеется русский перевод: Уилкинсон Дж. *Алгебраическая проблема собственных значений*. — М.: Наука, 1966).
23. Parlett B. P. *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs 1980.
24. Davidson E. R. *Matrix Eigenvector Methods in Methods of Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.
25. См. web-страницу <http://www.netlib.org/lapack>
26. del Re G. *The Non-Orthogonality Problem and Orthogonalization procedures in Quantum Science, Methods and Structure*, Plenum, New York 1976.



# Метод возмущений

## 1. Невырожденная теория возмущений Рэля—Шрёдингера

### 1.1. Формулировка задачи

Предположим, что нам известны решения (т. е. ортонормированные собственные функции  $\Psi_i^0$  и соответствующие собственные значения  $E_i^0$ ) некоторого «невозмущенного» уравнения Шрёдингера

$$\hat{H}^0 \Psi_i^0 = E_i^0 \Psi_i^0. \quad (4.1)$$

Мы хотим воспользоваться этим знанием для приближенного решения «возмущенного» уравнения Шрёдингера, содержащего гамильтониан  $\hat{H}$ , который считается «близким» к  $\hat{H}^0$ :

$$\hat{H}\Psi = (\hat{H}^0 + \hat{V})\Psi = E\Psi. \quad (4.2)$$

Возмущенный и «невозмущенный» гамильтонианы  $\hat{H}$  и  $\hat{H}^0$  «близки» друг к другу, если  $\hat{V}$  является «малым». Можно выразить степень малости  $\hat{V}$ , введя обозначение

$$\hat{V} = \lambda \hat{W}, \quad (4.3)$$

где оператор  $\hat{W}$  имеет тот же порядок величины  $\mathcal{O}(1)$ , что и  $\hat{H}^0$ , тогда как  $\lambda$  есть (малый) параметр «силы» возмущения. Получаем

$$(\hat{H}^0 + \lambda \hat{W})\Psi = E\Psi. \quad (4.4)$$

В теории возмущений (ТВ)  $\lambda$  обычно рассматривается как малый непрерывный параметр (формально  $\lambda \rightarrow 0$ ), и задача состоит в том, чтобы найти как волновую функцию, так и энергию в виде степенных рядов по  $\lambda$ :

$$\Psi_i = \sum_{l=0}^{\infty} \lambda^l \Psi_i^{(l)} \quad (4.5)$$

и

$$E_i = \sum_{k=0}^{\infty} \lambda^k E_i^{(k)}. \quad (4.6)$$

Существуют задачи, в которых  $\lambda$  на самом деле является непрерывной переменной, значение которой можно свободно менять. Это имеет место, например, когда рассматривается влияние на изучаемую систему внешнего электромагнитного поля, напряженность которого можно регулировать в эксперименте. Однако в большинстве случаев, интересных для квантовой химии,  $\lambda$  в действительности не является переменной, и рассмотрение ее как непрерывного параметра является только математической идеализацией. Тем не менее, если  $\lambda$  действительно мало, то классификация результатов по степеням  $\lambda$  часто оказывается очень полезной.

Во всяком случае, необходимо быть внимательным, так как даже очень слабое возмущение может изменить *общий вид* спектра гамильтониана. Например, если на потенциальной поверхности имеется яма конечной глубины, то могут существовать стационарные состояния, описывающие заряженную частицу, захваченную в этой яме. Однако, если мы введем сколь угодно слабое внешнее однородное электрическое поле  $\vec{E}$ , которое соответствует электростатическому потенциалу  $U = -\vec{E}\vec{r}$ , то на достаточно больших расстояниях (вдоль вектора поля) потенциальная энергия будет меньше, чем в исходном минимуме. Это обстоятельство изменит характер движения частицы со связанного на несвязанное; соответственно, спектр гамильтониана изменится с дискретного на непрерывный. Это не означает, что мы не сможем найти возмущенные связанные состояния; однако они будут не истинными стационарными состояниями, а только квазистационарными («резонансами»), распадающимися вследствие «туннельного эффекта» с постоянной времени, зависящей от высоты и ширины возникшего потенциального барьера. (Похожая картина туннелирования через высокий барьер используется в теории радиоактивности для описания явления  $\alpha$ -распада.)

Если возмущение вызвано не внешним полем, которое можно изменять, то сама физическая задача определяет только полный возмущенный гамильтониан  $\hat{H}$ , а не невозмущенный  $\hat{H}^0$ . Поэтому разделение  $\hat{H}$  на невозмущенную часть  $\hat{H}^0$  и возмущение  $\hat{V}$  является более или менее произвольным. Обычно выбор невозмущенной задачи производится на основании соображений (математического) удобства. Разумеется, выбор этого разделения является ключевым с точки зрения свойств сходимости ряда возмущений. В этой связи заметим, что фактически ряд возмущений часто обладает только так называемой «асимптотической сходимостью», которая означает, что приближение к точному решению, полученное по теории возмущений, улучшается только до некоторого (конечного) порядка ТВ, в то время как ряд теории возмущений в итоге расходится. Однако, при построении формализма теории возмущений

эта возможность обычно не рассматривается; она, в сущности, основывается на молчаливом предположении, что все рассматриваемые ряды являются абсолютно сходящимися — например, что их можно перемножать почленно. (Дело в том, что ряды возмущений сходятся для достаточно малых значений  $|\lambda|$ ; однако радиус сходимости может включать или не включать настоящие «физические» значения  $\lambda$ .) Мы будем рассматривать эту проблему более детально только в связи с обсуждением размерной согласованности.

Можно заметить, что многие авторы не используют два разных обозначения ( $\hat{V}$  и  $\hat{W}$ ) для операторов возмущения, отличающихся множителем  $\lambda$ . Вместо этого они формально вводят множитель  $\lambda$  у оператора  $\hat{V}$  в (4.2), производят разложение по степеням  $\lambda$ , и подставляют  $\lambda = 1$  в полученные окончательные выражения. Эта процедура полностью эквивалентна нашей; однако использование пары разных операторов, связанных формулой (4.3), нам представляется более наглядным.

Подставляя выражения (4.5) и (4.6) в уравнение Шрёдингера (4.4), получаем

$$(\hat{H}^0 + \lambda \hat{W}) \sum_{l=0}^{\infty} \lambda^l \Psi_i^{(l)} = \sum_{k=0}^{\infty} \lambda^k E_i^{(k)} \sum_{l=0}^{\infty} \lambda^l \Psi_i^{(l)}. \quad (4.7)$$

В выражении (4.7) параметр  $\lambda$  рассматривается (по крайней мере, формально) как независимая непрерывная переменная; поэтому ее различные степени должны иметь равные коэффициенты в обеих частях равенства.

Рассматривая слагаемые с  $\lambda^0$  (т. е. те, которые не содержат  $\lambda$  и поэтому не обращаются в нуль, когда  $\lambda \rightarrow 0$ ), получаем

$$\hat{H}^0 \Psi_i^{(0)} = E_i^{(0)} \Psi_i^{(0)}, \quad (4.8)$$

т. е. мы вернулись к невозмущенной задаче:

$$\Psi_i^{(0)} = \Psi_i^0; \quad E_i^{(0)} = E_i^0 \quad (4.9)$$

как и следовало ожидать ( $\hat{H} \rightarrow \hat{H}^0$  если  $\lambda \rightarrow 0$ ).

Удобно выделить слагаемые нулевого порядка в выражениях (4.5) и (4.6), и написать

$$\Psi_i = \Psi_i^0 + \sum_{l=1}^{\infty} \lambda^l \Psi_i^{(l)} \quad (4.10)$$

и

$$E_i = E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)}. \quad (4.11)$$

Равным образом, можно написать

$$\begin{aligned}\Psi_i &= \Psi_i^0 + \sum_{l=1}^{\infty} \psi_i^{(l)} \\ E_i &= E_i^0 + \sum_{k=1}^{\infty} \varepsilon_i^{(k)}\end{aligned}\tag{4.12}$$

где

$$\psi_i^{(l)} = \lambda^l \Psi_i^{(l)} ; \quad \varepsilon_i^{(k)} = \lambda^k E_i^{(k)} .\tag{4.13}$$

(Как ранее замечено, можно обойтись без введения дополнительных обозначений, формально положив  $\lambda = 1$  *после того*, как все выкладки завершены.)

## 1.2. «Алгебраическое» разложение

Каждую функцию  $\Psi_i^{(l)}$  можно разложить по решениям  $\Psi_k^0$  невозмущенной задачи (4.1):

$$\Psi_i^{(l)} = \sum_{\substack{k \\ (k \neq i)}} c_{ik}^{(l)} \Psi_k^0 \quad (l \geq 1) .\tag{4.14}$$

Мы ввели ограничение  $k \neq i$  в сумму, показывающее, что нет необходимости включать  $\Psi_i^0$  в ряд разложения поправок к волновой функции  $\Psi_i^{(l)}$ . Это и в самом деле так, потому что если бы мы ее включили, ее можно было бы скомбинировать с нулевой (невозмущенной) волновой функцией, что в итоге повлияло бы только на нормировку, которую, однако, можно выбрать произвольно. Волновая функция, записанная через (4.10) и (4.14), имеет вид

$$\Psi_i = \Psi_i^0 + \chi_i ,\tag{4.15}$$

где

$$\langle \Psi_i | \Psi_i^0 \rangle = 1 ; \quad \langle \Psi_i^0 | \chi_i \rangle = 0 .\tag{4.16}$$

и потому не нормирована на единицу. Этот тип нормировки обычно называется «промежуточной» (или «корреляционной») нормировкой.

Подставляя разложения (4.10), (4.11), и (4.14) в возмущенное уравнение Шрёдингера  $(H^0 + \lambda \hat{W})\Psi_i = E_i \Psi_i$ , мы можем написать

$$\begin{aligned} (\hat{H}^0 + \lambda \hat{W}) \left[ \Psi_i^0 + \sum_{l=1}^{\infty} \lambda^l \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l)} \Psi_q^0 \right] \\ = \left[ E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} \right] \left[ \Psi_i^0 + \sum_{l=1}^{\infty} \lambda^l \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l)} \Psi_q^0 \right]. \end{aligned} \quad (4.17)$$

Умножим уравнения (4.17) на  $\Psi_i^{0*}$  и проинтегрируем, учитывая, что

$$\langle \Psi_p^0 | \hat{H}^0 | \Psi_q^0 \rangle = E_q^0 \delta_{pq}, \quad (4.18)$$

— в частности,  $\langle \Psi_i^0 | \hat{H}^0 | \Psi_q^0 \rangle = 0$  для  $q \neq i$ . Получаем

$$E_i^0 + \lambda \langle \Psi_i^0 | \hat{W} | \Psi_i^0 \rangle + \lambda \langle \Psi_i^0 | \hat{W} | \sum_{l=1}^{\infty} \lambda^l \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l)} \Psi_q^0 \rangle = E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)}. \quad (4.19)$$

Слагаемые нулевого порядка в обеих частях сокращаются, в соответствии с уравнением (4.8). Вводя обозначение

$$W_{ij} = \langle \Psi_i^0 | \hat{W} | \Psi_j^0 \rangle, \quad (4.20)$$

можем записать

$$\sum_{k=1}^{\infty} \lambda^k E_i^{(k)} = \lambda W_{ii} + \sum_{l=1}^{\infty} \lambda^{l+1} \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l)} W_{iq}. \quad (4.21)$$

В правой части мы можем ввести новый индекс суммирования  $k = l + 1$ , так что  $l = k - 1$ . Когда  $l$  меняется от 1 до  $\infty$ , то  $k$  меняется от 2 до  $\infty$ :

$$\sum_{k=1}^{\infty} \lambda^k E_i^{(k)} = \lambda W_{ii} + \sum_{k=2}^{\infty} \lambda^k \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(k-1)} W_{iq}. \quad (4.22)$$

Приравняем коэффициенты при равных степенях  $\lambda$  в обеих частях равенства. Для  $\lambda^1$  получим

$$E_i^{(1)} = W_{ii}, \quad (4.23)$$

а для  $\lambda^k$  ( $k \geq 2$ )

$$E_i^{(k)} = \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(k-1)} W_{iq} . \quad (4.24)$$

Это означает, что  $k$ -я поправка к энергии определяется  $(k-1)$ -й поправкой к волновой функции (см. также разд. 1.4);

Чтобы получить поправки к волновой функции, умножим уравнение (4.17) на одну из функций  $\Psi_j^{0*}$  ( $j \neq i$ ) и проинтегрируем:

$$\begin{aligned} \lambda W_{ji} + E_j^0 \sum_{l=1}^{\infty} \lambda^l c_{ij}^{(l)} + \sum_{l=1}^{\infty} \lambda^{l+1} \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l)} W_{jq} \\ = \left[ E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} \right] \sum_{l=1}^{\infty} \lambda^l c_{ij}^{(l)} . \end{aligned} \quad (4.25)$$

Вводя индекс суммирования  $n = k + l$  (он пробегает значения от 2 до  $\infty$ ) и заменяя, соответственно,  $l$  на  $n - k$ , получим

$$\begin{aligned} (E_j^0 - E_i^0) \sum_{l=1}^{\infty} \lambda^l c_{ij}^{(l)} = \sum_{n=2}^{\infty} \lambda^n \sum_{k=1}^{n-1} E_i^{(k)} c_{ij}^{(n-k)} - \\ - \lambda W_{ij} - \sum_{l=1}^{\infty} \lambda^{l+1} \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l)} W_{jq} , \end{aligned} \quad (4.26)$$

т. е. вводя индекс суммирования  $k = l + 1$  в последнем слагаемом, получим

$$\begin{aligned} \sum_{l=1}^{\infty} \lambda^l c_{ij}^{(l)} = \frac{1}{E_j^0 - E_i^0} \left[ \sum_{n=2}^{\infty} \lambda^n \sum_{k=1}^{n-1} E_i^{(k)} c_{ij}^{(n-k)} - \lambda W_{ji} \right. \\ \left. - \sum_{k=2}^{\infty} \lambda^k \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(k-1)} W_{jq} \right] . \end{aligned} \quad (4.27)$$

Сравнивая коэффициенты при  $\lambda^1$  в двух частях равенства, получаем

$$c_{ij}^{(1)} = - \frac{W_{ji}}{E_j^0 - E_i^0} . \quad (4.28)$$

Согласно этому результату, можно ожидать, что теория возмущений Рэлея—Шрёдингера применима в тех случаях, когда полученные отношения малы (строго говоря, необходимо рассматривать дроби, умноженные на  $\lambda$ , т. е. дроби  $\frac{V_{ij}}{E_j^0 - E_i^0}$ , содержащие матричные элементы «исходного» оператора возмущения  $\hat{V} = \lambda \hat{W}$ ). ЛАНБ®

Для любого значения  $l \geq 2$  сравнение коэффициентов при  $\lambda^l$  в двух частях равенства (4.27) дает

$$c_{ij}^{(l)} = \frac{1}{E_j^0 - E_i^0} \left[ \sum_{k=1}^{l-1} E_i^{(k)} c_{ij}^{(l-k)} - \sum_{\substack{q \\ (q \neq i)}} c_{iq}^{(l-1)} W_{jq} \right]. \quad (4.29)$$

Этот результат, вместе с формулой (4.24), указывает на то, что теорию возмущений Рэлея—Шрёдингера можно рассматривать, как рекуррентную процедуру (что позволяет явно рассчитать ряды возмущений вплоть до очень высоких порядков для некоторых относительно простых задач).

Используя формулы (4.24) и (4.28), мы получаем явное выражение для поправки второго порядка к энергии:

$$E_i^{(2)} = \sum_{\substack{j \\ (j \neq i)}} \left( -\frac{W_{ji}}{E_j^0 - E_i^0} \right) W_{ij} = - \sum_{\substack{j \\ (j \neq i)}} \frac{|W_{ij}|^2}{E_j^0 - E_i^0}. \quad (4.30)$$

Можно видеть, что поправка второго порядка к энергии основного состояния всегда отрицательна (так как в этом случае все  $E_j^0 > E_i^0$ ).

В приложении П8 приведены явные формулы для ТВ Рэлея—Шрёдингера вплоть до четвертого порядка для энергии и третьего порядка для волновой функции (см. также и формулы в разд. 1.4).

Важно отметить, что разложение конечного порядка по теории возмущений не дает оценки точной энергии ни сверху, ни снизу. Частичные суммы рядов теории возмущений часто (но не обязательно) попеременно оказываются то выше, то ниже точного значения. К тому же, сумма поправок к энергии первого и второго порядков часто «дает перелет», т. е. приводит к значению энергии ниже точного, хотя это и не обязательно случается. Очевидно, можно получить оценку энергии сверху, если вычислить среднее значение гамильтониана с волновой функцией, найденной по теории возмущений. В разд. 1.4 мы покажем, что точность этого среднего значения намного больше, чем самой волновой функции теории возмущений. В самом деле, волновая функция до  $n$ -го порядка



содержит ошибку, пропорциональную  $\varepsilon = \lambda^{n+1}$ , поэтому ошибка в среднем значении энергии должна быть пропорциональна  $\varepsilon^2 = \lambda^{2n+2}$ . Это и есть краткая формулировка **2n + 1**-теоремы Вигнера, более детальное рассмотрение которой будет представлено в разд. 1.4.

### Компактные матричные обозначения

Предшествующие результаты можно представить очень компактно в матричных обозначениях. Введем единичный вектор  $\mathbf{e}^i$ , все элементы которого равны нулю, за исключением  $i$ -го, который равен 1, и векторы  $\mathbf{c}^{i(l)}$ , элементы которых совпадают с коэффициентами разложения  $c_{ij}^{(l)}$  для  $j \neq i$  и равны 0 для  $j = i$ . Далее, определим диагональную матрицу  $\Delta^i$ , с элементами  $\frac{1}{E_j - E_i}$  в  $j$ -х диагональных положениях, за исключением  $i$ -го, в котором стоит нуль. Легко проверить, что в этих обозначениях выражения (4.24), (4.28) и (4.29) можно компактно переписать:

$$\begin{aligned} E_i^{(k)} &= \mathbf{e}^{i\dagger} \mathbf{W} \mathbf{c}^{i(k-1)} \\ \mathbf{c}^{i(1)} &= -\Delta^i \mathbf{W} \mathbf{e}^i \\ \mathbf{c}^{i(l)} &= \Delta^i \left[ \sum_{k=1}^{l-1} E_i^{(k)} \mathbf{c}^{i(l-k)} - \mathbf{W} \mathbf{c}^{i(l-1)} \right] \end{aligned} \quad (4.31)$$

## 1.3. Использование приведенной резольвенты в теории возмущений Рэлея—Шрёдингера

### Приведенная резольвента

Определим разложение единицы в пространстве функций  $|\Psi_k^0\rangle$  как

$$\hat{P} + \hat{Q} = 1 \quad (4.32)$$

где

$$\hat{P} = |\Psi_i^0\rangle \langle \Psi_i^0| \quad (4.33)$$

и

$$\hat{Q} = 1 - \hat{P} = \sum_{\substack{k \\ (k \neq i)}} |\Psi_k^0\rangle \langle \Psi_k^0|. \quad (4.34)$$

Здесь  $|\Psi_i^0\rangle$  является  $i$ -м решением невозмущенной задачи  $\hat{H}^0 |\Psi_i^0\rangle = E_i^0 |\Psi_i^0\rangle$ , возмущение которой мы рассматриваем. Предположим, что эта задача невырождена. Выпишем спектральное разложение оператора  $\hat{H}^0 - E_i^0$  (более точно, оператора  $\hat{H}^0 - E_i^0 \hat{1}$ ) в терминах собственных векторов  $|\Psi_l^0\rangle$  оператора  $\hat{H}^0$ :

$$\hat{H}^0 - E_i^0 = \sum_l (E_l^0 - E_i^0) |\Psi_l^0\rangle \langle \Psi_l^0|. \quad (4.35)$$

Определим оператор  $\hat{R}^0$  его спектральным разложением:

$$\hat{R}^0 = \sum_{\substack{k \\ (k \neq i)}} \frac{1}{E_k^0 - E_i^0} |\Psi_k^0\rangle \langle \Psi_k^0|. \quad (4.36)$$

Ограничение на индекс суммирования гарантирует, что знаменатель не обращается в нуль ( $E_i^0$  предполагался невырожденным.)

Найдем произведение операторов  $\hat{R}^0(\hat{H}^0 - E_i^0)$ , используя определения (4.35) и (4.36):

$$\begin{aligned} \hat{R}^0(\hat{H}^0 - E_i^0) &= \sum_{\substack{k \\ (k \neq i)}} \frac{1}{E_k^0 - E_i^0} |\Psi_k^0\rangle \langle \Psi_k^0| \sum_l (E_l^0 - E_i^0) |\Psi_l^0\rangle \langle \Psi_l^0| \\ &= \sum_{\substack{k \\ (k \neq i)}} \sum_l \frac{E_l^0 - E_i^0}{E_k^0 - E_i^0} |\Psi_k^0\rangle \underbrace{\langle \Psi_k^0 | \Psi_l^0 \rangle}_{\delta_{kl}} \langle \Psi_l^0| = \sum_{\substack{k \\ (k \neq i)}} |\Psi_k^0\rangle \langle \Psi_k^0| = \hat{Q}. \end{aligned} \quad (4.37)$$

Этот расчет показывает, что в подпространстве, ортогональном к  $|\Psi_i^0\rangle$ , оператор  $\hat{R}^0$  ведет себя, как обратный к оператору  $\hat{H}^0 - E_i^0$ ; отсюда его название: «приведенная резольвента».

### Поправки к волновой функции

Рассмотрим разложения волновой функции и энергии по теории возмущений

$$\begin{aligned} |\Psi_i\rangle &= |\Psi_i^0\rangle + \sum_{j=1}^{\infty} \lambda^j |\Psi_i^{(j)}\rangle \\ E_i &= E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} \end{aligned} \quad (4.38)$$

и используем промежуточную нормировку

$$\begin{aligned} \langle \Psi_i^0 | \Psi_i^0 \rangle &= 1 \\ \langle \Psi_i^0 | \Psi_i^{(j)} \rangle &= 0 \quad j \geq 1, \end{aligned} \quad (4.39)$$

где  $|\Psi_i^0\rangle$  есть решение невозмущенной задачи, т. е.  $\hat{H}^0 |\Psi_i^0\rangle = E_i^0 |\Psi_i^0\rangle$ ;  $\langle \Psi_i^0 | \hat{H}^0 = E_i^0 \langle \Psi_i^0 |$ .

Использование промежуточной нормировки (4.39) приводит к тождествам

$$\langle \Psi_i^0 | \hat{H}^0 | \Psi_i^0 \rangle = E_i^0 \quad (4.40)$$

и

$$\langle \Psi_i^0 | \hat{H}^0 | \Psi_i^{(j)} \rangle = 0. \quad (4.41)$$

Подставим разложения (4.38) в уравнение Шрёдингера:

$$\begin{aligned} (\hat{H}^0 + \lambda \hat{W}) \left[ |\Psi_i^0\rangle + \sum_{j=1}^{\infty} \lambda^j |\Psi_i^{(j)}\rangle \right] \\ = \left[ E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} \right] \left[ |\Psi_i^0\rangle + \sum_{l=1}^{\infty} \lambda^l |\Psi_i^{(l)}\rangle \right] \end{aligned} \quad (4.42)$$

и перегруппируем слагаемые как

$$\begin{aligned} (\hat{H}^0 - E_i^0) \left[ |\Psi_i^0\rangle + \sum_{j=1}^{\infty} \lambda^j |\Psi_i^{(j)}\rangle \right] \\ = \left[ -\lambda \hat{W} + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} \right] \left[ |\Psi_i^0\rangle + \sum_{l=1}^{\infty} \lambda^l |\Psi_i^{(l)}\rangle \right]. \end{aligned} \quad (4.43)$$

Первое слагаемое с левой стороны уничтожается, так как  $(\hat{H}^0 - E_i^0)|\Psi_i^0\rangle = 0$ . С правой стороны введем новый индекс суммирования  $j = l + 1$  (тогда  $l = j - 1$ ) в слагаемом, содержащем произведение  $-\lambda \hat{W}$  с суммой по  $l$ , тогда как в слагаемом, содержащем произведение двух сумм, перейдем к индексам суммирования  $j = k + l$  и  $k$ , вместо  $l$  и  $k$ , пробегающим свои значения независимо (тогда  $l = j - k$ ). Получаем

$$\begin{aligned} (\hat{H}^0 - E_i^0) \sum_{j=1}^{\infty} \lambda^j |\Psi_i^{(j)}\rangle = -\lambda \hat{W} |\Psi_i^0\rangle - \sum_{j=2}^{\infty} \lambda^j \hat{W} |\Psi_i^{(j-1)}\rangle \\ + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} |\Psi_i^0\rangle + \sum_{j=2}^{\infty} \sum_{k=1}^{j-1} \lambda^j E_i^{(k)} |\Psi_i^{(j-k)}\rangle. \end{aligned} \quad (4.44)$$

Подействуем оператором  $\hat{R}^0$  на обе части равенства (4.44) и воспользуемся соотношением (4.37), а также тем фактом, что  $\hat{R}^0 |\Psi_i^0\rangle = 0$ . Также из условия промежуточной нормировки (4.39) следует, что  $\hat{Q} |\Psi_i^{(j)}\rangle = |\Psi_i^{(j)}\rangle$  для любого  $j \geq 1$ . Таким образом, находим, что

$$\begin{aligned} \sum_{j=1}^{\infty} \lambda^j |\Psi_i^{(j)}\rangle = -\lambda \hat{R}^0 \hat{W} |\Psi_i^0\rangle - \sum_{j=2}^{\infty} \lambda^j \hat{R}^0 \hat{W} |\Psi_i^{(j-1)}\rangle \\ + \sum_{j=2}^{\infty} \sum_{k=1}^{j-1} \lambda^j E_i^{(k)} \hat{R}^0 |\Psi_i^{(j-k)}\rangle. \end{aligned} \quad (4.45)$$

Теперь приравняем коэффициенты при равных степенях  $\lambda$  в двух частях равенства (4.44). Сравнивая коэффициенты при  $\lambda^1$ , получаем

$$|\Psi_i^{(1)}\rangle = -\hat{R}^0 \hat{W} |\Psi_i^0\rangle. \quad (4.46)$$

Для  $\lambda^j$  ( $j \geq 2$ ) имеем:

$$|\Psi_i^{(j)}\rangle = -\hat{R}^0 \hat{W} |\Psi_i^{(j-1)}\rangle + \sum_{k=1}^{j-1} E_i^{(k)} \hat{R}^0 |\Psi_i^{(j-k)}\rangle. \quad (4.47)$$

Очевидна близкая аналогия между этими выражениями и выражениями, выписанными в матричных обозначениях в конце разд. 1.2; то же самое имеет место для поправок к энергии, обсуждаемых ниже.

Подставляя определение (4.36) резольвенты  $\hat{R}^0$  в выражения (4.46) и (4.47), легко получаем выражение для коэффициентов  $c_{ip}^{(j)}$  разложений  $|\Psi_i^{(j)}\rangle = \sum_{\substack{p \\ (p \neq i)}} c_{ip}^{(j)} |\Psi_p^0\rangle$  поправок  $|\Psi_i^{(j)}\rangle$  разных порядков ТВ к волновой функции через невозмущенные функции. В самом деле,

$$|\Psi_i^{(1)}\rangle = -\hat{R}^0 \hat{W} |\Psi_i^0\rangle = - \sum_{\substack{j \\ (j \neq i)}} \frac{|\Psi_j^0\rangle \langle \Psi_j^0 | \hat{W} | \Psi_i^0\rangle}{E_j^0 - E_i^0} = \sum_j \frac{-W_{ji}}{E_j^0 - E_i^0} |\Psi_j^0\rangle, \quad (4.48)$$

что является хорошо известным выражением. Подобным же образом, для  $j \geq 2$  имеем

$$\begin{aligned} |\Psi_i^{(j)}\rangle &= \sum_{\substack{p \\ (p \neq i)}} c_{ip}^{(j)} |\Psi_p^0\rangle = -\hat{R}^0 \hat{W} |\Psi_i^{(j-1)}\rangle + \sum_{k=1}^{j-1} E_i^{(k)} \hat{R}^0 |\Psi_i^{(j-k)}\rangle \\ &= - \sum_{\substack{p \\ (p \neq i)}} \frac{|\Psi_p^0\rangle \langle \Psi_p^0 |}{E_p^0 - E_i^0} \hat{W} \sum_l c_{il}^{(j-1)} |\Psi_l^0\rangle \\ &\quad + \sum_{k=1}^{j-1} E_i^{(k)} \sum_{\substack{p \\ (p \neq i)}} \frac{|\Psi_p^0\rangle \langle \Psi_p^0 |}{E_p^0 - E_i^0} \sum_{\substack{l \\ (l \neq i)}} c_{il}^{(j-k)} |\Psi_l^0\rangle \end{aligned} \quad (4.49)$$

$$\begin{aligned}
 &= - \sum_p \sum_l |\Psi_p^0\rangle \frac{W_{pl} c_{il}^{(j-1)}}{E_p^0 - E_i^0} \\
 &\quad (p \neq i) \quad (l \neq i) \\
 &\quad + \sum_p \sum_l |\Psi_p^0\rangle \delta_{pl} \frac{1}{E_p^0 - E_i^0} \sum_{k=1}^{j-1} E_i^{(k)} c_{il}^{(j-k)} \\
 &= \sum_p |\Psi_p^0\rangle \frac{1}{E_p^0 - E_i^0} \left[ - \sum_l W_{pl} c_{il}^{(j-1)} + \sum_{k=1}^{j-1} E_i^{(k)} c_{ip}^{(j-k)} \right].
 \end{aligned}$$

Таким образом, мы воспроизвели выражение для  $c_{ip}^{(j)}$ , выведенное в предыдущем разделе «алгебраическим» способом.

### Поправки к энергии

Умножим уравнения (4.42) на  $\langle \Psi_i^0 |$  и учтем соотношения (4.40) и (4.41):

$$E_i^0 + \lambda \langle \Psi_i^0 | \hat{W} | \Psi_i^0 \rangle + \sum_{j=1}^{\infty} \lambda^{j+1} \langle \Psi_i^0 | \hat{W} | \Psi_i^{(j)} \rangle = E_i^0 + \sum_{k=1}^{\infty} \lambda^k E_i^{(k)} \quad (4.50)$$

Под знаком суммы в левой части перейдем к индексу суммирования  $k = j + 1$  (так что  $j = k - 1$ ):

$$\lambda \langle \Psi_i^0 | \hat{W} | \Psi_i^0 \rangle + \sum_{k=2}^{\infty} \lambda^k \langle \Psi_i^0 | \hat{W} | \Psi_i^{(k-1)} \rangle = \sum_{k=1}^{\infty} \lambda^k E_i^{(k)}. \quad (4.51)$$

Приравнявая коэффициенты при одинаковых степенях  $\lambda$  в двух частях равенства, получаем для  $\lambda^1$  хорошо известный результат:

$$E_i^{(1)} = \langle \Psi_i^0 | \hat{W} | \Psi_i^0 \rangle = W_{ii}. \quad (4.52)$$

Для  $\lambda^k$  ( $k \geq 2$ ) результатом будет

$$E_i^{(k)} = \langle \Psi_i^0 | \hat{W} | \Psi_i^{(k-1)} \rangle. \quad (4.53)$$

Подставляя сюда разложение  $|\Psi_i^{(k-1)}\rangle$  по функциям нулевого порядка, мы снова получаем «алгебраический» результат (4.24):

$$E_i^{(k)} = \langle \Psi_i^0 | \hat{W} | \sum_p c_{ip}^{(k-1)} \Psi_p^0 \rangle = \sum_p c_{ip}^{(k-1)} W_{ip}. \quad (4.54)$$

### Инвариантность

В соответствии с предшествующими выводами, поправки к волновой функции и энергии, найденные по теории возмущений, могут быть получены с использованием двух операторов: оператора возмущения  $\hat{W}$  и приведенной резольвенты  $\hat{R}^0$ . Это позволяет легко доказать, что результаты, получаемые по теории возмущений Рэлея—Шрёдингера, инвариантны относительно унитарных преобразований вырожденных решений нулевого порядка  $|\Psi_k^0\rangle$ , если таковые существуют. (Однако мы должны отметить, что  $i$ -е решение  $|\Psi_i^0\rangle$ , возмущение которого мы рассматриваем, предполагается невырожденным, так что вырожденными могут быть только некоторые другие функции нулевого порядка. Случай вырождения невозмущенной волновой функции будет рассмотрен ниже в разд. 3.)

Оператор  $\hat{W}$  не зависит от преобразований базисных функций, так что нам надо рассмотреть только приведенную резольвенту  $\hat{R}^0$ . Рассмотрев ее определение, легко заключить, что этот оператор также является инвариантным. В самом деле, предположим, что существует  $m$  решений нулевого порядка  $\Psi_l^0, \Psi_{l+1}^0, \dots, \Psi_{l+m-1}^0$ , имеющих одинаковую энергию  $E_l^0$ . Тогда эти  $m$  состояний дают следующий вклад в приведенную резольвенту (4.36):

$$\sum_{k=l}^{l+m-1} \frac{1}{E_l^0 - E_i^0} |\Psi_k^0\rangle \langle \Psi_k^0| = \frac{1}{E_l^0 - E_i^0} \sum_{k=l}^{l+m-1} |\Psi_k^0\rangle \langle \Psi_k^0|. \quad (4.55)$$

В правой части стоит проектор на подпространство вырожденных состояний; он инвариантен относительно унитарных преобразований растягивающих его функций (см. приложение П7).

## 1.4. $2n + 1$ -Теорема Вигнера

### 1.4.1. Точность среднего значения энергии, вычисленного с волновой функцией $n$ -го порядка теории возмущений

*Теорема:*

Среднее значение энергии, вычисленное с волновой функцией  $n$ -го порядка по возмущению является точным, вплоть до вкладов порядка  $2n + 1$  включительно (отличие от точной энергии имеет порядок  $2n + 2$  и выше).

*Доказательство:*

Точная волновая функция основного состояния  $\Psi_0$  может быть записана как

$$\Psi_0 = \Psi_0^{[n]} + \chi^{(n+1)}, \quad (4.56)$$

где

$$\Psi_0^{[n]} = \sum_{l=0}^n \lambda^l \Psi_0^{(l)} \quad (4.57)$$

является возмущенной волновой функцией, рассчитанной до  $n$ -го порядка теории возмущений включительно, тогда как  $\chi^{(n+1)}$  — это соответствующая ошибка:

$$\chi^{(n+1)} = \sum_{k=n+1}^{\infty} \lambda^k \Psi_0^{(k)} \sim \mathcal{O}(\lambda^{n+1}). \quad (4.58)$$

Можно переписать выражение (4.56) в виде

$$\Psi_0^{[n]} = \Psi_0 - \chi^{(n+1)}. \quad (4.59)$$

Поэтому среднее значение энергии, вычисленное с  $\Psi_0^{[n]}$ , равно

$$E = \frac{\langle \Psi_0^{[n]} | \hat{H} | \Psi_0^{[n]} \rangle}{\langle \Psi_0^{[n]} | \Psi_0^{[n]} \rangle} = \frac{\langle \Psi_0 - \chi^{(n+1)} | \hat{H} | \Psi_0 - \chi^{(n+1)} \rangle}{\langle \Psi_0 - \chi^{(n+1)} | \Psi_0 - \chi^{(n+1)} \rangle} \quad (4.60)$$

Здесь  $\Psi_0$  — точная волновая функция, т. е.

$$\hat{H}\Psi_0 = E_0\Psi_0, \quad (4.61)$$

где  $E_0$  — точная энергия. Из выражений (4.58) и (4.61) также имеем:

$$\begin{aligned} \langle \chi^{(n+1)} | \hat{H} | \Psi_0 \rangle &= E_0 \langle \chi^{(n+1)} | \Psi_0 \rangle \\ \langle \Psi_0 | \hat{H} | \chi^{(n+1)} \rangle &= E_0 \langle \Psi_0 | \chi^{(n+1)} \rangle \\ \langle \chi^{(n+1)} | \hat{H} | \chi^{(n+1)} \rangle &= \mathcal{O}(2n+2) \\ \langle \chi^{(n+1)} | \chi^{(n+1)} \rangle &= \mathcal{O}'(2n+2) \end{aligned} \quad (4.62)$$

где  $\mathcal{O}(2n+2)$ ,  $\mathcal{O}'(2n+2)$ , и т. д. обозначают величины порядка  $\lambda^{2n+2}$ . (Нет необходимости предполагать ортогональность  $\chi^{(n+1)}$  к  $\Psi_0$ .)

Подставляя результаты (4.61) и (4.62) в выражение (4.60), получаем:

$$\begin{aligned} E &= \frac{E_0(\langle \Psi_0 | \Psi_0 \rangle - \langle \chi^{(n+1)} | \Psi_0 \rangle - \langle \Psi_0 | \chi^{(n+1)} \rangle) + \mathcal{O}(2n+2)}{\langle \Psi_0 | \Psi_0 \rangle - \langle \chi^{(n+1)} | \Psi_0 \rangle - \langle \Psi_0 | \chi^{(n+1)} \rangle + \mathcal{O}'(2n+2)} \\ &= E_0 + \mathcal{O}''(2n+2). \text{ Ч. т. д.} \end{aligned} \quad (4.63)$$

#### 1.4.2. Вычисление точных поправок к энергии вплоть до порядка $2n + 1$ с использованием первых $n$ поправок к волновой функции

Пусть, как обычно,

$$\hat{H} = \hat{H}^0 + \lambda \hat{W} \quad (4.64)$$

$$\Psi_i = \Psi_i^0 + \sum_{l=1}^{\infty} \lambda^l \Psi_i^{(l)} \quad (4.65)$$

$$E_i = E_i^0 + \sum_{j=1}^{\infty} \lambda^j E_i^{(j)}, \quad (4.66)$$

где  $\hat{H}^0 \Psi_i^0 = E_i^0 \Psi_i^0$ , и используется промежуточная нормировка  $\langle \Psi_i^0 | \Psi_i^0 \rangle = 1$ ;  $\langle \Psi_i^0 | \Psi_i^{(l)} \rangle = 0$  ( $l \geq 1$ ).

Подставим выражения (4.64)-(4.66) в уравнение Шрёдингера  $\hat{H} \Psi_i = E_i \Psi_i$ . Так как в последующих выкладках каждое выражение относится к  $i$ -му уровню, для простоты опустим индексы  $i$ . Имеем:

$$(\hat{H}^0 + \lambda \hat{W}) \left[ \Psi^0 + \sum_{l=1}^{\infty} \lambda^l \Psi^{(l)} \right] = \left[ E^0 + \sum_{k=1}^{\infty} \lambda^k E^{(k)} \right] \left[ \Psi^0 + \sum_{j=1}^{\infty} \lambda^j \Psi^{(j)} \right]. \quad (4.67)$$

Соберем коэффициенты при  $\lambda^p$  ( $p \geq 1$ ) в обеих частях выражения (4.67). В произведении двух сумм справа мы должны учесть все слагаемые, для которых  $k + j = p$ . Поэтому можно записать  $k = p - j$  и просуммировать от  $j = 1$  до  $j = p - 1$ . Получаем:

$$\hat{H}^0 \Psi^{(p)} + \hat{W} \Psi^{(p-1)} = E^{(p)} \Psi^0 + E^0 \Psi^{(p)} + \sum_{j=1}^{p-1} E^{(p-j)} \Psi^{(j)}. \quad (4.68)$$

(Очевидно, последняя сумма пропадает, если  $p = 1$ .)

Умножая равенство (4.68) на  $\Psi^{0*}$  и интегрируя, мы получаем уже известный результат (4.53):

$$E^{(p)} = \langle \Psi^0 | \hat{W} | \Psi^{(p-1)} \rangle. \quad (4.69)$$

Теперь покажем, что можно понизить порядок  $p - 1$  в «кет»-скобке в правой части формулы (4.69) за счет использования некоторых поправок к волновой функции в «бра»-скобке, а также определенных дополнительных слагаемых, содержащих поправки низших порядков к волновой функции и энергии.



Перенесем  $E^0\Psi^{(p)}$  в левую часть выражения (4.68):

$$(\hat{H}^0 - E^0)\Psi^{(p)} + \hat{W}\Psi^{(p-1)} = E^{(p)}\Psi^0 + \sum_{j=1}^{p-1} E^{(p-j)}\Psi^{(j)}. \quad (4.70)$$

Это общий результат, верный для любого  $p \geq 1$ . (В случае  $p = 1$  последняя сумма должна быть отброшена.)

Умножим соотношение (4.70) на  $\Psi^{(1)*}$  и проинтегрируем (с учетом  $\langle \Psi^{(1)} | \Psi^0 \rangle = 0$ ):

$$\langle \Psi^{(1)} | \hat{H}^0 - E^0 | \Psi^{(p)} \rangle + \langle \Psi^{(1)} | \hat{W} | \Psi^{(p-1)} \rangle = \sum_{j=1}^{p-1} E^{(p-j)} \langle \Psi^{(1)} | \Psi^{(j)} \rangle. \quad (4.71)$$

Выпишем выражение (4.70) и для  $p = 1$  (сумму, как отмечено раньше, отбрасываем):

$$(\hat{H}^0 - E^0)\Psi^{(1)} + \hat{W}\Psi^0 = E^{(1)}\Psi^0, \quad (4.72)$$

умножим на  $\Psi^{(p)*}$  и проинтегрируем ( $\langle \Psi^{(p)} | \Psi^0 \rangle = 0$ ):

$$\langle \Psi^{(p)} | \hat{H}^0 - E^0 | \Psi^{(1)} \rangle + \langle \Psi^{(p)} | \hat{W} | \Psi^0 \rangle = 0. \quad (4.73)$$

Перейдем к комплексно сопряженному выражению и перегруппируем слагаемые:

$$\langle \Psi^{(1)} | \hat{H}^0 - E^0 | \Psi^{(p)} \rangle = -\langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle. \quad (4.74)$$

Подставляя равенство (4.74) в соотношение (4.71), получим

$$-\langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle + \langle \Psi^{(1)} | \hat{W} | \Psi^{(p-1)} \rangle = \sum_{j=1}^{p-1} E^{(p-j)} \langle \Psi^{(1)} | \Psi^{(j)} \rangle, \quad (4.75)$$

т. е. учитывая результат (4.69),

$$E^{(p+1)} = \langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle = \langle \Psi^{(1)} | \hat{W} | \Psi^{(p-1)} \rangle - \sum_{j=1}^{p-1} E^{(p-j)} \langle \Psi^{(1)} | \Psi^{(j)} \rangle \quad (4.76)$$

Выражение (4.76) показывает, что для вычисления  $E^{(p+1)}$  нам больше не нужна  $\Psi^{(p)}$ : поправка наивысшего порядка к волновой функции в правой части выражения (4.76) равна  $\Psi^{(p-1)}$ .

Можно повторить эту процедуру. Запишем соотношение (4.70) для  $p-1$  вместо  $p$  (так как при выводе выражения (4.70) предполагалось, что  $p \geq 1$ , то здесь мы должны положить  $p \geq 2$ ).

$$(\hat{H}^0 - E^0)\Psi^{(p-1)} + \hat{W}\Psi^{(p-2)} = E^{(p-1)}\Psi^0 + \sum_{j=1}^{p-2} E^{(p-1-j)}\Psi^{(j)}. \quad (4.77)$$

Умножая соотношение (4.77) на  $\Psi^{(2)*}$  и интегрируя, получим (с учетом того, что  $\langle \Psi^{(2)} | \Psi^0 \rangle = 0$ ):

$$\langle \Psi^{(2)} | \hat{H}^0 - E^0 | \Psi^{(p-1)} \rangle + \langle \Psi^{(2)} | \hat{W} | \Psi^{(p-2)} \rangle = \sum_{j=1}^{p-2} E^{(p-1-j)} \langle \Psi^{(2)} | \Psi^{(j)} \rangle. \quad (4.78)$$

Выпишем также частный вид выражения (4.70) для  $p = 2$ :

$$(\hat{H}^0 - E^0)\Psi^{(2)} + \hat{W}\Psi^{(1)} = E^{(2)}\Psi^0 + E^{(1)}\Psi^{(1)}, \quad (4.79)$$

умножим на  $\Psi^{(p-1)*}$  и проинтегрируем ( $\langle \Psi^{(p-1)} | \Psi^0 \rangle = 0$ , так как  $p \geq 2$ ). Получим

$$\langle \Psi^{(p-1)} | \hat{H}^0 - E^0 | \Psi^{(2)} \rangle + \langle \Psi^{(p-1)} | \hat{W} | \Psi^{(1)} \rangle = E^{(1)} \langle \Psi^{(p-1)} | \Psi^{(1)} \rangle. \quad (4.80)$$

После перехода к комплексно сопряженному и перегруппировки слагаемых из равенства (4.80) получаем

$$\langle \Psi^{(2)} | \hat{H}^0 - E^0 | \Psi^{(p-1)} \rangle = -\langle \Psi^{(1)} | \hat{W} | \Psi^{(p-1)} \rangle + E^{(1)} \langle \Psi^{(1)} | \Psi^{(p-1)} \rangle. \quad (4.81)$$

Подставляя выражение (4.81) в формулу (4.78) и перегруппировывая, получим

$$\begin{aligned} \langle \Psi^{(1)} | \hat{W} | \Psi^{(p-1)} \rangle &= \langle \Psi^{(2)} | \hat{W} | \Psi^{(p-2)} \rangle + E^{(1)} \langle \Psi^{(1)} | \Psi^{(p-1)} \rangle \\ &\quad - \sum_{j=1}^{p-2} E^{(p-1-j)} \langle \Psi^{(2)} | \Psi^{(j)} \rangle. \end{aligned} \quad (4.82)$$

Подставим равенство (4.82) в правую часть формулы (4.76):

$$\begin{aligned} E^{(p+1)} \langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle &= \langle \Psi^{(2)} | \hat{W} | \Psi^{(p-2)} \rangle + E^{(1)} \langle \Psi^{(1)} | \Psi^{(p-1)} \rangle \\ &\quad - \sum_{j=1}^{p-2} E^{(p-1-j)} \langle \Psi^{(2)} | \Psi^{(j)} \rangle - \sum_{j=1}^{p-1} E^{(p-j)} \langle \Psi^{(1)} | \Psi^{(j)} \rangle. \end{aligned} \quad (4.83)$$

В последней сумме выделим слагаемое с  $j = p - 1$ , которое сокращается со вторым слагаемым в правой части, и объединим остающиеся две суммы в одну:

$$\langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle = \langle \Psi^{(2)} | \hat{W} | \Psi^{(p-2)} \rangle - \sum_{j=1}^{p-2} \sum_{l=1}^2 E^{(p+1-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle. \quad (4.84)$$

Сопоставляя выражения (4.76) и (4.84), мы можем сформулировать следующее

*Предположение:*

$$\langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle = \langle \Psi^{(k)} | \hat{W} | \Psi^{(p-k)} \rangle - \sum_{j=1}^{p-k} \sum_{l=1}^k E^{(p+1-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle \quad (4.85)$$

для любого  $k$  в интервале  $1 < k < p$ .

*Доказательство по индукции*

Предположим, что равенство (4.85) выполняется для некоторого значения  $k$ , и повторим еще раз выкладки, уже сделанные нами дважды.

Выпишем равенство (4.70), заменив индекс  $p$  на  $p - k$  (предположим  $p > k$ ;  $p - k > 0$ ). Получим

$$(\hat{H}^0 - E^0) \Psi^{(p-k)} + \hat{W} \Psi^{(p-k-1)} = E^{(p-k)} \Psi^0 + \sum_{j=1}^{p-k-1} E^{(p-k-j)} \Psi^{(j)}. \quad (4.86)$$

Умножим на  $\Psi^{(k+1)*}$  и проинтегрируем ( $\langle \Psi^{(k+1)} | \Psi^0 \rangle = 0$ ):

$$\begin{aligned} \langle \Psi^{(k+1)} | \hat{H}^0 - E^0 | \Psi^{(p-k)} \rangle + \langle \Psi^{(k+1)} | \hat{W} | \Psi^{(p-k-1)} \rangle \\ = \sum_{j=1}^{p-k-1} E^{(p-k-j)} \langle \Psi^{(k+1)} | \Psi^{(j)} \rangle. \end{aligned} \quad (4.87)$$

Выпишем равенство (4.70) также для случая  $p = k + 1$ :

$$(\hat{H}^0 - E^0) \Psi^{(k+1)} + \hat{W} \Psi^{(k)} = E^{(k+1)} \Psi^0 + \sum_{j=1}^k E^{(k+1-j)} \Psi^{(j)}. \quad (4.88)$$

Умножим на  $\Psi^{(p-k)*}$  и проинтегрируем ( $\langle \Psi^{(p-k)} | \Psi^0 \rangle = 0$ , так как  $p - k > 0$ ):

$$\langle \Psi^{(p-k)} | \hat{H}^0 - E^0 | \Psi^{(k+1)} \rangle + \langle \Psi^{(p-k)} | \hat{W} | \Psi^{(k)} \rangle = \sum_{j=1}^k E^{(k+1-j)} \langle \Psi^{(p-k)} | \Psi^{(j)} \rangle. \quad (4.89)$$

Перейдем к выражению, комплексно сопряженному к выражению (4.89):

$$\langle \Psi^{(k+1)} | \hat{H}^0 - E^0 | \Psi^{(p-k)} \rangle + \langle \Psi^{(k)} | \hat{W} | \Psi^{(p-k)} \rangle = \sum_{j=1}^k E^{(k+1-j)} \langle \Psi^{(j)} | \Psi^{(p-k)} \rangle, \quad (4.90)$$

вычтем равенство (4.87) из равенства (4.90) и перенесем одно из слагаемых в правую часть:

$$\begin{aligned} \langle \Psi^{(k)} | \hat{W} | \Psi^{(p-k)} \rangle &= \langle \Psi^{(k+1)} | \hat{W} | \Psi^{(p-k-1)} \rangle + \sum_{j=1}^k E^{(k+1-j)} \langle \Psi^{(j)} | \Psi^{(p-k)} \rangle \\ &\quad - \sum_{j=1}^{p-k-1} E^{(p-k-j)} \langle \Psi^{(k+1)} | \Psi^{(j)} \rangle. \end{aligned} \quad (4.91)$$

Подставим это выражение для  $\langle \Psi^{(k)} | \hat{W} | \Psi^{(p-k)} \rangle$  в соотношение (4.85):

$$\begin{aligned} \langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle &= \langle \Psi^{(k+1)} | \hat{W} | \Psi^{(p-k-1)} \rangle + \sum_{j=1}^k E^{(k+1-j)} \langle \Psi^{(j)} | \Psi^{(p-k)} \rangle \\ &\quad - \sum_{j=1}^{p-k-1} E^{(p-k-j)} \langle \Psi^{(k+1)} | \Psi^{(j)} \rangle - \sum_{j=1}^{p-k} \sum_{l=1}^k E^{(p+1-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle. \end{aligned} \quad (4.92)$$

Выделим в двойной сумме слагаемое с  $j = p - k$ , соответствующее верхнему пределу первого суммирования  $[p + 1 - l - (p - k) = k + 1 - l]$ :

$$\begin{aligned} \langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle &= \langle \Psi^{(k+1)} | \hat{W} | \Psi^{(p-k-1)} \rangle + \sum_{j=1}^k E^{(k+1-j)} \langle \Psi^{(j)} | \Psi^{(p-k)} \rangle \\ &\quad - \sum_{j=1}^{p-k-1} E^{(p-k-j)} \langle \Psi^{(k+1)} | \Psi^{(j)} \rangle - \sum_{j=1}^{p-k-1} \sum_{l=1}^k E^{(p+1-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle \\ &\quad - \sum_{l=1}^k E^{(k+1-l)} \langle \Psi^{(l)} | \Psi^{(p-k)} \rangle. \end{aligned} \quad (4.93)$$

Первая и последняя суммы в правой части сокращаются (они отличаются только обозначениями индекса суммирования); остающиеся две суммы можно объединить в одну: первая дает лишь одно дополнитель-

ное слагаемое  $l = k + 1$  в последнюю  $[p + 1 - (k + 1) - j = p - k - j]$ .  
Получаем

$$\begin{aligned} \langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle &= \langle \Psi^{(k+1)} | \hat{W} | \Psi^{[p-(k+1)]} \rangle \\ &- \sum_{j=1}^{p-(k+1)} \sum_{l=1}^{k+1} E^{(p+1-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle. \end{aligned} \quad (4.94)$$

Таким образом, предположив, что соотношение (4.85) выполняется для некоторого  $k$ , мы доказали его и для  $k + 1$ . Так как мы видели, что гипотеза (4.85) верна для  $k = 1$  и  $k = 2$ , она должна выполняться и для любого  $k < p$ . Ч. т. д.

Так как  $\langle \Psi^0 | \hat{W} | \Psi^{(p)} \rangle = E^{(p+1)}$ , мы можем переписать соотношение (4.85) в виде

$$E^{(p+1)} = \langle \Psi^{(k)} | \hat{W} | \Psi^{(p-k)} \rangle - \sum_{j=1}^{p-k} \sum_{l=1}^k E^{p+1-l-j} \langle \Psi^{(l)} | \Psi^{(j)} \rangle, \quad (4.95)$$

где можно выбрать любое значение  $k$  в интервале  $0 < k < p$  (если  $k = 0$  или  $k = p$ , тогда сумму в правой части надо отбросить, так как сами пределы суммирования теряют смысл в этих случаях).

#### 1.4.3. Выражения для поправок к энергии $E^{(2n)}$ и $E^{(2n+1)}$

Предположим, что в выражении (4.95)  $p+1 = 2n$  (четное), т. е.  $p = 2n - 1$  (нечетное); выберем  $k = n - 1$ . Тогда  $p - k = 2n - 1 - (n - 1) = n$ , и

$$E^{(2n)} = \langle \Psi^{(n-1)} | \hat{W} | \Psi^{(n)} \rangle - \sum_{j=1}^n \sum_{l=1}^{n-1} E^{(2n-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle. \quad (4.96)$$

Теперь предположим, что  $p + 1 = 2n + 1$  (нечетное), т. е.  $p = 2n$  (четное); выберем  $k = n$ . Получаем  $p - k = 2n - n = n$  и

$$E^{(2n+1)} = \langle \Psi^{(n)} | \hat{W} | \Psi^{(n)} \rangle - \sum_{j=1}^n \sum_{l=1}^n E^{(2n+1-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle. \quad (4.97)$$

Видно, что поправкой наивысшего порядка к волновой функции, входящей в выражения (4.96) и (4.97), является  $\Psi^{(n)}$ . Нам нужны поправки к энергии более высоких порядков. Однако, если найдены поправки к волновой функции  $\Psi^{(1)}, \Psi^{(2)}, \dots, \Psi^{(n)}$ , то поправки к энергии вплоть до порядка  $2n + 1$  можно получить последовательно одну за другой, повторяя процедуру, выраженную формулами (4.96) и (4.97); при этом не

требуются поправки к волновой функции, имеющие порядок выше, чем  $\Psi^{(n)}$ . В частности,

$$\begin{aligned}
 n = 0 : \quad E^{(1)} &= \langle \Psi^0 | \hat{W} | \Psi^0 \rangle \\
 n = 1 : \quad E^{(2)} &= \langle \Psi^0 | \hat{W} | \Psi^{(1)} \rangle \\
 E^{(3)} &= \langle \Psi^{(1)} | \hat{W} | \Psi^{(1)} \rangle - E^{(1)} \langle \Psi^{(1)} | \Psi^{(1)} \rangle \\
 n = 2 : \quad E^{(4)} &= \langle \Psi^{(1)} | \hat{W} | \Psi^{(2)} \rangle - \sum_{j=1}^2 \sum_{l=1}^1 E^{(4-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle \\
 &= \langle \Psi^{(1)} | \hat{W} | \Psi^{(2)} \rangle - E^{(2)} \langle \Psi^{(1)} | \Psi^{(1)} \rangle - E^{(1)} \langle \Psi^{(1)} | \Psi^{(2)} \rangle, \\
 E^{(5)} &= \langle \Psi^{(2)} | \hat{W} | \Psi^{(2)} \rangle - \sum_{j=1}^2 \sum_{l=1}^2 E^{(5-l-j)} \langle \Psi^{(l)} | \Psi^{(j)} \rangle \\
 &= \langle \Psi^{(2)} | \hat{W} | \Psi^{(2)} \rangle - E^{(3)} \langle \Psi^{(1)} | \Psi^{(1)} \rangle - E^{(2)} \langle \Psi^{(2)} | \Psi^{(1)} \rangle \\
 &\quad - E^{(2)} \langle \Psi^{(1)} | \Psi^{(2)} \rangle - E^{(1)} \langle \Psi^{(2)} | \Psi^{(2)} \rangle, \text{ и т. д.}
 \end{aligned} \tag{4.98}$$

## 2. Вариационный метод и теория возмущений. Функционал Хиллерааса

Снова рассмотрим возмущенное уравнение Шрёдингера  $(\hat{H}^0 + \lambda \hat{W})\Psi = E\Psi$ , и разложение в ряд нулевого (наинизшего) уровня энергии по степеням параметра возмущения  $\lambda$ . (Так как в настоящем разделе все величины относятся к одному и тому же основному состоянию, мы опускаем нижний индекс «0».)

$$E = E^0 + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \tag{4.99}$$

Как уже говорилось в разд. 1.1, параметр возмущения  $\lambda$  можно рассматривать (по крайней мере формально) как непрерывную переменную.

Разложение энергии в степенной ряд по  $\lambda$  должно быть единственным. Это означает, что коэффициенты разложения  $E^0, E^{(1)}, E^{(2)}, \dots$  не должны зависеть от способа, которым они получены (разумеется, в предположении сходимости ряда теории возмущений). Мы можем получить разложение (4.99) не только методами, которые обсуждались в предыдущих разделах, но и с помощью следующего метода: вычислим среднее значение энергии для точной волновой функции, представленной в виде бесконечного ряда по  $\lambda$ , и потом переразложим это среднее значение в степенной ряд по  $\lambda$ .

Эти рассуждения показывают, что выражение (4.99) можно рассматривать также и как разложение в ряд среднего значения энергии, соответствующего точной волновой функции; последняя представляет собой

функцию как параметра  $\lambda$ , так и коэффициентов ряда теории возмущений.

Согласно вариационному принципу, энергия основного состояния является наименьшим возможным средним значением энергии,  $E = \min$ . Поэтому энергия, как функция коэффициентов разложения волновой функции, должна быть минимальной для любого  $\lambda$ , для которого ряд теории возмущений сходится. Так как  $\lambda$ , по крайней мере формально, является непрерывной переменной, энергия в выражении (4.99) должна быть минимальной при всех значениях  $\lambda$ . Это возможно только тогда, когда каждая из величин  $E^{(i)}$  достигает минимума как функция коэффициентов разложения или является константой. Следовательно, для данного  $\lambda$  можно рассматривать коэффициенты разложения волновой функции по теории возмущений как неизвестные величины и определить их значения, минимизируя одну за другой поправки ТВ к энергии  $E^{(i)}$ . (В пределе  $\lambda \rightarrow 0$  поправки к энергии высшего порядка, пропорциональные некоторой степени  $\lambda^j$ , пренебрежимо малы по сравнению с поправками более низкого порядка, пропорциональными  $\lambda^i$ , где  $i < j$ . Поэтому нужно сначала минимизировать поправки к энергии наименьшего порядка и переходить к высшим порядкам, зафиксировав уже найденные коэффициенты разложения).

В случае  $E^0$  и  $E^{(1)}$  волновая функция не содержит свободных параметров, однако  $E^{(2)}$  зависит от поправки первого порядка к волновой функции. Поэтому истинная волновая функция первого порядка — та, для которой вклад второго порядка в энергию достигает минимума.

Следует отметить, что все поправки к энергии нечетного порядка должны быть постоянными в том смысле, что они однозначно определяются параметрами волновой функции, оптимизированными для энергетических поправок более низких (четных) порядков. Иначе, минимизируя слагаемое  $\lambda^{2i+1}E^{(2i+1)}$  при положительных значениях  $\lambda$ , мы получили бы максимум для отрицательных значений, в противоречии с вариационным принципом. Очевидно, это заключение находится в тесной связи с  $2n + 1$ -теоремой Вигнера, обсуждавшейся в разд. 1.4.

Теперь перейдем к определению поправки первого порядка к волновой функции на основе изложенных вариационных соображений. Рассмотрим волновую функцию  $\Psi^0 + \lambda\chi$ , где  $\Psi^0$  — волновая функция нулевого порядка, удовлетворяющая уравнению Шрёдингера в нулевом порядке  $\hat{H}^0\Psi^0 = E^0\Psi^0$ , а  $\chi$  — волновая функция первого порядка, которую мы оптимизируем ( $\hat{H}^0$  эрмитов, так что  $E^0$  вещественно; в данном случае и не нужно предполагать ничего другого).

Разложим среднее значение энергии до второго порядка включительно:

$$\begin{aligned}
 E &= \frac{\langle \Psi^0 + \lambda \chi | \hat{H}^0 + \lambda \hat{W} | \Psi^0 + \lambda \chi \rangle}{\langle \Psi^0 + \lambda \chi | \Psi^0 + \lambda \chi \rangle} \\
 &= \left\{ \langle \Psi^0 | \hat{H}^0 | \Psi^0 \rangle + \lambda \langle \chi | \hat{H}^0 | \Psi^0 \rangle + \lambda \langle \Psi^0 | \hat{W} | \Psi^0 \rangle + \lambda \langle \Psi^0 | \hat{H}^0 | \chi \rangle \right. \\
 &\quad \left. + \lambda^2 \langle \chi | \hat{W} | \Psi^0 \rangle + \lambda^2 \langle \chi | \hat{H}^0 | \chi \rangle + \lambda^2 \langle \Psi^0 | \hat{W} | \chi \rangle + \mathcal{O}(\lambda^3) \right\} \\
 &\quad \times \left\{ \langle \Psi^0 | \Psi^0 \rangle \left[ 1 + \lambda \frac{\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle}{\langle \Psi^0 | \Psi^0 \rangle} + \lambda^2 \frac{\langle \chi | \chi \rangle}{\langle \Psi^0 | \Psi^0 \rangle} \right] \right\}^{-1}.
 \end{aligned} \tag{4.100}$$

Используя уравнение нулевого порядка и известное разложение  $(1 + x)^{-1} = 1 - x + x^2 - \dots$ , получим, сохраняя слагаемые до второго порядка малости:

$$\begin{aligned}
 E &= \frac{1}{\langle \Psi^0 | \Psi^0 \rangle} \left\{ E^0 \langle \Psi^0 | \Psi^0 \rangle + \lambda E^0 (\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle) \right. \\
 &\quad \left. + \lambda \langle \Psi^0 | \hat{W} | \Psi^0 \rangle + \lambda^2 \left( \langle \chi | \hat{W} | \Psi^0 \rangle + \langle \Psi^0 | \hat{W} | \chi \rangle \right) + \lambda^2 \langle \chi | \hat{H}^0 | \chi \rangle \right\} \\
 &\quad \times \left\{ 1 - \lambda \frac{\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle}{\langle \Psi^0 | \Psi^0 \rangle} + \lambda^2 \frac{(\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle)^2}{\langle \Psi^0 | \Psi^0 \rangle^2} \right. \\
 &\quad \left. - \lambda^2 \frac{\langle \chi | \chi \rangle}{\langle \Psi^0 | \Psi^0 \rangle} + \dots \right\} + \mathcal{O}(\lambda^3).
 \end{aligned} \tag{4.101}$$

Перемножим выражения в фигурных скобках, снова оставляя слагаемые до второго порядка малости:

$$\begin{aligned}
 E &= \frac{1}{\langle \Psi^0 | \Psi^0 \rangle} \left\{ E^0 \langle \Psi^0 | \Psi^0 \rangle + \lambda E^0 (\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle) + \lambda \langle \Psi^0 | \hat{W} | \Psi^0 \rangle \right. \\
 &\quad \left. + \lambda^2 \left( \langle \chi | \hat{W} | \Psi^0 \rangle + \langle \Psi^0 | \hat{W} | \chi \rangle + \langle \chi | \hat{H}^0 | \chi \rangle \right) - \lambda \frac{E^0 \langle \Psi^0 | \Psi^0 \rangle (\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle)}{\langle \Psi^0 | \Psi^0 \rangle} \right. \\
 &\quad \left. - \lambda^2 \frac{E^0 (\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle)^2}{\langle \Psi^0 | \Psi^0 \rangle} - \lambda^2 \frac{\langle \Psi^0 | \hat{W} | \Psi^0 \rangle (\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle)}{\langle \Psi^0 | \Psi^0 \rangle} \right. \\
 &\quad \left. + \lambda^2 \frac{E^0 \langle \Psi^0 | \Psi^0 \rangle (\langle \chi | \Psi^0 \rangle + \langle \Psi^0 | \chi \rangle)^2}{\langle \Psi^0 | \Psi^0 \rangle^2} - \lambda^2 E^0 \langle \chi | \chi \rangle \right\} + \mathcal{O}(\lambda^3).
 \end{aligned} \tag{4.102}$$



Второе слагаемое в скобках сокращается с пятым, а шестое — с восьмым. Учитывая, что энергия первого порядка равна

$$E^{(1)} = \frac{\langle \Psi^0 | \hat{W} | \Psi^0 \rangle}{\langle \Psi^0 | \Psi^0 \rangle}, \quad (4.103)$$

приводим уравнение (4.102) к виду

$$E = E^0 + \lambda E^{(1)} + \frac{\lambda^2}{\langle \Psi^0 | \Psi^0 \rangle} \left[ \langle \chi | \hat{W} - E^{(1)} | \Psi^0 \rangle + \langle \Psi^0 | \hat{W} - E^{(1)} | \chi \rangle + \langle \chi | \hat{H}^0 - E^0 | \chi \rangle \right] + \mathcal{O}(3). \quad (4.104)$$

Это значит, что

$$E = E^0 + \lambda E^{(1)} + \lambda^2 J_2 + \mathcal{O}(3), \quad (4.105)$$

где в случае нормированной на единицу  $\Psi^0$ :

$$J_2 = \langle \chi | \hat{W} - E^{(1)} | \Psi^0 \rangle + \langle \Psi^0 | \hat{W} - E^{(1)} | \chi \rangle + \langle \chi | \hat{H}^0 - E^0 | \chi \rangle \quad (4.106)$$

есть функционал Хиллерааса (второго порядка), минимум которого дает поправку второго порядка к энергии по теории возмущений

$$E^{(2)} = \min_{\chi} (J_2). \quad (4.107)$$

Уравнение (4.107) представляет собой вариационную задачу. Необходимое условие минимума — обращение в нуль вариации  $J_2$  при варьировании функции  $\chi$ :

$$\delta J_2 = 0. \quad (4.108)$$

Варьируя  $J_2$ , заданный выражением (4.106), получим

$$\langle \delta \chi | \hat{W} - E^{(1)} | \Psi^0 \rangle + \langle \delta \chi | \hat{H}^0 - E^0 | \chi \rangle + c.c. = 0. \quad (4.109)$$

Так как  $\delta \chi$  может быть комплексной и содержать произвольный фазовый множитель, условие (4.109) может выполняться для любой  $\delta \chi$ , только если слагаемое, выписанное явно, и его комплексное сопряженное (с.с.) обращаются в нуль по отдельности:

$$\langle \delta \chi | \hat{W} - E^{(1)} | \Psi^0 \rangle + \langle \delta \chi | \hat{H}^0 - E^0 | \chi \rangle = 0. \quad (4.110)$$

Если мы используем промежуточную нормировку  $\langle \Psi^0 | \Psi^0 \rangle = 1$ ;  $\langle \Psi^0 | \chi \rangle = 0$ , и разложим волновую функцию первого порядка  $\chi$  по собственным векторам  $\Psi_i^0$  ( $i \neq 0$ ) невозмущенного гамильтониана

$$\chi = \sum_{i=1} c_i \Psi_i^0, \quad (4.111)$$

то допустимые вариации  $\delta\chi$  имеют вид

$$\delta\chi = \eta \Psi_j^0; \quad j = 1, 2, \dots \quad (4.112)$$

где  $\eta \rightarrow 0$  есть произвольный вариационный параметр (мы также могли бы обозначить его  $\delta c_j$ , как делали в гл. 3, разд. 1). Подставив определение (4.112) в условие (4.110) и сократив общий множитель  $\eta^*$ , получим

$$\langle \Psi_j^0 | \hat{W} - E^{(1)} | \Psi^0 \rangle + \langle \Psi_j^0 | \hat{H}^0 - E^0 | \sum_{i=1} c_i \Psi_i^0 \rangle = 0. \quad (4.113)$$

Учитывая, что  $\langle \Psi_j^0 | \Psi^0 \rangle = 0$  ( $j \neq 0$ ),  $\hat{H}^0 \Psi_i^0 = E_i^0 \Psi_i^0$ , и  $\langle \Psi_j^0 | \Psi_i^0 \rangle = \delta_{ij}$ , сразу приводим уравнение (4.113) к виду

$$W_{j0} + (E_j^0 - E^0) c_j = 0, \quad (4.114)$$

где  $W_{j0} = \langle \Psi_j^0 | \hat{W} | \Psi^0 \rangle$ , т. е. это хорошо известная формула для коэффициентов волновой функции первого порядка в теории возмущений Рэлея—Шрёдингера:

$$c_j = -\frac{W_{j0}}{E_j^0 - E^0} \quad (4.115)$$

Предыдущий результат означает, что

$$\chi = \sum_{j=1} \frac{-W_{j0}}{E_j^0 - E^0} |\Psi_j^0\rangle \quad (4.116)$$

и именно этот результат надо подставить в выражение (4.106), чтобы получить поправку второго порядка к энергии:

$$\begin{aligned} J_2 = & \left\langle \sum_{j=1} \frac{-W_{j0}}{E_j^0 - E^0} \Psi_j^0 | \hat{W} - E^{(1)} | \Psi^0 \right\rangle \\ & + \langle \Psi^0 | \hat{W} - E^{(1)} | \sum_{j=1} \frac{-W_{j0}}{E_j^0 - E^0} \Psi_j^0 \rangle + \\ & + \left\langle \sum_{i=1} \frac{-W_{i0}}{E_i^0 - E^0} \Psi_i^0 | \hat{H}^0 - E^0 | \sum_{j=1} \frac{-W_{j0}}{E_j^0 - E^0} \Psi_j^0 \right\rangle. \end{aligned} \quad (4.117)$$

Воспользуемся тем, что  $W_{j0}^* = W_{0j}$ ,  $\langle \Psi_j^0 | \Psi^0 \rangle = 0$ , и тем, что в третьем слагаемом  $(\hat{H}^0 - E^0)\Psi_j^0 = (E_j^0 - E^0)\Psi_j^0$ , так что знаменатель в нем сокращается. Получаем

$$\begin{aligned} J_2 &= -2 \sum_{j=1} \frac{|W_{j0}|^2}{E_j^0 - E^0} + \sum_{i,j} \frac{W_{0i} W_{j0}}{E_i^0 - E^0} \underbrace{\langle \Psi_i^0 | \Psi_j^0 \rangle}_{\delta_{ij}} \\ &= - \sum_{j=1} \frac{|W_{j0}|^2}{E_j^0 - E^0}. \end{aligned} \quad (4.118)$$

Это — хорошо известная формула для поправки второго порядка к энергии.

Главное преимущество вариационной теории возмущений заключается не в независимом выводе выражений обычного формализма теории возмущений, а в возможности рассмотрения (по крайней мере приближенно) задач, в которых стандартные разложения теории возмущений не удастся использовать. Это имеет место, если не все собственные значения и векторы невозмущенного гамильтониана  $\hat{H}^0$  известны. В таком случае предполагают, что наилучшим приближением к волновой функции первого порядка является то, которое доставляет наименьшее значение  $J_2$ .

Интересно отметить, что форма функционала (4.106) такова, что необязательно требовать ортогональности волновой функции первого порядка  $\chi$  к  $\Psi^0$ . Как легко проверить, значение  $J_2$  не изменится, если заменить  $\chi$  на  $\chi + \alpha \Psi^0$ , где  $\alpha$  — произвольный коэффициент. Это значит, что искомое минимальное значение  $J_2$  можно получить оптимизацией, не накладывая каких-либо ограничений на  $\chi$ , но тогда  $\chi$  не будет определено однозначно. Точная поправка первого порядка к волновой функции получается при дополнительном условии  $\langle \Psi^0 | \chi \rangle = 0$ .

### 3. Вырожденная теория возмущений Рэля—Шрёдингера

Формулы, обсуждавшиеся в предыдущем разделе, становятся неприменимыми, если рассматриваемое невозмущенное состояние принадлежит к вырожденному собственному значению невозмущенного гамильтониана, так как в этом случае некоторые знаменатели обращаются в нуль.

Рассмотрим случай, когда собственное значение  $E_k^0$  невозмущенного гамильтониана  $f$ -кратно вырождено:

$$\hat{H}^0 \Psi_{k,i}^0 = E_k^0 \Psi_{k,i}^0; \quad i = 1, 2, \dots, f, \quad (4.119)$$

где мы обозначили как  $\Psi_{k,i}^0$  те  $f$  ортонормированных решений невозмущенной задачи, которые принадлежат вырожденному собственному значению  $E_k^0$ .

Пусть возмущенная задача, как обычно, имеет вид

$$(\hat{H}^0 + \lambda \hat{W})\Psi = E\Psi. \quad (4.120)$$

Разлагая  $\Psi$  и  $E$  в степенные ряды по  $\lambda$ , имеем

$$(\hat{H}^0 + \lambda \hat{W}) \sum_{l=0}^{\infty} \lambda^l \Psi_k^{(l)} = \sum_{j=0}^{\infty} \lambda^j E_k^{(j)} \sum_{l=0}^{\infty} \lambda^l \Psi_k^{(l)}. \quad (4.121)$$

Выпишем условия равенства коэффициентов при разных степенях  $\lambda$  в двух частях уравнения (4.121).

Для  $\lambda^0$  получаем

$$\hat{H}^0 \Psi_k^0 = E_k^0 \Psi_k^0 \quad (4.122)$$

т. е.  $E_k^{(0)} = E_k^0$  и уравнение (4.122) выполняется для произвольной линейной комбинации

$$\Psi_k^{(0)} = \sum_{i=1}^f c_i \Psi_{k,i}^0. \quad (4.123)$$

Важно отметить, что коэффициенты  $c_i$  не могут быть определены из решения одной только задачи нулевого порядка: только само возмущение  $\hat{W}$  определяет, какой вектор вырожденного подпространства является подходящей волновой функцией нулевого порядка.

Для коэффициентов при  $\lambda^1$  имеем

$$\hat{W}\Psi_k^{(0)} + \hat{H}^0\Psi_k^{(1)} = E_k^{(1)}\Psi_k^{(0)} + E_k^{(0)}\Psi_k^{(1)}. \quad (4.124)$$

Перегруппируем слагаемые в уравнении (4.124) и подставим в него разложение (4.123) для функции  $\Psi_k^{(0)}$ :

$$(\hat{W} - E_k^{(1)}) \sum_{i=1}^f c_i \Psi_{k,i}^0 = -(\hat{H}^0 - E_k^0) \Psi_k^{(1)}. \quad (4.125)$$

Домножим уравнения (4.125) на одну из функций  $\Psi_{k,m}^{0*}$  и проинтегрируем. Два слагаемых с правой стороны сокращаются в силу равенства (4.119) и эрмитовости  $\hat{H}^0$ :

$$\langle \Psi_{k,m}^0 | \hat{W} - E_k^{(1)} | \sum_{i=1}^f c_i \Psi_{k,i}^0 \rangle = 0; \quad m = 1, 2, \dots, f. \quad (4.126)$$

Вводя обозначение

$$W_{mi}^{(k)} = \langle \Psi_{k,m}^0 | \hat{W} | \Psi_{k,i}^{(0)} \rangle, \quad (4.127)$$

получим

$$\sum_{i=1}^f W_{mi}^{(k)} c_i = E_k^{(1)} c_m; \quad m = 1, 2, \dots, f. \quad (4.128)$$

Введем следующие матричные обозначения:  $\mathbf{W}^{(k)}$  — матрица размера  $f \times f$ , составленная из матричных элементов  $\hat{W}$  в вырожденном подпространстве, а  $\mathbf{c}$  —  $f$ -мерный вектор коэффициентов  $c_i$ . Тогда условие (4.128) можно переписать как *уравнение на собственные значения*:

$$\mathbf{W}^{(k)} \mathbf{c} = E_k^{(1)} \mathbf{c}. \quad (4.129)$$

Это уравнение имеет  $f$  различных ортонормированных решений, которые диагонализуют возмущение в подпространстве, растянутом вырожденными состояниями. Обычно (но не обязательно) вырождение энергии в первом порядке снимается (т. е. собственные значения задачи (4.129) различны и поэтому невырождены).

В соответствии с предшествующим обсуждением, для того, чтобы найти точные функции в нулевом порядке, надо решить линейную вариационную задачу (ср. гл. 3, разд. 1) для полного возмущенного гамильтониана в подпространстве, растянутом вырожденными собственными векторами невозмущенной задачи. Очевидно, эта схема применима также в «квазивырожденном» случае, когда невозмущенная задача имеет несколько близких — не обязательно строго вырожденных — состояний, отделенных заметной энергетической «щелью» от остальных состояний, хотя в таком случае подобная процедура не очень строга с точки зрения формальной теории возмущений.

В случае задач, вырожденных в нулевом порядке, лишь изредка рассматриваются поправки высших порядков к волновой функции и энергии. Интересно заметить, что волновую функцию первого порядка и энергию второго порядка можно найти, полностью отбросив остающиеся  $f - 1$  состояний подпространства, которое первоначально было вырожденным. По существу, если в качестве невозмущенного состояния использовать собственный вектор, найденный при решении задачи (4.129), то формулы невырожденной теории возмущений становятся применимыми: этот вектор взаимодействует с состояниями ортогонального дополнения к изначально вырожденному подпространству, но (поскольку была проведена диагонализация) не с оставшимися  $f - 1$  состояниями того же подпространства.

## 4. Теория возмущений Бриллюэна—Вигнера

Исторически эта теория рассматривалась как альтернатива теории возмущений Рэлея—Шрёдингера, так как она может давать более быструю сходимость — но за счет использования дополнительной итерационной процедуры. Итерации становятся необходимы, так как в рамках этого формализма искомая точная энергия появляется также и в правой части уравнений теории возмущений. В настоящее время эта схема уже не используется широко из-за проблемы размерной согласованности, которая обсуждается в следующем разделе<sup>1</sup>.

В ТВ Бриллюэна—Вигнера формальный параметр возмущения  $\lambda$  не применяется, а поиск волновой функции и энергии происходит итерационно. (Упомянутые итерации по энергии дополняют итерационный вывод самих формул теории.)

Запишем возмущенное уравнение Шрёдингера

$$(\hat{H}^0 + \hat{V})|\Psi^0 + \varphi\rangle = E|\Psi^0 + \varphi\rangle, \quad (4.130)$$

где  $|\Psi^0\rangle$  — решение невозмущенного уравнения Шрёдингера  $\hat{H}^0|\Psi^0\rangle = E^0|\Psi^0\rangle$ , и используется «промежуточная» нормировка:  $\langle\Psi^0|\Psi^0\rangle = 1$ ;  $\langle\Psi^0|\varphi\rangle = 0$ . (Мы предполагаем, что  $|\Psi^0\rangle$  является собственным вектором  $\hat{H}^0$  с наинизшей энергией, т. е.  $|\Psi^0\rangle = |\Psi_0^0\rangle$ ; для простоты, мы везде опустим нижний индекс 0.)

Умножая уравнение (4.130) на  $\langle\Psi^0|$ , легко получаем

$$E = \langle\Psi^0|\hat{H}^0 + \hat{V}|\Psi^0 + \varphi\rangle = E^0 + \langle\Psi^0|\hat{V}|\varphi\rangle, \quad (4.131)$$

так как

$$\langle\Psi^0|\hat{H}^0|\varphi\rangle = E^0\langle\Psi^0|\varphi\rangle = 0. \quad (4.132)$$

Уравнение Шрёдингера (4.130) можно перегруппировать

$$(E - \hat{H}^0)|\varphi\rangle = (\hat{H}^0 - E)|\Psi^0\rangle + \hat{V}|\Psi^0 + \varphi\rangle. \quad (4.133)$$

Введем проекционный оператор

$$\hat{Q} = 1 - |\Psi^0\rangle\langle\Psi^0| \quad (4.134)$$

на ортогональное дополнение к  $|\Psi^0\rangle$  и подействуем им на уравнение (4.133). Получим

$$\hat{Q}(E - \hat{H}^0)|\varphi\rangle = \hat{Q}\hat{V}|\Psi^0 + \varphi\rangle, \quad (4.135)$$

поскольку, как легко видеть,  $\hat{Q}(\hat{H}^0 - E)|\Psi^0\rangle = \hat{Q}(E^0 - E)|\Psi^0\rangle = 0$ .

<sup>1</sup> Однако в последние годы снова наблюдается некоторое возрастание интереса в связи с многоконфигурационными теориями.

Далее,  $|\varphi\rangle$  лежит в ортогональном дополнении к  $|\Psi^0\rangle$  и то же самое верно для  $\hat{H}^0|\varphi\rangle$  в соответствии с условием (4.132). Поэтому  $(E - \hat{H}^0)|\varphi\rangle$  ортогонально к  $|\Psi^0\rangle$ , и

$$\hat{Q}(E - \hat{H}^0)|\varphi\rangle = (E - \hat{H}^0)|\varphi\rangle. \quad (4.136)$$

Таким образом, оператор  $\hat{Q}$  в левой части уравнения (4.135) можно опустить:

$$(E - \hat{H}^0)|\varphi\rangle = \hat{Q}\hat{V}|\Psi^0 + \varphi\rangle. \quad (4.137)$$

Введем теперь обратный оператор  $(E - \hat{H}^0)^{-1}$  (резольвенту оператора  $\hat{H}^0$ ). Он имеет спектральное разложение по собственным векторам и собственным значениям  $\hat{H}^0$ :

$$(E - \hat{H}^0)^{-1} = \sum_{j=0} \frac{|\Psi_j^0\rangle\langle\Psi_j^0|}{E - E_j^0}. \quad (4.138)$$

(здесь мы предполагаем, что  $E$  не равно ни одному из собственных значений  $\hat{H}^0$ ). Действуя этим оператором на обе части уравнения (4.137), получаем

$$|\varphi\rangle = (E - \hat{H}^0)^{-1}\hat{Q}\hat{V}|\Psi^0 + \varphi\rangle. \quad (4.139)$$

Данное уравнение является точным; мы можем попытаться решить его итерационно, используя тот факт, что если  $\hat{V}$  является малым возмущением, то и  $|\varphi\rangle$  является малой поправкой. Получим первое приближение к  $|\varphi\rangle$ , пренебрегая им в правой части (т. е. используя  $|\varphi\rangle = 0$  как нулевое приближение). Это приводит нас к выражению

$$|\varphi^{(1)}\rangle = (E - \hat{H}^0)^{-1}\hat{Q}\hat{V}|\Psi^0\rangle. \quad (4.140)$$

Подставляя результат (4.140) в правую часть уравнения (4.139), получим второе приближение:

$$|\varphi^{(2)}\rangle = (E - \hat{H}^0)^{-1}\hat{Q}\hat{V}|\Psi^0\rangle + [(E - \hat{H}^0)^{-1}\hat{Q}\hat{V}]^2|\Psi^0\rangle \quad (4.141)$$

что можно повторять, в результате получая ряд:

$$\begin{aligned} |\varphi\rangle = & (E - \hat{H}^0)^{-1}\hat{Q}\hat{V}|\Psi^0\rangle + [(E - \hat{H}^0)^{-1}\hat{Q}\hat{V}]^2|\Psi^0\rangle \\ & + [(E - \hat{H}^0)^{-1}\hat{Q}\hat{V}]^3|\Psi^0\rangle + \dots \end{aligned} \quad (4.142)$$

Поправка второго порядка к энергии по теории возмущений Бриллюэна—Вигнера получается подстановкой волновой функции первого порядка (4.140) в последнее слагаемое в правой части формулы (4.131):

$$E^{(2)} = \langle\Psi^0|\hat{V}|\varphi^{(1)}\rangle = \langle\Psi^0|\hat{V}(E - \hat{H}^0)^{-1}\hat{Q}\hat{V}|\Psi^0\rangle. \quad (4.143)$$

Если остановиться на втором порядке, то  $E$  необходимо заменить на  $E^0 + V_{00} + E^{(2)}$  также и в правой части. Подставим резольвенту (4.138) в выражение (4.143) (учитывая, что  $\langle \Psi^0 | \hat{Q} = 0$ ) и получим

$$\begin{aligned} E^{(2)} &= \langle \Psi^0 | \hat{V} \sum_{j=0} \frac{|\Psi_j^0\rangle \langle \Psi_j^0|}{E - E_j^0} \hat{Q} \hat{V} | \Psi^0 \rangle = \langle \Psi^0 | \hat{V} \sum_{j=1} \frac{|\Psi_j^0\rangle \langle \Psi_j^0|}{E - E_j^0} \hat{V} | \Psi^0 \rangle \\ &= \sum_{j=1} \frac{|V_{0j}|^2}{E - E_j^0}. \end{aligned} \quad (4.144)$$

Тогда полная энергия с точностью до второго порядка ТВ Бриллюэна—Вигнера дается выражением:

$$E = E^0 + V_{00} + \sum_{j=1} \frac{|V_{0j}|^2}{E - E_j^0}. \quad (4.145)$$

Этот результат отличается от формулы второго порядка Рэлея—Шрёдингера присутствием в правой части неизвестной энергии  $E = E^0 + V_{00} + E^{(2)}$  (вместо  $E^0$ ). Поэтому выражение (4.145), в отличие от аналогичной формулы ТВ Рэлея—Шрёдингера представляет собой уравнение, которое должно решаться итерационно.

Метод, аналогичный тому, который приводит к выражению (4.145), позволяет получить и формулу для волновой функции первого порядка:

$$|\Psi\rangle = |\Psi^0\rangle + \sum_{i=1} \frac{V_{i0}}{E - E_i^0} |\Psi_i^0\rangle \quad (4.146)$$

которая снова отличается от соответствующей формулы теории возмущений Рэлея—Шрёдингера заменой  $E^0$  на  $E$ .

### Пример

Рассмотрим простейший пример системы, имеющей только два невозмущенных состояния,  $\Psi_0^0$  и  $\Psi_1^0$ , и предположим, что диагональные элементы матрицы возмущения равны нулю ( $V_{00} = V_{11} = 0, V_{01} \neq 0$ ). Матрица гамильтониана в этом случае примет вид

$$\mathbf{H} = \mathbf{H}^0 + \mathbf{V} = \begin{pmatrix} E^0 & V_{01} \\ V_{10} & E_1^0 \end{pmatrix}, \quad (4.147)$$

а вековое уравнение

$$\begin{vmatrix} E^0 - E & V_{01} \\ V_{10} & E_1^0 - E \end{vmatrix} = (E - E^0)(E - E_1^0) - |V_{01}|^2 = 0. \quad (4.148)$$



Формула энергии по ТВ Бриллюэна—Вигнера во втором порядке дает в этом случае

$$E = E^0 + \frac{|V_{01}|^2}{E - E_1^0} \quad (4.149)$$

так как сумма по  $j$  в выражении (4.145) сводится к единственному слагаемому. Очевидно, что формулу (4.149) можно преобразовать так, что получится исходное квадратное вековое уравнение (4.148). Этот результат показывает, что в этом частном случае энергия, получающаяся во втором порядке теории возмущений Бриллюэна—Вигнера, совпадает с точной. Можно легко проверить, что в этой весьма частной модели ряд Бриллюэна—Вигнера обрывается после первой поправки к волновой функции, из-за того, что в системе имеется всего лишь два уровня и  $V_{11} = 0$ .

#### 4.1. Проблема размерной согласованности

Теория возмущений Бриллюэна—Вигнера второго порядка дает точное решение для простейшей двухуровневой модельной задачи. Однако ситуация становится совершенно другой, если имеются две не взаимодействующие двухуровневые системы  $A$  и  $B$ , и каждая из них описывается по отдельности матрицей гамильтониана, подобной матрице (4.147). В таком случае объединенная система  $AB$  имеет четыре состояния,  $\Psi_{00}$ ,  $\Psi_{01}$ ,  $\Psi_{10}$  и  $\Psi_{11}$ , где  $\Psi_{ij} = \Psi_i^A \Psi_j^B$  (базис системы  $AB$  есть прямое произведение базисов систем  $A$  и  $B$ ). Гамильтониан объединенной системы равен  $\hat{H} = \hat{H}^A + \hat{H}^B$ , где  $\hat{H}^A$  и  $\hat{H}^B$  действуют только на функции подсистем  $A$  и  $B$ , соответственно. Поэтому

$$\begin{aligned} \langle \Psi_{ij} | \hat{H} | \Psi_{kl} \rangle &= \langle \Psi_i^A \Psi_j^B | \hat{H}^A + \hat{H}^B | \Psi_k^A \Psi_l^B \rangle \\ &= H_{ik}^A \delta_{jl} + \delta_{ik} H_{jl}^B. \end{aligned} \quad (4.150)$$

Верно и в общем случае, что матрица гамильтониана  $\mathbf{H}$  системы, состоящей из двух не взаимодействующих подсистем  $A$  и  $B$  записывается как

$$\mathbf{H} = \mathbf{H}^A \otimes \mathbf{1}^B + \mathbf{1}^A \otimes \mathbf{H}^B, \quad (4.151)$$

где  $\mathbf{H}^A$  и  $\mathbf{H}^B$  являются матрицами гамильтонианов отдельных систем  $A$  и  $B$ , соответственно,  $\mathbf{1}^A$  и  $\mathbf{1}^B$  — единичные матрицы таких же размерностей, как и матрицы  $\mathbf{H}^A$  и  $\mathbf{H}^B$ , а символ  $\mathbf{X} \otimes \mathbf{Y}$  обозначает прямое произведение матриц  $\mathbf{X}$  и  $\mathbf{Y}$  (свойства прямого произведения матриц кратко описаны в приложении П9).

Явный вид матрицы гамильтониана системы, состоящей из двух не взаимодействующих двухуровневых подсистем  $A$  и  $B$ , есть

$$\mathbf{H} = \begin{array}{cccc} \begin{array}{c} 00 \\ 01 \\ 10 \\ 0 \end{array} & \begin{array}{c} 01 \\ 0 \\ 0 \\ V_{10}^A \end{array} & \begin{array}{c} 10 \\ V_{01}^B \\ 0 \\ V_{10}^B \end{array} & \begin{array}{c} 11 \\ 0 \\ V_{01}^A \\ E_1^A + E_1^B \end{array} & \begin{array}{c} \leftarrow \text{уровни} \\ \downarrow \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \end{array} \cdot \quad (4.152)$$

Прямым вычислением можно убедиться, что собственные значения такой матрицы — это суммы  $\varepsilon_i^A + \varepsilon_j^B$  ( $i, j = 0, 1$ ), где  $\varepsilon_i^A, \varepsilon_j^B$  являются точными энергиями отдельных подсистем. Чтобы проверить это, достаточно составить прямое произведение

$$\mathbf{c} = \mathbf{c}^{Ai} \otimes \mathbf{c}^{Bj} = \begin{pmatrix} c_0^{Ai} c_0^{Bj} \\ c_0^{Ai} c_1^{Bj} \\ c_1^{Ai} c_0^{Bj} \\ c_1^{Ai} c_1^{Bj} \end{pmatrix} \quad (4.153)$$

решений уравнений Шрёдингера для подсистем  $A$  и  $B$

$$\mathbf{c}^{Ai} = \begin{pmatrix} c_0^{Ai} \\ c_1^{Ai} \end{pmatrix} \quad \text{и} \quad \mathbf{c}^{Bj} = \begin{pmatrix} c_0^{Bj} \\ c_1^{Bj} \end{pmatrix}, \quad (4.154)$$

подставить его в уравнение на собственные значения  $\mathbf{H}\mathbf{c} = \varepsilon\mathbf{c}$ , и соответствующим образом сгруппировать слагаемые. (Ясно, что можно использовать и формулы для прямого произведения, приведенные в приложении П9.)

В соответствии с формулой (4.145), мы получаем формулу для энергии объединенной системы  $AB$ , описываемой матрицей гамильтониана (4.152), во втором порядке теории возмущений Бриллюэна—Вигнера:

$$E = E_0^A + E_0^B + \frac{|V_{01}^A|^2}{E - (E_1^A + E_0^B)} + \frac{|V_{01}^B|^2}{E - (E_0^A + E_1^B)}. \quad (4.155)$$

Она не совпадает с точным собственным значением  $\varepsilon_0^A + \varepsilon_0^B$  гамильтониана (4.152); таким образом, она не равна и сумме энергий ТВ Бриллюэна—Вигнера второго порядка подсистем  $A$  и  $B$  по отдельности. Это означает, что энергия, полученная во втором порядке теории возмущений Бриллюэна—Вигнера не удовлетворяет очень важному требованию «размерной согласованности».

Размерная согласованность крайне важна, если рассматриваются относительные энергии (изменения энергии). Если рассматривается взаимодействие двух или более подсистем некоторой большой системы, а расстояние между ними увеличивается, то правильная энергия должна стремиться к сумме энергий изолированных подсистем. Поэтому расчет объединенной системы, состоящей из подсистем, находящихся на бесконечном расстоянии друг от друга, должен дать точно ту же энергию, что и сумма энергий, получающаяся в расчете изолированных подсистем на том же уровне теории. Методы, удовлетворяющие этому требованию, называются размерно согласованными. Методы же, не являющиеся размерно согласованными, вряд ли дадут разумное значение для энергии взаимодействия.

Важно в этой связи отметить разницу в свойствах теорий возмущений Бриллюэна—Вигнера и Рэлея—Шрёдингера. Снова рассмотрим недиагональные элементы матрицы (4.152) как возмущение. Тогда энергия во втором порядке теории возмущений Рэлея—Шрёдингера дается выражением

$$E = E_0^A + E_0^B - \frac{|V_{01}^A|^2}{E_1^A - E_0^A} - \frac{|V_{01}^B|^2}{E_1^B - E_0^B}, \quad (4.156)$$

которое как раз совпадает с суммой соответствующих энергий второго порядка теории возмущений Рэлея—Шрёдингера для изолированных систем  $A$  и  $B$  по отдельности. Ни одна из этих энергий не равна соответствующей точной энергии, но они удовлетворяют условию размерной согласованности и поэтому их можно использовать для вычисления относительных энергий. Размерная согласованность выполняется для теории возмущений Рэлея—Шрёдингера в любом порядке и для произвольных систем.

Размерная согласованность теории возмущений Рэлея—Шрёдингера в каждом порядке по отдельности заслуживает детального исследования. Следующий раздел целиком посвящен этой проблеме.

## 5. Размерная согласованность теории возмущений Рэлея—Шрёдингера

Мы рассмотрим размерную согласованность теории возмущений Рэлея—Шрёдингера в двух аспектах. Во-первых, мы покажем, что размерная согласованность должна соблюдаться в каждом порядке по отдельности, как формальное следствие того факта, что мы работаем со степенным рядом; затем покажем, как именно эта размерная согласованность проявляется на уровне реальных формул теории возмущений.

### 5.1. Формальное рассмотрение, основанное на свойствах степенного ряда

Сначала приведем некоторую качественную аргументацию, показывающую, что теория возмущений Рэлея—Шрёдингера дает размерно согласованные результаты, если ряд теории возмущений (абсолютно) сходится.

Энергия системы  $AB$ , состоящей из двух не взаимодействующих подсистем  $A$  и  $B$ , равна, очевидно, сумме энергий изолированных подсистем:

$$E^{AB}(\lambda) = E^A(\lambda) + E^B(\lambda), \quad (4.157)$$

где эти энергии рассматриваются как функции одного и того же параметра возмущения  $\lambda$ . В теории возмущений Рэлея—Шрёдингера каждая из этих энергий разлагается в ряд по степеням  $\lambda$  — ср. (4.11). Поэтому сумма рядов, найденных для  $E^A(\lambda)$  и  $E^B(\lambda)$ , соответственно, равна ряду для  $E^{AB}(\lambda)$ . Это равенство может выполняться для произвольного значения  $\lambda$ , только если коэффициенты при каждой степени  $\lambda^l$  одинаковы в обеих частях равенства:

$$E^{AB(l)} = E^{A(l)} + E^{B(l)}. \quad (4.158)$$

Это означает размерную согласованность каждого порядка разложения возмущений Рэлея—Шрёдингера. (Теория возмущений Бриллюэна—Вигнера не дает энергию в виде простого степенного ряда по  $\lambda$ ; так что неудивительно, что она не размерно согласована. Однако, степенной характер разложения является достаточным, но не необходимым условием размерной согласованности метода: энергия, полученная в рамках подхода «связанных кластеров» — см. гл. 8, разд. 3.4 — не является простым степенным рядом, однако размерно согласована.)

Приведем здесь также и более детальное рассмотрение равенства (4.158), которое позволит заключить, что размерная согласованность каждого порядка разложения теории возмущений Рэлея—Шрёдингера выполняется даже в тех случаях, когда ряд возмущений в конечном счете не сходится. Разложения по теории возмущений часто представляют так называемые «асимптотические ряды», для которых результаты приближаются к точному решению до определенного порядка, но затем начинают расходиться. На практике мы проделываем вычисления по теории возмущений до какого-то конечного порядка, и обычно не можем определить, является ли ряд действительно сходящимся или только асимптотическим. Поэтому размерная согласованность важна и в том общем случае, когда неизвестно, сходится ли ряд в конечном итоге или же расходится.

В теории возмущений мы разлагаем каждый гамильтониан на невозмущенную часть и возмущение:

$$\begin{aligned}\hat{H}^A &= \hat{H}^{A0} + \hat{V}^A = \hat{H}^{A0} + \lambda \hat{W}^A \\ \hat{H}^B &= \hat{H}^{A0} + \hat{V}^B = \hat{H}^{B0} + \lambda \hat{W}^B \\ \hat{H}^{AB} &= \hat{H}^A + \hat{H}^B = \hat{H}^{A0} + \hat{H}^{B0} + \lambda(\hat{W}^A + \hat{W}^B) = \hat{H}^{AB0} + \lambda \hat{W}^{AB}.\end{aligned}\quad (4.159)$$

Как обсуждалось в разд. 1.1, параметр возмущения  $\lambda$  можно рассматривать, по крайней мере формально, как непрерывную переменную; мы всегда можем отмасштабировать  $\hat{W}^A$  и  $\hat{W}^B$ , чтобы иметь один и тот же параметр  $\lambda$  как для  $\hat{H}^A$ , так и для  $\hat{H}^B$ , и поэтому также для  $\hat{H}^{AB}$ . Так как гамильтонианы (4.159) зависят от  $\lambda$ , то же самое верно для энергий, как и указано в обеих частях выражения (4.157). Из математики известно, что если энергии  $E^A(\lambda)$  и  $E^B(\lambda)$  возможно разложить в (абсолютно) сходящийся ряд по  $\lambda$ , то сумма этих рядов также является сходящимся рядом, который сходится к сумме  $E^A(\lambda) + E^B(\lambda)$ . Из единственности степенного ряда следует, что не существует разложения в степенной ряд  $E^{AB}(\lambda)$ , отличного от суммы разложений, полученных для  $E^A(\lambda)$  и  $E^B(\lambda)$ . Разложение энергий в рамках теории возмущений Рэлея—Шрёдингера дает ничто иное, как эти степенные ряды  $E^A(\lambda)$ ,  $E^B(\lambda)$ , и  $E^{AB}(\lambda)$ ; поэтому вклады возмущения в  $E^{AB}(\lambda)$  в *каждом порядке* теории возмущений должны равняться сумме соответствующих вкладов в  $E^A(\lambda)$  и  $E^B(\lambda)$ , т. е. (4.158) должно выполняться для каждого  $l$ .

Если  $E(\lambda)$  является гладкой (т. е. непрерывной и бесконечно дифференцируемой) функцией в окрестности  $\lambda = 0$ , тогда ее можно разложить в степенной ряд (ряд Маклорена) с конечными коэффициентами. В свою очередь, любой степенной ряд с конечными коэффициентами имеет ненулевой радиус сходимости; поэтому для любой возмущенной задачи рассматриваемого типа все три интересующих нас ряда должны сходиться при достаточно малых значениях  $\lambda$ , а тогда их коэффициенты подчиняются условию (4.158). Но равенство (4.158) не содержит параметра  $\lambda$ . Поэтому, размерная согласованность теории возмущений Рэлея—Шрёдингера выполняется в каждом порядке по отдельности, безотносительно к тому, сходится ли ряд *при данном* значении  $\lambda$ . Из этих рассуждений следует, что размерная согласованность соблюдается также и для альтернативного описания, упомянутого в разд. 1.1, в котором не используются промежуточные обозначения  $\hat{V} = \lambda \hat{W}$ , а сначала формально вводится параметр  $\lambda$  и затем подставляют  $\lambda = 1$  после того, как разложение в ряд было завершено.

Заметим, что размерная согласованность для двух невзаимодействующих подсистем ( $A$  и  $B$ ) предполагает также размерную согласованность для произвольного числа таких подсистем ( $A, B, C \dots$ ): сначала можно показать размерную согласованность для  $AB = A + B$ , затем (рассматривая  $AB$  как одну подсистему) для  $ABC = AB + C$  и т. д.

## 5.2. Размерная согласованность разложений в ряд теории возмущений

Теперь рассмотрим проблему размерной согласованности на уровне фактических выражений, полученных для поправок теории возмущений к волновой функции и энергии.

Мы рассматриваем объединенную систему  $AB$ , состоящую из двух не взаимодействующих подсистем  $A$  и  $B$ . Гамильтониан объединенной системы равен сумме гамильтонианов двух подсистем:

$$\hat{H}^{AB} = \hat{H}^A + \hat{H}^B \quad (4.160)$$

а волновая функция  $\Psi^{AB}$  определена в гильбертовом пространстве, являющемся прямым произведением гильбертовых пространств двух подсистем. Это значит, что в гильбертовом пространстве объединенной системы может быть выбран базис  $\{\Psi_I\}$ , каждый элемент которого является произведением базисных функций изолированных подсистем

$$\Psi_I^{AB} = \Psi_{ij}^{AB} = \Psi_i^A \Psi_j^B, \quad (4.161)$$

где индекс  $I$  представляет собой пару индексов  $i$  и  $j$ ; будем обозначать это как  $I = (ij)$ .

Так как подсистемы не взаимодействуют, в любом выражении можно переставить гамильтониан  $\hat{H}^A$  и функцию  $\Psi_j^B$ , и то же самое верно и для гамильтониана  $\hat{H}^B$  и функции  $\Psi_i^A$ . В частности, если  $\Psi_i^A$  и  $\Psi_j^B$  являются собственными функциями  $\hat{H}^A$  и  $\hat{H}^B$ , соответственно, тогда имеем, как и можно было ожидать:

$$\begin{aligned} \hat{H}^{AB} \Psi_I^{AB} &= (\hat{H}^A + \hat{H}^B) \Psi_i^A \Psi_j^B = (\hat{H}^A \Psi_i^A) \Psi_j^B + \Psi_i^A (\hat{H}^B \Psi_j^B) \\ &= (E_i^A + E_j^B) \Psi_i^A \Psi_j^B = (E_i^A + E_j^B) \Psi_I^{AB} = E_I^{AB} \Psi_I^{AB} \end{aligned} \quad (4.162)$$

(ср. с обсуждением в гл. 1, разд. 1.1).

Здесь мы обсуждаем общий формализм без ссылок на физическую природу рассматриваемых систем. Если две подсистемы содержат электроны (и/или другие фермионы), то волновые функции должны быть антисимметричны относительно их перестановок. Очевидно, волновая функция вида (4.161) не подчиняется условию антисимметричности по перестановкам электронов *между* двумя подсистемами. Однако (что будет далее обсуждаться в гл. 5, разд. 4.2) антисимметризация волновой функции двух не взаимодействующих систем не имеет практического значения, так что мы можем пренебречь этим затруднением, пока рассматриваются строго не взаимодействующие подсистемы.

Так, если имеем разложение

$$\hat{H}^A = \hat{H}^{A0} + \lambda \hat{W}^A \quad (4.163)$$

и

$$\hat{H}^B = \hat{H}^{B0} + \lambda \hat{W}^B \quad (4.164)$$

то гамильтониан объединенной системы будет

$$\hat{H}^{AB} = \hat{H}^{AB0} + \lambda \hat{W}^{AB} \quad (4.165)$$

с

$$\hat{H}^{AB0} = \hat{H}^{A0} + \hat{H}^{B0} \quad (4.166)$$

и

$$\hat{W}^{AB} = \hat{W}^A + \hat{W}^B. \quad (4.167)$$

Для упрощения обозначим невозмущенные состояния изолированных подсистем  $A$  и  $B$  нижним индексом  $0$ . Они могут быть основными состояниями, которые в большинстве случаев нас и интересуют, но это не обязательно — мы всегда можем изменить нумерацию состояний. В том случае, когда речь идет о возмущенных волновых функциях и энергиях, будем опускать индекс « $0$ ».

Ряды теории возмущений Рэлея—Шрёдингера для энергий и волновых функций изолированных подсистем имеют вид

$$\begin{aligned} E^A &= E_0^{A0} + \sum_{l=1}^{\infty} \lambda^l E^{A(l)} \\ E^B &= E_0^{B0} + \sum_{l=1}^{\infty} \lambda^l E^{B(l)} \end{aligned} \quad (4.168)$$

и

$$\begin{aligned} \Psi^A &= \Phi_0^A + \sum_{l=1}^{\infty} \lambda^l \sum_{i=1}^{\infty} c_i^{A(l)} \Phi_i^A \\ \Psi^B &= \Phi_0^B + \sum_{l=1}^{\infty} \lambda^l \sum_{j=1}^{\infty} c_j^{B(l)} \Phi_j^B, \end{aligned} \quad (4.169)$$

где введены обозначения  $\Phi_i^A$  и  $\Phi_j^B$  для решений невозмущенных уравнений Шрёдингера соответствующих изолированных подсистем.

Состояние объединенной системы  $AB$  с подсистемами  $A$  и  $B$  в состояниях  $\Psi^A$  и  $\Psi^B$  обозначим  $\Psi^{AB} = \Psi^A \Psi^B$ . Произведение невозмущенных волновых функций  $\Phi_i^A \Phi_j^B$  можно обозначить как

$$\Phi_I^{AB} = \Phi_{ij}^{AB} = \Phi_i^A \Phi_j^B. \quad (4.170)$$

Эта функция является собственным вектором невозмущенного уравнения Шрёдингера

$$\begin{aligned}\hat{H}^{AB0}\Phi_I^{AB} &= (\hat{H}^{A0} + \hat{H}^{B0})\Phi_i^A\Phi_j^B = (E_i^{A0} + E_j^{B0})\Phi_i^A\Phi_j^B \\ &= (E_i^{A0} + E_j^{B0})\Phi_I^{AB}.\end{aligned}\quad (4.171)$$

Невозмущенная волновая функция объединенной системы равна

$$\Phi_{(00)}^{AB} = \Phi_0^A\Phi_0^B \quad (4.172)$$

а соответствующая невозмущенная энергия

$$E_{(00)}^{AB0} = E_0^{A0} + E_0^{B0}. \quad (4.173)$$

Ряды возмущений для объединенной системы имеют вид

$$E^{AB} = E_{(00)}^{AB0} + \sum_{l=1}^{\infty} \lambda^l E^{AB(l)} \quad (4.174)$$

и

$$\Psi^{AB} = \Phi_{(00)}^{AB} + \sum_{l=1}^{\infty} \lambda^l \sum_{I \neq (00)} c_I^{(l)} \Phi_I^{AB}. \quad (4.175)$$

С другой стороны, точная волновая функция  $\Psi^{AB}$  объединенной системы является произведением точных решений  $\Psi^A$  и  $\Psi^B$  не взаимодействующих подсистем, в соответствии с выражением (4.162). Если ряды теории возмущений (4.169) сходятся абсолютно, то эта волновая функция равна

$$\Psi^{AB} = \Psi^A \Psi^B = \left( \Phi_0^A + \sum_{l=1}^{\infty} \lambda^l \sum_{i=1} c_i^{A(l)} \Phi_i^A \right) \left( \Phi_0^B + \sum_{m=1}^{\infty} \lambda^m \sum_{j=1} c_j^{B(m)} \Phi_j^B \right). \quad (4.176)$$

Перемножая ряды и собирая коэффициенты при равных степенях  $\lambda$ , получим

$$\begin{aligned}\Psi^{AB} &= \Phi_0^A\Phi_0^B + \sum_{l=1}^{\infty} \lambda^l \left( \sum_{i=1} c_i^{A(l)} \Phi_i^A\Phi_0^B + \sum_{j=1} c_j^{B(l)} \Phi_0^A\Phi_j^B \right) \\ &+ \sum_{l=2}^{\infty} \lambda^l \sum_{i,j=1}^{l-1} \sum_{k=1} c_i^{A(l-k)} c_j^{B(k)} \Phi_i^A\Phi_j^B.\end{aligned}\quad (4.177)$$



В соответствии с формулой (4.177), произведение волновых функций (4.169) приводит к следующим значениям коэффициентов  $c_I^I = c_{ij}^I$  в разложении (4.175):

$$c_{(ij)}^{(1)} = 0 \quad (i, j \neq 0) \quad (4.178)$$

$$c_{(i0)}^{(l)} = c_i^{A(l)} c_{(0j)}^{(l)} = c_j^{B(l)} \quad (i, j \neq 0; l \geq 1) \quad (4.179)$$

и

$$c_{(ij)}^{(l)} = \sum_{k=1}^{l-1} c_i^{A(l-k)} c_j^{B(k)} \equiv \sum_{k=1}^{l-1} c_i^{A(k)} c_j^{B(l-k)} \quad (i, j \neq 0; l \geq 2). \quad (4.180)$$

Размерная согласованность означает, что вклады в энергию удовлетворяют критерию (4.158), т. е.  $E^{AB(l)} = E^{A(l)} + E^{B(l)}$ . Для того, чтобы доказать размерную согласованность, мы покажем, что равенства (4.178)-(4.180) могут быть получены непосредственно из формул теории возмущений, выписанных в разд. 1.2, независимо от того, сходится ряд (4.169) или нет.

В соответствии с формулой (4.173), размерная согласованность выполняется для энергий нулевого порядка. Для энергии в первом порядке нам понадобится среднее значение  $\langle \Phi_{(00)}^{AB} | \hat{W}^{AB} | \Phi_{(00)}^{AB} \rangle$ , поэтому рассмотрим матричные элементы  $\hat{W}^{AB}$ .

Общий вид матричного элемента  $\hat{W}^{AB}$  между невозмущенными волновыми функциями объединенной системы  $\Phi_I^{AB} = \Phi_{(ij)}^{AB} = \Phi_i^A \Phi_j^B$  и  $\Phi_J^{AB} = \Phi_{(kl)}^{AB} = \Phi_k^A \Phi_l^B$  есть

$$\begin{aligned} W_{IJ}^{AB} &= \langle \Phi_I^{AB} | \hat{W}^{AB} | \Phi_J^{AB} \rangle = \langle \Phi_{(ij)}^{AB} | \hat{W}^{AB} | \Phi_{(kl)}^{AB} \rangle \\ &= \langle \Phi_i^A \Phi_j^B | \hat{W}^A + \hat{W}^B | \Phi_k^A \Phi_l^B \rangle = W_{ik}^A \delta_{jl} + W_{jl}^B \delta_{ik} \end{aligned} \quad (4.181)$$

так как можно переставлять  $\hat{W}^A$  и  $\Phi_l^B$ , и также  $\hat{W}^B$  и  $\Phi_k^A$ , а волновые функции объединенной системы в нулевом порядке предполагаются ортонормированными. Это означает, что  $\hat{W}^{AB}$  имеет ненулевые матричные элементы только между состояниями, которые являются возбужденными по отношению друг к другу в одной подсистеме, но соответствуют одному и тому же состоянию в другой подсистеме. Заметим также, что соотношения (4.181) можно выразить в матричных обозначениях как

$$\mathbf{W}^{AB} = \mathbf{W}^A \otimes \mathbf{1}^B + \mathbf{1}^A \otimes \mathbf{W}^B, \quad (4.182)$$

где  $\mathbf{1}^A$  и  $\mathbf{1}^B$  являются единичными матрицами в соответствующих гильбертовых пространствах функций подсистем  $A$  и  $B$ , соответственно (см. приложение П9).

С помощью определений (4.181) получаем поправку первого порядка к энергии

$$E^{AB(1)} = \langle \Phi_{(00)}^{AB} | \hat{W}^{AB} | \Phi_{(00)}^{AB} \rangle = W_{00}^A + W_{00}^B = E^{A(1)} + E^{B(1)}. \quad (4.183)$$

Значит, энергия первого порядка тоже размерно согласованна.

Формула (4.30) для энергии во втором порядке в этом случае примет вид

$$\begin{aligned} E^{AB(2)} &= \sum_{I \neq (00)} \frac{-|W_{(00)I}^{AB}|^2}{E_I^{AB0} - E_{(00)}^{AB0}} \\ &= \sum_{(ij) \neq (00)} \frac{-|\langle \Phi_{(00)}^{AB} | \hat{W}^{AB} | \Phi_{(ij)}^{AB} \rangle|^2}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}}. \end{aligned} \quad (4.184)$$

В соответствии с определениями (4.181), матричный элемент в числителе равен

$$\begin{aligned} W_{(00)I}^{AB} &= \langle \Phi_{(00)}^{AB} | \hat{W}^{AB} | \Phi_{(ij)}^{AB} \rangle = \langle \Phi_0^A \Phi_0^B | \hat{W}^A + \hat{W}^B | \Phi_i^A \Phi_j^B \rangle \\ &= W_{0i}^A \delta_{j0} + W_{0j}^B \delta_{i0}. \end{aligned} \quad (4.185)$$

В выражении (4.184) сумма по  $I = (ij)$  пробегает по всем парам индексов  $i$  и  $j$ , т. е. сводится к двойной сумме по  $i$  и  $j$ . Однако символы Кронекера в правой части формулы (4.185) указывают на то, что только возбужденные конфигурации  $\Phi_i^A \Phi_j^B$ , в которых либо  $i = 0$ , либо  $j = 0$ , дадут вклад в поправку второго порядка к энергии; т. е. мы должны рассматривать возбуждения либо в подсистеме  $A$ , либо в подсистеме  $B$ , но не в обеих одновременно. Поэтому двойная сумма по  $i$  и  $j$  сведется к двум отдельным суммам — по  $i \neq 0$ , в которой необходимо положить  $j = 0$ , и по  $j \neq 0$ , в которой  $i = 0$ . В обоих случаях в энергетическом знаменателе два слагаемых уничтожаются, так что получаем

$$E^{AB(2)} = \sum_{i \neq 0} \frac{-|W_{0i}^A|^2}{E_i^{A0} - E_0^{A0}} + \sum_{j \neq 0} \frac{-|W_{0j}^B|^2}{E_j^{B0} - E_0^{B0}} = E^{A(2)} + E^{B(2)} \quad (4.186)$$

Это значит, что поправка второго порядка к энергии тоже размерно согласована.

Для вычисления поправок высших порядков к энергии мы воспользуемся формулой (4.24), которая для данного случая имеет вид

$$E^{AB(l)} = \sum_{I \neq (00)} c_I^{(l-1)} W_{(00)I}^{AB} = \sum_{(ij) \neq (00)} c_{(ij)}^{(l-1)} W_{(00)(ij)}^{AB}, \quad (4.187)$$

где  $c_I^{(l-1)} = c_{(ij)}^{(l-1)}$  — коэффициент при  $\Phi_I^{AB} = \Phi_{(ij)}^{AB} = \Phi_i^A \Phi_j^B$  в разложении поправки  $(l-1)$ -го порядка к волновой функции объединенной системы.

Вследствие равенства (4.185), двойная сумма в формуле (4.187) снова сводится к двум отдельным суммам — одной по  $i$  (в которую нужно подставить  $j = 0$ ) и другой по  $j$  (с  $i = 0$ ). Получаем

$$E^{AB(l)} = \sum_{i \neq 0} c_{(i0)}^{(l-1)} W_{0i}^A + \sum_{j \neq 0} c_{(0j)}^{(l-1)} W_{0j}^B. \quad (4.188)$$

Сравнивая с формулой (4.154), мы видим, что  $E^{AB(l)} = E^A(l) + E^B(l)$  выполняется, если

$$c_{(i0)}^{(l-1)} = c_i^{A(l-1)}; \quad c_{(0j)}^{(l-1)} = c_j^{B(l-1)} \quad (4.189)$$

что является достаточным условием размерной согласованности поправки  $l$ -го порядка к энергии. Это условие есть ничто иное, как условие (4.179), записанное для  $(l-1)$ -го, а не для  $l$ -го порядка.

Чтобы доказать выполнение этих условий, необходимо систематически рассмотреть коэффициенты разложения волновой функции. Как будет ясно, чтобы обеспечить выполнение условия (4.179), должны также выполняться соотношения (4.178) и (4.180).

Уравнения (4.28) и (4.29), определяющие коэффициенты разложения поправки первого порядка к волновой функции и поправки  $l$ -го порядка, в данном случае имеют вид

$$c_I^{(1)} = c_{(ij)}^{(1)} = -\frac{W_{I0}^{AB}}{E_I^{AB0} - E_{(00)}^{AB0}} \quad (4.190)$$

и

$$c_I^{(l)} = -\frac{1}{E_I^{AB0} - E_{(00)}^{AB0}} \left[ \sum_{k=1}^{l-1} E^{AB(k)} c_I^{(l-k)} - \sum_{J \neq (00)} c_J^{(l-1)} W_{IJ}^{AB} \right]. \quad (4.191)$$

Подставив  $W_{(00)I}^{AB}$  в выражение (4.190) (оно комплексно сопряжено к  $W_{I(00)}^{AB}$ , выписанному в формуле (4.185)), получим

$$\begin{aligned} c_I^{(1)} = c_{(ij)}^{(1)} &= \frac{-W_{i0}^A \delta_{j0} - W_{j0}^B \delta_{i0}}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \\ &= -\frac{W_{i0}^A}{E_i^{A0} - E_0^{A0}} \delta_{j0} - \frac{W_{j0}^B}{E_j^{B0} - E_0^{B0}} \delta_{i0} = c_i^{A(1)} \delta_{j0} + c_j^{B(1)} \delta_{i0}. \end{aligned} \quad (4.192)$$

Это согласуется с равенствами (4.178)-(4.180). Обозначая в формуле (4.191) составной индекс  $J = (kl)$  и используя выражения (4.183),

(4.192), и (4.181), получим коэффициенты разложения поправки второго порядка к волновой функции:

$$c_I^{(2)} = c_{(ij)}^{(2)} = \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ (W_{00}^A + W_{00}^B)(c_i^{A(1)}\delta_{j0} + c_j^{B(1)}\delta_{i0}) - \sum_{(kl) \neq (00)} (c_k^{A(1)}\delta_{l0} + c_l^{B(1)}\delta_{k0})(W_{ik}^A\delta_{jl} + W_{jl}^B\delta_{ik}) \right]. \quad (4.193)$$

Следует рассмотреть отдельно три случая в формуле (4.193):

а)  $i \neq 0, j = 0$ ; б)  $i = 0, j \neq 0$  и в)  $i, j \neq 0$ .

В случае а)  $i \neq 0$  и  $j = 0$ , поэтому  $\delta_{j0} = 1, \delta_{i0} = 0; E_j^{B0} = E_0^{B0}$ . Получаем

$$c_{(i0)}^{(2)} = \frac{1}{E_i^{A0} + E_0^{A0}} \left[ (W_{00}^A + W_{00}^B)c_i^{A(1)} - \sum_{(kl) \neq (00)} (c_k^{A(1)}W_{ik}^A\delta_{l0}\delta_{0l} + c_l^{B(1)}W_{ik}^A\delta_{k0}\delta_{0l} + c_k^{A(1)}W_{0l}^B\delta_{l0}\delta_{ik} + c_l^{B(1)}W_{0l}^B\delta_{k0}\delta_{ik}) \right]. \quad (4.194)$$

Второе слагаемое в последней сумме исчезает, так как  $k$  и  $l$  не могут равняться нулю одновременно; четвертое слагаемое также пропадает, так как из суммы по  $k$  остается только слагаемое с  $k = i$ , но по условию  $i \neq 0$ . В первом и третьем слагаемом в сумме по  $l$  остаются только слагаемые с  $l = 0$ :

$$c_{(i0)}^{(2)} = \frac{1}{E_i^{A0} - E_0^{A0}} \left[ (W_{00}^A + W_{00}^B)c_i^{A(1)} - \sum_{k \neq 0} c_k^{A(1)}W_{ik}^A - \sum_{k \neq 0} c_k^{A(1)}W_{00}^B\delta_{ik} \right] \quad (4.195)$$

$$= \frac{1}{E_i^{A0} - E_0^{A0}} \left[ W_{00}^A c_i^{A(1)} - \sum_{k \neq 0} c_k^{A(1)}W_{ik}^A \right] = c_i^{A(2)}.$$

Случай б) совершенно аналогичен случаю а) и отличается лишь перестановкой подсистем  $A$  и  $B$  (и индексов  $i$  и  $j$ ). Поэтому тем же способом получаем

$$c_{(0j)}^{(2)} = c_j^{B(2)}. \quad (4.196)$$

Уравнения (4.195) и (4.196) показывают, что поправки второго порядка к волновой функции удовлетворяют условию (4.189), записанному

для  $l = 3$ , поэтому энергия в третьем порядке также размерно согласована.

Чтобы иметь возможность обсуждать размерную согласованность поправок энергии, получаемых в более высоких порядках, нам необходимо рассмотреть также и случай  $\epsilon$ , для которого ни  $i$ , ни  $j$  не равны нулю. В этом случае  $\delta_{j0} = \delta_{i0} = 0$  и необходимо расписать только двойную сумму в формуле (4.193):

$$c_I^{(2)} = c_{(ij)}^{(2)} = \frac{-1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \sum_{(kl) \neq (00)} (c_k^{A(1)} W_{ik}^A \delta_{l0} \delta_{jl} + c_l^{B(1)} W_{ik}^A \delta_{k0} \delta_{jl} + c_k^{A(1)} W_{jl}^B \delta_{l0} \delta_{ik} + c_l^{B(1)} W_{jl}^B \delta_{k0} \delta_{ik}). \quad (4.197)$$

Первый и четвертый слагаемые исчезают, так как суммы по  $l$  и  $k$  сводятся к  $\delta_{j0}$  и  $\delta_{i0}$ , соответственно, но мы рассматриваем случай  $i, j \neq 0$ . Во втором и третьем слагаемом в сумме остаются только вклады с  $k = 0$ ,  $l = j$  и  $k = i$ ,  $l = 0$ , соответственно, что дает

$$c_I^{(2)} = c_{(ij)}^{(2)} = \frac{-1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left( c_j^{B(1)} W_{i0}^A + c_i^{A(1)} W_{j0}^B \right) \quad (4.198)$$

Подставляя сюда уже известные выражения  $c_j^{B(1)} = -\frac{W_{j0}^B}{E_j^{B0} - E_0^{B0}}$  и  $c_i^{A(1)} = -\frac{W_{i0}^A}{E_i^{A0} - E_0^{A0}}$ , получим, приводя к общему знаменателю

$$c_I^{(2)} = c_{(ij)}^{(2)} = \frac{W_{i0}^A W_{j0}^B}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left( \frac{1}{E_j^{B0} - E_0^{B0}} + \frac{1}{E_i^{A0} - E_0^{A0}} \right) = \frac{W_{i0}^A W_{j0}^B}{(E_j^{B0} - E_0^{B0})(E_i^{A0} - E_0^{A0})} = c_i^{A(1)} c_j^{B(1)}. \quad (4.199)$$

что является частным случаем выражения (4.180) для  $l = 2$ .

Чтобы доказать, что равенства (4.179) и (4.180) выполняются в любом порядке теории возмущений, воспользуемся индукцией.

#### Доказательство по индукции

Мы убедились, что поправки первого и второго порядков к волновой функции подчиняются равенствам (4.178)-(4.180). Здесь мы покажем, что если эти равенства выполняются вплоть до некоторого порядка возмущений  $t - 1$ , то они также верны в  $t$ -м порядке; поэтому следует по индукции, что они должны выполняться и в любом порядке. Таким образом окажется, что разложение теории возмущений Рэлея—Шрёдингера размерно согласованно в любом порядке.

Проследим за коэффициентами  $c_{(ij)}^{(t)}$  с помощью формулы разложения (4.29) и предполагая выполнение соотношений (4.178)-(4.180) — т. е. размерной согласованности, — во всех порядках до  $(t-1)$ -го включительно. Общая формула (4.29) в данном случае выглядит так:

$$\begin{aligned} c_I^{(t)} &= c_{(ij)}^{(t)} = \frac{1}{E_I^{AB0} - E_0^{AB0}} \left[ \sum_{s=1}^{t-1} E^{AB(s)} c_I^{(t-s)} - \sum_{J \neq (00)} c_J^{(t-1)} W_{IJ}^{AB} \right] \\ &= \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-1} E^{AB(s)} c_{(ij)}^{(t-s)} \right. \\ &\quad \left. - \sum_{(kl) \neq (00)} c_{(kl)}^{(t-1)} W_{(ij)(kl)}^{AB} \right]. \end{aligned} \quad (4.200)$$

Имеем те же три случая а), б) и в), что и ранее для  $c_{(ij)}^{(2)}$ .

Случай а)  $j = 0$ ;  $i \neq 0$

$$\begin{aligned} c_{(i0)}^{(t)} &= \frac{1}{E_i^{A0} - E_0^{A0}} \left[ \sum_{s=1}^{t-1} \left( E^{A(s)} + E^{B(s)} \right) c_{(i0)}^{(t-s)} \right. \\ &\quad \left. + \sum_{(kl) \neq (00)} c_{(kl)}^{(t-1)} W_{(i0)(kl)}^{AB} \right]. \end{aligned} \quad (4.201)$$

Двойная сумма распадается на три: одна для  $k \neq 0, l = 0$ ; одна для  $k = 0, l \neq 0$ ; и одна для  $k, l \neq 0$ :

$$\begin{aligned} c_{(i0)}^{(t)} &= \frac{1}{E_i^{A0} - E_0^{A0}} \left[ \sum_{s=1}^{t-1} E^{A(s)} c_{(i0)}^{(t-s)} + \sum_{s=1}^{t-1} E^{B(s)} c_{(i0)}^{(t-s)} \right. \\ &\quad - \sum_{k \neq 0} c_{(k0)}^{(t-1)} (W_{ik}^A + W_{00}^B \delta_{ik}) - \sum_{l \neq 0} c_{(0l)}^{(t-1)} (W_{i0}^A \delta_{0l} + W_{0l}^B \delta_{i0}) \\ &\quad \left. - \sum_{k, l \neq 0} c_{(kl)}^{(t-1)} (W_{ik}^A \delta_{0l} + W_{0l}^B \delta_{ik}) \right]. \end{aligned} \quad (4.202)$$

Так как мы рассматриваем случай  $j = 0$ , в выражениях (4.201)-(4.202)  $j$  везде заменено на 0. Некоторые слагаемые в формуле (4.202) исчезают из-за символов Кронекера, а каждая из двух сумм по  $k$  сводятся к единственному слагаемому с  $k = i$ . Слагаемое, содержащее  $W_{00}^B = E^{B(1)}$  сокращается со слагаемым  $s = 1$  во второй сумме. Подста-

виз  $c_{(i0)}^{(p)} = c_i^{A(p)}$ ,  $c_{(0l)}^{(p)} = c_l^{B(p)}$ , и  $c_{(il)}^{(t-1)} = \sum_{q=1}^{t-2} c_i^{A(t-1-q)} c_l^{B(q)}$  в соответствии с условиями (4.179) и (4.180), имеем

$$c_{(i0)}^{(t)} = \frac{1}{E_i^{A0} - E_0^{A0}} \left[ \sum_{s=1}^{t-1} E^{A(s)} c_i^{A(t-s)} - \sum_{k \neq 0} c_k^{A(t-1)} W_{ik}^A + \sum_{s=2}^{t-1} E^{B(s)} c_i^{A(t-s)} - \sum_{l \neq 0} \sum_{q=1}^{t-2} c_i^{A(t-1-q)} c_l^{B(q)} W_{0l}^B \right]. \quad (4.203)$$

Используя формулу (4.29) и заменив индекс суммирования в последней сумме на  $s = q + 1$ , получим

$$c_{(i0)}^{(t)} = c_i^{A(t)} + \frac{1}{E_i^{A0} - E_0^{A0}} \left[ \sum_{s=2}^{t-1} E^{B(s)} c_i^{A(t-s)} - \sum_{s=2}^{t-1} c_i^{A(t-s)} \sum_{l \neq 0} c_l^{B(s-1)} W_{0l}^B \right] = c_i^{A(t)}, \quad (4.204)$$

где мы учли, что сумма по  $l$  дает  $E^{B(s)}$  в соответствии с формулой (4.24) и тогда две последние суммы сокращаются.

Случай б)  $i = 0$ ;  $j \neq 0$

Случай б) совершенно аналогичен случаю а) и отличается только перестановкой подсистем  $A$  и  $B$  (и индексов  $i$  и  $j$ ). Получаем

$$c_{(0j)}^{(t)} = c_j^{B(t)} \quad (4.205)$$

Случай в)  $i \neq 0$ ;  $j \neq 0$

Согласно формуле (4.200), имеем

$$\begin{aligned} c_{(ij)}^{(t)} &= \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-2} (E^{A(s)} + E^{B(s)}) c_{(ij)}^{(t-s)} - \sum_{(kl) \neq (00)} c_{(kl)}^{(t-1)} W_{(ij)(kl)}^{AB} \right] \\ &= \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-2} (E^{A(s)} + E^{B(s)}) c_{(ij)}^{(t-s)} - c_i^{A(t-1)} W_{j0}^B - c_j^{B(t-1)} W_{i0}^A - \sum_{k \neq 0} c_{(kj)}^{(t-1)} W_{ik}^A - \sum_{l \neq 0} c_{(il)}^{(t-1)} W_{jl}^B \right], \end{aligned} \quad (4.206)$$

где учтено, что для  $s \leq t-1$  уже имеет место размерная согласованность энергии  $E^{AB}(s) = E^{A(s)} + E^{B(s)}$ , а слагаемое с  $s = t-1$  в первой сумме содержит  $c_{(ij)}^{(t-t+1)} = c_{(ij)}^{(1)} = 0$ , так что его можно опустить ( $i, j \neq 0$ ). Так же, как и в случае  $a$ ), мы разделили двойную сумму на три суммы, подставили далее выражение (4.179) для коэффициентов  $c_{(i0)}^{(t-1)}$  и  $c_{(0j)}^{(t-1)}$ , выражение (4.181) для  $W_{(ij)(kl)}^{AB}$ , и просуммировали с учетом символов Кронекера.

Теперь подставим значения (4.180) для коэффициентов  $c_{(ij)}^{(t-s)}$  и  $c_{(ij)}^{(t-1)}$ , выбрав подходящие из двух эквивалентных представлений:



$$\begin{aligned}
 c_{(ij)}^{(t)} = & \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-2} E^{A(s)} \sum_{q=1}^{t-s-1} c_i^{A(t-s-q)} c_j^{B(q)} \right. \\
 & + \sum_{s=1}^{t-2} E^{B(s)} \sum_{p=1}^{t-s-1} c_j^{B(t-s-p)} c_i^{A(p)} - W_{j0}^B c_i^{A(t-1)} - W_{i0}^A c_j^{B(t-1)} \\
 & \left. - \sum_{k \neq 0} W_{ik}^A \sum_{q=1}^{t-2} c_k^{A(t-1-q)} c_j^{B(q)} - \sum_{l \neq 0} W_{jl}^B \sum_{p=1}^{t-2} c_l^{B(t-1-p)} c_i^{A(p)} \right]. \quad (4.207)
 \end{aligned}$$

Выделим слагаемое с  $q = 1$  или  $p = 1$  из первых двух сумм; тогда верхний предел оставшихся сумм можно заменить на  $s = t-3$ , так как для  $s = t-2$  индекс  $q$  ( $p$ ) принимает только значение  $q = 1$  ( $p = 1$ ), которое мы уже выделили. Подставим также разложения  $c_i^{A(t-1)}$  и  $c_j^{B(t-1)}$  согласно результату (4.29):

$$\begin{aligned}
 c_{(ij)}^{(t)} = & \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left\{ \sum_{s=1}^{t-2} E^{A(s)} c_i^{A(t-1-s)} c_j^{B(1)} \right. \\
 & + \sum_{s=1}^{t-3} E^{A(s)} \sum_{q=2}^{t-s-1} c_i^{A(t-s-q)} c_j^{B(q)} \\
 & + \sum_{s=1}^{t-2} E^{B(s)} c_j^{B(t-1-s)} c_i^{A(1)} + \sum_{s=1}^{t-3} E^{B(s)} \sum_{p=2}^{t-s-1} c_j^{B(t-s-p)} c_i^{A(p)} \\
 & \left. - \frac{W_{j0}^B}{E_i^{A0} - E_0^{A0}} \left[ \sum_{s=1}^{t-2} E^{A(s)} c_i^{A(t-1-s)} - \sum_{k \neq 0} c_k^{A(t-2)} W_{ik}^A \right] \right\} \quad (4.208)
 \end{aligned}$$



$$\begin{aligned}
& - \frac{W_{i0}^A}{E_j^{B0} - E_0^{B0}} \left[ \sum_{s=1}^{t-2} E^{B(s)} c_j^{B(t-1-s)} - \sum_{l \neq 0} c_l^{B(t-2)} W_{jl}^B \right] \\
& - \sum_{k \neq 0} W_{ik}^A \sum_{q=1}^{t-2} c_k^{A(t-1-q)} c_j^{B(q)} - \sum_{l \neq 0} W_{jl}^B \sum_{p=1}^{t-2} c_l^{B(t-1-p)} c_i^{A(p)} \Bigg\}.
\end{aligned}$$

ЛАНЬ®

Первое и третье слагаемое содержат те же суммы, что и первая сумма в пятом и шестом слагаемом, соответственно. После подстановки явных выражений  $c_i^{A(1)} = \frac{-W_{i0}^A}{E_i^{A0} - E_0^{A0}}$  и  $c_j^{B(1)} = \frac{-W_{j0}^B}{E_j^{B0} - E_0^{B0}}$ , можно привести их к общему знаменателю. Далее мы отделяем слагаемое  $q = 1$  и  $p = 1$  в первых двух суммах и подставляем в них выражения для  $c_i^{A(1)}$  и  $c_j^{B(1)}$ :

$$\begin{aligned}
c_{(ij)}^{(t)} &= \frac{-1}{(E_i^{A0} - E_0^{A0})(E_j^{B0} - E_0^{B0})} \left[ \sum_{s=1}^{t-2} E^{A(s)} c_i^{A(t-1-s)} W_{j0}^B \right. \\
& \quad \left. + \sum_{s=1}^{t-2} E^{B(s)} c_j^{B(t-1-s)} W_{i0}^A \right] \\
& + \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-3} E^{A(s)} \sum_{q=2}^{t-s-1} c_i^{A(t-s-q)} c_j^{B(q)} \right. \\
& \quad \left. + \sum_{s=1}^{t-3} E^{B(s)} \sum_{p=2}^{t-s-1} c_j^{B(t-s-p)} c_i^{A(p)} \right] \quad (4.209) \\
& + \frac{W_{i0}^A}{E_j^{B0} - E_0^{B0}} \sum_{l \neq 0} c_l^{B(t-2)} W_{jl}^B + \frac{W_{j0}^B}{E_i^{A0} - E_0^{A0}} \sum_{k \neq 0} c_k^{A(t-2)} W_{ik}^A \\
& - \sum_{k \neq 0} W_{ik}^A c_k^{A(t-2)} \frac{-W_{j0}^B}{E_j^{B0} - E_0^{B0}} - \sum_{k \neq 0} W_{ik}^A \sum_{q=2}^{t-2} c_k^{A(t-1-q)} c_j^{B(q)} \\
& - \sum_{l \neq 0} W_{jl}^B c_l^{B(t-2)} \frac{-W_{i0}^A}{E_i^{A0} - E_0^{A0}} - \sum_{l \neq 0} W_{jl}^B \sum_{p=2}^{t-2} c_l^{B(t-1-p)} c_i^{A(p)} \Bigg].
\end{aligned}$$

ЛАНЬ®

Теперь заметим, что первое и второе слагаемое содержат определения  $c_j^{B(1)}$  и  $c_i^{A(1)}$ , соответственно. Вклады, содержащие произведения  $W_{i0}^A W_{jl}^B$

и  $W_{j0}^B W_{ik}^A$ , можно привести к общему знаменателю:

$$\begin{aligned}
 c_{(ij)}^{(t)} = & \sum_{s=1}^{t-2} E^{A(s)} c_i^{A(t-1-s)} \frac{c_j^{B(1)}}{E_i^{A0} - E_0^{A0}} + \sum_{s=1}^{t-2} E^{B(s)} c_j^{B(t-1-s)} \frac{c_i^{A(1)}}{E_j^{B0} - E_0^{B0}} \\
 & + \frac{1}{(E_i^{A0} - E_0^{A0})(E_j^{B0} - E_0^{B0})} \left[ W_{i0}^A \sum_{l \neq 0} c_l^{B(t-2)} W_{jl}^B + W_{j0}^B \sum_{k \neq 0} c_k^{A(t-2)} W_{ik}^A \right] \\
 & + \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-3} E^{A(s)} \sum_{q=2}^{t-s-1} c_i^{A(t-s-q)} c_j^{B(q)} \right. \\
 & + \sum_{s=1}^{t-3} E^{B(s)} \sum_{p=2}^{t-s-1} c_j^{B(t-s-p)} c_i^{A(p)} - \sum_{k \neq 0} W_{ik}^A \sum_{q=2}^{t-2} c_k^{A(t-1-q)} c_j^{B(q)} \\
 & \left. - \sum_{l \neq 0} W_{jl}^B \sum_{p=2}^{t-2} c_l^{B(t-1-p)} c_i^{A(p)} \right]. \quad (4.210)
 \end{aligned}$$

В третьем слагаемом мы снова видим выражения для коэффициентов  $c_i^{A(1)}$  и  $c_j^{B(1)}$ , которые можно скомбинировать со вторым и первым слагаемыми, соответственно:

$$\begin{aligned}
 c_{(ij)}^{(t)} = & \frac{c_j^{B(1)}}{E_i^{A0} - E_0^{A0}} \left[ \sum_{s=1}^{t-2} E^{A(s)} c_i^{A(t-1-s)} - \sum_{k \neq 0} c_k^{A(t-2)} W_{ik}^A \right] \\
 & + \frac{c_i^{A(1)}}{E_j^{B0} - E_0^{B0}} \left[ \sum_{s=1}^{t-2} E^{B(s)} c_j^{B(t-1-s)} - \sum_{l \neq 0} c_l^{B(t-2)} W_{jl}^B \right] \\
 & + \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ \sum_{s=1}^{t-3} E^{A(s)} \sum_{q=2}^{t-s-1} c_i^{A(t-s-q)} c_j^{B(q)} \right. \\
 & - \sum_{k \neq 0} W_{ik}^A \sum_{q=2}^{t-2} c_k^{A(t-1-q)} c_j^{B(q)} \\
 & \left. + \sum_{s=1}^{t-3} E^{B(s)} \sum_{p=2}^{t-s-1} c_j^{B(t-s-p)} c_i^{A(p)} - \sum_{l \neq 0} W_{jl}^B \sum_{p=2}^{t-2} c_l^{B(t-1-p)} c_i^{A(p)} \right]. \quad (4.211)
 \end{aligned}$$

Первые два слагаемых содержат определение  $c_i^{A(t-1)}$  и  $c_j^{B(t-1)}$ , в соответствии с результатом (4.29). В третьем слагаемом суммирование идет по всем значениям  $1 \leq s \leq t-3$  и  $2 \leq q \leq t-2$ , которые удовлетворяют условию  $t-s-q \geq 1$ ; можно переписать его, как сумму по  $2 \leq q \leq t-2$

и  $1 \leq s \leq t - q - 1$ . (То же самое верно для пятого слагаемого, если заменить  $p$  на  $q$ .) Получаем

$$\begin{aligned}
 c_{(ij)}^{(t)} &= c_j^{B(1)} c_i^{A(t-1)} + c_i^{A(1)} c_j^{B(t-1)} \\
 &+ \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left\{ \sum_{q=2}^{t-2} \left[ \sum_{s=1}^{t-q-1} E^{A(s)} c_i^{A(t-s-q)} \right. \right. \\
 &\quad \left. \left. - \sum_{k \neq 0} W_{ik}^A c_k^{A(t-1-q)} \right] c_j^{B(q)} \right. \\
 &\left. + \sum_{q=2}^{t-2} \left[ \sum_{s=1}^{t-q-1} E^{B(s)} c_j^{B(t-s-q)} - \sum_{l \neq 0} W_{jl}^B c_l^{B(t-1-q)} \right] c_i^{A(q)} \right\}.
 \end{aligned} \tag{4.212}$$

Умножив и разделив на  $E_i^{A0} - E_0^{A0}$  и  $E_j^{B0} - E_0^{B0}$ , можно выделить формулу (4.29) для  $c_i^{A(t-q)}$  и  $c_j^{B(t-q)}$ :

$$\begin{aligned}
 c_{(ij)}^{(t)} &= c_j^{B(1)} c_i^{A(t-1)} + c_i^{A(1)} c_j^{B(t-1)} \\
 &+ \frac{1}{E_i^{A0} + E_j^{B0} - E_0^{A0} - E_0^{B0}} \left[ (E_i^{A0} - E_0^{A0}) \sum_{q=2}^{t-2} c_i^{A(t-q)} c_j^{B(q)} \right. \\
 &\quad \left. + (E_j^{B0} - E_0^{B0}) \sum_{q=2}^{t-2} c_j^{B(t-q)} c_i^{A(q)} \right].
 \end{aligned} \tag{4.213}$$

Две суммы в квадратных скобках на самом деле равны — одна из них идет в восходящем порядке, другая — в нисходящем, и знаменатель сокращается:

$$\begin{aligned}
 c_{(ij)}^{(t)} &= c_j^{B(1)} c_i^{A(t-1)} + c_i^{A(1)} c_j^{B(t-1)} + \sum_{q=2}^{t-2} c_i^{A(t-q)} c_j^{B(q)} \\
 &= \sum_{q=1}^{t-1} c_i^{A(t-q)} c_j^{B(q)}. \text{ Ч. т. д.}
 \end{aligned} \tag{4.214}$$

## 6. Метод разбиения по Лёвдину

Рассмотрим вековое уравнение в  $m$ -мерном линейном пространстве ортонормированных базисных функций  $\{\varphi_i\}$ :

$$\mathbf{H}\mathbf{c} = E\mathbf{c}.$$

Здесь  $\mathbf{H}$  является эрмитовой матрицей размера  $m \times m$ , а  $\mathbf{c}$  —  $m$ -мерный вектор. Метод разбиения по Лёвдину (Löwdin partition) позволяет, по крайней мере формально, понизить размерность векового уравнения (4.215). Эта возможность представляет концептуальный интерес; кроме того, результаты, получаемые при разбиении, можно использовать как единую исходную точку для большинства схем теории возмущений.

Разделим  $m$  базисных функций  $\varphi_i$  на две группы. Первая группа — которую мы будем называть «первичной», — состоит из  $n$  функций, по той или иной причине более интересных нам (подпространство этих функций часто называют «модельным пространством»). Оставшиеся  $m - n$  функций можно, соответственно, назвать «вторичными» (или «внешними»).

Из математики известно, что матрицы можно разбить на блоки и обычные правила матричного умножения будут пригодны для таких блочных матриц. Мы припишем индекс «1» набору функций в первичном подпространстве и индекс «2» для функций вторичного подпространства. Тогда матричное уравнение (4.215) преобразуется к виду

$$\begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix} = E \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}. \quad (4.216)$$

Здесь  $\mathbf{c}_1$  —  $n$ -мерный вектор, составленный из первых  $n$  элементов вектора  $\mathbf{c}$ ; оставшиеся  $m - n$  элементов  $\mathbf{c}$  образуют вектор  $\mathbf{c}_2$ . Матрицы  $\mathbf{H}_{11}$  и  $\mathbf{H}_{22}$  содержат матричные элементы  $\mathbf{H}_{ij}$  верхнего левого и нижнего правого блоков исходной матрицы  $\mathbf{H}$ , соответственно, и описывают взаимодействия между функциями внутри одного и того же (первичного или вторичного) пространств. Очевидно,  $\mathbf{H}_{11}$  и  $\mathbf{H}_{22}$  эрмитовы и имеют размерности  $n \times n$  и  $(m - n) \times (m - n)$ , соответственно. Матричные элементы  $\mathbf{H}$ , связывающие первичное и вторичное подпространства, образуют блоки  $\mathbf{H}_{12}$  и  $\mathbf{H}_{21}$  размерностей  $n \times (m - n)$  и  $(m - n) \times n$ . Вследствие эрмитовости исходной матрицы  $\mathbf{H}$ , эти матрицы связаны соотношением

$$\mathbf{H}_{21} = \mathbf{H}_{12}^\dagger. \quad (4.217)$$

Теперь перепишем уравнение (4.216) в развернутом виде

$$\begin{aligned} \mathbf{H}_{11}\mathbf{c}_1 + \mathbf{H}_{12}\mathbf{c}_2 &= E\mathbf{c}_1 \\ \mathbf{H}_{21}\mathbf{c}_1 + \mathbf{H}_{22}\mathbf{c}_2 &= E\mathbf{c}_2 \end{aligned} \quad (4.218)$$

и решим формально второе уравнение относительно  $\mathbf{c}_2$ . Имеем

$$(\mathbf{H}_{22} - E)\mathbf{c}_2 = -\mathbf{H}_{21}\mathbf{c}_1, \quad (4.219)$$

$$\text{т. е.} \quad \mathbf{c}_2 = -(\mathbf{H}_{22} - E)^{-1} \mathbf{H}_{21} \mathbf{c}_1. \quad (4.220)$$

Подставим этот результат в первое уравнение системы (4.218):

$$\mathbf{H}_{11} \mathbf{c}_1 - \mathbf{H}_{12} (\mathbf{H}_{22} - E)^{-1} \mathbf{H}_{21} \mathbf{c}_1 = E \mathbf{c}_1, \quad (4.221)$$

$$\text{т. е.} \quad [\mathbf{H}_{11} - \mathbf{H}_{12} (\mathbf{H}_{22} - E)^{-1} \mathbf{H}_{21}] \mathbf{c}_1 = E \mathbf{c}_1. \quad (4.222)$$

Уравнение (4.222) представляет собой задачу на собственные значения матрицы *эффективного* гамильтониана, имеющей размер  $n \times n$ ,

$$\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{eff}}(E) = \mathbf{H}_{11} - \mathbf{H}_{12} (\mathbf{H}_{22} - E)^{-1} \mathbf{H}_{21}. \quad (4.223)$$

Важно подчеркнуть, что уравнение (4.222) полностью эквивалентно исходному вековому уравнению (4.215) для всех собственных векторов последнего, которые не строго ортогональны первичному подпространству, т. е. для которых  $\mathbf{c}_1 \neq \mathbf{0}$ . (Легко видеть, что равенства  $\mathbf{c}_1 = \mathbf{0}$  или  $\mathbf{c}_2 = \mathbf{0}$  возможны, только если матрица  $\mathbf{H}$  блочно-диагональна, т. е.  $\mathbf{H}_{12} = \mathbf{H}_{21}^\dagger = \mathbf{0}$ . Тогда решения можно найти из уравнений на собственные значения в первичном и вторичном подпространствах независимо друг от друга.)

Имеется две трудности, связанные с эффективным гамильтонианом (4.223). Первая та, что он явно зависит от искомого собственного значения  $E$ . Поэтому уравнение (4.222) не является настоящим уравнением на собственные значения, а лишь уравнением на «псевдособственные значения» — его можно решить только итерационно. Еще большей помехой является то, что надо уметь вычислять обратную матрицу  $(\mathbf{H}_{22} - E)^{-1}$  — точнее, обратную к матрице  $\mathbf{H}_{22} - E \mathbf{1}_{22}$ , где  $\mathbf{1}_{22}$  есть единичная матрица размера  $(m - n) \times (m - n)$ , соответствующая блоку 2, — необходимую для описания влияния вторичного подпространства на задачу на собственные значения в первичном подпространстве; более того, эта обратная матрица является функцией  $E$ .

Эффективный гамильтониан  $\mathbf{H}_{\text{eff}}$  является матрицей размера  $n \times n$ , так что для любого значения  $E$ , для которого матрица  $\mathbf{H}_{22} - E \mathbf{1}$  не сингулярна, получаем  $n$  собственных векторов и собственных значений. Очевидно, только те решения уравнения (4.222) имеют смысл, для которых найденные собственные значения равны величине  $E$ , использованной для построения матрицы  $\mathbf{H}_{\text{eff}}(E)$ . Если мы рассматриваем невырожденное решение задачи (4.215), то имеется лишь одно такое решение. Однако, как отмечено ранее, все решения задачи (4.215) могут быть получены, по крайней мере в принципе, подстановкой разных значений  $E$  в уравнения на псевдособственные значения (4.222) и (4.220).

В зависимости от выбора первичного подпространства и различных разложений (приближений) для  $(\mathbf{H}_{22} - E)^{-1}$ , схема разбиения по Лёвдину может быть использована при выводе разнообразных выражений теории возмущений, предназначенных для приближенного решения задач на собственные значения (4.215).

### Пример

В качестве простого примера рассмотрим случай  $n = 1$ , т. е. первичное подпространство состоит из единственной функции  $\varphi_1$ . Тогда можно предположить, что  $c_1 = 1$  (это эквивалентно использованию промежуточной нормировки в ТВ Рэлея—Шрёдингера) и, расписывая матричное произведение, получаем из (4.222)

$$E = H_{11} - \sum_{\substack{j,k \\ (j,k \neq 1)}} H_{1j} [(\mathbf{H}_{22} - E\mathbf{1})^{-1}]_{jk} H_{k1}. \quad (4.224)$$

Теперь аппроксимируем обратную матрицу в выражении (4.224), учитывая только диагональную часть  $\mathbf{H}_{22}$  и считая  $E$  приближенно равным  $H_{11}$ . Тогда

$$[(\mathbf{H}_{22} - E\mathbf{1})^{-1}]_{jk} \approx (H_{jj} - H_{11})^{-1} \delta_{jk} \quad (4.225)$$

и получим

$$E \approx H_{11} - \sum_{\substack{j,k \\ (j,k \neq 1)}} \frac{H_{1j} \delta_{jk} H_{k1}}{H_{jj} - H_{11}} = H_{11} - \sum_{\substack{j \\ (j \neq 1)}} \frac{|H_{1j}|^2}{H_{jj} - H_{11}}. \quad (4.226)$$

Это есть ничто иное, как формула ТВ Рэлея—Шрёдингера второго порядка для случая, когда невозмущенный гамильтониан состоит из диагональной части  $\mathbf{H}$ , а недиагональная часть рассматривается, как возмущение. (Это так называемое разбиение гамильтониана на невозмущенную и возмущенную части по схеме Эпштейна—Несбета.)

В качестве другого примера, воспользуемся тождеством

$$(\mathbf{A} + \mathbf{B})^{-1} = (\mathbf{1} + \mathbf{A}^{-1}\mathbf{B})^{-1} \mathbf{A}^{-1} \quad (4.227)$$

которое можно элементарно проверить умножением правой части справа на  $(\mathbf{A} + \mathbf{B})$ :

$$(\mathbf{1} + \mathbf{A}^{-1}\mathbf{B})^{-1} \mathbf{A}^{-1} (\mathbf{A} + \mathbf{B}) = (\mathbf{1} + \mathbf{A}^{-1}\mathbf{B})^{-1} (\mathbf{1} + \mathbf{A}^{-1}\mathbf{B}) = \mathbf{1} \quad (4.228)$$

Обозначим  $\Delta E = E - E_0$ , где  $E_0$  есть (любое) приближение нулевого порядка к энергии  $E$ , и определим  $\mathbf{A}$  и  $\mathbf{B}$  как

$$\mathbf{A} = \mathbf{H}_{22} - E_0 \mathbf{1}_{22} \quad (4.229)$$

$$\mathbf{B} = -\Delta E \mathbf{1}_{22} ,$$

и используем тождество (4.227), чтобы выразить  $(\mathbf{H}_{22} - E)^{-1} = (\mathbf{A} + \mathbf{B})^{-1}$ . Учитывая, что  $\mathbf{B}$  пропорционально единичной матрице, получим

$$(\mathbf{H}_{22} - E)^{-1} = [\mathbf{1}_{22} - \Delta E (\mathbf{H}_{22} - E_0 \mathbf{1}_{22})^{-1}] (\mathbf{H}_{22} - E_0 \mathbf{1}_{22})^{-1} . \quad (4.230)$$

Обозначив

$$\mathbf{R} = (\mathbf{H}_{22} - E_0 \mathbf{1}_{22})^{-1} , \quad (4.231)$$

приводим матрицу эффективного гамильтониана (4.223) к виду

$$\mathbf{H}_{\text{eff}} = \mathbf{H}_{11} - \mathbf{H}_{12} (\mathbf{1}_{22} - \Delta E \mathbf{R})^{-1} \mathbf{R} \mathbf{H}_{21} . \quad (4.232)$$

Можно видеть, что второе слагаемое в скобках является малой поправкой, при условии, что приближение нулевого порядка  $E_0$  не очень грубое. Такая форма удобна для итераций и также может быть использована как исходное выражение для вывода различных разложений в ряды ТВ.

В частном случае одномерного модельного пространства и выбора  $E_0 = H_{11}$ , получаем уравнение для сдвига энергии  $\Delta E$

$$\Delta E = -\mathbf{H}_{12} (\mathbf{1}_{22} - \Delta E \mathbf{R})^{-1} \mathbf{R} \mathbf{H}_{21} . \quad (4.233)$$

## Библиографические заметки

### Раздел 1.

Классическая техника теории возмущений, развитая лордом Рэлеем, была приспособлена Шрёдингером к квантовой механике [1]. (В качестве другого раннего — и, по-видимому, независимого — использования разложения по степеням возмущений упомянем статью Борна и Оппенгеймера [2].)

Изложение, предложенное в этой главе, в основном является прямым обобщением подхода, с помощью которого самого автора обучали теории возмущений по классическому учебнику Блохинцева [3]. Это, в частности, относится и к способу введения параметра возмущения  $\lambda$  в уравнении (4.3).

### Раздел 1.2.

Рекурсивный характер теории возмущений Рэля—Шрёдингера был использован Ноулсом и соавторами [4] для вычислений высших порядков теории возмущений для некоторых малых модельных систем. Они использовали сокращенные обозначения, напоминающие наши «компактные матричные обозначения».

Явные выражения для коэффициентов  $c_{ij}^{(l)}$  понадобятся нам для изучения проблемы размерной согласованности в разд. 5.2 (эти формулы имеются также, например, в [5]).

### Раздел 1.3.

Стандартный подход, см. например, [6].

### Раздел 1.4.

Как отмечено в разд. 1.2, теорема о точности среднего значения также следует из свойств стационарности точной волновой функции. Доказательство, приводимое здесь, следует изложению в [7].

Возможность выразить поправки к точной энергии до порядка  $2n+1$  с использованием поправок к волновой функции до  $n$ -го порядка известна с 50-х годов [8, 9] (также см. [6, 10]). Доказательство, приведенное здесь, в основном следует [11] (автор признателен Д-ру Бесалу за предоставление ему копии своей неопубликованной работы).

### Раздел 2.

Впервые предложено Хиллераасом [12]. Здесь мы приводим стандартное изложение, см. например [13]. Для обобщения на случай неэрмитова невозмущенного гамильтониана см. [14].

### Раздел 3.

Стандартный материал, см. например в [15]

### Раздел 4.

Впервые предложено в [16,17]. Здесь мы приводим стандартное изложение, следующее, например, [18].

### Раздел 5.

Размерная согласованность теории возмущений Рэлея—Шрёдингера обычно доказывается с привлечением диаграммных методов [20]. Наши результаты показывают, что это не обязательно.

Рассмотрение, основанное на единственности ряда теории возмущений (разд. 5.1), возможно, является оригинальным (но не претендует на то, чтобы быть таковым). Доказательство по индукции, приведенное в разд. 5.2, скорее всего, оригинально.

### Раздел 6.

Впервые введено в [19]. Может рассматриваться, как теоретическая основа любого формализма эффективного гамильтониана, включая теорию магнетизма Гейзенберга.

## Для дальнейшего изучения

Рекомендуется исчерпывающая серия работ Лёвдина «*Studies in Perturbation Theory*», части с I по XIV [21-33], равно как и замечательный обзор Хиршфельдера с соавт. [6].

## Литература

1. Schrödinger E. Ann. Phys. **80**, 437 (1926).
2. Born M., Oppenheimer J. R. Ann. Phys. **84**, 457 (1927).



3. Блохинцев Д. И. *Основы квантовой механики*. — М.: Высшая школа, 1963.
4. Knowles P. J., Somasundram S., Handy N. C., Hirao K. *Chem. Phys. Letters* **113**, 8 (1985).
5. Mayer I. *Fejezetek a Kvantumkémiából*, BME Mérnöktoábbképző Int., Budapest 1987.
6. Hirschfelder J. O., Byers Brown W., Epstein S. T. *Adv. Quantum Chem.* **1**, 255 (1964).
7. Zülicke L. *Quantenchemie*, Deutcher Verlag der Wissenschaften, Berlin 1973. (Имеется русский перевод: Цюлике Л. *Квантовая химия*. — М.: Мир, 1976).
8. Dalgarno A., Stewart A. L. *Proc. Roy. Soc (London)* **A 238**, 269 (1956).
9. Dupont-Bourdelet F., Tillieu J., Guy J. *J. Phys. Radium* **21**, 776 (1960).
10. Brändas E., Goscinski O. *Phys. Rev.* **A 1** 552 (1970).
11. Besalú E. Institute of Computational Chemistry, University of Girona, Girona, Spain (неопубликовано).
12. Hylleraas E. A. *Z. Physik* **65**, 209 (1930).
13. Epstein S. T. *The Variation Method in Quantum Chemistry*, Academic Press, New York 1974. (Имеется русский перевод: Эпштейн С. *Вариационный метод в квантовой химии*. — М.: Мир, 1976).
14. Mayer I. *Mol. Phys.* **89**, 515 (1996).
15. Ландау Л. Д., Лифшиц Е. М. *Квантовая механика*. — М.: Наука, 1974.
16. Brillouin L. *J. Phys. Radium (7)* **3**, 373 (1932).
17. Wigner E. *Math. u. naturw. Anz. ungar Akad. Wiss.* **53**, 477 (1935).
18. Ziman J. M. *Elements of Advanced Quantum Theory*, University Press, Cambridge 1969. (Имеется русский перевод: Займан Дж. *Современная квантовая теория*. — М.: Мир, 1971).
19. Löwdin P.-O. *J. Chem. Phys.* **19**, 1396 (1951).
20. Jørgensen P., Simons J. *Second Quantization-Based Methods in Quantum Chemistry*, Academic Press, Mew York 1981.
21. Löwdin P.-O. *J. Mol. Spectry*, **10**, 12 (1963).
22. Löwdin P.-O. *J. Mol. Spectry*, **13**, 326 (1964). (Части II и III серии работ.)
23. Löwdin P.-O. *J. Math. Phys.* **3**, 969 (1962).
24. Löwdin P.-O. *J. Math. Phys.* **3**, 1171 (1962).
25. Löwdin P.-O. *J. Mol. Spectry*, **14**, 112 (1964).
26. Löwdin P.-O. *J. Mol. Spectry*, **14**, 119 (1964).
27. Löwdin P.-O. *J. Mol. Spectry*, **14**, 131 (1964).
28. Löwdin P.-O. *J. Math. Phys.* **6** 1341 (1965).
29. Löwdin P.-O. *Phys. Rev.* **139** A357 (1965).
30. Löwdin P.-O. *J. Chem. Phys.* **43** S175 (1965).
31. Löwdin P.-O. in *Perturbation Theory and its Application in Quantum Mechanics* ed. C. H. Wilcox, John Wiley, 1966.
32. Löwdin P.-O. *Int. J. Quantum Chem.* **2**, 867 (1968).
33. Löwdin P.-O., Goscinski O. *Int. J. Quantum Chem. Symp.* **5**, 665 (1971).

# Детерминантные волновые функции

ЛАНЬ®

## 1. Спин-орбитали

В квантовой химии мы обычно имеем дело с многоэлектронными системами, чьи волновые функции зависят от координат всех электронов. За исключением очень частных случаев (таких, как атом He и молекула  $H_2$ ), используемые многоэлектронные волновые функции строятся из одноэлектронных функций как строительных блоков, так, чтобы удовлетворять принципу Паули (и доставлять настолько хорошее приближение к точному решению, насколько это возможно).

Как известно, электрон имеет собственный момент импульса, спин, у которого может быть только два разных значения проекции на выделенную пространственную ось  $z$ :  $+\frac{1}{2}\hbar$  и  $-\frac{1}{2}\hbar$ . В случае одноэлектронных задач — если используется обычный в квантовой химии бесспиновый гамильтониан — нет необходимости рассматривать эту степень свободы явно. Однако мы не можем обойтись без спиновых переменных при рассмотрении многоэлектронных волновых функций.

Одноэлектронная волновая функция, зависящая как от пространственных координат  $x, y, z$  (радиус-вектора  $\vec{r}$ ), так и от спиновой координаты  $\sigma$ , обычно называется «спин-орбиталью», чтобы отличать ее от (пространственной) орбитали, зависящей только от  $\vec{r}$ :

$$\psi = \psi(x, y, z, \sigma) = \psi(\vec{r}, \sigma). \quad (5.1)$$

Так как проекция спина электрона может принимать только два значения  $s_z = \pm\frac{1}{2}$  (в единицах  $\hbar$ , т. е. атомных единицах), и то же самое верно для независимой переменной  $\sigma$ , принимающей значения  $\pm\frac{1}{2}$ , получаем две функции: одну для  $\sigma = \frac{1}{2}$  и другую для  $\sigma = -\frac{1}{2}$ :

$$\psi_1(\vec{r}) = \psi(\vec{r}, +\frac{1}{2}); \quad \psi_2(\vec{r}) = \psi(\vec{r}, -\frac{1}{2}). \quad (5.2)$$

В этом случае одноэлектронную волновую функцию можно рассматривать как вектор из двух строк

$$\psi = \begin{pmatrix} \psi_1(\vec{r}) \\ \psi_2(\vec{r}) \end{pmatrix}, \quad (5.3)$$

причем первая строка соответствует  $\sigma = \frac{1}{2}$ , а вторая —  $\sigma = -\frac{1}{2}$ .

Зависимость от спина обычно выражают, вводя состояния («спиновые функции») «спин вверх» ( $\sigma_z = +\frac{1}{2}$ ) и «спин вниз» ( $\sigma_z = -\frac{1}{2}$ )  $\alpha$  и  $\beta$  по правилам:

$$\alpha = \alpha(\sigma) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \beta = \beta(\sigma) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (5.4)$$

(Это значит, что, как функция  $\sigma$ ,  $\alpha(\frac{1}{2}) = 1$ ,  $\alpha(-\frac{1}{2}) = 0$ , тогда как  $\beta(\frac{1}{2}) = 0$ ,  $\beta(-\frac{1}{2}) = 1$ ). Тогда спин-орбиталь  $\psi$  можно представить как

$$\psi(\vec{r}, \sigma) = \psi_1(\vec{r})\alpha(\sigma) + \psi_2(\vec{r})\beta(\sigma). \quad (5.5)$$

Однако использование таких «спин-орбиталей общего вида» слишком громоздко, и обычно вводят спин-орбитали, соответствующие чистым одноэлектронным спиновым состояниям со спином вверх или вниз:  $\psi(\vec{r})\alpha(\sigma)$  и  $\psi(\vec{r})\beta(\sigma)$ .

Спин электрона описывается операторами, действующими в двумерном пространстве спиновых функций  $\alpha$  и  $\beta$ ; они представляются матрицами размера  $2 \times 2$ :

$$\hat{s}_z = \frac{1}{2}\sigma_z; \quad \hat{s}_x = \frac{1}{2}\sigma_x; \quad \hat{s}_y = \frac{1}{2}\sigma_y \quad (5.6)$$

(в атомных единицах; иначе нужно писать  $\hat{s}_x = \frac{\hbar}{2}\sigma_x$ , и т. д.), где использованы так называемые «матрицы Паули»:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (5.7)$$

Легко проверить следующие основные свойства операторов (5.6):

$$a) \quad \hat{s}^2 = \hat{s}_x^2 + \hat{s}_y^2 + \hat{s}_z^2 = \frac{3}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5.8)$$

б) Спиновые операторы (5.6) удовлетворяют коммутационным соотношениям, которым в квантовой механике подчиняются все операторы момента импульса:

$$\hat{s}_x\hat{s}_y - \hat{s}_y\hat{s}_x = i\hat{s}_z, \quad (5.9)$$

причем в формуле (5.9) можно производить циклические перестановки  $x \rightarrow y \rightarrow z \rightarrow x$ .

в) Спин-функции  $\alpha$  и  $\beta$  являются собственными функциями оператора  $\hat{s}_z$  с собственными значениями  $+\frac{1}{2}$  и  $-\frac{1}{2}$ , соответственно; они также являются собственными функциями оператора  $\hat{s}^2$  с собственным значением  $S(S+1) = \frac{1}{2}(\frac{1}{2}+1) = \frac{3}{4}$ , как этого и следовало ожидать. ( $S$  — спиновое квантовое число системы; для единственного электрона  $S = \frac{1}{2}$ .)

Для вычисления средних значений матричных элементов различных операторов необходимо интегрировать по всем переменным, от которых зависит волновая функция. Так как спиновая переменная  $\sigma$  принимает только дискретные значения  $+\frac{1}{2}$  и  $-\frac{1}{2}$ , практически это означает, что необходимо *суммировать* по проекциям спина. Такое суммирование часто включается в интегрирование, и тогда

$$\int d\tau \equiv \sum_{\sigma} \int dv ; \quad dv = dx dy dz \quad (5.10)$$

для общего обозначения суммирования по спиновым и интегрирования по пространственным переменным.

Спиновые функции  $\alpha$  и  $\beta$  ортонормированы:

$$\begin{aligned} \sum_{\sigma} \alpha^*(\sigma) \alpha(\sigma) &= \sum_{\sigma} \beta^*(\sigma) \beta(\sigma) = 1 \\ \sum_{\sigma} \alpha^*(\sigma) \beta(\sigma) &= \sum_{\sigma} \beta^*(\sigma) \alpha(\sigma) = 0 \end{aligned} \quad (5.11)$$

(По определению (5.4) спиновых функций  $\alpha$  и  $\beta$ , как двумерных векторов, эти равенства также можно записать компактно, используя матричные обозначения:  $\alpha^\dagger \alpha = \beta^\dagger \beta = 1$ ;  $\alpha^\dagger \beta = \beta^\dagger \alpha = 0$ .)

В силу условия ортонормировки (5.11), использование спин-орбиталей, построенных из чистых спиновых функций  $\alpha$  и  $\beta$ , значительно упрощает суммирование по спину: мы получаем ненулевое значение, только если имеем ту же самую спиновую функцию как в «бра», так и в «кет»-части для каждого электрона по отдельности, — и тогда суммирование по проекциям спина дает просто единичный множитель.

Полагая  $\psi_i = \varphi_i(\vec{r}) \gamma_i(\sigma)$ , где  $\gamma_i$  — либо  $\alpha$ , либо  $\beta$ , и используя предыдущие соображения, получаем для матричного элемента бесспинового оператора  $\hat{L}$  следующее выражение:

$$\begin{aligned} \langle \psi_1(\vec{r}, \sigma) | \hat{L} | \psi_2(\vec{r}, \sigma) \rangle &= \langle \varphi_1(\vec{r}) \gamma_1(\sigma) | \hat{L} | \varphi_2(\vec{r}) \gamma_2(\sigma) \rangle \\ &= \int d\tau \varphi_1^*(\vec{r}) \gamma_1^*(\sigma) \hat{L} \varphi_2(\vec{r}) \gamma_2(\sigma) \\ &= \sum_{\sigma} \gamma_1^*(\sigma) \gamma_2(\sigma) \int dv \varphi_1^*(\vec{r}) \hat{L} \varphi_2(\vec{r}) \\ &= \delta_{\gamma_1 \gamma_2} \int dv \varphi_1^*(\vec{r}) \hat{L} \varphi_2(\vec{r}) \\ &= \langle \varphi_1(\vec{r}) | \hat{L} | \varphi_2(\vec{r}) \rangle \delta_{\gamma_1 \gamma_2} . \end{aligned} \quad (5.12)$$

Аналогично, для двухэлектронного матричного элемента (интеграла) двухчастичного оператора  $\hat{g}(1, 2)$ , не зависящего от спина, имеем

$$\begin{aligned} \langle \psi_1(1)\psi_2(2) | \hat{g}(1, 2) | \psi_3(1)\psi_4(2) \rangle \\ = \delta_{\gamma_1\gamma_3} \delta_{\gamma_2\gamma_4} \langle \varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) | \hat{g}(1, 2) | \varphi_3(\vec{r}_1)\varphi_4(\vec{r}_2) \rangle, \end{aligned} \quad (5.13)$$

где мы использовали сокращенные обозначения «1» и «2» для всех пространственных и спиновых переменных электрона 1 и 2:  $\psi_i(1) \equiv \psi_i(\vec{r}_1, \sigma_1)$ ,  $\psi_j(2) \equiv \psi_j(\vec{r}_2, \sigma_2)$ .



## 2. Многоэлектронные спиновые состояния

Гамильтониан Борна–Оппенгеймера не действует на спиновые переменные и поэтому коммутирует с операторами спина. Тогда собственные функции гамильтониана можно классифицировать по собственным значениям квадрата полного спина системы ( $\hat{S}^2$ ) и проекции полного спина ( $\hat{S}_z$ ) на выбранную ось (обычно ось  $z$ ). Обозначим через  $S$  наибольшее значение проекции, которое возможно для данного собственного состояния  $\hat{S}^2$ :  $|S_z|_{\max} = S$ . (Квантовое число  $S$  является полуцелым, если число электронов нечетно, и целым, если число электронов четно.) Тогда в соответствии с общими правилами квантовой механики для момента импульса, собственным значением  $\hat{S}^2$  является  $S(S+1)$ , и соответствующее состояние имеет спиновую мультиплетность  $2S+1$ , так как для данного  $S$  существует  $2S+1$  возможных проекций  $S_z$ , а именно  $-S, -S+1, \dots, S-1, S$ . Соответственно, разные спиновые состояния обычно называются как

$$\begin{aligned} S = 0 & \quad \text{синглет} \\ S = \frac{1}{2} & \quad \text{дублет} \\ S = 1 & \quad \text{триплет} \\ S = \frac{3}{2} & \quad \text{квартет} \\ S = 2 & \quad \text{квинтет и т. д.} \end{aligned}$$

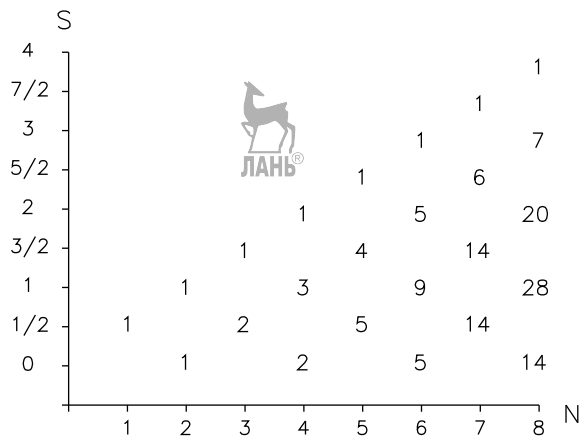
Использование спиновых функций осложнено тем, что в случае, когда число электронов, распределенных по однократно занятым пространственным орбиталям, превышает два, обычно можно построить несколько различных многоэлектронных спиновых функций, соответствующих одним и тем же собственным значениям операторов  $\hat{S}^2$  и  $\hat{S}_z$ . (Дважды занятые орбитали соответствуют синглетным подсистемам и их можно не рассматривать явно.) Существование нескольких независимых спиновых функций иногда называется «спиновым вырождением»,

но это лишь традиционное название, так как истинного вырождения нет — разные спинорные функции обычно соответствуют разным энергиям.

Возможные значения  $S$  для случая  $N$  электронов и  $N$  однократно занятых пространственных орбиталей изменяются от 0 (если  $N$  четно) или от  $\frac{1}{2}$  (если  $N$  нечетно) до максимального значения  $S = \frac{N}{2}$ . Для данного значения  $S_z$  числа  $n_\alpha$  и  $n_\beta$  электронов со спинами  $\alpha$  и  $\beta$ , очевидно равны  $n_\alpha = \frac{N}{2} + S_z$  и  $n_\beta = \frac{N}{2} - S_z$ . Число различных элементарных спинорных функций (простых произведений  $\alpha$ - и  $\beta$ -функций), в которых  $n_\alpha$  электронов имеют спин  $\alpha$  и  $n_\beta$  электронов имеют спин  $\beta$ , равно  $\binom{N}{n_\alpha} = \binom{N}{n_\beta}$  — числу способов выбрать  $n_\alpha$  позиций из их общего числа  $N$ , и разместить в них проекцию спина  $\alpha$  — или, что то же самое, числу способов выбрать  $n_\beta$  мест, и разместить в них проекцию спина  $\beta$ . (Равенство следует из того, что  $n_\alpha + n_\beta = N$ ). Отдельные элементарные спинорные функции, в общем случае, не являются «чистыми спинорными функциями», т. е. они не являются собственными функциями оператора  $\hat{S}^2$ ; однако, образуя из них линейные комбинации, можно построить чистые спинорные функции для всех значений  $S$ , при которых данное  $S_z = \frac{1}{2}(n_\alpha - n_\beta)$  вообще возможно (т. е.  $|S_z| \leq S$ ).

Число линейно независимых собственных спинорных функций, соответствующих разным числам электронов  $N$  на однократно занятых орбиталях и разным спинорным квантовым числам  $S$ , устанавливается так называемой «диаграммой ветвления», показанной на нижеследующем рисунке, в некоторой степени напоминающем известный из математики треугольник Паскаля. На диаграмме ветвления каждая цифра задает число независимых собственных спинорных функций для значений  $N$  и  $S$ , указанных на осях. Если рассмотреть систему  $N$  электронов в собственном спинорном состоянии со спинорным квантовым числом  $S$  и связать ее с дополнительным электроном, то спин последнего может быть либо параллельным, либо антипараллельным полному спину первых  $N$  электронов; таким образом можно получить две различные  $N + 1$ -электронные спинорные функции: одну с  $S' = S + \frac{1}{2}$ , а другую с  $S' = S - \frac{1}{2}$ . Отсюда можно вывести правило, по которому формируется диаграмма ветвления: каждое число равно сумме двух ближайших чисел, которые можно найти слева от него: данное  $N$ -электронное состояние с квантовым числом  $S$  можно сформировать добавлением одного

электрона к каждому из  $(N-1)$ -электронных состояний с квантовыми числами  $S \pm \frac{1}{2}$ .



Полное число состояний с разными спинами для данного  $N$  можно определить следующим образом. Элементарные спиновые функции, содержащие  $n_\alpha$  функций  $\alpha$  и  $n_\beta = N - n_\alpha$  функций  $\beta$ , образуют базис  $N$ -электронных спиновых функций с  $S_z = \frac{1}{2}(n_\alpha - n_\beta)$ . Спиновые собственные функции являются линейными комбинациями элементарных спиновых функций, и поэтому число разных спиновых собственных функций, которые можно образовать для данных значений  $N$  и  $S_z$ , как раз равно числу соответствующих элементарных спиновых функций. Каждое спиновое состояние имеет компоненту с наименьшей возможной проекцией  $S_z = 0$  (если  $N$  четно) или  $S_z = \frac{1}{2}$  (если  $N$  нечетно), поэтому число разных  $N$ -электронных спиновых состояний со всеми возможными значениями  $S$  равно числу  $\binom{N}{n_\alpha}$  элементарных спиновых функций с этим наименьшим значением  $S_z$ . (При этом  $n_\alpha = \frac{N}{2}$ , если  $N$  четно, и  $n_\alpha = \frac{N+1}{2}$ , если  $N$  нечетно.) Это число определяет сумму чисел в колонке, соответствующей данному  $N$ , на диаграмме ветвления.

Приведенные рассуждения относятся к случаю, когда каждая из  $N$  пространственных орбиталей однократно занята. На практике часто необходимо рассматривать случаи, в которых некоторые орбитали могут заполняться дважды, тогда как какие-то другие остаются пустыми. Тогда число конфигураций (детерминантов — см. следующий раздел) можно определить, распределяя независимо друг от друга  $n_\alpha$  спинов  $\alpha$  и  $n_\beta$  спинов  $\beta$  между  $N$  пространственными орбиталями (нет необходимо-

сти предполагать, что  $N = n_\alpha + n_\beta$ . Это приводит к числу  $\binom{N}{n_\alpha} \binom{N}{n_\beta}$ , которое растет очень быстро с увеличением  $N, n_\alpha$  и  $n_\beta$ .

Существует несколько методов для построения разных спиновых собственных функций. Их систематическое изучение является темой для отдельной монографии и требует развернутого использования теории групп. Мы не будем рассматривать эту проблему подробно, упомянем только один способ построения линейно независимого (но неортогонального) набора спиновых собственных функций. Необходимо выбрать столько элементарных спиновых функций, сколько должно быть различных спиновых функций для данного значения  $S$  и  $N$ , согласно диаграмме ветвления, и спроектировать их по спину, используя оператор спиновой проекции Лёвдина, который обсуждается в гл. 6, разд. 4.3.

### 3. Детерминанты Слэтера

Многоэлектронная волновая функция должна быть антисимметричной по всем электронам, т. е. менять знак при одновременной перестановке пространственных и спиновых координат любой пары электронов:

$$\Psi(1, 2, \dots, i, \dots, j, \dots, N) = -\Psi(1, 2, \dots, j, \dots, i, \dots, N) . \quad (5.14)$$

В принципе всегда возможно представить антисимметричную волновую функцию, по крайней мере формально, в виде произведения пространственной и спиновой составляющей, но обычно это весьма сложно сделать и приходится использовать изощренные теоретико-групповые методы. Ситуация проста только в двухэлектронном случае, в котором требуется либо симметричность пространственной части волновой функции и антисимметричность ее спиновой части по отношению к перестановкам двух электронов, что соответствует синглетному состоянию, либо, наоборот, требуется симметричность спиновой части и антисимметричность пространственной, что соответствует триплетному состоянию. Столь простая конструкция, вообще говоря, невозможна для большего числа электронов: в этих случаях пространственная часть получается симметризацией по отношению к некоторым парам электронов, и антисимметризацией по отношению к другим парам, тогда как для спиновой части роли меняются, что обеспечивает полную антисимметричность произведения пространственной и спиновой частей. (Мы не будем вдаваться в подробности этого метода.)

На практике более удобно описывать пространственные и спиновые координаты отдельных электронов совместно, и строить волновую функцию в виде одного или нескольких «детерминантов Слэтера», обеспечивающих требуемую антисимметричность по построению.



Если рассматривать  $N$  электронов и  $N$  спин-орбиталей  $\psi_k$ , то можно построить волновую функцию в виде определителя (детерминанта Слэтера):

$$\Psi = \Psi(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \psi_3(1) & \dots & \psi_N(1) \\ \psi_1(2) & \psi_2(2) & \psi_3(2) & \dots & \psi_N(2) \\ \vdots & \vdots & \vdots & & \vdots \\ \psi_1(N) & \psi_2(N) & \psi_3(N) & \dots & \psi_N(N) \end{vmatrix} \quad (5.15)$$

Здесь  $\psi_k(i) \equiv \psi_k(\vec{r}_i, \sigma_i)$  —  $k$ -я спин-орбиталь, как функция пространственных и спиновых координат  $i$ -го электрона.

По известным свойствам определителей, волновая функция  $\Psi$ , заданная формулой (5.15), меняет знак при перестановке любой пары электронов, так как это равносильно перестановке двух строк детерминанта. Как будет показано в разд. 6, множитель  $\frac{1}{\sqrt{N!}}$  обеспечивает нормировку многоэлектронной волновой функции (5.15) на единицу, если только она построена из ортонормированных одноэлектронных спин-орбиталей.

Слэтеровская детерминантная волновая функция  $\Psi$  антисимметрична не только по перестановкам электронов (строк детерминанта), но и по перестановкам спин-орбиталей (столбцов детерминанта). Поэтому  $\Psi$  не может содержать одну и ту же спин-орбиталь дважды — иначе было бы  $\Psi \equiv 0$ , — в согласии с наглядной формулировкой принципа Паули («в любом состоянии может существовать только один электрон»). Из данной пространственной орбитали можно получить две независимых спин-орбитали (одну со спином  $\alpha$ , другую со спином  $\beta$ ), так что каждая пространственная орбиталь может быть занята максимум дважды. В более общем смысле, из свойств определителей следует, что все участвующие в детерминантной волновой функции спин-орбитали должны быть линейно независимыми, иначе  $\Psi \equiv 0$ .

### 3.1. Двухэлектронные примеры

В частном случае двух электронов детерминант Слэтера, построенный подходящим образом, может быть представлен как произведение пространственной и спиновой частей, причем одна из них будет симметричной, а другая — антисимметричной. Одновременно он является собственной функцией также и полного спина  $\hat{S}^2$  и может быть охарактеризован своей спиновой мультиплетностью.

*Синглет:* детерминант, построенный из дважды занятой пространственной орбитали  $\varphi$ . Ей соответствуют две спин-орбитали:

$$\psi_1(\vec{r}, \sigma) = \varphi(\vec{r})\alpha(\sigma) ; \quad \psi_2(\vec{r}, \sigma) = \varphi(\vec{r})\beta(\sigma) . \quad (5.16)$$

С помощью сокращенных обозначений  $\varphi(i) = \varphi(\vec{r}_i)$ ,  $\alpha(i) = \alpha(\sigma_i)$ , и  $\beta(i) = \beta(\sigma_i)$ , можно написать детерминантную волновую функцию как

$$\begin{aligned} {}^1\Psi(1, 2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi(1)\alpha(1) & \varphi(1)\beta(1) \\ \varphi(2)\alpha(2) & \varphi(2)\beta(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} [\varphi(1)\alpha(1)\varphi(2)\beta(2) - \varphi(1)\beta(1)\varphi(2)\alpha(2)] \\ &= \underbrace{\varphi(1)\varphi(2)}_{\text{симметричная}} \underbrace{\frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \beta(1)\alpha(2)]}_{\text{антисимметричная}} . \end{aligned} \quad (5.17)$$

*Триплет:* следует заметить, что для построения триплетной двухэлектронной волновой функции необходимы две различные пространственные орбитали. В простейшем случае они ортогональны, но это не обязательно. Имеем две спин-орбитали спина  $\alpha$   $\psi_1(\vec{r}, \varphi) = \varphi_1(\vec{r})\alpha(\sigma)$  и  $\psi_2(\vec{r}, \sigma) = \varphi_2(\vec{r})\alpha(\sigma)$ , и триплетная детерминантная волновая функция с  $S_z = 1$  равна

$$\begin{aligned} {}^3\Psi(1, 2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_1(1)\alpha(1) & \varphi_2(1)\alpha(1) \\ \varphi_1(2)\alpha(2) & \varphi_2(2)\alpha(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} [\varphi_1(1)\alpha(1)\varphi_2(2)\alpha(2) - \varphi_2(1)\alpha(1)\varphi_1(2)\alpha(2)] \\ &= \underbrace{\frac{1}{\sqrt{2}} [\varphi_1(1)\varphi_2(2) - \varphi_2(1)\varphi_1(2)]}_{\text{антисимметричная}} \underbrace{\alpha(1)\alpha(2)}_{\text{симметричная}} . \end{aligned} \quad (5.18)$$

Заменяя обе проекции спина  $\alpha$  на  $\beta$ , получаем компоненту с  $S_z = -1$  для той же триплетной конфигурации; однако компонента  $S_z = 0$  триплета не может быть задана одним детерминантом, а только лишь суммой двух детерминантов. Аналогично, для записи синглетной волновой функции, если она построена из двух различных пространственных орбиталей, понадобятся два детерминанта.

Эти двухдетерминантные двухэлектронные волновые функции все же можно представить в виде произведения симметричной пространственной части и антисимметричной спиновой части в синглетном случае, и наоборот — в триплетном случае. В предположении, что  $\varphi_1$  и  $\varphi_2$

являются ортогональными пространственными орбиталями, эти нормированные функции таковы:

$${}^{1,3}\Psi = \frac{1}{2} \left\{ \begin{vmatrix} \varphi_1(1)\alpha(1) & \varphi_2(1)\beta(1) \\ \varphi_1(2)\alpha(2) & \varphi_2(2)\beta(2) \end{vmatrix} \mp \begin{vmatrix} \varphi_1(1)\beta(1) & \varphi_2(1)\alpha(1) \\ \varphi_1(2)\beta(2) & \varphi_2(2)\alpha(2) \end{vmatrix} \right\}, \quad (5.19)$$

где верхний знак соответствует синглету, а нижний — триплету.

Раскрыв определители в формулах (5.19), легко получим

$${}^{1,3}\Psi = \frac{1}{2} \{ \varphi_1(1)\alpha(1)\varphi_2(2)\beta(2) - \varphi_2(1)\beta(1)\varphi_1(2)\alpha(2) \mp \varphi_1(1)\beta(1)\varphi_2(2)\alpha(2) \pm \varphi_2(1)\alpha(1)\varphi_1(2)\beta(2) \} \quad (5.20)$$

т. е.

$${}^1\Psi = \frac{1}{2} \underbrace{[\varphi_1(1)\varphi_2(2) + \varphi_1(2)\varphi_2(1)]}_{\text{симметричная}} \underbrace{[\alpha(1)\beta(2) - \beta(1)\alpha(2)]}_{\text{антисимметричная}} \quad (5.21)$$

и

$${}^3\Psi = \frac{1}{2} \underbrace{[\varphi_1(1)\varphi_2(2) - \varphi_2(1)\varphi_1(2)]}_{\text{антисимметричная}} \underbrace{[\alpha(1)\beta(2) + \beta(1)\alpha(2)]}_{\text{симметричная}} \quad (5.22)$$

#### 4. Оператор антисимметризации

Детерминантная волновая функция  $\Psi$ , определенная формулой (5.15), может быть записана компактно с помощью оператора антисимметризации («антисимметризатора»)  $\hat{\mathcal{A}}$ :

$$\Psi = \hat{\mathcal{A}}[\psi_1(1)\psi_2(2) \dots \psi_N(N)], \quad (5.23)$$

где  $\psi_1(1)\psi_2(2) \dots \psi_N(N)$  — так называемое произведение Хартри, а оператор  $\hat{\mathcal{A}}$  определяется как

$$\hat{\mathcal{A}} = \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^{\hat{P}} \hat{P}. \quad (5.24)$$

Суммирование в правой части выражения (5.24) ведется по всем перестановкам  $\hat{P}$ , которые возможны для  $N$  различных чисел (номеров электронов)  $1, 2, \dots, N$  — эти перестановки образуют так называемую «симметрическую группу»  $S_N$ . Множитель  $(-1)^{\hat{P}}$  есть *четность* перестановки  $\hat{P}$  (обзор основных свойств перестановок приведен в приложении П10). Отметим, что определения (5.23)-(5.24) являются

лишь повторением алгебраического выражения (5.15) для определителя. В самом деле, действие оператора  $\hat{A}$  на произведение Хартри  $\psi_1(1)\psi_2(2)\dots\psi_N(N)$  дает

$$\Psi = \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^p \psi_1(P_1)\psi_2(P_2)\dots\psi_N(P_N) \, , \quad (5.25)$$

что есть выражение определителя (5.15) в виде суммы по разным перестановкам.

#### 4.1. Оператор антисимметризации как проекционный оператор

Оператор антисимметризации  $\hat{A}$  является, с точностью до постоянного коэффициента, проекционным оператором, который извлекает из какой-либо волновой функции составляющую, антисимметричную по перестановкам любых пар электронов.

Известно, что оператор является проекционным (приложение П7), тогда и только тогда, когда он является а) *идемпотентным*, т. е. квадрат оператора равен самому оператору, и б) *эрмитовым*. Здесь мы покажем, что оператор антисимметризации  $\hat{A}$  обладает этими свойствами. (Так как идемпотентность выполняется с точностью до постоянного множителя, оператор  $\hat{A}$  лишь пропорционален соответствующему проекционному оператору. Как уже отмечено, коэффициент  $\frac{1}{\sqrt{N!}}$  в определении  $\hat{A}$  выбран так, чтобы обеспечить нормированность на единицу  $N$ -электронной детерминантной волновой функции, построенной из ортонормированных спин-орбиталей.)

*Идемпотентность* с точностью до постоянного множителя  $\sqrt{N!}$

Рассмотрим операторное произведение  $\hat{A}\hat{A}$ :

$$\begin{aligned} \hat{A}\hat{A} &= \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^p \hat{P} \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^q \hat{Q} \\ &= \frac{1}{N!} \sum_{\hat{P} \in S_N} \sum_{\hat{Q} \in S_N} (-1)^{p+q} \hat{P}\hat{Q} . \end{aligned} \quad (5.26)$$

Если  $\hat{Q}$  пробегает по всем перестановкам, то — для данного  $\hat{P}$  — перестановка  $\hat{R} = \hat{P}\hat{Q}$  также пробегает по всей совокупности перестановок, только в другом порядке. Поэтому вторую сумму в правой части выражения (5.26), которая ведется по  $\hat{Q} \in S_N$ , можно заменить на сумму по

$\hat{R} \in S_N$ . Четность перестановки  $\hat{R}$  равна произведению четностей двух перестановок, так что  $(-1)^{p+q} = (-1)^r$ . Тогда мы можем записать

$$\hat{A}\hat{A} = \frac{1}{N!} \sum_{\hat{P} \in S_N} \left[ \sum_{\hat{R} \in S_N} (-1)^r \hat{R} \right]. \quad (5.27)$$

Внешняя сумма по  $\hat{P} \in S_N$  показывает, что необходимо сложить  $N!$  тождественных операторов (число различных перестановок  $\hat{P}$  равно  $N!$ , а оператор в скобках не зависит от  $\hat{P}$ ), так что получаем:

$$\hat{A}\hat{A} = \sum_{\hat{R} \in S_N} (-1)^r \hat{R} = \sqrt{N!} \hat{A}. \text{ Ч. т. д.} \quad (5.28)$$

Из этого результата следует, что строго идемпотентным проектором является  $\frac{1}{\sqrt{N!}} \hat{A}$ , т. е. оператор, получающийся заменой множителя  $\frac{1}{\sqrt{N!}}$  на  $\frac{1}{N!}$  в определении (5.24).

### Эрмитовость

Оператор  $\hat{L}$  является эрмитовым (самосопряженным) если  $\hat{L}^\dagger = \hat{L}$ , что означает  $\langle \hat{L}\psi | \varphi \rangle = \langle \psi | \hat{L}\varphi \rangle$  для любых  $\psi$  и  $\varphi$ . (Эрмитово сопряженное  $\hat{L}^\dagger$  оператора  $\hat{L}$  определяется как оператор, для которого  $\langle \hat{L}^\dagger \psi | \varphi \rangle \equiv \langle \psi | \hat{L}\varphi \rangle$ .) Соответственно, мы должны показать, что

$$\langle \hat{A}\Phi(1, 2, \dots, N) | \Psi(1, 2, \dots, N) \rangle = \langle \Phi(1, 2, \dots, N) | \hat{A}\Psi(1, 2, \dots, N) \rangle \quad (5.29)$$

для любых  $N$ -электронных функций  $\Phi$  и  $\Psi$ .

Очевидно, что

$$\langle \Phi(1, 2, \dots, N) | \Psi(1, 2, \dots, N) \rangle = \langle \hat{Q}\Phi(1, 2, \dots, N) | \hat{Q}\Psi(1, 2, \dots, N) \rangle \quad (5.30)$$

для любого оператора перестановки  $\hat{Q} \in S_N$ , так как правая часть отличается от левой только изменением обозначений переменных интегрирования. Используя этот факт, рассмотрим выражение

$$\begin{aligned} & \langle \hat{A}\Phi(1, 2, \dots, N) | \Psi(1, 2, \dots, N) \rangle \\ &= \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^p \langle \hat{P}\Phi(1, 2, \dots, N) | \Psi(1, 2, \dots, N) \rangle. \end{aligned} \quad (5.31)$$

Поддействуем на «бра-» и «кет-» функции перестановкой  $\hat{Q} = \hat{P}^{-1}$  (обратной к  $\hat{P}$ ). В этом случае получаем

$$\begin{aligned} & \langle \hat{A}\Phi(1, 2, \dots, N) | \Psi(1, 2, \dots, N) \rangle \\ &= \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^p \langle \hat{P}^{-1} \hat{P} \Phi(1, 2, \dots, N) | \hat{P}^{-1} \Psi(1, 2, \dots, N) \rangle \\ &= \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^p \langle \Phi(1, 2, \dots, N) | \hat{P}^{-1} \Psi(1, 2, \dots, N) \rangle. \end{aligned} \quad (5.32)$$

Если  $\hat{P}$  пробегает все возможные перестановки, тогда то же самое верно и для обратной к ней  $\hat{Q} = \hat{P}^{-1}$ ; четности  $\hat{P}$  и  $\hat{P}^{-1} = \hat{Q}$  также одинаковы, так что получаем

$$\begin{aligned} & \langle \hat{A}\Phi(1, 2, \dots, N) | \Psi(1, 2, \dots, N) \rangle \\ &= \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^q \langle \Phi(1, 2, \dots, N) | \hat{Q} \Psi(1, 2, \dots, N) \rangle \\ &= \langle \Phi(1, 2, \dots, N) | \hat{A} \Psi(1, 2, \dots, N) \rangle. \quad \text{Ч. т. д.} \end{aligned} \quad (5.33)$$

*Двухэлектронный пример*, иллюстрирующий проекционный характер оператора антисимметризации.

Произведение Хартри  $\Psi_H(1, 2) = \varphi(1)\psi(2)$  (разных) спин-орбиталей  $\varphi$  и  $\psi$  можно рассматривать как сумму антисимметричного слагаемого

$$\Phi_A(1, 2) = \frac{1}{2}[\varphi(1)\psi(2) - \varphi(2)\psi(1)] \quad (5.34)$$

и симметричного слагаемого


$$\Phi_S(1, 2) = \frac{1}{2}[\varphi(1)\psi(2) + \varphi(2)\psi(1)]. \quad (5.35)$$

В самом деле, легко видеть, что  $\Psi_H = \Phi_A + \Phi_S$ . Функция  $\Phi_A(1, 2)$  отличается от детерминантной волновой функции  $\hat{A}[\varphi(1)\psi(2)] = \frac{1}{\sqrt{2}}[\varphi(1)\psi(2) - \varphi(2)\psi(1)]$  только коэффициентом  $\sqrt{2}$ . Следовательно, в случае двух электронов оператором, проектирующим на антисимметричное подпространство, является оператор

$$\hat{P}_A = \frac{1}{\sqrt{2}}\hat{A} = \frac{1}{2}[\hat{I} - \hat{P}_{12}], \quad (5.36)$$

где  $\hat{I}$  — единичный оператор (тождественная перестановка), а  $\hat{P}_{12}$  является оператором, переставляющим координаты двух электронов. Это находится в полном согласии с условием идемпотентности (5.28).


Если действовать оператором  $\hat{P}_A$  на антисимметричную функцию  $\Phi_A(1, 2)$ , то она останется неизменной, так как уже находится в антисимметричном подпространстве и поэтому является собственным вектором  $\hat{P}_A$  с собственным значением единица:  $\hat{P}_A \Phi_A(1, 2) = \Phi_A(1, 2)$ . В самом деле



$$\begin{aligned}
 \hat{P}_A \Phi_A(1, 2) &= \frac{1}{2}(\hat{I} - \hat{P}_{12}) \frac{1}{2}[\varphi(1)\psi(2) - \varphi(2)\psi(1)] \\
 &= \frac{1}{4} \left\{ \hat{I}[\varphi(1)\psi(2) - \varphi(2)\psi(1)] - \hat{P}_{12}[\varphi(1)\psi(2) - \varphi(2)\psi(1)] \right\} \\
 &= \frac{1}{4} \{ \varphi(1)\psi(2) - \varphi(2)\psi(1) - \varphi(2)\psi(1) + \varphi(1)\psi(2) \} \\
 &= \frac{1}{2}[\varphi(1)\psi(2) - \varphi(2)\psi(1)] = \Phi_A(1, 2) .
 \end{aligned} \tag{5.37}$$

В случае большего числа электронов произведение Хартри  $\varphi_1(1)\varphi_2(2)\dots\varphi_N(N)$  невозможно представить в виде суммы полностью симметричного и полностью антисимметричного слагаемых, так как имеются также вклады, симметричные по одним электронам и антисимметричные по другим. Однако *если спин-орбитали  $\varphi_i$  линейно независимы*, то произведение Хартри всегда имеет ненулевую составляющую, антисимметричную по всем электронам. Этот вклад выделяется (с точностью до постоянного множителя) при составлении детерминантной волновой функции  $\hat{A}[\varphi_1(1)\varphi_2(2)\dots\varphi_N(N)]$ . Так как электроны являются фермионами, нас интересуют только такие полностью антисимметричные волновые функции.

Согласно известным свойствам определителей, в них можно переставлять строки со столбцами без изменения определителя. Применительно к определению (5.25) это значит, что его правую часть можно также переписать как



$$\hat{A}[\psi_1(1)\psi_2(2)\dots\psi_N(N)] = \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^q \psi_{Q_1}(1)\psi_{Q_2}(2)\dots\psi_{Q_N}(N) . \tag{5.38}$$

Полезно проследить за доказательством эквивалентности выражений (5.25) и (5.38) в рамках применяемого здесь формализма.

Для этого определим оператор  $\hat{Q}'$ , который производит перестановку  $\hat{Q}$  нижних индексов спин-орбиталей:

$$\hat{Q}'\psi_1(1)\psi_2(2)\dots\psi_N(N) = \psi_{Q_1}(1)\psi_{Q_2}(2)\dots\psi_{Q_N}(N) . \tag{5.39}$$

Имеем очевидное равенство

$$\begin{aligned}\hat{Q}'\hat{Q}\psi_1(1)\psi_2(2)\dots\psi_N(N) &= \hat{Q}'\psi_1(Q_1)\psi_2(Q_2)\dots\psi_N(Q_N) \\ &= \psi_{Q_1}(Q_1)\psi_{Q_2}(Q_2)\dots\psi_{Q_N}(Q_N) \\ &= \psi_1(1)\psi_2(2)\dots\psi_N(N)\end{aligned}\quad (5.40)$$

для каждой  $\hat{Q} \in S_N$ , так как одновременная перестановка электронов и спин-орбиталей меняет только порядок сомножителей в произведении. Поэтому для произведений Хартри  $\hat{Q}'\hat{Q} = \hat{I}$  или  $\hat{Q}' = \hat{Q}^{-1}$ .

Теперь рассмотрим

$$\Psi = \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^{\hat{P}} \hat{P} \psi_1(1)\psi_2(2)\dots\psi_N(N) \quad (5.41)$$

и вставим оператор  $\hat{I} = \hat{Q}\hat{Q}'$  после  $\hat{P}$ :

$$\Psi = \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^{\hat{P}} \hat{P} \hat{Q} \hat{Q}' \psi_1(1)\psi_2(2)\dots\psi_N(N). \quad (5.42)$$

Это равенство верно для каждой перестановки  $\hat{Q}$ , так что можем положить  $\hat{Q} = \hat{P}^{-1}$ . Получим

$$\Psi = \frac{1}{\sqrt{N!}} \sum_{\hat{P} \in S_N} (-1)^{\hat{P}} \hat{Q}' \psi_1(1)\psi_2(2)\dots\psi_N(N). \quad (5.43)$$

Если  $\hat{P}$  пробегает по всем перестановкам, тогда то же самое верно и для  $\hat{Q} = \hat{P}^{-1}$ ; их четности также равны,  $(-1)^{\hat{P}} = (-1)^{\hat{Q}}$ , поэтому

$$\begin{aligned}\Psi &= \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^{\hat{Q}} \hat{Q}' \psi_1(1)\psi_2(2)\dots\psi_N(N) \\ &= \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^{\hat{Q}} \psi_{Q_1}(1)\psi_{Q_2}(2)\dots\psi_{Q_N}(N). \quad \text{Ч. т. д.}\end{aligned}\quad (5.44)$$

## 4.2. Коммутационные свойства оператора антисимметризации

Оператор  $\hat{\mathcal{A}}$  коммутирует с любым оператором, симметричным по координатам всех электронов. Наиболее важными примерами являются гамильтониан Борна–Оппенгеймера и операторы полного спина  $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$  и его проекции  $\hat{S}_z$ , где  $\hat{S}_x = \sum_i \hat{s}_x(i)$ ;  $\hat{S}_y = \sum_i \hat{s}_y(i)$ ;  $\hat{S}_z = \sum_i \hat{s}_z(i)$ .



Чтобы доказать это утверждение, сначала рассмотрим какой-нибудь оператор  $\hat{L}$ , представляющий собой симметричную сумму одноэлектронных операторов  $\hat{l}(i)$ :

$$\hat{L} = \sum_{i=1}^N \hat{l}(i) . \quad (5.45)$$

Мы должны сравнить  $\hat{\mathcal{A}}\hat{L}\Psi(1, 2, \dots, N)$  с  $\hat{L}\hat{\mathcal{A}}\Psi(1, 2, \dots, N)$ . Для доказательства того, что они равны, достаточно показать, что  $\sum_{i=1}^N \hat{l}(i)$  коммутирует с *любым* оператором перестановки  $\hat{P}$ . Поэтому рассмотрим

$$\hat{P} \sum_{i=1}^N \hat{l}(i) \Psi(1, 2, \dots, N) = \sum_{i=1}^N \hat{l}(P_i) \Psi(P_1, P_2, \dots, P_N) \quad (5.46)$$

и

$$\sum_{i=1}^N \hat{l}(i) \hat{P} \Psi(1, 2, \dots, N) = \sum_{i=1}^N \hat{l}(i) \Psi(P_1, P_2, \dots, P_N) . \quad (5.47)$$

Но каждое значение  $P_i$ , как аргумент  $\hat{l}$  в правой части выражения (5.46), есть элемент совокупности  $\{1, 2, \dots, N\}$ , который стоит на  $i$ -м месте в последовательности  $P_1, P_2, \dots, P_N$ , и каждый элемент этой совокупности встречается в наборе  $\{P_1, P_2, \dots, P_N\}$ , причем только один раз. Поэтому

$$\sum_{i=1}^N \hat{l}(i) = \sum_{i=1}^N \hat{l}(P_i), \text{ и правые части выражений (5.46) и (5.47) равны. Ч. Д.}$$

Доказательство для двухэлектронных операторов очень похоже. Такие операторы обычно записываются в виде

$$\hat{G} = \sum_{i < j} \hat{g}(i, j) = \frac{1}{2} \sum_{\substack{i, j \\ (i \neq j)}} \hat{g}(i, j) . \quad (5.48)$$

Две указанных формы эквивалентны, так как оператор  $\hat{G}$  предполагается симметричным по координатам электронов:  $\hat{g}(i, j) = \hat{g}(j, i)$ . Благодаря этой симметрии можно переписать определение (5.48) также в виде

$$\hat{G} = \sum_{\{i, j\}} \hat{g}(i, j) , \quad (5.49)$$

где суммирование ведется по различным *парам*  $\{i, j\}$  электронов без указания на то, какой индекс больше,  $i$  или  $j$ . Тогда при действии перестановки на  $\hat{G}$  сумма поменяется на  $\sum_{\{i, j\}} \hat{g}(P_i, P_j)$ . Но как  $P_i$ , так и  $P_j$ ,

являются элементами совокупности  $\{1, 2, \dots, N\}$ , и они различны для  $i \neq j$ . Кроме того, пары  $\{P_i, P_j\}$  различны для разных пар  $\{i, j\}$ . Поэтому суммирование происходит по всем  $\binom{N}{2}$  различным сочетаниям элементов  $\{1, 2, \dots, N\}$ , и это снова означает, что *суммы* до и после произведенной перестановки одинаковы:

$$\sum_{\{i,j\}} \hat{g}(P_i, P_j) = \sum_{\{i,j\}} \hat{g}(i, j) \quad (5.50)$$

и поэтому

$$\hat{P} \sum_{\{i,j\}} \hat{g}(i, j) \Psi(1, 2, \dots, N) = \sum_{\{i,j\}} \hat{g}(i, j) \hat{P} \Psi(1, 2, \dots, N). \quad \text{Ч. т. д.} \quad (5.51)$$

### 4.3. Антисимметризация координат электронов в пространственно удаленных подсистемах

В соответствии с принципом Паули, волновая функция должна быть антисимметричной по отношению к перестановкам любых пар электронов. Рассмотрим две системы  $A$  и  $B$ , содержащие  $N_A$  и  $N_B$  электронов, и их антисимметризованные волновые функции  $\Psi_A$  и  $\Psi_B$ , являющиеся собственными функциями соответствующих гамильтонианов  $\hat{H}_A$  и  $\hat{H}_B$ , с собственными значениями  $E_A$  и  $E_B$ . Мы предполагаем, что  $A$  и  $B$  находятся на столь большом расстоянии друг от друга, что любыми взаимодействиями между ними можно полностью пренебречь. Тем не менее, мы можем формально рассматривать  $A$  и  $B$  как части объединенной системы  $AB$ , содержащей  $N_A + N_B$  электронов.

Если рассматривать  $A$  и  $B$  отдельно, то аргументами  $\Psi_A$  являются пространственные и спинные координаты электронов с номерами  $1, 2, \dots, N_A$ , а для  $\Psi_B$  — электронов, занумерованных как  $1, 2, \dots, N_B$ . В случае объединенной системы  $AB$  необходимо нумеровать электроны последовательно от 1 до  $N_A + N_B$ . Рассматривая уравнение Шрёдингера для объединенной системы, вначале мы используем схему, упомянутую в гл. 1, разд. 1.1, не требуя антисимметричности волновой функции по отношению к перестановкам электронов между двумя подсистемами. При этом мы можем приписать первые  $N_A$  электронов подсистеме  $A$ , а оставшиеся  $N_B$  — подсистеме  $B$ . Таким образом, аргументами  $\Psi_A$  являются координаты электронов  $1, 2, \dots, N_A$ , а для  $\Psi_B$  — координаты электронов  $N_A + 1, N_A + 2, \dots, N_A + N_B$ :  $\Psi_A = \Psi_A(1, 2, \dots, N_A)$ ;  $\Psi_B = \Psi_B(N_A + 1, N_A + 2, \dots, N_A + N_B)$ .

В этом случае гамильтонианы  $\hat{H}_A$  и  $\hat{H}_B$  в атомных единицах (приложение П6) имеют вид

$$\hat{H}_A = \sum_{i=1}^{N_A} \left[ -\frac{1}{2} \Delta_i + V_A^N(\vec{r}_i) \right] + \frac{1}{2} \sum_{\substack{i,j=1 \\ (i \neq j)}}^{N_A} \frac{1}{r_{ij}} \quad (5.52)$$

и

$$\hat{H}_B = \sum_{i=N_A+1}^{N_A+N_B} \left[ -\frac{1}{2} \Delta_i + V_B^N(\vec{r}_i) \right] + \frac{1}{2} \sum_{\substack{i,j=N_A+1 \\ (i \neq j)}}^{N_A+N_B} \frac{1}{r_{ij}} \quad (5.53)$$

где  $V_A^N(\vec{r}_i)$  и  $V_B^N(\vec{r}_i)$  являются потенциальными энергиями взаимодействия  $i$ -го электрона с ядрами из систем  $A$  и  $B$ , соответственно. Очевидно (ср. гл. 1, разд. 1.1):

$$(\hat{H}_A + \hat{H}_B) \Psi_A \Psi_B = (E_A + E_B) \Psi_A \Psi_B. \quad (5.54)$$

Волновая функция  $\Psi_A \Psi_B$  не антисимметрична по парам электронов, относящихся к разным подсистемам. Можно было бы непосредственно ввести оператор антисимметризации (5.24) для  $N = N_A + N_B$  электронов, но тогда волновая функция уже не была бы собственной функцией  $\hat{H}_A + \hat{H}_B$ .

В такой ситуации можно поступить следующим образом. Если подсистемы строго невзаимодействующие, тогда  $V_A^N(\vec{r}_i)$ ,  $V_B^N(\vec{r}_j)$ , и  $\frac{1}{r_{ij}}$  должны обратиться в нуль, если электроны  $i$  и  $j$  принадлежат к подсистемам  $B$  и  $A$ , соответственно. Поэтому, объединяя все такие «перекрестные слабые» в

$$\hat{H}_{\text{cross}} = \sum_{i=N_A+1}^{N_A+N_B} V_A^N(\vec{r}_i) + \sum_{j=1}^{N_A} V_B^N(\vec{r}_j) + \sum_{j=1}^{N_A} \sum_{i=N_A+1}^{N_B} \frac{1}{r_{ij}}, \quad (5.55)$$

получаем для невзаимодействующих подсистем

$$\begin{aligned} \hat{H}_{\text{cross}} \Psi_A \Psi_B \\ = \hat{H}_{\text{cross}} \Psi_A(1, 2, \dots, N_A) \Psi_B(N_A + 1, N_A + 2, \dots, N_A + N_B) = 0. \end{aligned} \quad (5.56)$$

Можем сложить равенства (5.56) и (5.54) и получить

$$(\hat{H}_A + \hat{H}_B + \hat{H}_{\text{cross}}) \Psi_A \Psi_B = (E_A + E_B) \Psi_A \Psi_B \quad (5.57)$$

Теперь легко видеть (проведя тривиальные алгебраические выкладки), что сумма  $\hat{H}_A + \hat{H}_B + \hat{H}_{\text{cross}}$  симметрична по всем перестановкам электронов, включая их перестановки между системами  $A$  и  $B$ , как и должно

быть — это ничто иное, как гамильтониан Борна–Оппенгеймера, записанный для  $N_A + N_B$  электронов объединенной системы  $AB$ :

$$\begin{aligned} \hat{H}_{AB} = \hat{H}_A + \hat{H}_B + \hat{H}_{\text{cross}} = & \sum_{i=1}^{N_A+N_B} \left[ -\frac{1}{2} \Delta_i + V_A^N(\vec{r}_i) + V_B^N(\vec{r}_i) \right] \\ & + \frac{1}{2} \sum_{\substack{i,j=1 \\ (i \neq j)}}^{N_A+N_B} \frac{1}{r_{ij}}. \end{aligned} \quad (5.58)$$

Поэтому имеем

$$\hat{H}_{AB} \Psi_A \Psi_B = (E_A + E_B) \Psi_A \Psi_B. \quad (5.59)$$

Теперь используем тот факт, что гамильтониан  $\hat{H}_{AB}$  симметричен по всем электронам, поэтому он коммутирует со всеми их перестановками (разд. 4.2), включая те, которые переставляют электроны между двумя подсистемами. Так как антисимметризатор  $\hat{A}$  равен линейной комбинации перестановок, приходим к выражению

$$\hat{A} \hat{H}_{AB} \Psi_A \Psi_B = \hat{H}_{AB} \hat{A} \Psi_A \Psi_B = (E_A + E_B) \hat{A} \Psi_A \Psi_B, \quad (5.60)$$

т. е. антисимметризация электронов в случае двух не взаимодействующих подсистем никак не влияет на разделяемость уравнения Шрёдингера: полностью *антисимметричная* волновая функция  $\Psi_{AB} = \hat{A} \Psi_A \Psi_B$  является собственной функцией гамильтониана объединенной системы  $\hat{H}_{AB}$  с собственным значением  $E = E_A + E_B$ , в полном согласии с представлением о не взаимодействующих системах.

Здесь можно добавить следующее замечание. Если подсистемы находятся на конечном расстоянии друг от друга, то  $\hat{H}_{\text{cross}} \Psi_A \Psi_B$  уже не обращается в нуль и рассуждения, приведенные выше, неприменимы. Это объясняет происхождение дилеммы, возникающей при изучении межмолекулярных взаимодействий с использованием теории возмущений, в которой для определения невозмущенной задачи применяют не взаимодействующие молекулы: волновая функция, собственная для  $\hat{H}^0 = \hat{H}_A + \hat{H}_B$ , не полностью антисимметрична; если провести ее антисимметризацию, то она перестанет быть собственной функцией  $\hat{H}^0$ . В этой книге мы не рассматриваем межмолекулярную теорию возмущений; отметим лишь очень кратко, что в литературе представлены оба подхода к решению этой проблемы: либо включение антисимметризации в процедуру теории возмущений, как это сделано в разных вариантах так называемой «несимметричной по гамильтониану теории возмуще-

ний»<sup>1</sup>, либо полная переформулировка всей задачи с помощью метода вторичного квантования, в котором полная антисимметричность любой волновой функции гарантирована самим формализмом.

## 5. Инвариантность детерминантной волновой функции по отношению к «смешиванию» занятых орбиталей

Следующее свойство детерминантных волновых функций крайне важно. Если провести «смешивание» занятых спин-орбиталей в волновой функции детерминанта Слэтера  $\Psi = \hat{\mathcal{A}}[\varphi_1(1)\varphi_2(2)\dots\varphi_N(N)]$ , т. е. подвергнуть их любому неособенному (имеющему ненулевой определитель) линейному преобразованию путем замены спин-орбиталей  $\varphi_i$  на спин-орбитали

$$\psi_j = \sum_{k=1}^N T_{kj}\varphi_k, \quad (5.61)$$

то детерминантная волновая функция изменится лишь на физически несущественный нормировочный множитель, который равен определителю линейного преобразования:

$$\Psi' = \hat{\mathcal{A}}[\psi_1(1)\psi_2(2)\dots\psi_N(N)] = \text{Det}(\mathbf{T})\Psi. \quad (5.62)$$

Если преобразование орбиталей  $T$  унитарно, то множитель оказывается равным по модулю единице, т. е. может меняться лишь фаза волновой функции.

*Доказательство*

Подставим разложение (5.61) в выражение для волновой функции  $\Psi'$ :

$$\begin{aligned} \Psi' &= \hat{\mathcal{A}} \left[ \sum_{k_1=1}^N T_{k_1 1} \varphi_{k_1}(1) \sum_{k_2=1}^N T_{k_2 2} \varphi_{k_2}(2) \dots \sum_{k_N=1}^N T_{k_N N} \varphi_{k_N}(N) \right] \\ &= \sum_{k_1, k_2, \dots, k_N=1}^N T_{k_1 1} T_{k_2 2} \dots T_{k_N N} \hat{\mathcal{A}}[\varphi_{k_1}(1) \varphi_{k_2}(2) \dots \varphi_{k_N}(N)]. \end{aligned} \quad (5.63)$$

Известно, что определитель обращается в нуль, если два (или больше) столбца в нем одинаковы. Поэтому в последней сумме в выражении (5.63) останутся только те слагаемые, в которых все орбитали  $\varphi_{k_j}$  различны. Тогда набор  $\{k_1, k_2, \dots, k_N\}$  совпадает с набором  $\{1, 2, \dots, N\}$ , т. е.

<sup>1</sup> В английской литературе применяется термин «symmetry adapted perturbation theory» (SAPT).

каждое ненулевое слагаемое является детерминантом, построенным из исходных спин-орбиталей  $\varphi_i$ , варьируется только порядок последних. (В сумме будут встречаться все возможные перестановки  $\hat{K}$ , и каждая из них встретится только один раз.) Такой детерминант с переставленными орбиталями равен исходному  $\Psi$ , если перестановка  $\hat{K}^{-1}$ , действием которой восстанавливается исходный порядок, является четной, и равен  $-\Psi$ , если перестановка нечетная. Поэтому правая часть выражения (5.63) сводится к сумме по перестановкам  $\hat{K}$  с соответствующими факторами четности. (Четность  $\hat{K}$  и  $\hat{K}^{-1}$  одна и та же.) Получаем

$$\Psi' = \sum_{\hat{K} \in S_N} (-1)^k T_{K_1 1} T_{K_2 2} \dots T_{K_N N} \Psi . \quad (5.64)$$

Множитель в правой части выражения (5.64) есть ничто иное как определитель матрицы  $\mathbf{T}$

$$\sum_{\hat{K}} (-1)^k T_{K_1 1} T_{K_2 2} \dots T_{K_N N} \equiv \text{Det}(\mathbf{T}) . \quad (5.65)$$

Поэтому

$$\Psi' = \text{Det}(\mathbf{T}) \Psi . \text{ Ч. т. д. } \quad (5.66)$$

Этот же самый результат можно получить и с помощью теоремы об определителе произведения матриц: если  $\mathbf{A}$  и  $\mathbf{B}$  являются матрицами размера  $m \times m$ , то

$$\text{Det}(\mathbf{AB}) = \text{Det}(\mathbf{A}) \text{Det}(\mathbf{B}) . \quad (5.67)$$

Волновая функция  $\Psi$  является детерминантом Слэтера; она может быть записана как определитель матрицы  $\mathbf{A}$ , матричные элементы которой имеют вид  $A_{ij} = \varphi_j(i)$ .

$$\Psi = \frac{1}{\sqrt{N!}} \text{Det}\{A_{ij}\} = \frac{1}{\sqrt{N!}} \text{Det}(\mathbf{A}) . \quad (5.68)$$

Точно так же детерминант  $\Psi'$ , построенный из спин-орбиталей  $\psi_j$ , можно записать, как определитель матрицы  $\mathbf{B}$ , имеющей элементы  $B_{ij} = \psi_j(i)$ .

$$\Psi' = \frac{1}{\sqrt{N!}} \text{Det}\{B_{ij}\} = \frac{1}{\sqrt{N!}} \text{Det}(\mathbf{B}) . \quad (5.69)$$

Однако

$$B_{ij} = \psi_j(i) = \sum_{k=1}^N T_{kj} \varphi_k(i) = \sum_{k=1}^N A_{ik} T_{kj} = (\mathbf{AT})_{ij} , \quad (5.70)$$

$$\mathbf{B} = \mathbf{A}\mathbf{T}, \quad (5.71)$$

поэтому

$$\Psi' = \frac{1}{\sqrt{N!}} \text{Det}(\mathbf{B}) = \frac{1}{\sqrt{N!}} \text{Det}(\mathbf{A}\mathbf{T}) = \text{Det}(\mathbf{T})\Psi. \quad \text{Ч. т. д.} \quad (5.72)$$

Если преобразование унитарное,  $\mathbf{T} = \mathbf{U}$ ,  $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ , тогда  $|\text{Det}(\mathbf{T})| = |\text{Det}(\mathbf{U})| = 1$ , так как определитель любой унитарной матрицы равен единице по абсолютной величине. Поэтому детерминантная волновая функция *инвариантна* при унитарном преобразовании занятых орбиталей — за исключением, возможно, физически несущественного фазового множителя, по модулю равного единице.

Отметим, что указанные свойства инвариантности можно считать обобщением хорошо известного свойства определителя, согласно которому его значение не меняется, если к каждому элементу столбца добавить элементы другого столбца, умноженные на произвольную константу.

*Важным следствием* предыдущего результата является то, что каждую детерминантную волновую функцию можно представить (пренебрегая постоянным множителем) с помощью ортонормированных спин-орбиталей. Последние можно найти, подвергая занятые орбитали какой-либо процедуре ортонормирования. Например, симметричной ортогонализации по Лёвдину (гл. 3, разд. 2) отвечает использование в выражении (5.61) матрицы преобразования  $\mathbf{T} = \mathbf{S}^{-1/2}$ . (Здесь  $S_{ij} = \langle \varphi_i | \varphi_j \rangle$  — интеграл перекрывания спин-орбиталей, из которых построен детерминант  $\Psi$ .) Вообще, ортонормированные орбитали получаются с использованием любой матрицы преобразования, которую можно записать в виде  $\mathbf{T} = \mathbf{S}^{-1/2}\mathbf{V}$ , где  $\mathbf{V}$  — произвольная *унитарная* матрица. В самом деле

$$\begin{aligned} \langle \psi_i | \psi_j \rangle &= \left\langle \sum_{l=1}^N T_{li} \varphi_l \right| \left| \sum_{k=1}^N T_{kj} \varphi_k \right\rangle = \sum_{k,l=1}^N T_{li}^* \langle \varphi_l | \varphi_k \rangle T_{kj} = \sum_{k,l} (\mathbf{T}^\dagger)_{il} S_{lk} T_{kj} \\ &= (\mathbf{T}^\dagger \mathbf{S} \mathbf{T})_{ij} = (\mathbf{V}^\dagger \underbrace{\mathbf{S}^{-1/2} \mathbf{S} \mathbf{S}^{-1/2}}_{\mathbf{1}} \mathbf{V})_{ij} = (\mathbf{V}^\dagger \mathbf{V})_{ij} = \delta_{ij}, \end{aligned} \quad (5.73)$$

где использовано тождество  $(\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$ , а также то, что  $\mathbf{V}$  является унитарной матрицей ( $\mathbf{V}^\dagger = \mathbf{V}^{-1}$ ), а  $\mathbf{S}^{-1/2}$  — эрмитовой матрицей ( $\mathbf{S}^{-1/2\dagger} = \mathbf{S}^{-1/2}$ ).

На основании предыдущих результатов можно также прийти к следующему важному выводу: детерминантная волновая функция однозначно (с точностью до физически несущественного постоянного множителя) определяется *подпространством*, растянутым занятыми спин-

орбиталями. Можно выбрать *любой* набор  $N$  линейно независимых векторов (одноэлектронных функций), лежащих полностью в этом подпространстве, и построить из них детерминантную волновую функцию. Любые такие наборы  $N$  линейно независимых функций связаны друг с другом линейным преобразованием; следовательно, все такие наборы дают (с точностью до постоянного множителя) *одну и ту же*  $N$ -электронную детерминантную волновую функцию.

## 6. Матричные элементы между детерминантными волновыми функциями: общие формулы Лёвдина для неортогональных орбиталей

Слэтер получил выражения для матричных элементов различных операторов между волновыми функциями, заданными в виде детерминантов Слэтера для случая ортонормированных спин-орбиталей. Его результаты были обобщены Лёвдиным, который допустил использование также и неортогональных орбиталей.

### 6.1. Перекрывание

В квантовой химии «перекрыванием» или «интегралом перекрывания» обычно называется скалярное произведение  $\langle \varphi | \psi \rangle$  любой пары функций  $\varphi$  и  $\psi$ .

#### 6.1.1. Общий случай

Пусть  $U$  и  $V$  являются двумя однодетерминантными волновыми функциями, построенными из спин-орбиталей  $u_i$  и  $v_j$ :

$$U = \hat{\mathcal{A}}[u_1(1)u_2(2) \dots u_N(N)] \quad (5.74)$$

и

$$V = \hat{\mathcal{A}}[v_1(1)v_2(2) \dots v_N(N)] . \quad (5.75)$$

Перекрывание  $U$  и  $V$  равно

$$\langle U | V \rangle = \langle \hat{\mathcal{A}}[u_1(1)u_2(2) \dots u_N(N)] | \hat{\mathcal{A}}[v_1(1)v_2(2) \dots v_N(N)] \rangle . \quad (5.76)$$

Перенесем оператор антисимметризации  $\hat{\mathcal{A}}$  из «бра»- в «кет»-часть, воспользовавшись его эрмитовостью и свойством идемпотентности:  $\hat{\mathcal{A}}^\dagger = \hat{\mathcal{A}}$  и  $\hat{\mathcal{A}}^2 = \sqrt{N!}\hat{\mathcal{A}}$ , рассмотренными в разд. 4. Уравнение (5.76) при этом переписывается как

$$\langle U | V \rangle = \sqrt{N!} \langle u_1(1)u_2(2) \dots u_N(N) | \hat{\mathcal{A}}[v_1(1)v_2(2) \dots v_N(N)] \rangle . \quad (5.77)$$



Чтобы вычислить интеграл (5.77), мы воспользуемся представлением оператора антисимметризации (5.38), в котором он действует на индексы спин-орбиталей:

$$\begin{aligned}
 \langle U|V \rangle &= \sqrt{N!} \langle u_1(1)u_2(2) \dots u_N(N) | \text{ЛАНЬ}^{\circledast} \\
 &\quad \times \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^q v_{Q_1}(1)v_{Q_2}(2) \dots v_{Q_N}(N) \rangle \\
 &= \sum_{\hat{Q} \in S_N} (-1)^q \langle u_1(1)u_2(2) \dots u_N(N) | v_{Q_1}(1)v_{Q_2}(2) \dots v_{Q_N}(N) \rangle .
 \end{aligned} \tag{5.78}$$

Мы должны проинтегрировать по координатам всех электронов и просуммировать по проекциям спинов. Как легко видеть, каждый интеграл в правой части выражения (5.78) распадается на произведение интегралов по отдельным электронам; переменные интегрирования тогда не нужно указывать:

$$\langle U|V \rangle = \sum_{\hat{Q} \in S_N} (-1)^q \langle u_1|v_{Q_1} \rangle \langle u_2|v_{Q_2} \rangle \dots \langle u_N|v_{Q_N} \rangle . \tag{5.79}$$

Правая часть выражения (5.79) есть ничто иное, как формула, по которой в математике вводится определитель  $D$ , составленный из перекрываний  $\langle u_i|v_j \rangle$  отдельных пар спин-орбиталей. Таким образом, получаем общую формулу Лёвдина для перекрывания детерминантных волновых функций  $U$  и  $V$ :

$$\langle U|V \rangle = D = \text{Det}\{\langle u_i|v_j \rangle\} = \text{Det}(\mathbf{S}^{uv}) . \tag{5.80}$$

Этот результат означает, что интеграл перекрывания для двух многоэлектронных детерминантных волновых функций (детерминантов Слэтера) равен определителю матрицы взаимного перекрывания  $\mathbf{S}^{uv}$  спин-орбиталей, использованных для построения рассматриваемых детерминантных волновых функций ( $S_{ij}^{uv} = \langle u_i|v_j \rangle$ ).

### 6.1.2. Факторизация

В следующем разделе нам потребуется такое представление интеграла перекрывания  $\langle U|V \rangle$ , в котором выделено интегрирование по некоторым переменным.

а) Рассмотрим выражение (5.79), выберем какой-то индекс  $i$  ( $1 \leq i \leq N$ ) и выделим сомножитель, появляющийся при интегрировании по координатам  $i$ -го электрона:

$$\begin{aligned}
 \langle U|V \rangle &= \sum_{\hat{Q} \in S_N} (-1)^q \langle u_i | v_{Q_i} \rangle \langle u_1 | v_{Q_1} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \langle u_n | v_{Q_n} \rangle \\
 &= \sum_{j=1}^N \langle u_i | v_j \rangle \sum_{\hat{Q} \in S_N} \delta_{jQ_i} (-1)^q \langle u_1 | v_{Q_1} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \langle u_n | v_{Q_n} \rangle .
 \end{aligned} \tag{5.81}$$

В первой строке, по сравнению с выражением (5.79), изменен только порядок сомножителей, тогда как во второй строке мы использовали тот тривиальный факт, что индекс  $Q_i$  равен одному из целых чисел между 1 и  $N$ , и поэтому  $\sum_{j=1}^N \langle u_i | v_j \rangle \delta_{jQ_i} = \langle u_i | v_{Q_i} \rangle$ . Сравнивая выражения (5.81) и (5.80), можно легко узнать в (5.81) разложение определителя  $D$  по  $i$ -й строке, согласно известной теореме Лапласа:

$$\langle U|V \rangle = D = \text{Det}\{\langle u_i | v_j \rangle\} = \sum_{j=1}^N \langle u_i | v_j \rangle D(i|j) , \tag{5.82}$$

где  $D(i|j) = (-1)^{i+j} M(i|j) - i, j$ -е «алгебраическое дополнение» определителя  $D$ , т. е. минор  $M(i|j)$  (определитель, полученный вычеркиванием  $i$ -й строки и  $j$ -го столбца из исходного детерминанта), умноженный на  $(-1)^{i+j}$ . Поэтому коэффициент при  $\langle u_i | v_j \rangle$  в формуле (5.81) равен:

$$\sum_{\hat{Q} \in S_N} \delta_{jQ_i} (-1)^q \langle u_1 | v_{Q_1} \rangle \dots \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \langle u_n | v_{Q_n} \rangle = D(i|j) . \tag{5.83}$$

б) Теперь выберем два индекса  $i$  и  $j$  ( $1 \leq i < j \leq N$ ) и выделим сомножители, отвечающие интегрированию по координатам  $i$ -го и  $j$ -го электронов. Сначала мы проводим те же преобразования, что и в предыдущем случае, но теперь введем дополнительное суммирование и сим-

волы Кронекера для двух индексов:

$$\begin{aligned}
 \langle U|V \rangle &= \sum_{\hat{Q} \in S_N} (-1)^q \langle u_i | v_{Q_i} \rangle \langle u_j | v_{Q_j} \rangle \langle u_1 | v_{Q_1} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \\
 &\quad \times \langle u_{j-1} | v_{Q_{j-1}} \rangle \langle u_{j+1} | v_{Q_{j+1}} \rangle \dots \langle u_N | v_{Q_N} \rangle \\
 &= \sum_{k, l=1}^N \langle u_i | v_k \rangle \langle u_j | v_l \rangle \sum_{\hat{Q} \in S_N} \delta_{kQ_i} \delta_{lQ_j} (-1)^q \langle u_1 | v_{Q_1} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \\
 &\quad \times \langle u_{j-1} | v_{Q_{j-1}} \rangle \langle u_{j+1} | v_{Q_{j+1}} \rangle \dots \langle u_N | v_{Q_N} \rangle .
 \end{aligned} \tag{5.84}$$

В правой части выражения (5.84) индексы суммирования  $k$  и  $l$  изменяются независимо друг от друга. (Только слагаемые с  $k \neq l$  дают вклад в сумму, так как для каждой перестановки  $\hat{Q}$  имеем  $Q_i \neq Q_j$ ; тогда  $\delta_{kQ_i} \delta_{lQ_j} = 0$ , если  $k = l$ .) Для данной пары чисел  $k$  и  $l$  в сумме имеются два слагаемых, которые отличаются лишь перестановкой  $k$  и  $l$ . Эта переменна индексов означает, что соответствующие перестановки  $\hat{Q}$  отличаются одной транспозицией, поэтому их четности противоположны:  $(-1)^q$  для одной и  $-(-1)^q$  для другой. В остальном коэффициенты при этих слагаемых одинаковы. Тогда получаем

$$\begin{aligned}
 \langle U|V \rangle &= \sum_{k < l} (\langle u_i | v_k \rangle \langle u_j | v_l \rangle - \langle u_i | v_l \rangle \langle u_j | v_k \rangle) \\
 &\quad \times \sum_{\hat{Q} \in S_N} \delta_{kQ_i} \delta_{lQ_j} (-1)^q \langle u_1 | v_{Q_1} \rangle \dots \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \\
 &\quad \times \langle u_{j-1} | v_{Q_{j-1}} \rangle \langle u_{j+1} | v_{Q_{j+1}} \rangle \dots \langle u_N | v_{Q_N} \rangle .
 \end{aligned} \tag{5.85}$$

Слагаемое в круглых скобках является определителем второго порядка:

$$\begin{vmatrix} \langle u_i | v_k \rangle & \langle u_i | v_l \rangle \\ \langle u_j | v_k \rangle & \langle u_j | v_l \rangle \end{vmatrix}$$

(минор второго порядка для определителя  $D$ ). Таким образом, интеграл перекрывания  $\langle U|V \rangle$  можно записать, как сумму таких определителей с коэффициентами, происходящими из интегрирования по координатам других электронов, отличных от  $i$ -го и  $j$ -го. Уравнение (5.85) необходимо сравнить со следующей обобщенной теоремой Лапласа. Для любой пары строк  $i < j$ , любой определитель может быть разложен в сумму всех его миноров второго порядка, образованных элементами строк  $i$  и  $j$ , и всех

возможных столбцов  $k < l$ , умноженных на их алгебраические дополнения, т. е. миноров порядка  $N - 2$ , полученных исключением строк  $i, j$  и столбцов  $k, l$  из исходного определителя и умноженных на  $(-1)^{i+j+k+l}$ ,  $D(ij|kl) = (-1)^{i+j+k+l} M(ij|kl)$ :

$$\begin{aligned} \langle U|V \rangle &= D = \text{Det}\{\langle u_i|v_j \rangle\} \\ &= \sum_{k < l} (\langle u_i|v_k \rangle \langle u_j|v_l \rangle - \langle u_i|v_l \rangle \langle u_j|v_k \rangle) D(ij|kl). \end{aligned} \quad (5.86)$$

(Теорему Лапласа можно обобщить на миноры любого порядка и их соответствующие алгебраические дополнения.)

Сравнение с формулой (5.86) показывает, что коэффициент при  $(\langle u_i|v_k \rangle \langle u_j|v_l \rangle - \langle u_i|v_l \rangle \langle u_j|v_k \rangle)$  в формуле (5.85) равен

$$\begin{aligned} \sum_{Q \in S_N} \delta_{kQ_i} \delta_{lQ_j} (-1)^q \langle u_1|v_{Q_1} \rangle \dots \langle u_{i-1}|v_{Q_{i-1}} \rangle \langle u_{i+1}|v_{Q_{i+1}} \rangle \dots \\ \times \langle u_{j-1}|v_{Q_{j-1}} \rangle \langle u_{j+1}|v_{Q_{j+1}} \rangle \dots \langle u_N|v_{Q_N} \rangle = D(ij|kl). \end{aligned} \quad (5.87)$$

### 6.1.3. Частный случай ортонормированных орбиталей (правила Слэтера)

При вычислении матричных элементов орбитали двух детерминантов обычно располагают так, чтобы две последовательности орбиталей максимально совпадали, изменяя если нужно их порядок и вводя множитель  $\pm 1$ . Здесь и далее мы будем рассматривать частные случаи, предполагая, что эти перестановки орбиталей уже выполнены.

1. Рассмотрим волновую функцию  $\Psi = \hat{\mathcal{A}}[\varphi_1(1)\varphi_2(2)\dots\varphi_N(N)]$  и найдём интеграл нормировки  $\langle \Psi|\Psi \rangle$ . В этом случае мы должны использовать формулу (5.80) для случая  $U = V = \Psi$  и  $u_i = v_i = \varphi_i$ . Тогда  $S_{ij}^{uv} = S_{ij} = \langle \varphi_i|\varphi_j \rangle$  и

$$\langle \Psi|\Psi \rangle = \text{Det}\{S_{ij}\} = \text{Det}(\mathbf{S}). \quad (5.88)$$

Если орбитали  $\varphi_i$  ортонормированы, то  $S_{ij} = \langle \varphi_i|\varphi_j \rangle = \delta_{ij}$ ,  $\mathbf{S} = \mathbf{1}$ , и детерминант в правой части выражения (5.88) равен единице: *волновая функция, заданная в виде слэтеровского детерминанта, построенного из ортонормированных спин-орбиталей, нормирована на единицу.* (Данное условие нормировки мотивирует выбор множителя  $\frac{1}{\sqrt{N!}}$  в определениях детерминанта Слэтера (5.15) и оператора антисимметризации (5.24).)

2. Если одна из детерминантных волновых функций  $U$  или  $V$  содержит одну или более спин-орбиталей, которые ортогональны всем спин-орбиталям другой волновой функции, то  $\mathbf{S}^{uv}$  будет содержать одну или более строк или столбцов, все элементы которых равны нулю. Тогда  $\text{Det}(\mathbf{S}^{uv}) = 0$  согласно известному свойству определителей; в этом случае многоэлектронные детерминантные волновые функции  $U$  и  $V$  также ортогональны.

Поэтому, если используется ортонормированный набор спин-орбиталей, то ненулевые интегралы перекрывания имеются лишь между теми детерминантными волновыми функциями, которые содержат те же самые спин-орбитали. Такой интеграл перекрывания  $\langle U|V \rangle$  равен единице, если порядок орбиталей в двух детерминантах  $U$  и  $V$  одинаков или если порядок орбиталей связан четной перестановкой (т. е. если порядок орбиталей становится одинаковым после четного числа транспозиций); перекрывание  $\langle U|V \rangle$  равно  $-1$ , если порядок орбиталей в двух детерминантах связан нечетной перестановкой.

### Иллюстрация

Уравнение (5.80) можно также использовать для доказательства инвариантности детерминантных волновых функций по отношению к унитарным преобразованиям спин-орбиталей.

Пусть  $\Psi$  есть детерминант Слэтера, построенный из ортонормированных спин-орбиталей  $\varphi_i$  ( $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$ ):

$$\Psi = \hat{\mathcal{A}}[\varphi_1(1)\varphi_2(2)\dots\varphi_N(N)] \quad (5.89)$$

и пусть

$$\Psi' = \hat{\mathcal{A}}[\psi_1(1)\psi_2(2)\dots\psi_N(N)] \quad (5.90)$$

есть детерминант Слэтера, построенный из спин-орбиталей  $\psi_j$ , полученных унитарным преобразованием орбиталей  $\varphi_i$ :

$$\psi_j = \sum_{k=1}^N U_{kj} \varphi_k. \quad (5.91)$$

Так как  $\mathbf{U}$  унитарна ( $\mathbf{U}^{-1} = \mathbf{U}^\dagger$ ), спин-орбитали  $\psi_j$  также ортонормированы. Так как  $\Psi$  и  $\Psi'$  обе построены из ортонормированных спин-орбиталей, то они обе нормированы на единицу:

$$\langle \Psi | \Psi \rangle = \langle \Psi' | \Psi' \rangle = 1. \quad (5.92)$$

Вычислим перекрывание  $\langle \Psi | \Psi' \rangle$ . В соответствии с результатом (5.80)

$$\begin{aligned} \langle \Psi | \Psi' \rangle &= \text{Det} \{ \langle \varphi_i | \psi_j \rangle \} = \text{Det} \left\{ \langle \varphi_i | \sum_{k=1}^N U_{kj} \varphi_k \rangle \right\} \\ &= \text{Det} \left\{ \sum_k U_{kj} \underbrace{\langle \varphi_i | \varphi_k \rangle}_{\delta_{ik}} \right\} = \text{Det} \{ U_{ij} \}. \end{aligned} \quad (5.93)$$

Определитель любой унитарной матрицы равен по модулю единице, поэтому  $|\langle \Psi | \Psi' \rangle| = 1$ . Если перекрывание двух нормированных волновых функций равно по модулю единице, то эти волновые функции тождественны, если не принимать во внимание физически несущественный фазовый множитель.

## 6.2. Одноэлектронные операторы

### 6.2.1. Общий случай

Пусть  $\hat{H}_1$  является суммой одноэлектронных операторов  $\hat{h}(i)$ , причем каждый  $\hat{h}(i)$  действует на функции, зависящие от координат  $i$ -го электрона:

$$\hat{H}_1 = \sum_{i=1}^N \hat{h}(i). \quad (5.94)$$

Эта сумма симметрична по отношению к перестановкам отдельных электронов  $i$ . Например,  $\hat{H}_1$  может быть одноэлектронной частью гамильтониана Борна–Оппенгеймера; в этом случае  $\hat{h}(i) = -\frac{1}{2}\Delta_i - \sum_{\alpha} \frac{Z_{\alpha}}{r_{\alpha i}}$  (в атомных единицах; см. приложение П6).

Найдем матричный элемент  $\langle U | \hat{H}_1 | V \rangle$ , где  $U$  и  $V$  снова являются слэтеровскими детерминантами, заданными выражениями (5.74) и (5.75), построенными из спин-орбиталей  $u_i, v_j$ . Вывод аналогичен тому, который использовался для вычисления интеграла перекрывания  $\langle U | V \rangle$ . Снова перенесем антисимметризатор из «бра»-части в «кет»-часть, воспользовавшись тем, что оператор  $\hat{H}_1$ , будучи симметричным по электронам, коммутирует с оператором антисимметризации  $\hat{A}$  (разд. 4.2), и применим свойства эрмитовости  $\hat{A}^\dagger = \hat{A}$  и идемпотентности  $\hat{A}^2 = \sqrt{N!} \hat{A}$  для оператора антисимметризации:

$$\begin{aligned} \langle U | \hat{H}_1 | V \rangle &= \langle \hat{A} u_1(1) u_2(2) \dots u_N(N) | \hat{H}_1 | \hat{A} [v_1(1) v_2(2) \dots v_N(N)] \rangle \\ &= \sqrt{N!} \langle u_1(1) u_2(2) \dots u_N(N) | \hat{H}_1 | \hat{A} [v_1(1) v_2(2) \dots v_N(N)] \rangle. \end{aligned} \quad (5.95)$$

Используем тот вид оператора антисимметризации, который действует на индексы спин-орбиталей, и получим

$$\begin{aligned}
 \langle U | \hat{H}_1 | V \rangle &= \sqrt{N!} \langle u_1(1) u_2(2) \dots u_N(N) | \sum_{i=1}^N h(i) | \frac{1}{\sqrt{N!}} \sum_{\hat{Q} \in S_N} (-1)^q v_{Q_1}(1) \\
 &\quad \times v_{Q_2}(2) \dots v_{Q_N}(N) \rangle \\
 &= \sum_{i=1}^N \sum_{\hat{Q} \in S_N} (-1)^q \langle u_1(1) u_2(2) \dots u_i(i) \dots u_N(N) | \hat{h}(i) | v_{Q_1}(1) \\
 &\quad \times v_{Q_2}(2) \dots v_{Q_i}(i) \dots v_{Q_N}(N) \rangle .
 \end{aligned} \tag{5.96}$$

Каждый из интегралов в правой части выражения (5.96) в свою очередь расписывается в виде произведения интегралов; все они, за исключением одного, относящегося к  $i$ -му электрону, являются теми же самыми интегралами перекрывания между спин-орбиталями, которые встречались при расчете перекрывания  $\langle U | V \rangle$ . Перенесем интеграл  $\langle u_i(i) | \hat{h}(i) | v_{Q_i}(i) \rangle$  на первое место (указывать явно переменные интегрирования при этом уже не нужно):

$$\begin{aligned}
 \langle U | \hat{H}_1 | V \rangle &= \sum_{i=1}^N \sum_{\hat{Q} \in S_N} (-1)^q \langle u_i | \hat{h} | v_{Q_i} \rangle \langle u_1 | v_{Q_1} \rangle \langle u_2 | v_{Q_2} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \langle u_N | v_{Q_N} \rangle .
 \end{aligned} \tag{5.97}$$

Так же, как и в случае интеграла перекрывания (5.81), воспользуемся тождеством  $\sum_{j=1}^N \langle u_i | \hat{h} | v_j \rangle \delta_{jQ_i} = \langle u_i | \hat{h} | v_{Q_i} \rangle$  и получим

$$\begin{aligned}
 \langle U | \hat{H}_1 | V \rangle &= \sum_{i=1}^N \sum_{\hat{Q} \in S_N} (-1)^q \sum_{j=1}^N \langle u_i | \hat{h} | v_j \rangle \delta_{jQ_i} \langle u_1 | v_{Q_1} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \langle u_N | v_{Q_N} \rangle \\
 &= \sum_{i,j=1}^N \langle u_i | \hat{h} | v_j \rangle \sum_{\hat{Q} \in S_N} (-1)^q \delta_{jQ_i} \langle u_1 | v_{Q_1} \rangle \dots \\
 &\quad \times \langle u_{i-1} | v_{Q_{i-1}} \rangle \langle u_{i+1} | v_{Q_{i+1}} \rangle \dots \langle u_N | v_{Q_N} \rangle .
 \end{aligned} \tag{5.98}$$

Сравнение с результатом (5.83) показывает, что коэффициент при интеграле  $\langle u_i | \hat{h} | v_j \rangle$  является алгебраическим дополнением  $D(i|j)$ , т. е. мино-

ром  $M(i|j)$  матрицы перекрывания  $\mathbf{S}^{uv}$ , умноженным на  $(-1)^{i+j}$ . Наконец, получаем результат:

$$\langle U|\hat{H}_1|V\rangle = \sum_{i,j=1}^N \langle u_i|\hat{h}|v_j\rangle D(i|j) . \quad (5.99)$$

### 6.2.2. Частный случай ортонормированных орбиталей (правила Слэтера)

#### 1. Среднее значение

Пусть  $\Psi = \hat{\mathcal{A}}[\psi_1(1)\psi_2(2)\dots\psi_N(N)]$  является волновой функцией в виде слэтеровского детерминанта, построенного из ортонормированных спин-орбиталей  $\psi_i = \psi_i(\vec{r}, \sigma) = \varphi_i(\vec{r})\gamma_i(\sigma)$ , где  $\varphi_i(\vec{r})$  и  $\gamma_i(\sigma)$  обозначают пространственную и спиновую части  $\psi_i$ , соответственно ( $\gamma_i = \alpha$  или  $\beta$ ). Тогда мы должны воспользоваться формулой (5.99), полагая в ней  $U = V = \Psi$ ;  $u_i = v_i = \psi_i$ . Матрица взаимного перекрывания равна  $\mathbf{S}^{uv} = \mathbf{S} = \mathbf{1}$ ; поэтому  $D(i|j) = \delta_{ij}$ , так как только «главные миноры» единичной матрицы отличны от нуля, и они равны единице. Тогда имеем

$$\begin{aligned} \langle \Psi|\hat{H}_1|\Psi\rangle &= \langle \Psi|\sum_{i=1}^N \hat{h}(i)|\Psi\rangle = \sum_{i,j=1}^N \langle \psi_i|\hat{h}|\psi_j\rangle \underbrace{D(i|j)}_{\delta_{ij}} \\ &= \sum_{i=1}^N \langle \psi_i|\hat{h}|\psi_i\rangle = \sum_{i=1}^N \langle \varphi_i|\hat{h}|\varphi_i\rangle , \end{aligned} \quad (5.100)$$

где в последнем выражении мы выполнили суммирование по спиновым переменным, так что оставшиеся интегралы берутся только по пространственным переменным. Таким образом, получаем результат:

$$\langle \Psi|\sum_{i=1}^N \hat{h}(i)|\Psi\rangle = \sum_{i=1}^N \langle \varphi_i|\hat{h}|\varphi_i\rangle . \quad (5.101)$$

Заметим, что волновая функция  $\Psi$ , построенная из ортонормированных спин-орбиталей, нормирована на единицу, поэтому среднее значение  $\hat{H}_1$  по  $\Psi$  равно матричному элементу  $\langle \Psi|\hat{H}_1|\Psi\rangle$  в выражении (5.101).

Благодаря ортонормированности спиновых функций  $\alpha$  и  $\beta$ , требование ортонормированности спин-орбиталей  $\psi_i$  можно сформулировать, как  $\langle \psi_i|\psi_j\rangle = \langle \varphi_i|\varphi_j\rangle\delta_{\gamma_i\gamma_j} = \delta_{ij}$ . Поэтому достаточно потребовать того, чтобы пространственные части спин-орбиталей, имеющих *одинаковый спин*, были ортогональны, тогда как на пространственные части спин-орбиталей, относящихся к разным проекциям спина ограничений



нет и они часто полагаются попарно идентичными. В случае использования таких *дважды занятых орбиталей* имеем  $\psi_{2i-1} = \varphi_i(\vec{r})\alpha(\sigma)$ ,  $\psi_{2i} = \varphi_i(\vec{r})\beta(\sigma)$ , и выражение (5.101) сводится к

$$\langle \Psi_R | \sum_{i=1}^N \hat{h}(i) | \Psi_R \rangle = 2 \sum_{i=1}^{N/2} \langle \varphi_i | \hat{h}(i) | \varphi_i \rangle, \quad (5.102)$$

где индекс « $R$ » относится к использованию дважды занятых орбиталей, как обычно бывает в «ограниченном» (restricted) методе Хартри–Фока. В формуле (5.102) сумма ведется только по разным *пространственным орбиталям*.

2. *Матричный элемент между детерминантами, различающимися на одну спин-орбиталь.*

Рассмотрим детерминант

$$V = \Psi = \hat{\mathcal{A}}[\psi_1(1)\psi_2(2)\dots\psi_N(N)] \quad (5.103)$$

— тот же, что и в предыдущем случае, — и детерминант

$$\begin{aligned} U &= \Psi_1(\psi_k \rightarrow \psi'_k) \\ &= \hat{\mathcal{A}}[\psi_1(1)\psi_2(2)\dots\psi_{k-1}(k-1)\psi'_k(k)\psi_{k+1}(k+1)\dots\psi_N(N)], \end{aligned} \quad (5.104)$$

отличающийся от  $\Psi$  заменой спин-орбитали  $\psi_k$  на спин-орбиталь  $\psi'_k$ , которая ортогональна ко всем спин-орбиталям  $\psi_i$ ,  $i = 1, 2, \dots, N$ . (Конечно, имеет смысл рассматривать только спин-орбитали  $\psi'_k$ , имеющие тот же спин, что и  $\psi_k$ , иначе  $U$  и  $V$  соответствовали бы разным проекциям полного спина  $S_z$  и любой матричный элемент бесспинового оператора между ними обратился бы в нуль. Это означает, что пространственная часть  $\psi'_k$  должна быть ортогональна к пространственным частям всех спин-орбиталей того же спина:  $\langle \varphi'_k | \varphi_i \rangle = 0$  если  $\gamma_k = \gamma_i$ .)

В этом случае матрица взаимного перекрывания  $\mathbf{S}^{uv}$  отличается от единичной матрицы в одном месте:  $k$ -й диагональный элемент равен нулю, так как  $\langle \psi'_k | \psi_k \rangle = 0$ . Поэтому все миноры  $M(i|j)$  матрицы  $\mathbf{S}^{uv}$  равны нулю, за исключением единственного, который получается вычеркиванием  $k$ -й строки и  $k$ -го столбца, причем этот минор равен единице. Тогда  $D(i|j) = (-1)^{i+j} M(i|j) = \delta_{ik} \delta_{jk}$ . Подставляя этот результат в общую формулу (5.99) и проводя также суммирование по спиновым переменным, получим

$$\langle \Psi_1(\psi_k \rightarrow \psi'_k) | \sum_{i=1}^N \hat{h}(i) | \Psi \rangle = \langle \psi'_k | \hat{h} | \psi_k \rangle = \langle \varphi'_k | \hat{h} | \varphi_k \rangle. \quad (5.105)$$

3. Если используются ортонормированные спин-орбитали, *матричный элемент*  $\langle U|\hat{H}_1|V\rangle$  *исчезает, если определители*  $U$  *и*  $V$  *отличаются на две или больше спин-орбитали*: все миноры  $M(i|j)$  матрицы  $\mathbf{S}^{uv}$  будут нулевыми, так как они содержат одну или более строк и столбцов, все элементы которых равны нулю.

### 6.3. Двухэлектронные операторы

#### 6.3.1. Общий случай

Пусть  $\hat{H}_2$  является суммой двухэлектронных операторов  $\hat{g}(i, j) = \hat{g}(j, i)$ , симметричной относительно перестановок всех электронов:

$$\hat{H}_2 = \sum_{i < j}^N \hat{g}(i, j). \quad (5.106)$$

Каждый оператор  $\hat{g}(i, j)$  действует на функции координат электронов  $i$  и  $j$ . Наиболее важным оператором такого типа является двухэлектронная часть гамильтониана Борна–Оппенгеймера; в этом случае (в атомных единицах)  $\hat{g}(i, j) = \frac{1}{r_{ij}}$ .

Будем действовать аналогично двум предыдущим случаям. Снова воспользуемся эрмитовостью и идемпотентностью оператора антисимметризации  $\hat{A}$ , а также тем фактом, что он коммутирует с оператором  $\hat{H}_2$ , симметричным по перестановкам электронов:

$$\begin{aligned} \langle U|\hat{H}_2|V\rangle &= \langle \hat{A}[u_1(1)u_2(2)\dots u_N(N)]|\hat{H}_2|\hat{A}[v_1(1)v_2(2)\dots v_N(N)]\rangle \\ &= \sqrt{N!}\langle [u_1(1)u_2(2)\dots u_N(N)]|\hat{H}_2|\hat{A}[v_1(1)v_2(2)\dots v_N(N)]\rangle \\ &= \sqrt{N!}\langle u_1(1)u_2(2)\dots u_N(N)|\sum_{i < j}^N \hat{g}(i, j)|\frac{1}{\sqrt{N!}}\sum_{\hat{Q}\in S_N} (-1)^q v_{Q_1}(1) \\ &\quad \times v_{Q_2}(2)\dots v_{Q_N}(N)\rangle \\ &= \sum_{i < j}^N \sum_{\hat{Q}\in S_N} (-1)^q \langle u_1(1)\dots u_i(i)\dots u_j(j)\dots u_N(N)|\hat{g}(i, j)|v_{Q_1}(1)\dots \\ &\quad \times v_{Q_i}(i)\dots v_{Q_j}(j)\dots v_{Q_N}(N)\rangle. \end{aligned} \quad (5.107)$$

Интегралы в правой части опять распадаются на произведения интегралов перекрывания, содержащих интегрирование по координатам отдельных электронов, за исключением координат  $i$ -го и  $j$ -го электронов. Вы-

нося соответствующие интегралы  $\langle u_i(i)u_j(j)|\hat{g}(i,j)|v_{Q_i}(i)v_{Q_j}(j)\rangle$  на первое место, получим

$$\begin{aligned} \langle U|\hat{H}_2|V\rangle &= \sum_{i<j}^N \sum_{\hat{Q}\in S_N} (-1)^q \langle u_i(i)u_j(j)|\hat{g}(i,j)|v_{Q_i}(i)v_{Q_j}(j)\rangle \\ &\times \langle u_1|v_{Q_1}\rangle \dots \langle u_{i-1}|v_{Q_{i-1}}\rangle \langle u_{i+1}|v_{Q_{i+1}}\rangle \dots \\ &\times \langle u_{j-1}|v_{Q_{j-1}}\rangle \langle u_{j+1}|v_{Q_{j+1}}\rangle \dots \langle u_N|v_{Q_N}\rangle. \end{aligned} \quad (5.108)$$

Введем два символа Кронекера, как и при выводе формулы (5.84), что дает

$$\begin{aligned} \langle U|\hat{H}_2|V\rangle &= \sum_{i<j}^N \sum_{k,l=1}^N \sum_{\hat{Q}\in S_N} \delta_{kQ_i}\delta_{lQ_j} (-1)^q \langle u_i(i)u_j(j)|\hat{g}(i,j)|v_k(i)v_l(j)\rangle \\ &\times \langle u_1|v_{Q_1}\rangle \dots \langle u_{i-1}|v_{Q_{i-1}}\rangle \langle u_{i+1}|v_{Q_{i+1}}\rangle \dots \\ &\times \langle u_{j-1}|v_{Q_{j-1}}\rangle \langle u_{j+1}|v_{Q_{j+1}}\rangle \dots \langle u_N|v_{Q_N}\rangle. \end{aligned} \quad (5.109)$$

В последующих уравнениях мы будем использовать в качестве обозначений переменных интегрирования символы «1» и «2» вместо  $i$  и  $j$ . (Здесь «1» и «2» обозначают совокупность  $\{\vec{r}_1, \sigma_1\}$  и  $\{\vec{r}_2, \sigma_2\}$  пространственных и спиновых координат каждого из двух электронов.) Так же, как и в выкладках, ведущих от выражения (5.84) к выражению (5.85), мы можем объединить два случая, для которых  $k$  и  $l$  переставлены, а все другие индексы  $Q_p$  одинаковы; они соответствуют перестановкам с противоположной четностью, но в остальном имеют один и тот же коэффициент. Тогда

$$\begin{aligned} \langle U|\hat{H}_2|V\rangle &= \sum_{i<j}^N \sum_{k<l}^N [\langle u_i(1)u_j(2)|\hat{g}(1,2)|v_k(1)v_l(2)\rangle \\ &\quad - \langle u_i(1)u_j(2)|\hat{g}(1,2)|v_l(1)v_k(2)\rangle] \\ &\times \sum_{\hat{Q}\in S_N} \delta_{kQ_i}\delta_{lQ_j} (-1)^q \langle u_1|v_{Q_1}\rangle \dots \langle u_{i-1}|v_{Q_{i-1}}\rangle \langle u_{i+1}|v_{Q_{i+1}}\rangle \dots \\ &\quad \times \langle u_{j-1}|v_{Q_{j-1}}\rangle \langle u_{j+1}|v_{Q_{j+1}}\rangle \dots \langle u_N|v_{Q_N}\rangle. \end{aligned} \quad (5.110)$$

Сравнение с выражением (5.87) показывает, что коэффициент при скобках, содержащих разность двух двухэлектронных интегралов, равен ал-

гебраическому дополнению  $D(ij|kl)$ , т. е. минору  $M(ij|kl)$  матрицы перекрывания  $\mathbf{S}^{uv}$ , домноженному на  $(-1)^{i+j+k+l}$ . Итак, получаем результат:

$$\langle U|\hat{H}_2|V\rangle = \sum_{\substack{i < j \\ k < l}}^N [\langle u_i(1)u_j(2)|\hat{g}(1,2)|v_k(1)v_l(2)\rangle - \langle u_i(1)u_j(2)|\hat{g}(1,2)|v_l(1)v_k(2)\rangle] D(ij|kl) . \quad (5.111)$$

Другой, но эквивалентный, вид этой формулы можно найти, если воспользоваться симметрией интегралов по отношению к перестановке переменных интегрирования:

$$\langle u_i(1)u_j(2)|\hat{g}(1,2)|v_k(1)v_l(2)\rangle = \langle u_j(1)u_i(2)|\hat{g}(1,2)|v_l(1)v_k(2)\rangle \quad (5.112)$$

а также обобщить определение величин  $D(ij|kl)$  на случай  $i > j$  и/или  $k > l$ , полагая их антисимметричными по обеим парам индексов  $i, j$  и  $k, l$ . Тогда (это проверяется тривиально) уравнение (5.111) можно также переписать и в более компактном виде

$$\langle U|\hat{H}_2|V\rangle = \frac{1}{2} \sum_{i,j,k,l=1}^N \langle u_i(1)u_j(2)|\hat{g}(1,2)|v_k(1)v_l(2)\rangle D(ij|kl) . \quad (5.113)$$

Этот вид был первоначально получен Лёвдиным. В то же время, запись (5.111) более удобна в большинстве практических приложений.

### 6.3.2. Частный случай ортонормированных орбиталей (правила Слэтера)

#### 1. Среднее значение

Это тот самый случай, который обсуждался выше, когда мы рассматривали среднее значение одноэлектронного оператора  $\hat{H}_1$ :  $U = V = \Psi$ ;  $u_i = v_i = \psi_i = \varphi_i \gamma_i$ . Снова верно, что только главные миноры матрицы  $\mathbf{S}^{uv} = \mathbf{1}$  отличны от нуля, и что они равны 1. Следовательно, если мы вычеркиваем строку  $i$ , то необходимо вычеркнуть и столбец  $i$ , и наоборот, т. е.  $D(ij|kl) = \delta_{ik}\delta_{jl}$  (предполагается, что  $i < j, k < l$ ). Уравнение (5.111) сводится к выражению

$$\langle \Psi|\hat{H}_2|\Psi\rangle = \sum_{i < j}^N [\langle \psi_i(1)\psi_j(2)|\hat{g}(1,2)|\psi_i(1)\psi_j(2)\rangle - \langle \psi_i(1)\psi_j(2)|\hat{g}(1,2)|\psi_j(1)\psi_i(2)\rangle] . \quad (5.114)$$

Здесь легко явным образом просуммировать по спиновым переменным. В первом интеграле это просто дает единичный множитель, так как для обоих электронов имеем одну и ту же спин-орбиталь в «бра»- и в «кет»-частях. Суммирование по спиновым переменным во втором интеграле дает множитель  $\delta_{\gamma_i \gamma_j}$ , т. е. единицу, если спин-орбитали  $\psi_i$  и  $\psi_j$  имеют одинаковые проекции спина, и нуль, если их проекции спина различны. Таким образом, записывая результат через пространственные орбитали, получаем

$$\langle \Psi | \hat{H}_2 | \Psi \rangle = \sum_{i < j}^N [\langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_j(2) \rangle - \langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_j(1) \varphi_i(2) \rangle \delta_{\gamma_i \gamma_j}] . \quad (5.115)$$

Иногда удобнее использовать выражение (5.115) в таком виде, чтобы индексы  $i$  и  $j$  входили симметрично:

$$\langle \Psi | \hat{H}_2 | \Psi \rangle = \frac{1}{2} \sum_{i, j=1}^N [\langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_j(2) \rangle - \langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_j(1) \varphi_i(2) \rangle \delta_{\gamma_i \gamma_j}] . \quad (5.116)$$

Заметим, что слагаемое с  $i = j$  сокращается автоматически.

В случае *дважды занятых* орбиталей (индекс  $R$ ), перейдем к суммам по разным *пространственным* орбиталям, и тогда выражения (5.115) и (5.116) переписываются:

$$\begin{aligned} \langle \Psi_R | \hat{H}_2 | \Psi_R \rangle &= \sum_{i < j}^{N/2} [4 \langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_j(2) \rangle - 2 \langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_j(1) \varphi_i(2) \rangle] \\ &+ \sum_{i=1}^{N/2} \langle \varphi_i(1) \varphi_i(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_i(2) \rangle \\ &= \sum_{i, j=1}^{N/2} [2 \langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_j(2) \rangle - \langle \varphi_i(1) \varphi_j(2) | \hat{g}(1, 2) | \varphi_j(1) \varphi_i(2) \rangle] . \end{aligned} \quad (5.117)$$

*2. Матричные элементы между детерминантами, отличающимися одной спин-орбиталью*

Снова сошлемся на обсуждение похожего случая, рассмотренного в связи с одноэлектронным оператором  $\hat{H}_1$ , и использованные там обозначения. Рассматривая выражение (5.111), мы видим, что одна из строк

и один из столбцов матрицы  $\mathbf{S}^{uv}$ , вычеркиваемых при формировании алгебраического дополнения  $D(ij|kl)$ , должны совпасть со строкой и столбцом, соответствующим той орбитали, которой два детерминанта отличаются друг от друга (мы предполагаем, что это  $k$ -я орбиталь), иначе остались бы строка и столбец, содержащие только нули. По этой же причине и вторые строка и столбец, вычеркнутые при образовании минора, должны иметь одинаковый номер. Тогда мы получаем ненулевые алгебраические дополнения  $D(ik|ik)$  для  $i < k$  и  $D(ki|ki)$  для  $i > k$ , которые, очевидно, равны единице. Так как двухэлектронные интегралы симметричны по перестановкам переменных интегрирования, нет необходимости различать далее эти случаи. Получаем

$$\begin{aligned} \langle \Psi_1(\psi_k \rightarrow \psi'_k) | \hat{H}_2 | \Psi \rangle = & \sum_{\substack{i \\ (i \neq k)}}^N [\langle \psi'_k(1) \psi_i(2) | \hat{g}(1, 2) | \psi_k(1) \psi_i(2) \rangle \\ & - \langle \psi'_k(1) \psi_i(2) | \hat{g}(1, 2) | \psi_i(1) \psi_k(2) \rangle] . \end{aligned} \quad (5.118)$$

Ограничение  $i \neq k$  в сумме, очевидно, можно опустить, так как соответствующее слагаемое с  $i = k$  автоматически равно нулю. Можно переписать формулу (5.118) через пространственные орбитали:

$$\begin{aligned} \langle \Psi_1(\psi_k \rightarrow \psi'_k) | \hat{H}_2 | \Psi \rangle = & \sum_{\substack{i=1 \\ (i \neq k)}}^N [\langle \varphi'_k(1) \varphi_i(2) | \hat{g}(1, 2) | \varphi_k(1) \varphi_i(2) \rangle \\ & - \langle \varphi'_k(1) \varphi_i(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_k(2) \rangle \delta_{\gamma_i \gamma_k}] . \end{aligned} \quad (5.119)$$

При использовании дважды занятых орбиталей уравнение (5.119) можно переписать как

$$\begin{aligned} \langle \Psi_{R1}(\psi_k \rightarrow \psi'_k) | \hat{H}_2 | \Psi_R \rangle = & \sum_{i=1}^{N/2} [2 \langle \varphi'_k(1) \varphi_i(2) | \hat{g}(1, 2) | \varphi_k(1) \varphi_i(2) \rangle \\ & - \langle \varphi'_k(1) \varphi_i(2) | \hat{g}(1, 2) | \varphi_i(1) \varphi_k(2) \rangle] , \end{aligned} \quad (5.120)$$

где сумма берется по различным пространственным орбиталям и для простоты мы опустили ограничение  $i \neq k$ . (Иначе понадобилось бы рассматривать отдельно слагаемое  $\langle \varphi'_k(1) \varphi_k(2) | \hat{g}(1, 2) | \varphi_k(1) \varphi_k(2) \rangle$ , появляющееся в случае, когда  $\psi_k$  и  $\psi_i$  в выражении (5.118) соответствуют одной и той же пространственной орбитали, занятой электронами с разными проекциями спина.)

3. Матричные элементы между детерминантами, отличающимися двумя спин-орбиталями

В этом случае  $U = \Psi_2(\psi_k \rightarrow \psi'_k, \psi_l \rightarrow \psi'_l)$ , и имеется единственное ненулевое алгебраическое дополнение  $D(kl|kl)$ , причем оно равно единице. Таким образом получаем, проводя также суммирование по спиновым переменным:

$$\begin{aligned} & \langle \Psi_2(\psi_k \rightarrow \psi'_k, \psi_l \rightarrow \psi'_l) | \hat{H}_2 | \Psi \rangle \\ &= \langle \psi'_k(1) \psi'_l(2) | \hat{g}(1, 2) | \psi_k(1) \psi_l(2) \rangle - \langle \psi'_k(1) \psi'_l(2) | \hat{g}(1, 2) | \psi_l(1) \psi_k(2) \rangle \\ &= \langle \varphi'_k(1) \varphi'_l(2) | \hat{g}(1, 2) | \varphi_k(1) \varphi_l(2) \rangle - \langle \varphi'_k(1) \varphi'_l(2) | \hat{g}(1, 2) | \varphi_l(1) \varphi_k(2) \rangle \delta_{\gamma_k \gamma_l}. \end{aligned} \quad (5.121)$$

4. Если используются ортонормированные спин-орбитали, то матричный элемент  $\langle U | \hat{H}_2 | V \rangle$  обращается в нуль, если детерминанты  $U$  и  $V$  отличаются тремя или более спин-орбиталями: ненулевых миноров  $M(ij|kl)$  в этом случае не существует.

## 7. Теорема парности Лёвдина и ее обобщение

Теорема парности Лёвдина использует инвариантность детерминантных волновых функций по отношению к унитарным преобразованиям орбиталей (разд. 5). Эта теорема применима в тех случаях, когда можно разделить спин-орбитали на два отдельных набора (или несколько пар различных наборов) по какому-либо признаку. В исходной формулировке Лёвдин рассматривал два набора, образованные спин-орбиталями со спинами  $\alpha$  и  $\beta$ , соответственно, и применял теорему парности в методах, использующих проектирование на определенные значения спина (см. гл. 6, разд. 4.3). В этих случаях не использовались дважды занятые орбитали, и нужно было обеспечить ту максимальную степень ортогональности между пространственными частями спин-орбиталей, занятых электронами с различными проекциями спина, которую только можно получить без изменения многоэлектронной детерминантной волновой функции. Другой областью применения этой теоремы является теория межмолекулярных взаимодействий. В ней использование теоремы парности позволило нам вывести и проанализировать явное выражение для энергии двух слабо взаимодействующих молекул в рамках следующей упрощенной модели: каждая молекула описывается наборами дважды занятых орбиталей и в первом приближении мы пренебрегаем их изменениями при образовании комплекса, однако явно учитываем более важные эффекты, вызванные межмолекулярным перекрытием. В таком случае двумя наборами орбиталей являются (дважды

занятые) молекулярные орбитали двух слабо взаимодействующих молекул, и использование теоремы парности обеспечивает — без изменения многоэлектронной волновой функции, — то, что только некоторые из межмолекулярных интегралов перекрывания будут отличны от нуля.

## 7.1. Теорема парности Лёвдина

Рассмотрим два набора орбиталей  $\{|a_i\rangle; i = 1, 2, \dots, n_a\}$  and  $\{|b_j\rangle; j = 1, 2, \dots, n_b\}$ , каждый из которых предполагается ортонормированным:

$$\langle a_i | a_j \rangle = \langle b_i | b_j \rangle = \delta_{ij} \quad . \quad (5.122)$$

(В этом разделе используется формализм «бра-кет».) Теорема парности утверждает, что всегда возможно найти такие два *унитарных* преобразования, смешивающих орбитали только *в пределах* одного набора ( $\{|a_i\rangle\}$  или  $\{|b_j\rangle\}$ ), с помощью которых можно добиться того, что наборы преобразованных орбиталей  $\{|a'_i\rangle\}$  и  $\{|b'_j\rangle\}$  будут взаимно ортогональны также и между собой, за исключением пар «соответственных орбиталей»

$$\langle a'_i | b'_j \rangle = \lambda_i \delta_{ij} \quad . \quad (5.123)$$

Заметим, что унитарные преобразования сохраняют также и свойство ортонормированности самих наборов (5.122).

### Доказательство<sup>1</sup>

Предположим, что  $n_a \geq n_b$  (иначе два набора должны быть переставлены) и образуем оператор  $\hat{P}^b$  проекции на подпространство, растянутое орбиталями  $\{|b_j\rangle\}$

$$\hat{P}^b = \sum_{l=1}^{n_b} |b_l\rangle \langle b_l| \quad . \quad (5.124)$$

Тогда можно построить матрицу  $\mathbf{Q}^b$  с элементами

$$Q_{ij}^b = \langle a_i | \hat{P}^b | a_j \rangle \quad . \quad (5.125)$$

Матрица  $\mathbf{Q}^b$  эрмитова, так что ее можно диагонализировать при помощи некоторой унитарной матрицы  $\mathbf{U}$ :

$$\mathbf{U}^\dagger \mathbf{Q}^b \mathbf{U} = \text{diag}\{\eta_i\} \quad . \quad (5.126)$$

<sup>1</sup> Приведенное доказательство воспроизводится по моей работе «Simple Proof of the Pairing Theorem,» I. Mayer, Int. J. Quantum Chem. **63**, 31–33 (1997); Copyright 1997 Wiley Periodicals.



Выберем преобразованный набор орбиталей  $\{|a'_j\rangle\}$  согласно

$$|a'_j\rangle = \sum_{l=1}^{n_a} U_{lj} |a_l\rangle . \quad (5.127)$$

Легко видеть, что в силу (5.126) они удовлетворяют условию

$$\langle a'_i | \hat{P}^b | a'_j \rangle = \eta_i \delta_{ij} . \quad (5.128)$$

Оператор  $\hat{P}^b$  идемпотентен и эрмитов, так что уравнение (5.128) можно тривиально преобразовать в

$$\langle a'_i | \hat{P}^b \hat{P}^b | a'_j \rangle = \langle \hat{P}^b a'_i | \hat{P}^b a'_j \rangle = \eta_i \delta_{ij} . \quad (5.129)$$

Из последнего равенства следует, что  $\eta_i \geq 0$ , так что можно выбрать их как квадраты



$$\eta_i = \lambda_i^2 , \quad (5.130)$$

где также  $\lambda_i \geq 0$ .

Функции  $|\hat{P}^b a'_i\rangle \equiv \hat{P}^b |a'_i\rangle$ , для которых  $\lambda_i \neq 0$ , полностью лежат в подпространстве, растянутом набором исходных орбиталей  $\{|b_j\rangle\}$ , тогда как функции с  $\lambda_i = 0$  исчезают в соответствии с условием (5.129). Функции с  $\lambda_i \neq 0$  также и ортогональны (но еще не нормированы), поэтому они линейно независимы. Так как невозможно образовать больше, чем  $n_b$ , линейно независимых орбиталей, лежащих исключительно в подпространстве, растянутом орбиталями  $\{|b_j\rangle\}$ , по крайней мере  $n_a - n_b$  собственных значений в выражении (5.126) должны быть равны нулю. (Если  $n_a = n_b$ , нулевые собственные значения обычно не появляются, за исключением некоторых весьма особых случаев систем с высокой симметрией.) Нормируя орбитали, о которых шла речь, получим преобразованный набор орбиталей  $\{|b'_i\rangle\}$ :

$$|b'_i\rangle = \frac{1}{\lambda_i} \hat{P}^b |a'_i\rangle . \quad (5.131)$$

Из условий (5.129) и (5.130) следует, что эти орбитали действительно ортонормированы:

$$\langle b'_i | b'_j \rangle = \frac{1}{\lambda_i \lambda_j} \langle \hat{P}^b a'_i | \hat{P}^b a'_j \rangle = \frac{1}{\lambda_i \lambda_j} \lambda_i^2 \delta_{ij} = \delta_{ij} . \quad (5.132)$$

Вследствие соотношений (5.128) и (5.130), два полученных набора орбиталей,  $\{|a'_i\rangle\}$  и  $\{|b'_j\rangle\}$ , удовлетворяют теореме парности (5.123):

$$\langle a'_i | b'_j \rangle = \langle a'_i | \frac{1}{\lambda_j} \hat{P}^b a'_j \rangle = \frac{1}{\lambda_j} \langle a'_i | \hat{P}^b | a'_j \rangle = \frac{1}{\lambda_j} \lambda_i^2 \delta_{ij} = \lambda_i \delta_{ij}, \quad (5.133)$$

что завершает доказательство в случае всех пар орбиталей, для которых значения  $\lambda_i$  отличны от нуля.

Орбитали  $\{|a'_i\rangle\}$ , полученные унитарным преобразованием, растягивают то же подпространство, что и исходные орбитали  $\{|a_i\rangle\}$ , и удовлетворяют условиям теоремы. Однако мы не можем использовать формулу (5.131) для получения тех преобразованных орбиталей  $|b'_i\rangle$ , которые соответствуют нулевым значениям  $\lambda_i$  (если таковые существуют), так как  $\lambda_i$  появляется в знаменателе. В таком случае мы должны дополнить набор  $\{|b'_j\rangle\}$  некоторым числом орбиталей, удовлетворяющих условиям теоремы. Легко видеть, однако, что можно просто выбрать произвольные ортонормированные векторы в ортогональном дополнении к орбиталям  $\{|b'_j\rangle\}$  (уже найденным с помощью формулы (5.131)), определенном по отношению к исходному подпространству  $\{|b_j\rangle\}$ . Их можно задать как собственные векторы (с собственными значениями, равными единице) оператора проекции на это ортогональное дополнение  $\hat{P}_{ort}^b$ ,

$$\hat{P}_{ort}^b = \hat{P}^b - \sum_{l=1}^{n'_b} |b'_l\rangle\langle b'_l| ; \quad (5.134)$$

здесь  $n'_b$  есть число ненулевых значений  $\lambda_i$ .

Чтобы завершить доказательство также и для этого случая, мы должны показать, что орбитали  $\{|b'_k\rangle; k = n'_b, n'_b + 1, \dots, n_b\}$ , найденные таким образом, ортогональны ко всем орбиталям  $\{|a'_i\rangle\}$ . Ввиду свойств проекционных операторов, оператор  $\hat{P}_{ort}^b$  можно выразить в виде спектрального разложения:

$$\hat{P}_{ort}^b = \sum_{l=n'_b+1}^{n_b} |b'_l\rangle\langle b'_l| . \quad (5.135)$$

Расписывая матричный элемент  $\langle a'_i | \hat{P}_{ort}^b | a'_i \rangle$ , получаем

$$\langle a'_i | \hat{P}_{ort}^b | a'_i \rangle = \sum_{l=n'_b+1}^{n_b} \langle a'_i | b'_l \rangle \langle b'_l | a'_i \rangle = \sum_{l=n'_b+1}^{n_b} |\langle a'_i | b'_l \rangle|^2 . \quad (5.136)$$

В то же время, согласно определению (5.134):

$$\langle a'_i | \hat{P}_{ort}^b | a'_i \rangle = \langle a'_i | \hat{P}^b | a'_i \rangle - \sum_{l=1}^{n'_b} \langle a'_i | b'_l \rangle \langle b'_l | a'_i \rangle = 0 , \quad (5.137)$$

где последнее равенство следует из соотношений (5.128), (5.130), и (5.133). Подставляя этот результат в формулу (5.136), получаем

$$\sum_{l=n'_b+1}^{n_b} |\langle a'_i | b'_l \rangle|^2 = 0 , \quad (5.138)$$

Сумма положительных чисел может равняться нулю, только если каждое из этих чисел равно нулю. Ч. т. д.

На практике используют орбитали, разлагаемые по конечному, обычно неортогональному, набору базисных орбиталей. Так как составление линейных комбинаций вида (5.127) не представляет проблемы, особое внимание необходимо уделять влиянию перекрывания на выражения матричных представлений проекционных операторов (ср. приложение П7).

Предполагая, что орбитальные коэффициенты образуют вектор-столбцы  $\mathbf{a}^i$ ,  $\mathbf{a}^{i'}$ ,  $\mathbf{b}^i$ , и  $\mathbf{b}^{i'}$ , набор  $\mathbf{b}^{i'}$ , вычисленный в соответствии с рецептом (5.131), будет иметь вид

$$\mathbf{b}^{i'} = \frac{1}{\lambda_i} \mathbf{P}^b \mathbf{S} \mathbf{a}^{i'} , \quad (5.139)$$

где  $\mathbf{S}$  — матрица интегралов перекрывания базисных орбиталей, и

$$\mathbf{P}^b = \sum_{l=1}^{n_b} \mathbf{b}_l \mathbf{b}_l^\dagger . \quad (5.140)$$

Если имеются орбитали, которые должны быть определены в ортогональном дополнении, как обсуждалось выше, то необходимо решить неэрмитову задачу на собственные значения:

$$(\mathbf{P}^b - \mathbf{P}^{b'}) \mathbf{S} \mathbf{b}'_k = \varepsilon_k \mathbf{b}'_k , \quad (5.141)$$

с матрицей  $\mathbf{P}^b$ , определенной выражением (5.140),

$$\mathbf{P}^{b'} = \sum_{l=1}^{n'_b} \mathbf{b}'_l \mathbf{b}'_l{}^\dagger , \quad (5.142)$$

и  $n'_b < k \leq n_b$ . Умножая уравнение (5.141) на матрицу  $\mathbf{S} = \mathbf{S}^\dagger$  слева, приходим к обобщенной эрмитовой задаче на собственные значения:

$$\mathbf{S}(\mathbf{P}^b - \mathbf{P}^{b'}) \mathbf{S} \mathbf{b}'_k = \varepsilon_k \mathbf{S} \mathbf{b}'_k . \quad (5.143)$$

Вычислительные трудности при ее решении могут возникать, если некоторые значения  $\lambda_i$  очень малы, хотя и отличны от нуля. В этом случае можно провести итерационное спаривание по различным парам орбиталей, отвечающим малым или нулевым значениям  $\lambda_i$ .

## 7.2. Расширенная теорема парности Карадакова

Обычно одноэлектронные (спин)орбитали, используемые для построения данной детерминантной волновой функции, называют «занятыми орбиталями» («орбиталями, занятыми в данном детерминанте»), тогда как функции в ортогональном дополнении к «занятому подпространству» (к подпространству, растянутому занятыми орбиталями) называются «виртуальными орбиталями», а их подпространство является «виртуальным подпространством». (Последнее является, в принципе, бесконечномерным, но часто мы рассматриваем только ту его часть, которую можно разложить по какому-то конечному базисному набору.)

Когда речь идет о двух наборах орбиталей, рассматриваемых в теореме парности Лёвдина, можно определить два виртуальных подпространства — одно, представляющее ортогональное дополнение к подпространству орбиталей  $\{|a_i\rangle\}$ , и другое, которое является ортогональным дополнением к орбиталям  $\{|b_i\rangle\}$ . В последующих уравнениях мы будем использовать верхние индексы « $o$ » (occupied) и « $v$ » (virtual), чтобы различать занятые и виртуальные наборы орбиталей. Очевидно, наборы орбиталей  $\{|a_i^o\rangle\}$  и  $\{|a_i^v\rangle\}$  совместно растягивают то же самое — конечномерное или бесконечномерное, — линейное пространство одноэлектронных функций, что и наборы  $\{|b_i^o\rangle\}$  и  $\{|b_i^v\rangle\}$  вместе взятые. Все четыре набора предполагаются ортонормированными по отдельности, и каждая виртуальная орбиталь, в согласии с представлением об ортогональном дополнении, ортогональна к орбиталям соответствующего занятого набора:

$$\begin{aligned} \langle a_i^o | a_j^o \rangle &= \langle a_i^v | a_j^v \rangle = \delta_{ij} \\ \langle b_i^o | b_j^o \rangle &= \langle b_i^v | b_j^v \rangle = \delta_{ij} \\ \langle a_i^o | a_j^v \rangle &= \langle b_i^o | b_j^v \rangle = 0, \end{aligned} \tag{5.144}$$

но, вообще говоря, не ортогональна к орбиталям другого занятого или виртуального набора.

Карадаков обобщил теорему парности Лёвдина на эти два виртуальных подпространства. Конечно, мы можем провести их спаривание, независимо от занятых орбиталей, с помощью метода, описанного в предыдущем разделе. Обобщенная теорема парности Карадакова, однако, утверждает большее: можно *одновременно* выбрать такие базисные наборы орбиталей в каждом виртуальном подпространстве (скажем,  $\{|a_i^v\rangle\}$ ), для которых имеется только по одной неортогональной орбитали как среди занятых, так и среди виртуальных орбиталей другого набора ( $\{|b_i^o\rangle\}$  и  $\{|b_i^v\rangle\}$  в данном случае).

Теорему можно легко доказать, исходя из предположения, что спаривание занятых подпространств уже было проведено, как описано в предыдущем разделе. Тогда мы можем построить удовлетворяющие условиям теоремы базисные наборы в двух виртуальных подпространствах.

Это делается так: занятые орбитали  $\{|a_i^o\rangle\}$  и  $\{|b_i^o\rangle\}$ , взятые вместе, растягивают подпространство размерности, не превышающей  $n_a + n_b$ . (Если все эти орбитали линейно независимы, то эта размерность равна  $n_a + n_b$ ; если существует  $k$  дважды занятых орбиталей, соответствующих значениям  $\lambda_i = 1$ , тогда размерность этого подпространства равна  $n_a + n_b - k$ .) Все орбитали, ортогональные к этому объединенному подпространству, образуют пересечение двух виртуальных подпространств. Выбирая любой набор ортонормированных векторов в этом подпространстве, получим орбитали, удовлетворяющие условиям теоремы: каждая из них имеет ненулевой интеграл перекрывания (с единичным значением) только с одной орбиталью другого виртуального подпространства и ортогональна к обоим занятым подпространствам. Нет необходимости рассматривать эти орбитали более детально.

Теперь спроектируем каждую занятую орбиталь  $|a_i^o\rangle$  и  $|b_j^o\rangle$  на виртуальное подпространство, соответствующее другому набору. Эти проекции представляют собой векторы в соответствующих виртуальных подпространствах. Виртуальные подпространства  $\{|a_i^v\rangle\}$  и  $\{|b_i^v\rangle\}$  являются ортогональными дополнениями к занятым подпространствам  $\{|a_i^o\rangle\}$  и  $\{|b_i^o\rangle\}$ ; поэтому соответствующие проекторы равны  $\hat{P}_a^{\text{virt}} = 1 - \sum_{l=1}^{n_a} |a_l^o\rangle\langle a_l^o|$

и  $\hat{P}_b^{\text{virt}} = 1 - \sum_{l=1}^{n_b} |b_l^o\rangle\langle b_l^o|$ . Так получают векторы

$$\begin{aligned}\hat{P}_b^{\text{virt}}|a_i^o\rangle &= |a_i^o\rangle - \sum_{l=1}^{n_b} |b_l^o\rangle\langle b_l^o|a_i^o\rangle \\ &= |a_i^o\rangle - \lambda_i^* |b_i^o\rangle ; \\ \hat{P}_a^{\text{virt}}|b_j^o\rangle &= |b_j^o\rangle - \sum_{l=1}^{n_a} |a_l^o\rangle\langle a_l^o|b_j^o\rangle \\ &= |b_j^o\rangle - \lambda_j |a_j^o\rangle ,\end{aligned}\tag{5.145}$$

где мы приняли во внимание тот факт, что занятые орбитали предполагаются уже спаренными:  $\langle b_l^o|a_i^o\rangle = \lambda_i^* \delta_{li}$ ;  $\langle a_l^o|b_j^o\rangle = \lambda_j \delta_{lj}$ . После норми-

ровки полученных орбиталей находим

$$\begin{aligned} |b_i^v\rangle &= \frac{1}{\sqrt{1-|\lambda_i|^2}}(|a_i^o\rangle - \lambda_i^*|b_i^o\rangle); \\ |a_j^v\rangle &= \frac{1}{\sqrt{1-|\lambda_j|^2}}(|b_j^o\rangle - \lambda_j|a_j^o\rangle). \end{aligned} \quad (5.146)$$

Легко видеть, что орбитали, определенные выражениями (5.146), удовлетворяют условиям теоремы. Так,  $|b_i^v\rangle$  и  $|a_j^v\rangle$  по построению являются ортогональными ко всем занятым орбиталям  $|b_i^o\rangle$  и  $|a_i^o\rangle$ , соответственно. Интеграл перекрывания  $\langle a_j^o | b_i^v \rangle$  отличен от нуля только для одной орбитали  $|a_i^o\rangle$ :

$$\begin{aligned} \langle a_j^o | b_i^v \rangle &= \frac{1}{\sqrt{1-|\lambda_i|^2}}(\langle a_j^o | a_i^o \rangle - \lambda_i^* \langle a_j^o | b_i^o \rangle) \\ &= \frac{1}{\sqrt{1-|\lambda_i|^2}}(\delta_{ji} - |\lambda_i|^2 \delta_{ij}) = \sqrt{1-|\lambda_i|^2} \delta_{ji}. \end{aligned} \quad (5.147)$$

Точно так же получаем

$$\langle b_j^o | a_i^v \rangle = \sqrt{1-|\lambda_i|^2} \delta_{ji}. \quad (5.148)$$

(Если  $n_a > n_b$ , то существует  $n_a - n_b$  орбиталей  $|b_i^v\rangle$ , в которых слагаемое  $\lambda_i^*|b_i^o\rangle$  отсутствует; эти занятые орбитали  $|a_i^o\rangle$  ( $i > n_b$ ) совпадают с первыми  $n_a - n_b$  виртуальными орбиталями  $|b_{i-n_b}^o\rangle$ .)

Чтобы закончить доказательство, рассмотрим перекрывания между виртуальными орбиталями, образованными по формулам (5.146):

$$\begin{aligned} \langle a_i^v | a_j^v \rangle &= \frac{1}{\sqrt{(1-|\lambda_i|^2)(1-|\lambda_j|^2)}} \langle b_i^o - \lambda_i a_i^o | b_j^o - \lambda_j a_j^o \rangle \\ &= \frac{1}{\sqrt{(1-|\lambda_i|^2)(1-|\lambda_j|^2)}} (\delta_{ij} - \lambda_i^* \lambda_i \delta_{ij} - \lambda_j \lambda_i^* \delta_{ij} + \lambda_i^* \lambda_j \delta_{ij}) = \delta_{ij}, \end{aligned} \quad (5.149)$$

где мы воспользовались тем фактом, что занятые орбитали ортонормированы и спарены. Точно так же получаем  $\langle b_i^v | b_j^v \rangle = \delta_{ij}$ ; в то же время

$$\begin{aligned} \langle a_i^v | b_j^v \rangle &= \frac{1}{\sqrt{(1-|\lambda_i|^2)(1-|\lambda_j|^2)}} \langle b_i^o - \lambda_i a_i^o | a_j^o - \lambda_j^* b_j^o \rangle \\ &= \frac{1}{\sqrt{(1-|\lambda_i|^2)(1-|\lambda_j|^2)}} (\lambda_i^* \delta_{ij} - \lambda_i^* \delta_{ij} - \lambda_j^* \delta_{ij} + \lambda_i^* \lambda_j^* \delta_{ij}) \\ &= -\lambda_i^* \delta_{ij}. \quad \text{Ч. т. д.} \end{aligned} \quad (5.150)$$

## 8. Теорема о существовании орбиталей специальной структуры

В этом разделе мы обсудим теорему, которая первоначально появилась в связи с локализованными орбиталями, но может пригодиться при изучении самых разнообразных задач, относящихся к однодетерминантным волновым функциям.

При качественном обсуждении молекулярной структуры часто используют понятия остовных орбиталей и неподеленных пар, локализованных на отдельных атомах, двухатомных связывающих и разрыхляющих орбиталей и т.п. Локализованные орбитали, которые можно получить путем *a posteriori* анализа уже найденных однодетерминантных волновых функций, обычно в основном подтверждают эти представления, однако значительная часть локализованных орбиталей имеет также (малые) вклады, делокализованные по другим частям молекулы. Обычно локализованные орбитали получают унитарным преобразованием канонических орбиталей, требуя выполнения какого-либо условия локализации. Эти локализованные орбитали ортонормированы, так же, как и канонические орбитали.

Мы видели в разд. 5, что каждая детерминантная волновая функция может быть выражена (с точностью до физически несущественного постоянного множителя) через ортонормированные орбитали, но это, в свою очередь, не обязательно: можно использовать любой набор линейно независимых одноэлектронных орбиталей, которые растягивают «занятое подпространство». Было бы естественно выяснить, можно ли, отбросив требование ортогональности, описать молекулу с помощью еще более локализованных орбиталей, не меняя при этом многоэлектронную детерминантную волновую функцию. Положительный ответ на этот вопрос был дан теоремой, рассматриваемой ниже.

Рассмотрим однодетерминантную волновую функцию системы, описываемой конечным базисным набором, и предположим, что имеется некий набор приближений нулевого порядка как к занятым, так и виртуальным орбиталям. (Происхождение этого нулевого приближения несуществен для дальнейшего.) Ортонормированность орбиталей нулевого порядка не нужна, а достаточно потребовать их линейной независимости.

Обозначим занятые орбитали нулевого порядка как  $\varphi_i^0$ ,  $i = 1, 2, \dots, n$ , а виртуальные орбитали нулевого порядка как  $\psi_j^0$ ,  $j = 1, 2, \dots, m - n$ , где  $m$  и  $n$  — размерность полного базиса и число занятых орбиталей, соответственно. Так как орбитали нулевого порядка предполагаются линейно независимыми, их можно использовать как базис, по

которому можно разложить «истинные» занятые орбитали, составляющие детерминантную волновую функцию.

Назовем приближения нулевого порядка  $\{\varphi_i^0\}$ ,  $\{\psi_j^0\}$  «сингулярно плохими», если в истинном занятом подпространстве существует хотя бы одна такая орбиталь, которая может быть *точно* представлена в виде линейной комбинации приближенных *виртуальных* орбиталей. Для наборов, не являющихся сингулярно плохими, ни одной такой орбитали, очевидно, не существует. Имеет место

### 8.1. Теорема существования

Если приближение нулевого порядка  $\{\varphi_i^0\}$ ,  $\{\psi_j^0\}$  не является сингулярно плохим, то истинную детерминантную волновую функцию всегда можно построить (с точностью до несущественного постоянного множителя) с использованием  $n$  (в общем случае неортогональных) занятых орбиталей  $\varphi_i$ , имеющих специальную структуру:

$$\varphi_i = \varphi_i^0 + \sum_{j=1}^{m-n} \kappa_{ij} \psi_j^0, \quad (5.151)$$

где  $\kappa_{ij}$  — некоторые коэффициенты, и каждая точная занятая орбиталь может быть выражена через *только одну* из приближенных занятых орбиталей нулевого порядка. (Поэтому любое смешивание между занятыми орбиталями нулевого порядка можно рассматривать, как всего лишь эффект ортогонализации.)

#### Доказательство

Рассмотрим для данного  $i$  ( $i = 1, 2, \dots, n$ ) функции следующего вида:

$$\varphi_i' = \alpha_i \varphi_i^0 + \sum_{j=1}^{m-n} \lambda_{ij} \psi_j^0. \quad (5.152)$$

Орбитали вида (5.152) с произвольными коэффициентами  $\alpha_i$  и  $\lambda_{ij}$  образуют подпространство размерности  $m-n+1$ , и это значит, что мы можем выбрать  $m-n+1$  линейно независимых векторов в этом подпространстве. Истинное занятое подпространство имеет размерность  $n$ , т. е. мы можем выбрать в нем  $n$  линейно независимых векторов. Если объединить эти два подпространства, то мы получим  $m-n+1+n = m+1$  векторов. Однако они не могут быть все линейно независимы, так как полное число базисных векторов (размерность) равно всего лишь  $m$ . Следовательно, эти подпространства должны иметь общий вектор: он имеет вид (5.152) и полностью лежит в истинном занятом подпространстве. Для



этого вектора коэффициент  $\alpha_i$  не может быть 0, так как тогда приближение нулевого порядка было бы сингулярно плохим, что противоречит предположению теоремы. Разделив выражение (5.152) на  $\alpha_i \neq 0$ , и обозначив  $\varphi_i = \frac{\varphi'_i}{\alpha_i}$  и  $\kappa_{ij} = \frac{\lambda_{ij}}{\alpha_i}$ , мы приходим к выражению (5.151). Такие орбитали можно построить для каждого значения  $i = 1, 2, \dots, n$ ; любая из них лежит в истинном занятом подпространстве и они линейно независимы, благодаря линейной независимости орбиталей  $\varphi_i^0$ ; поэтому из них можно построить требуемую детерминантную волновую функцию. Ч. т. д.

## 8.2. A posteriori определение

Покажем теперь, что орбитали  $\varphi_i$  вида (5.151) существуют не только в абстрактно-математическом смысле, но также можно предложить алгоритм для их практического вычисления.

Введем общее обозначение  $\vartheta_i$  для базисных функций  $\varphi_i^0$  и  $\psi_j^0$

$$\vartheta_i = \begin{cases} \varphi_i^0 & (1 \leq i \leq n) \\ \psi_{i-n}^0 & (n < i \leq m) \end{cases} . \quad (5.153)$$

Функции  $\vartheta_i$  предполагаются линейно независимыми, поэтому матрица их интегралов перекрывания  $\langle \vartheta_i | \vartheta_j \rangle$  не сингулярна. Мы воспользуемся обозначением **Q** для матрицы, обратной к матрице интегралов перекрывания.

Рассмотрим какой-либо (например, канонический) набор истинных занятых орбиталей, обозначим их  $\varphi_i^{SCF}$ , и введем прямоугольную матрицу **D** размера  $m \times n$ , содержащую интегралы перекрывания:

$$D_{kj} = \langle \vartheta_k | \varphi_j^{SCF} \rangle . \quad (5.154)$$

Разложим  $\varphi_j^{SCF}$  по базису  $\{\vartheta_i\}$ :

$$\varphi_j^{SCF} = \sum_{l=1}^m f_l^j \vartheta_l . \quad (5.155)$$

Домножив разложение (5.155) на  $\vartheta_k^*$  и проинтегрировав, получим

$$D_{kj} = \sum_{l=1}^m \langle \vartheta_k | \vartheta_l \rangle f_l^j , \quad (5.156)$$

откуда, после умножения на элемент  $Q_{qk}$  обратной матрицы перекрывания **Q** и суммируя по  $k$ , получаем

$$f_q^j = \sum_{k=1}^m Q_{qk} D_{kj} = (\mathbf{QD})_{qj} , \quad (5.157)$$



т. е.

$$\varphi_j^{SCF} = \sum_{l=1}^m (\mathbf{QD})_{lj} \vartheta_l . \quad (5.158)$$

С другой стороны, искомая функция  $\varphi_i$  находится в точном занятом подпространстве, поэтому ее можно представить, как линейную комбинацию функций  $\varphi_j^{SCF}$ , растягивающих это подпространство:

$$\varphi_i = \sum_{j=1}^m A_{ji} \varphi_j^{SCF} . \quad (5.159)$$

Подставляя сюда разложение (5.158), можем написать

$$\varphi_i = \sum_{l=1}^m \sum_{j=1}^n (\mathbf{QD})_{lj} A_{ji} \vartheta_l . \quad (5.160)$$

Теперь вернемся к исходным обозначениям  $\varphi_i^0$  и  $\psi_j^0$ ; разложение (5.159) тогда принимает вид

$$\varphi_i = \sum_{j,l=1}^n (\mathbf{QD})_{lj} A_{ji} \varphi_l^0 + \sum_{l=1}^{m-n} \sum_{j=1}^n (\mathbf{QD})_{l+n,j} A_{ji} \psi_l^0 . \quad (5.161)$$

Потребуем, чтобы  $\varphi_i$  имела вид (5.151), а это означает, что в первой сумме только одна  $\varphi_l^0$  должна иметь ненулевой коэффициент:

$$\sum_{j=1}^n (\mathbf{QD})_{lj} A_{ji} = \delta_{li} . \quad (5.162)$$

Смысл этого результата в том, что если матрицу  $\mathbf{A}$  определить, как обратную к верхнему левому блоку размера  $n \times n$  матрицы  $\mathbf{QD}$ , то мы получим орбитали, удовлетворяющие теореме.

После того, как матрица  $\mathbf{A}$  определена, требуемые орбитали, удовлетворяющие условию (5.151), можно найти из представления (5.159). Или же, можно определить явные значения коэффициентов из соотношения (5.161) как

$$\kappa_{il} = \sum_{j=1}^n (\mathbf{QD})_{l+n,j} A_{ji} . \quad (5.163)$$

## Библиографические заметки

### Разделы 1–5.

Стандартный материал (например [1, 2]); изложение в основном следует [3]. (Детерминант Слэтера был введен в [4], где уже и было установлено большинство его важных свойств.)

### Раздел 6.

Правила Слэтера были введены в [4, 5]. (См. также [6] — иногда также используется название «правила Слэтера—Кондона».) Матричные элементы для неортогональных орбиталей впервые были выведены Лёвдиным [7] с использованием формализма матрицы плотности. Также смотри, например, [2]. Возможно, что представленная техника «факторизации» оригинальна (но не претендует на это).

### Раздел 7.1

Теорема парности была впервые сформулирована Лёвдиным как предположение [8] и была доказана Амосом и Холлом [9]. Также см. [10 – 12]. Представленное доказательство было опубликовано в [13]. Теорема парности была больше всего использована в спин-спроектированной «расширенной» теории Хартри-Фока (см. обзор [12]); о недавнем применении теоремы в теории межмолекулярных взаимодействий см. [14].

### Раздел 7.2

Расширенная теорема парности предложена и доказана Карадаковым [15], использовавшим так называемый метод «сингулярного разложения» (обобщение диагонализации на случай прямоугольных матриц). Представленное доказательство является упрощенной версией доказательства в [16].

### Раздел 8

Авторские результаты [17]. Об их использовании в теории локализованных орбиталей и делокализованных взаимодействий см. [18, 19], а в теории «суперпозиционной ошибки базиса» при вычислении межмолекулярных взаимодействий см. [20].

## Для дальнейшего изучения

О построении спиновых собственных функций упомянем, например, [2, 10, 21]. Уникальным источником сведений о разных свойствах детерминантных волновых функций и матрицах плотности, не рассматриваемых в данной книге, является серия статей Лёвдина [7, 22].

Здесь уместно отметить, что многие результаты этой главы можно гораздо легче получить с помощью формализма *вторичного квантования*, который позволяет «встроить» требование антисимметричности в любую волновую функцию. Книжки [23], и особенно [24], дают хорошее введение в предмет (также см. [2]).



## Литература

1. Zülicke L. *Quantenchemie*, Deutscher Verlag der Wissenschaften, Berlin 1973. (Имеется русский перевод: Цюлике Л. *Квантовая химия*. – М.: Мир, 1976).
2. McWeeny R. *Methods of Molecular Quantum Mechanics*, Academic Press, London 1992. (Имеется русский перевод первого издания: МакВини Р., Сатклиф С. *Квантовая механика молекул*. – М.: Мир, 1973).
3. Mayer I. *Fejezetek a Kvantumkémiából*, BME Mérnök-távoképző Int., Budapest 1987.
4. Slater J. C. Phys. Rev. **34**, 1293 (1929).
5. Slater J. C. Phys. Rev. **38**, 1109 (1931).
6. Condon E. U. Phys. Rev. **36**, 1121 (1930).
7. Löwdin P.-O. Phys. Rev. **97**, 1470 (1955).
8. Löwdin P.-O. J. Appl. Phys. Suppl. **33**, 251 (1962).
9. Amos A. T., Hall G. G. Proc. Roy. Soc. (London) **A263**, 483 (1961).
10. Pauncz R. *Alternant Molecular Orbital Method* Saunders, Philadelphia 1967.
11. Горлов Ю. И., Украинский И. И. Институт теоретической физики УССР, Препринты ИТФ-73-138 (1973) и ИТФ-74-69Е (1974).
12. Mayer I. Adv. Quantum Chem. **12**, 189 (1980).
13. Mayer I. Int. J. Quantum Chem. **63**, 31 (1997).
14. Hamza A., Mayer I. Int. J. Quantum Chem. **82**, 53, 105 (2001).
15. Karadakov P. Int. J. Quantum Chem. **27**, 699 (1985).
16. Mayer I. Int. J. Quantum Chem. **29**, 31 (1986).
17. Mayer I. Chem. Phys. Letters, **89**, 390 (1982).
18. Surján P. R., Mayer I. and Kertész M. J. Chem. Phys. **77**, 2454 (1982).
19. Mayer I., Surján P. R. J. Chem. Phys. **80**, 5649 (1984).
20. Mayer I., Turi L. J. Molec. Struct. (Theochem), **227**, 43 (1991).
21. Pauncz R. *Spin Eigenfunctions*, Plenum Press, New York 1979.
22. Löwdin P.-O. Phys. Rev. **97**, 1490, 1510 (1955).
23. Jørgensen P., Simons J. *Second Quantization-Based Methods in Quantum Chemistry*, Academic Press, New York 1981.
24. Surján P. R. *Second Quantized Approach to Quantum Chemistry*, Springer, Berlin 1989.



# Метод Хартри—Фока



Метод Хартри—Фока (ХФ) имеет центральное значение для квантовой химии — как в концептуальном, так и в вычислительном смысле. В методе ХФ волновая функция приближенно описывается одним детерминантом Слэтера, параметры которого оптимизируют согласно вариационному принципу. Помимо того, что метод ХФ важен сам по себе, он также используется в большинстве более точных методов, учитывающих «электронную корреляцию», как нулевое приближение (гл. 8).

Метод ХФ часто называют приближением (моделью) «независимых частиц» — или «независимых электронов». Это не значит, что межэлектронным отталкиванием пренебрегают, а только то, что его истинное значение заменяется некоторым «усредненным». Каждому электрону при этом сопоставляется одна спин-орбиталь (т. е. применяют столько же спин-орбиталей, сколько в системе имеется электронов) и оптимизируют параметры этих спин-орбиталей. Такая вариационная процедура определяет эффективное среднее поле, в котором движутся электроны. Наглядно такая ситуация может быть описана следующим образом. Каждый электрон движется в эффективном поле, являющемся суммой полей, наведенных ядрами, и среднего поля, наведенного другими<sup>1</sup> электронами. Тогда поле, определяющее орбитали, само зависит от этих орбиталей: необходимо найти орбитали, порождающие такое поле, в котором решениями уравнений будут именно те орбитали, которые и порождают это поле. В этом случае говорят о «самосогласованном поле» (ССП), и процедура нахождения таких орбиталей называется процедурой СПП.

Чаще всего, при описании молекул с замкнутой оболочкой используют дважды занятые *пространственные* орбитали, т. е. спин-орбитали, получающиеся парами из каждой пространственной орбитали — одна со спином  $\alpha$ , а другая со спином  $\beta$ . Это — «ограниченный по спину метод Хартри—Фока» (ОХФ, RHF — restricted HF). Если требование двукратной занятости снимается, то получается более общая схема «неограни-

---

<sup>1</sup> Здесь, конечно, под «данном электроном» понимается «электрон, занимающий данную орбиталь», поскольку сами электроны неразличимы.

ченного по спине Хартри—Фока» (НХФ, UNF — unrestricted HF); ее решения часто совпадают с решениями ОХФ, если рассматриваются системы с замкнутой оболочкой (см. обсуждение в разд. 4.3).

Так как общая форма волновой функции Хартри—Фока уже задана (один детерминант либо с ограничением на двукратное заполнение, либо без него), то использование вариационного принципа означает, что мы должны оптимизировать включаемые в этот детерминант орбитали. Это эквивалентно решению уравнений Хартри—Фока (которые получаются из условия стационарности энергии) методом последовательных приближений (итераций). Обычно процедуру начинают с некоторой приближенной «оценки» орбиталей, определяющей эффективное поле, порожаемое этими орбиталями, и затем находят новые орбитали, соответствующие стационарным одноэлектронным состояниям в этом поле. Затем с новыми орбиталями пересчитывают поле, из него — снова орбитали, и повторяют всю эту процедуру до выполнения некоторого выбранного критерия сходимости, указывающего на то, что самосогласование достигнуто с требуемой точностью.

## 1. Вариационный принцип для одnodетерминантных волновых функций: теорема Бриллюэна

Теорема Бриллюэна обычно рассматривается как следствие уравнений Хартри—Фока. Однако мы предпочтем двигаться в противоположном направлении: сначала покажем, что теорема Бриллюэна эквивалентна вариационному принципу для одnodетерминантных волновых функций, и затем выведем из нее уравнения Хартри—Фока, расписав теорему Бриллюэна через одноэлектронные функции.

Рассмотрим волновую функцию, имеющую вид слэтеровского детерминанта:

$$\Psi_0 = \hat{A}[\psi_1(1)\psi_2(2) \dots \psi_N(N)] \quad (6.1)$$

которая по предположению удовлетворяет вариационному принципу, и какой-либо «однократно возбужденный» детерминант  $\Psi_1 = \Psi_1(\psi_i \rightarrow \psi'_i)$ , который можно получить из  $\Psi_0$  замещением спин-орбитали  $\psi_i$  на произвольную спин-орбиталь  $\psi'_i$ , ортогональную ко всем занятым спин-орбиталям  $\psi_j$ ,  $j = 1, 2, \dots, N$ . Теорема Бриллюэна утверждает, что полученные таким образом однократно возбужденные детерминанты  $\Psi_1$  не взаимодействуют с вариационной (или «хартри-фоковской») детерминантной волновой функцией  $\Psi_0$ , т. е.:

$$H_{01} = \langle \Psi_0 | \hat{H} | \Psi_1 \rangle = 0, \quad (6.2)$$

где  $\hat{H}$  есть гамильтониан Борна—Оппенгеймера данной системы.

Сначала мы докажем эту теорему для детерминанта НХФ, имеющего *абсолютный минимум* энергии, а затем для детерминанта ОХФ или НХФ со *стационарной энергией*, т.е. удовлетворяющего условию  $\delta E = 0$ .

### 1.1. Теорема Бриллюэна для детерминанта, дающего абсолютный минимум энергии

*Доказательство от противного*

Предположим, что детерминантная волновая функция  $\Psi_0$  доставляет энергии абсолютный минимум. Мы докажем, что теорема Бриллюэна (6.2) должна выполняться для нее, иначе можно было бы построить другую однодетерминантную волновую функцию с еще меньшей энергией, что противоречило бы предположениям теоремы.

Образует линейную комбинацию  $\Psi = c_0 \Psi_0 + c_1 \Psi_1$  и решим линейную вариационную задачу, чтобы определить коэффициенты  $c_0$  и  $c_1$ , дающие минимум энергии для функции  $\Psi$ . Детерминанты Слэтера  $\Psi_0$  и  $\Psi_1$  отличаются только одним столбцом. Поэтому, согласно свойствам определителей, их линейную комбинацию  $\Psi$  также можно представить в виде одного детерминанта Слэтера. Из результатов гл. 5, разд. 6.1 следует, что детерминантные волновые функции  $\Psi_0$  и  $\Psi_1$  ортогональны. Без потери общности мы можем предположить, что они также нормированы. Поэтому для определения коэффициентов  $c_1$  и  $c_2$  и соответствующей им энергии  $E$  можно использовать вековое уравнение (3.15). Вековой определитель равен

$$\begin{vmatrix} H_{00} - E & H_{01} \\ H_{10} & H_{11} - E \end{vmatrix} = 0. \quad (6.3)$$

Тривиальными алгебраическими выкладками можно привести меньший из двух корней уравнения (6.3) к виду

$$E = H_{00} + \frac{1}{2}(H_{00} - H_{11}) \left[ \sqrt{1 + \frac{4|H_{01}|^2}{(H_{00} - H_{11})^2}} - 1 \right], \quad (6.4)$$

где  $H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle$ . Разность  $H_{00} - H_{11} < 0$ , так как  $\Psi_0$  доставляет абсолютный минимум энергии. Тогда, если мы предположим, что  $H_{01} \neq 0$ , то получим из (6.4)  $E < H_{00}$  для энергии детерминантной волновой функции  $\Psi$ . Однако  $H_{00}$  соответствует энергии состояния  $\Psi_0$  и предполагалось, что она является *абсолютным минимумом*. Мы приходим к противоречию, поэтому  $H_{01}$  должен быть равен нулю и тогда  $E = H_{00}$ .  
Ч. т. д.

## 1.2. Теорема Бриллюэна для детерминанта, имеющего стационарную энергию

Для того, чтобы проварьировать детерминантную волновую функцию  $\Psi = \hat{\mathcal{A}}[\psi_1(1)\psi_2(2)\dots\psi_N(N)]$ , необходимо проварьировать одноэлектронные спин-орбитали  $\psi_i$ , заменив каждую  $\psi_i$  на  $\psi_i + \delta\psi_i = \psi_i + \eta\psi'_i$ . Здесь  $\psi'_i$  является произвольной спин-орбиталью, а  $\eta$  есть произвольный стремящийся к нулю комплексный вариационный параметр. Тогда многоэлектронная волновая функция, полученная после вариации, примет вид:

$$\Psi + \delta\Psi = \hat{\mathcal{A}}\{[\psi_1(1) + \eta\psi'_1(1)][\psi_2(2) + \eta\psi'_2(2)]\dots[\psi_N(N) + \eta\psi'_N(N)]\} . \quad (6.5)$$

В этом детерминанте каждый столбец является суммой двух слагаемых  $\psi_i$  и  $\eta\psi'_i$ . Согласно известным свойствам определителей, его можно записать как сумму  $2^N$  детерминантов. Большая часть этих слагаемых содержит квадраты или еще более высокие степени  $\eta$  и ими можно пренебречь по сравнению с вкладами первого порядка по  $\eta$ . Таким образом, можно записать

$$\Psi + \delta\Psi = \Psi + \eta \sum_{i=1}^N \Psi_1(\psi_i \rightarrow \psi'_i) . \quad (6.6)$$

Здесь  $\Psi_1(\psi_i \rightarrow \psi'_i)$  отличается от  $\Psi$  заменой спин-орбитали  $\psi_i$  на произвольную спин-орбиталь  $\psi'_i$ . (Мы пока не налагаем никаких ограничений на спин-орбитали  $\psi'_i$ .) Тогда получаем наиболее общий вид вариации первого порядка детерминантной волновой функции  $\Psi$ :

$$\delta\Psi = \eta \sum_{i=1}^N \Psi_1(\psi_i \rightarrow \psi'_i) . \quad (6.7)$$

Вариацию (6.7) детерминантной волновой функции нужно подставить в выражение (2.13), что даст условие стационарности энергии ( $\delta E = 0$ ). Получаем

$$\eta^* \sum_{i=1}^N \langle \Psi_1(\psi_i \rightarrow \psi'_i) | \hat{H} - E | \Psi \rangle = 0 . \quad (6.8)$$

Вспомним, что  $\eta \rightarrow 0$ , но  $\eta \neq 0$ , так что можно разделить уравнение (6.8) на  $\eta^*$ . Так как вариации  $\delta\psi_i = \eta\psi'_i$  отдельных орбиталей независимы, мы можем выбрать  $\psi'_i \equiv 0$  для всех орбиталей, кроме одной. Поэтому равенство

$$\langle \Psi_1(\psi_i \rightarrow \psi'_i) | \hat{H} - E | \Psi \rangle = 0 \quad (6.9)$$

должно выполняться отдельно для каждого  $i$ . Вариационный принцип, записанный в виде (6.9), обычно называется «обобщенной теоремой Бриллюэна».



Как мы видели, каждый однократно возбужденный детерминант  $\Psi_1(\psi_i \rightarrow \psi'_i)$  связан с вариацией  $\delta\psi_i = \eta\psi'_i$  спин-орбитали  $\psi_i$ . Эту вариацию всегда можно записать как

$$\delta\psi_i = \delta\psi_{i1} + \delta\psi_{i2} = \eta(\psi'_{i1} + \psi'_{i2}), \quad (6.10)$$

где  $\psi'_{i1}$  ортогональна ко всем спин-орбиталям  $\psi_j$ , занятым в  $\Psi$ , тогда как  $\psi'_{i2}$  можно записать в виде линейной комбинации занятых спин-орбиталей. (Если  $\hat{P}$  — оператор проекции на подпространство занятых орбиталей, то  $\psi'_{i1} = (1 - \hat{P})\psi'_i$  и  $\psi'_{i2} = \hat{P}\psi'_i$ .) Из свойств определителей следует, что  $\Psi_1(\psi_i \rightarrow \psi'_i) = \Psi_1(\psi_i \rightarrow \psi'_{i1}) + \Psi_1(\psi_i \rightarrow \psi'_{i2})$ , и поэтому необходимо потребовать выполнения условия (6.9) как для  $\Psi_1(\psi_i \rightarrow \psi'_{i1})$ , так и для  $\Psi_1(\psi_i \rightarrow \psi'_{i2})$ . Однако вариация  $\delta\psi_{i2} = \eta\psi'_{i2}$  оставляет инвариантным подпространство занятых орбиталей, а следовательно не изменяет детерминантную волновую функцию или меняет только ее нормировку; поэтому она не приводит ни к какому изменению энергии. (Другими словами,  $\Psi_1(\psi_i \rightarrow \psi'_{i2})$  либо равно нулю, либо пропорционально  $\Psi$ ; в обоих случаях условие (6.9) выполняется автоматически.)

Из приведенного рассуждения следует, что для получения необходимых и достаточных условий стационарности энергии, вычисленной с детерминантной волновой функцией  $\Psi$ , достаточно рассмотреть случай, в котором все  $\delta\psi_{i2} = \eta\psi'_{i2} = 0$ , т. е. ограничиться вариациями  $\delta\psi_i = \eta\psi'_i = \eta\psi'_{i1}$ , которые ортогональны всем занятым спин-орбиталям  $\psi_j$ . В таком случае  $\Psi_1(\psi_i \rightarrow \psi'_i)$  также ортогональна  $\Psi$ , т. е.  $\langle \Psi_1(\psi_i \rightarrow \psi'_i) | \Psi \rangle = 0$ , и обобщенная теорема Бриллюэна сводится к

$$\langle \Psi_1(\psi_i \rightarrow \psi'_i) | \hat{H} | \Psi \rangle = 0. \quad (6.11)$$

Эта теорема Бриллюэна должна выполняться для всех занятых спин-орбиталей  $\psi_i$  и произвольных спин-орбиталей  $\psi'_i$ , ортогональных занятым, и в этом случае она эквивалентна вариационному принципу  $\delta E = 0$  для однодетерминантных волновых функций.

Равенство (6.11) выполняется тривиально, если  $\psi_i$  и  $\psi'_i$  имеют разные проекции спина, поэтому необходимо потребовать выполнения теоремы Бриллюэна для случая замены спин-орбитали  $\psi_i(\vec{r}, \sigma) = \varphi_i(\vec{r})\gamma_i(\sigma)$  на произвольную спин-орбиталь  $\psi'_i = \varphi'_i(\vec{r})\gamma_i(\sigma)$  с той же самой проекцией спина  $\gamma_i$  ( $\gamma_i = \alpha$  или  $\beta$ ), пространственная часть  $\varphi'_i$  которой ортогональна ко всем пространственным орбиталям, занятым в детерминанте  $\Psi$  электронами с той же проекцией спина. Тогда теорему Бриллюэна можно также записать в таком виде:

$$\langle \Psi_1(\varphi_i\gamma_i \rightarrow \varphi'_i\gamma_i) | \hat{H} | \Psi \rangle = 0 \quad (6.12)$$

для всех  $\varphi'_i$ , таких, что  $\langle \varphi'_i | \varphi_j \rangle \delta_{\gamma_i\gamma_j} = 0$  ( $i, j = 1, 2, \dots, N$ ).

### 1.3. Алгоритм для решения задачи Хартри—Фока, основанный на теореме Бриллюэна

В принципе, теорема Бриллюэна не только дает условия стационарности энергии, но ее также можно использовать для построения процедуры поиска волновых функций, удовлетворяющих этим условиям. Мы обсудим эту возможность очень кратко; она основана на рассуждениях, сходных с теми, которые применялись при доказательстве теоремы Бриллюэна в разд. 1.1.

Предположим, что имеется ортонормированный набор приближенных занятых и виртуальных спин-орбиталей  $\{\psi_i^o\}$  и  $\{\psi_j^v\}$  и хотим улучшить это приближение. Мы можем вычислить заново спин-орбитали циклически одну за другой, пересчитывая все орбитали в каждом цикле. Когда данная орбиталь  $\psi_i^o$  улучшается, все остальные занятые орбитали остаются фиксированными. Так же, как и в случае, рассмотренном в разд. 1.1, сформулируем линейную вариационную задачу, потребовав, чтобы энергия линейной комбинации детерминантных функций

$$\Psi = c_0 \Psi_0 + \sum_{j=1}^{m_v} c_j \Psi_1(\psi_j^o \rightarrow \psi_j^v) \quad (6.13)$$

была минимальна. Для этого надо решить уравнение на собственные значения вида (3.15), имеющее размерность  $m_v + 1$ , где  $m_v$  — число виртуальных орбиталей. (Формально наши рассуждения верны также в пределе  $m_v \rightarrow \infty$ .) Теперь воспользуемся тем фактом, что как функция  $\Psi_0$ , так и различные функции  $\Psi_1$  все являются определителями, отличающимися друг от друга одним и тем же столбцом. Поэтому волновую функцию  $\Psi$ , определенную выражением (6.13), тоже можно записать, как один детерминант, в котором занятая спин-орбиталь  $\psi_i^o$  заменена улучшенной орбиталью:

$$\psi_i^{o'} = c_0 \psi_i^o + \sum_{j=1}^{m_v} c_j \psi_j^v. \quad (6.14)$$

Аналогично этому, другие собственные векторы векового уравнения можно использовать для получения новых наборов виртуальных орбиталей. Очевидно, если используются дважды занятые орбитали, то две спин-орбитали, имеющие одинаковые пространственные части, можно скорректировать одновременно. (Обратим внимание, что последующие рассуждения, касающиеся сходимости, не обязательно верны в этом случае.)

На каждом шаге этой процедуры новая орбиталь получается решением линейной вариационной задачи и взятием наименьшего корня. Повторяя эту процедуру циклически для всех орбиталей, получаем убывающую (или по крайней мере невозрастающую) последовательность значений энергии. Эта последовательность, в принципе, бесконечна и ограничена снизу (что следует из вариационного принципа). Из математики известно, что такая последовательность сходится к конечному пределу (минимальному значению энергии). При достижении сходимости теорема Бриллюэна будет выполнена (иначе мы могли бы получить дальнейшее уменьшение энергии при решении линейной вариационной задачи (6.13) — ср. разд. 1.1) — поэтому такая схема является безусловно сходящимся алгоритмом решения задачи ССП, по крайней мере в случае НХФ. Однако эта схема требует больших вычислительных затрат. Предполагая, что мы находимся вблизи от искомого решения, можно попробовать уменьшить объем вычислений, используя оценки теории возмущений вместо точного решения задачи на собственные значения. (В последнем случае особое внимание следует уделить обеспечению ортогональности новых виртуальных орбиталей — ср. приложение П11.) Также можно попробовать определить поправки для всех орбиталей одновременно, что может уменьшить вычислительные затраты. Конечно, в этом случае сходимость уже не гарантирована.

## 2. Уравнения Хартри—Фока

Теорема Бриллюэна дает необходимые и достаточные условия стационарности энергии, вычисленной с детерминантной волновой функцией  $\Psi_0$ , выраженные через многоэлектронные волновые функции — саму  $\Psi_0$  и все возможные однократно возбужденные детерминанты  $\Psi_1(\psi_i \rightarrow \psi'_i)$ , в которых спин-орбиталь  $\psi_i$  заменена произвольной спин-орбиталью  $\psi'_i$ , ортогональной ко всем занятым орбиталям. Выражая утверждение теоремы Бриллюэна с помощью одноэлектронных орбиталей, можно вывести уравнения, которым они должны удовлетворять.

### 2.1. Неограниченные по спину уравнения Хартри—Фока

Сначала мы рассмотрим случай НХФ, т. е. мы пока не будем делать предположений о двукратной занятости пространственных орбиталей. Без потери общности мы можем считать, что детерминантная волновая функция  $\Psi$ , соответствующая стационарной энергии, построена из ортонормированных спин-орбиталей. (Как обсуждалось в гл. 5, разд. 5, ортонормировка орбиталей оставляет неизменной детерминантную вол-

новую функцию с точностью до физически несущественного постоянного множителя.)

Распишем выражение для теоремы Бриллюэна (6.12) с использованием результатов гл. 5, разд. 6.2 и 6.3 для матричных элементов одно- и двухэлектронной частей гамильтониана, вычисленных между определителями  $\Psi_0$  и  $\Psi_1$ , которые построены из ортонормированных спин-орбиталей  $\psi_i(\vec{r}, \sigma) = \varphi_i(\vec{r})\gamma_i(\sigma)$  и отличаются одной орбиталью — в  $\Psi_1$   $\varphi_i$  заменена на  $\varphi'_i$ . В этом случае имеем

$$\langle \Psi_1(\varphi_i\gamma_i \rightarrow \varphi'_i\gamma_i) | \hat{H} | \Psi \rangle \quad (6.15)$$

$$= \langle \varphi'_i | \hat{h} | \varphi_i \rangle + \sum_{\substack{j=1 \\ (j \neq i)}}^N ([\varphi'_i\varphi_j | \varphi_i\varphi_j] - [\varphi'_i\varphi_j | \varphi_j\varphi_i] \cdot \delta_{\gamma_i\gamma_j}) = 0 ,$$

где мы использовали обозначение

$$[\varphi_i\varphi_j | \varphi_k\varphi_l] = \iint \frac{\varphi_i^*(\vec{r}_1)\varphi_j^*(\vec{r}_2)\varphi_k(\vec{r}_1)\varphi_l(\vec{r}_2)}{r_{12}} dv_1 dv_2 \quad (6.16)$$

для двухэлектронных интегралов межэлектронного отталкивания. (Далее мы везде будем заменять общий двухэлектронный оператор  $\hat{g}(1, 2)$ , использованный в гл. 5, разд. 6.3, на явный вид оператора межэлектронного взаимодействия  $\frac{1}{r_{12}}$ , входящего в гамильтониан Борна—Оппенгеймера.) Уравнение (6.15) можно также переписать в виде:

$$\int \varphi_i'^*(\vec{r}_1) \left\{ \left[ \hat{h} + \sum_{\substack{j=1 \\ (j \neq i)}}^N (\hat{J}_j - \hat{K}_j \cdot \delta_{\gamma_i\gamma_j}) \right] \varphi_i(\vec{r}_1) \right\} dv_1 = 0 , \quad (6.17)$$

где введены «кулоновский» и «обменный» операторы, определенные каждый своим действием на произвольную функцию  $\varphi(\vec{r})$ :

$$\hat{J}_j\varphi(\vec{r}_1) = \int \frac{|\varphi_j(\vec{r}_2)|^2}{r_{12}} dv_2 \varphi(\vec{r}_1) = \int \frac{\varphi_j^*(\vec{r}_2)\varphi_j(\vec{r}_2)}{r_{12}} dv_2 \varphi(\vec{r}_1) \quad (6.18)$$

и

$$\hat{K}_j\varphi(\vec{r}_1) = \int \frac{\varphi_j^*(\vec{r}_2)\varphi(\vec{r}_2)}{r_{12}} dv_2 \varphi_j(\vec{r}_1) . \quad (6.19)$$

Следует обратить внимание на «нелокальный» характер обменного оператора  $\hat{K}_j$ .

Если бы функция  $\varphi'_i$  в выражении (6.17) была совершенно произвольной, то из основной леммы вариационного исчисления (функция,

ортогональная к произвольной функции, равна нулю «почти всюду») следовало бы, что выражение в фигурных скобках обращается в нуль «почти всюду». Но  $\varphi_i'$  ограничена условием ортогональности ко всем орбиталям  $\varphi_j$ , которые заняты в  $\Psi$ , и имеют ту же проекцию спина, что и  $\varphi_i$ . Поэтому функция в фигурных скобках может иметь вклады, пропорциональные орбиталям, занятым с теми же самыми проекциями спина, но не содержит вклады из ортогонального дополнения к ним. С учетом этого мы получаем уравнения метода НХФ в следующем общем виде:

$$\left[ \hat{h} + \sum_{\substack{j=1 \\ (j \neq i)}}^N (\hat{J}_j - \hat{K}_j \cdot \delta_{\gamma_i \gamma_j}) \right] \varphi_i = \sum_{j=1}^N \lambda_{ji} \varphi_j \cdot \delta_{\gamma_i \gamma_j} \quad (i = 1, 2, \dots, N), \quad (6.20)$$

где коэффициенты  $\lambda_{ij}$  могут быть произвольными.

Легко видеть, что эти выкладки можно проделать и в противоположном направлении, а потому выполнение уравнения НХФ (6.20) дает не только необходимые, но также и достаточные условия стационарности энергии ( $\delta E = 0$ ) для однодетерминантной волновой функции.

Теперь выпишем более подробно уравнения для случая, когда  $n_a$  орбиталей  $a_i = a_i(\vec{r})$  заполнены электронами со спином  $\alpha$ , а  $n_b$  орбиталей  $b_i = b_i(\vec{r})$  заполнены электронами со спином  $\beta$  (т. е.  $\varphi_i = a_i$  или  $b_i$ ). Тогда волновую функцию  $\Psi$  можно записать как

$$\Psi = \hat{\mathcal{A}}[a_1(\vec{r}_1)\alpha(\sigma_1)b_1(\vec{r}_2)\beta(\sigma_2)a_2(\vec{r}_3)\alpha(\sigma_3)\dots]. \quad (6.21)$$

Определим кулоновские и обменные операторы  $\hat{J}_j^a, \hat{J}_j^b, \hat{K}_j^a, \hat{K}_j^b$ , соответствующие орбиталям  $a_j, b_j$  аналогично определениям (6.18), (6.19), приведенным выше; например:

$$\hat{J}_j^a \varphi(\vec{r}_1) = \int \frac{a_j^*(\vec{r}_2)a_j(\vec{r}_2)}{r_{12}} dv_2 \varphi(\vec{r}_1) \quad (6.22)$$

и

$$\hat{K}_j^b \varphi(\vec{r}_1) = \int \frac{b_j^*(\vec{r}_2)\varphi(\vec{r}_2)}{r_{12}} dv_2 b_j(\vec{r}_1). \quad (6.23)$$

Тогда уравнения метода НХФ можно переписать так:

$$\left[ \hat{h} + \sum_{\substack{j=1 \\ (j \neq i)}}^{n_a} (\hat{J}_j^a - \hat{K}_j^a) + \sum_{j=1}^{n_b} \hat{J}_j^b \right] a_i = \sum_{j=1}^{n_a} \varepsilon_{ji}^a a_j \quad (i = 1, 2, \dots, n_a) \quad (6.24)$$

Аналогичный набор уравнений для орбиталей  $b_i$  можно получить из уравнений (6.24), поменяв всюду  $a$  на  $b$  и наоборот. Легко проверить непосредственно, что  $\hat{J}_i^a a_i = \hat{K}_i^a a_i$ , поэтому в левую часть можно добавить, так называемое, слагаемое «самоотталкивания»  $(\hat{J}_i^a - \hat{K}_i^a) a_i = 0$ , позволяющее отбросить ограничение  $j \neq i$  при суммировании. В результате оператор в скобках становится одним и тем же для всех орбиталей.

Введем операторы Фока («фокианы»)  $\hat{F}^a$  и  $\hat{F}^b$ :

$$\begin{aligned}\hat{F}^a &= \hat{h} + \sum_{j=1}^{n_a} (\hat{J}_j^a - \hat{K}_j^a) + \sum_{j=1}^{n_b} \hat{J}_j^b \\ \hat{F}^b &= \hat{h} + \sum_{j=1}^{n_b} (\hat{J}_j^b - \hat{K}_j^b) + \sum_{j=1}^{n_a} \hat{J}_j^a\end{aligned}\quad (6.25)$$

с использованием которых уравнения НХФ (условия стационарности энергии) могут быть записаны более компактно:

$$\begin{aligned}\hat{F}^a a_i &= \sum_{j=1}^{n_a} \varepsilon_{ji}^a a_j \quad (i = 1, 2, \dots, n_a) ; \\ \hat{F}^b b_i &= \sum_{j=1}^{n_b} \varepsilon_{ji}^b b_j \quad (i = 1, 2, \dots, n_b) .\end{aligned}\quad (6.26)$$

Операторы  $\hat{F}^a$  и  $\hat{F}^b$  зависят от всех орбиталей, занятых электронами с обеими проекциями спина в волновой функции  $\Psi$ , но зависимость эта такова, что они определяются однозначно не точным видом отдельных орбиталей, а лишь двумя *подпространствами*, растянутыми наборами занятых орбиталей  $\{a_i\}$  и  $\{b_i\}$ . Это следует из того факта, что  $\hat{F}^a$  и  $\hat{F}^b$  инвариантны при унитарных преобразованиях орбиталей, занятых электронами с той же самой проекцией спина. (Смешивание занятых орбиталей с разными проекциями спинов не рассматривается, так как это привело бы к «спин-орбиталям общего вида», не отвечающим определенной проекции спина.)

*Доказательство инвариантности*

Рассмотрим, например, унитарное преобразование орбиталей  $a_i$ , и вычислим сумму

$$\sum_{i=1}^{n_a} \hat{J}_i^a = \sum_{i=1}^{n_a} \int \frac{a_i^*(\vec{r}_2) a_i(\vec{r}_2)}{r_{12}} dv_2 =$$

$$\begin{aligned}
&= \sum_{i=1}^{n_a} \int \frac{\sum_{j=1}^{n_a} U_{ji}^* a_j^{I*}(\vec{r}_2) \sum_{k=1}^{n_a} U_{ki} a'_k(\vec{r}_2)}{r_{12}} dv_2 \\
&= \sum_{i,j,k=1}^{n_a} U_{ki} U_{ji}^* \int \frac{a_j^{I*}(\vec{r}_2) a'_k(\vec{r}_2)}{r_{12}} dv_2.
\end{aligned} \tag{6.27}$$

Так как  $U_{ji}^* = (\mathbf{U}^\dagger)_{ij} = (\mathbf{U}^{-1})_{ij}$ , имеем  $\sum_{i=1}^{n_a} U_{ki} U_{ji}^* = \delta_{kj}$ , так что правая часть (6.27) упрощается, и мы получаем

$$\sum_{i=1}^{n_a} \hat{J}_i^a = \sum_{j=1}^{n_a} \int \frac{a_j^{I*}(\vec{r}_2) a'_j(\vec{r}_2)}{r_{12}} dv_2 = \sum_{j=1}^{n_a} \hat{J}_j^{a'}. \tag{6.28}$$

Инвариантность сумм операторов  $\hat{J}_j^b$ ,  $\hat{K}_j^a$ , и  $\hat{K}_j^b$  можно проверить точно также. Ч. т. д.

Сама волновая функция  $\Psi$  также инвариантна по отношению к унитарным преобразованиям занятых орбиталей (за исключением, возможно, несущественного фазового множителя, равного по модулю единице); ее энергия, конечно, тоже инвариантна. Предшествующий вывод уравнений НХФ, который дает необходимые и достаточные условия стационарности энергии, равным образом применим к любому набору ортонормированных орбиталей, растягивающему подпространство занятых орбиталей; следовательно, каждый такой набор должен удовлетворять уравнениям НХФ (6.26). Необходимые и достаточные условия стационарности энергии можно поэтому сформулировать, сказав, что занятые орбитали  $a_i$  и  $b_i$  образуют *инвариантное подпространство* операторов Фока  $\hat{F}^a$  и  $\hat{F}^b$ , соответственно. (Следовательно, действуя оператором  $\hat{F}^a$  на любую орбиталь в подпространстве занятых орбиталей  $\{a_i\}$ , мы получим функцию, которая лежит в том же самом подпространстве.)

Кулоновский и обменный операторы являются эрмитовыми. Например, для произвольных функций  $\varphi$  и  $\psi$  имеем

$$\begin{aligned}
\langle \varphi | \hat{J}^a \psi \rangle &= \int \varphi^*(\vec{r}_1) dv_1 \left[ \int \frac{|a_i(\vec{r}_2)|^2}{r_{12}} dv_2 \psi(\vec{r}_1) \right] \\
&= \iint \frac{\varphi^*(\vec{r}_1) \psi(\vec{r}_1) |a_i(\vec{r}_2)|^2}{r_{12}} dv_1 dv_2.
\end{aligned} \tag{6.29}$$

Чтобы вычислить  $\langle \hat{J}_i^a \varphi | \psi \rangle$ , мы должны подставить функцию  $[\hat{J}_i^a \varphi(\vec{r})]^*$  в подинтегральное выражение. Получаем

$$\begin{aligned} \langle \hat{J}_i^a \varphi | \psi \rangle &= \int \left[ \int \frac{|a_i(\vec{r}_2)|^2}{r_{12}} dv_2 \varphi(\vec{r}_1) \right]^* \psi(\vec{r}_1) dv_1 \\ &= \iint \frac{|a_i(\vec{r}_2)|^2 \varphi^*(\vec{r}_1) \psi(\vec{r}_1)}{r_{12}} dv_1 dv_2 = \langle \varphi | \hat{J}_i^a \psi \rangle. \end{aligned} \quad (6.30)$$

Таким же образом имеем

$$\begin{aligned} \langle \varphi | \hat{K}_i^a \psi \rangle &= \int \varphi^*(\vec{r}_1) \left[ \int \frac{a_i^*(\vec{r}_2) \psi(\vec{r}_2)}{r_{12}} dv_2 a_i(\vec{r}_1) \right] dv_1 \\ &= \iint \frac{\varphi^*(\vec{r}_1) a_i^*(\vec{r}_2) a_i(\vec{r}_1) \psi(\vec{r}_2)}{r_{12}} dv_1 = [\varphi a_i | a_i \psi] \end{aligned} \quad (6.31)$$

и

$$\begin{aligned} \langle \hat{K}_i^a \varphi | \psi \rangle &= \int \left[ \int \frac{a_i^*(\vec{r}_2) \varphi(\vec{r}_2)}{r_{12}} dv_2 a_i(\vec{r}_1) \right]^* \psi(\vec{r}_1) dv_1 \\ &= \iint \frac{a_i^*(\vec{r}_1) \varphi^*(\vec{r}_2) \psi(\vec{r}_1) a_i(\vec{r}_2)}{r_{12}} dv_1 dv_2 = [a_i \varphi | \psi a_i]. \end{aligned} \quad (6.32)$$

Перестановка переменных интегрирования в определении (6.16) приводит к тождеству

$$[ab|cd] \equiv [ba|dc]. \quad (6.33)$$

Используя это, мы можем преобразовать правую часть выражения (6.32) и получить

$$\langle \hat{K}_i^a \varphi | \psi \rangle = [\varphi a_i | a_i \psi] = \langle \varphi | \hat{K}_i^a \psi \rangle. \quad \text{Ч. т. д.} \quad (6.34)$$

Операторы Фока  $\hat{F}^a$  и  $\hat{F}^b$  выражаются через суммы эрмитовых операторов, поэтому они сами также эрмитовы. Величины  $\varepsilon_{ji}^a$  и  $\varepsilon_{ji}^b$  можно рассматривать, как элементы матриц  $\varepsilon^a$  и  $\varepsilon^b$ , имеющих размеры  $n_a \times n_a$  и  $n_b \times n_b$ , соответственно. Эти матрицы также эрмитовы.

В самом деле, рассмотрим

$$\hat{F}^a a_i = \sum_{j=1}^{n_a} \varepsilon_{ji}^a a_j, \quad (6.35)$$

умножим на  $a_k^*$ , и проинтегрируем. Получим

$$\langle a_k | \hat{F}^a | a_i \rangle = \sum_{j=1}^{n_a} \varepsilon_{ji}^a \underbrace{\langle a_k | a_j \rangle}_{\delta_{kj}} \quad (6.36)$$



т. е.

$$\varepsilon_{ki}^a = \langle a_k | \hat{F}^a | a_i \rangle . \quad (6.37)$$

В то же время, учитывая эрмитовость оператора  $\hat{F}^a$ , имеем

$$\varepsilon_{ik}^a = \langle a_i | \hat{F}^a | a_k \rangle = (\langle a_k | \hat{F}^a | a_i \rangle)^* = \varepsilon_{ki}^{a*} . \quad (6.38)$$

Эрмитовость матрицы  $\varepsilon^b$  можно доказать таким же образом. Ч. т. д.

Согласно результату (6.37), матрица  $\varepsilon^a$  представляет собой матрицу оператора Фока  $\hat{F}^a$  в подпространстве занятых орбиталей  $\{a_i\}$ . (Точно так же,  $\varepsilon^b$  дает представление  $\hat{F}^b$  в подпространстве  $\{b_i\}$ .) Так как матрица  $\varepsilon^a$  эрмитова, ее можно привести к диагональному виду унитарным преобразованием

$$\mathbf{U}^\dagger \varepsilon^a \mathbf{U} = \text{diag}\{\varepsilon_i^{a'}\} . \quad (6.39)$$

Подвергнем орбитали  $a_i$  унитарному преобразованию  $\mathbf{U}$ , диагонализующему матрицу  $\varepsilon^a$ . Если

$$a'_i = \sum_{j=1}^{n_a} U_{ji} a_j , \quad (6.40)$$

то, используя инвариантность  $\hat{F}^a$  по отношению к унитарным преобразованиям орбиталей, получим

$$\begin{aligned} \varepsilon_{ik}^{a'} &= \langle a'_i | \hat{F}^a | a'_k \rangle = \left\langle \sum_{j=1}^{n_a} U_{ji} a_j \right| \hat{F}^a \left| \sum_{l=1}^{n_a} U_{lk} a_l \right\rangle = \sum_{j,l=1}^{n_a} U_{ji}^* \underbrace{\langle a_j | \hat{F}^a | a_l \rangle}_{\varepsilon_{jl}^a} U_{lk} \\ &= \sum_{j,l=1}^{n_a} (\mathbf{U}^\dagger)_{ij} \varepsilon_{jl}^a U_{lk} = (\mathbf{U}^\dagger \varepsilon^a \mathbf{U})_{ik} = \varepsilon_i^{a'} \delta_{ik} \end{aligned} \quad (6.41)$$

в полном согласии с предшествующим утверждением, что матрица  $\varepsilon^a$  является матрицей, представляющей оператор  $\hat{F}^a$  в подпространстве занятых орбиталей.

Можно проделать аналогичные выкладки с матрицей  $\varepsilon^b$  и орбиталями  $b_i$ , и тогда, опуская штрихи, в результате получим уравнения НХФ в так называемом «каноническом» виде

$$\begin{aligned} \hat{F}^a a_i &= \varepsilon_i^a a_i & (i = 1, 2, \dots, n_a) ; \\ \hat{F}^b b_i &= \varepsilon_i^b b_i & (i = 1, 2, \dots, n_b) . \end{aligned} \quad (6.42)$$

Эти уравнения представляют собой систему зацепляющихся «псевдозадач на собственные значения». Приставку «псевдо» используют потому,

что хотя уравнения и имеют вид уравнений на собственные значения, операторы  $\hat{F}^a$ ,  $\hat{F}^b$  сами зависят от искоемых собственных функций; уравнения «сцеплены», так как для того, чтобы найти решение уравнений для данной орбитали, необходимо знать и другие орбитали. Таким образом, в методе ХФ  $3N$ -мерное уравнение Шрёдингера приближенно сводится к системе  $N$  сцепленных трехмерных уравнений, которые решать гораздо легче. Собственные значения  $\varepsilon_i^a$ ,  $\varepsilon_i^a$  обычно называются «орбитальными энергиями» по причине, которая будет обсуждаться в разд. 3.

Уравнения этого типа обычно решаются итерационно: начинают с первого приближения орбиталей, строят с их использованием операторы Фока, решая уравнения, получают новые орбитали, и образуют следующий оператор Фока с этими новыми орбиталями и т. д. Символически этот процесс самосогласования можно отобразить так:

$$\{\varphi^{[0]}\} \rightarrow \hat{F}^{[1]} \rightarrow \{\varphi^{[1]}\} \rightarrow \hat{F}^{[2]} \rightarrow \dots \quad (6.43)$$

Процедура должна повторяться до тех пор, пока не будет достигнута сходимость с требуемой точностью. Критерий сходимости может относиться к самим орбиталям или (что лучше) к подпространству, растянутому занятыми орбиталями; обычно следят также за значениями энергии. Если удалось выбрать удачное начальное приближение  $\{\varphi_i^{[0]}\}$ , то можно быстро достичь сходимости; в случае худшего выбора процедура ССП может сходиться медленно или совсем не сходиться (расходиться).

В разд. 5.3 мы выведем специальный вариант уравнений НХФ, который позволяет последовательно оптимизировать орбитали в духе алгоритма, обсуждавшегося в разд. 1.3. И вновь проблема будет состоять в том, что такой метод сходится безусловно, но требует непомерно большого объема вычислений.

## 2.2. Альтернативный вывод с использованием множителей Лагранжа

При наиболее распространенном способе вывода уравнений Хартри—Фока элементы матрицы  $\epsilon$  возникают как множители Лагранжа, соответствующие условию ортонормировки одноэлектронных орбиталей. Однако этот метод, который можно найти во многих учебниках, вызывает некоторые вопросы математического характера, так что мы вновь внимательно обсудим эту задачу.

В этом выводе исходным является выражение для энергии в приближении НХФ, полученное для случая ортонормированных орбиталей; его

можно написать, комбинируя уравнения (5.101) и (5.116) и используя упрощенное обозначение (6.16) для интегралов, в виде

$$E = \sum_{i=1}^{n_a} \langle a_i | \hat{h} | a_i \rangle + \sum_{i=1}^{n_b} \langle b_i | \hat{h} | b_i \rangle + \frac{1}{2} \sum_{i,j=1}^{n_a} ([a_i a_j | a_i a_j] - [a_i a_j | a_j a_i]) + \frac{1}{2} \sum_{i,j=1}^{n_b} ([b_i b_j | b_i b_j] - [b_i b_j | b_j b_i]) + \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} [a_i b_j | a_i b_j] . \quad (6.44)$$

Условие стационарности для *этой*  $E$  должно быть выполнено при дополнительном условии ортонормировки

$$\begin{aligned} \langle a_i | a_j \rangle &= \delta_{ij} & (i, j = 1, 2, \dots, n_a) ; \\ \langle b_i | b_j \rangle &= \delta_{ij} & (i, j = 1, 2, \dots, n_b) . \end{aligned} \quad (6.45)$$

Для этого составляют вспомогательный функционал

$$F = E - \sum_{i,j=1}^{n_a} \varepsilon_{ji}^a (\langle a_i | a_j \rangle - \delta_{ij}) - \sum_{i,j=1}^{n_b} \varepsilon_{ji}^b (\langle b_i | b_j \rangle - \delta_{ij}) \quad (6.46)$$

включающий множители Лагранжа  $-\varepsilon_{ji}^a$  и  $-\varepsilon_{ji}^b$  и требуют выполнения условия  $\delta F = 0$ . (Для дальнейшего удобства множители Лагранжа вводятся со знаком минус.)

Уместно следующее замечание относительно функционала (6.46). Каждое из равенств (6.45) с  $i \neq j$  содержит два независимых условия — одно для вещественной части, другое для мнимой части интеграла  $\langle a_i | a_j \rangle$  (или  $\langle b_i | b_j \rangle$ ); поэтому мы должны ввести два независимых множителя Лагранжа  $\varepsilon_{ij}^a$  и  $\varepsilon_{ji}^a$  для интеграла  $\langle a_i | a_j \rangle$  и его комплексно сопряженного  $\langle a_j | a_i \rangle$ . В самом деле, выпишем условие ортогональности для орбиталей  $a_i$  и  $a_j$  через вещественную и мнимую части интегралов перекрывания:

$$\begin{aligned} \operatorname{Re} \langle a_i | a_j \rangle &= \frac{1}{2} (\langle a_i | a_j \rangle + \langle a_j | a_i \rangle) = 0 ; \\ \operatorname{Im} \langle a_i | a_j \rangle &= \frac{1}{2i} (\langle a_i | a_j \rangle - \langle a_j | a_i \rangle) = 0 . \end{aligned} \quad (6.47)$$

Мы можем ввести множители Лагранжа  $\mu_{ij}$  и  $\nu_{ij}$ , соответствующие этим двум условиям. Тогда получим следующее слагаемое во вспомогательном функционале

$$\begin{aligned} &\mu_{ij} \frac{1}{2} (\langle a_i | a_j \rangle + \langle a_j | a_i \rangle) + \nu_{ij} \frac{1}{2i} (\langle a_i | a_j \rangle - \langle a_j | a_i \rangle) \\ &= \frac{1}{2} [(\mu_{ij} - i\nu_{ij}) \langle a_i | a_j \rangle + (\mu_{ij} + i\nu_{ij}) \langle a_j | a_i \rangle] . \end{aligned} \quad (6.48)$$

Обозначая  $\varepsilon_{ji}^a = -\frac{1}{2}(\mu_{ij} - i\nu_{ij})$  и  $\varepsilon_{ij}^a = -\frac{1}{2}(\mu_{ij} + i\nu_{ij})$ , приходим к уравнению вида (6.46), а также можем заключить, что в задачах такого типа матрица множителей Лагранжа должна быть эрмитовой. (Разумеется, подобные рассуждения верны и для множителей  $\varepsilon_{ij}^b$ .)

Теперь проварируем орбиталь  $a_k$ . Получим

$$\begin{aligned} \delta F &= \langle \delta a_k | \hat{h} | a_k \rangle + \frac{1}{2} \sum_{j=1}^{n_a} ([\delta a_k a_j | a_k a_j] - [\delta a_k a_j | a_j a_k]) \\ &+ \frac{1}{2} \sum_{i=1}^{n_a} ([a_i \delta a_k | a_i a_k] - [a_i \delta a_k | a_k a_i]) \\ &+ \sum_{j=1}^{n_b} ([\delta a_k b_j | a_k b_j] - \sum_{j=1}^{n_a} \varepsilon_{jk}^a \langle \delta a_k | a_j \rangle + c.c. = 0. \end{aligned} \quad (6.49)$$

По свойству симметрии двухэлектронных интегралов, первые две суммы можно собрать в одну. Тогда, воспользовавшись определением (6.25) оператора  $\hat{F}^a$ , получим

$$\begin{aligned} \delta F &= \langle \delta a_k | \hat{F}^a a_k \rangle - \sum_{j=1}^{n_a} \varepsilon_{jk}^a \langle \delta a_k | a_j \rangle + c.c. \\ &= \int \delta a_k^* \left[ \hat{F}^a a_k - \sum_{j=1}^{n_a} \varepsilon_{jk}^a a_j \right] dv + c.c. = 0. \end{aligned} \quad (6.50)$$

Здесь вариация  $\delta a_k$  совершенно произвольна (и содержит произвольный фазовый множитель), поэтому функции в скобках должны обратиться в нуль «почти всюду», что приводит к уравнениям НХФ

$$\hat{F}^a a_k = \sum_{j=1}^{n_a} \varepsilon_{jk}^a a_j \quad k = 1, 2, \dots, n_a. \quad (6.51)$$

Уравнения для орбиталей  $b_k$  можно вывести аналогично.

Необходимо отметить, что предшествующий вывод не вполне соответствует стандартному методу использования множителей Лагранжа, принятому в вариационном исчислении. Там показано, что необходимо образовать вспомогательный функционал  $F'$  в виде суммы интеграла  $\langle \Psi | \hat{H} | \Psi \rangle$  по детерминантной волновой функции  $\Psi$  и ограничений, умноженных на соответствующие множители Лагранжа:

$$F' = \langle \Psi | \hat{H} | \Psi \rangle + \sum_{i,j=1}^{n_a} \lambda_{ij}^a (\langle a_i | a_j \rangle - \delta_{ij}) + \sum_{i,j=1}^{n_b} \lambda_{ij}^b (\langle b_i | b_j \rangle - \delta_{ij}), \quad (6.52)$$

и затем проварьировать его, как для вариационной задачи *без ограничений*; дополнительные условия (ограничения) должны быть подставлены только *после этого*. Это значит, что интегралы перекрывания между орбиталями и миноры матрицы интегралов перекрывания, которые неявно присутствуют в формуле для энергии (6.44) с нулевыми или единичными значениями, также должны варьироваться. Рассматривая специальные вариации одной орбитали, легко проверить, что эта стандартная математическая процедура приводит к результатам, сильно отличающимся от результата предыдущих выкладок:

а) Пусть  $\delta a_i = \eta a_k$  для  $i \neq k$ , т. е. мы умножаем один столбец определителя  $\Psi$  на  $\eta$  и добавляем его к другому столбцу. Это не меняет волновую функцию, согласно свойствам определителей, поэтому  $\delta \langle \Psi | \hat{H} | \Psi \rangle = 0$ , и требование  $\delta F' = 0$  приводит к  $\eta^* \lambda_{ik}^a + \eta \lambda_{ki}^a = 0$ . Так как  $\eta$  содержит произвольный фазовый множитель, такое равенство возможно, только если  $\lambda_{ik}^a = \lambda_{ki}^a = 0$ . Нулевое значение, полученное для недиагональных матричных элементов  $\lambda_{ik}^a$ , вполне согласуется с тем фактом, что всегда можно добиться ортогональности орбиталей, и поэтому она не представляет истинного ограничения для вариационной задачи.

б) Пусть  $\delta a_i = \eta a_i$ . Это означает, что после вариации один из столбцов определителя  $\Psi$  умножается на  $1 + \eta$ . По свойствам определителей, эта вариация меняет волновую функцию  $\Psi$  на тот же множитель, поэтому  $\delta \Psi = \eta \Psi$ . Накладывая условие  $\delta F' = 0$ , получаем  $\delta F' = \eta \langle \Psi | \hat{H} | \Psi \rangle + \lambda_{ii}^a \eta \langle a_i | a_i \rangle + \text{с.с.} = 0$ . Используя условие  $\langle a_i | a_i \rangle = 1$ , приходим к равенству  $\lambda_{ii}^a = -\langle \Psi | \hat{H} | \Psi \rangle$ . Если удовлетворяются все дополнительные условия в (6.45), то  $\langle \Psi | \Psi \rangle = 1$  и  $\lambda_{ii}^a = -E$ , где  $E$  есть энергия детерминантной волновой функции  $\Psi$ . Этот результат находится в полном согласии с полученным при других способах вывода, проведенных с использованием множителей Лагранжа (гл. 2, разд. 1.3 и 1.5 и гл. 3, разд. 1), так как условия (6.45) для  $i = j$  непосредственно связаны с нормировкой волновой функции  $\Psi$ .

Очевидно, что множители Лагранжа  $\lambda_{ij}^a, \lambda_{ij}^b$ , использованные в рассуждениях а) и б), никак не связаны с множителями Лагранжа  $\varepsilon_{ij}^a, \varepsilon_{ij}^b$ , которые использованы в функционале (6.46).

Вариации интегралов перекрывания и разных миноров матрицы интегралов перекрывания, необходимые для полного вывода с помощью функционала (6.52), очень громоздки. В принципе этот вывод можно провести (автор делал это однажды в молодости) и, конечно, он в конце концов приводит к уравнениям, эквивалентным выведенным другими методами.

Хотелось бы все же понять, почему при таких обстоятельствах простой вывод (6.46)-(6.51) с использованием множителей Лагранжа, все-таки приводит к корректным уравнениям. Ответ состоит в том, что вариацию величины  $E$ , являющейся частью функционала (6.46), нельзя рассматривать как вариацию энергии и рассмотренную процедуру надо понимать следующим образом: величина  $E$ , определенная выражением (6.44), получается из выражения для энергии приближения НХФ с использованием дополнительных условий. Ее нужно определять как некоторый новый вспомогательный функционал орбиталей, который, строго говоря, не имеет непосредственного физического смысла. Однако, он *равен* энергии *тогда* (и только тогда), когда выполнены дополнительные условия ортонормировки. При такой постановке задачи нужно искать стационарное значение этого вспомогательного функционала при тех же условиях ортонормировки. В классе пробных функций, удовлетворяющих дополнительным условиям, эта величина и энергия имеют, разумеется, идентичные стационарные значения, реализуемые для тех же самых волновых функций. Однако с формальной математической точки зрения мы имеем две разные вариационные задачи (одну для истинной энергии, другую для вспомогательной величины), поэтому их множители Лагранжа, конечно, будут различными.

Это различие может показаться несущественным с практической точки зрения. Однако его надо иметь в виду при использовании метода множителей Лагранжа, чтобы избежать возможных ошибок при исследовании более сложных вариационных задач с ограничениями.

### 2.3. Альтернативный вывод с использованием специальных вариаций

Теперь у нас есть все необходимое для еще одного альтернативного вывода уравнений метода НХФ. Рассмотрим снова формулу (6.44) для энергии, соответствующей ортонормированным одноэлектронным орбиталям, и рассмотрим только такие их вариации, которые не нарушают условие ортонормировки (по крайней мере, в первом порядке). Это значит, что формула для энергии остается применимой и после варьирования орбиталей. При таких условиях вариация энергии определяется вариацией выражения (6.44).

Сначала рассмотрим вариацию орбитали  $a_k$ , имеющую вид

$$\delta a_k = \eta c, \quad (6.53)$$

где  $\eta \rightarrow 0$  и орбиталь  $c$  удовлетворяет условию

$$\langle c | a_i \rangle = 0; \quad i = 1, 2, \dots, n_a. \quad (6.54)$$

Такая вариация сохраняет ортонормировку орбиталей в нулевом и первом порядках по вариационному параметру  $\eta$ .

Вариация энергии (6.44) в выражении (6.50) была приведена к виду  $\langle \delta a_k | \hat{F} | a_k \rangle + c.c.$  Поэтому условие  $\delta E = 0$  при вариации вида (6.53) приводит к условию

$$\eta^* \langle c | \hat{F} | a_k \rangle = 0. \quad (6.55)$$

Так как  $\eta \rightarrow 0$ , но  $\eta \neq 0$ , то равенство (6.55) можно переписать, как:

$$\int c^* \hat{F} a_k dv = 0. \quad (6.56)$$

Функция  $c$  произвольна, но требуется ее ортогональность ко всем занятым орбиталям  $a_i$ . Из обсуждения разд. 2.1 следует, что функция  $\hat{F} a_k$  в выражении (6.56) может содержать только вклады, пропорциональные орбиталям, занятым электронами с той же проекцией спина, что и сама  $a_k$ :

$$\hat{F} a_k = \sum_{j=1}^{n_a} \varepsilon_{jk} a_j; \quad k = 1, 2, \dots, n_a. \quad (6.57)$$

Точно так же можно вывести уравнения НХФ для орбиталей  $b_k$ .

Теперь рассмотрим некоторые другие вариации. Вариация  $\delta a_k = \eta a_i$  ( $i \neq k$ ) нарушает условия ортонормировки, но их можно восстановить, рассматривая совместную вариацию

$$\begin{aligned} \delta a_k &= \eta a_i \\ \delta a_i &= -\eta^* a_k. \end{aligned} \quad (6.58)$$

Пренебрегая слагаемыми, квадратичными по  $\eta$ , выражение для вариации энергии (6.44) при таком одновременном варьировании равно сумме вариаций, вызванных изменением  $\delta a_k$  и  $\delta a_i$  соответственно:

$$\langle \delta a_k | \hat{F} | a_k \rangle + \langle \delta a_i | \hat{F} | a_i \rangle + c.c. = 0. \quad (6.59)$$

Подставим вариации (6.58) и выпишем явно комплексно сопряженную часть:

$$\eta^* \langle a_i | \hat{F} | a_k \rangle - \eta \langle a_k | \hat{F} | a_i \rangle + \eta \langle a_k | \hat{F}^\dagger | a_i \rangle - \eta^* \langle a_i | \hat{F}^\dagger | a_k \rangle = 0. \quad (6.60)$$

Это условие выполняется автоматически за счет эрмитовости фокиана  $\hat{F}$ . Этого и можно было ожидать, так как вариации (6.58) представляют собой всего лишь линейное преобразование занятых орбиталей, и поэтому не приводят к какому-либо реальному изменению волновой функции и энергии.

### 3. Теорема Купманса

Собственные значения канонических уравнений Хартри—Фока обычно называются «орбитальными энергиями», так как, до некоторой степени, они напоминают энергии одноэлектронной системы во внешнем поле. Эту связь выражает «теорема Купманса» (Т. Koormans, 1933).

Мы исходим из выражения (6.44) для энергии системы, содержащей  $N = n_a + n_b$  электронов, в приближении НХФ (обозначим ее, как  $E_N$ ) и выделяем в этом выражении слагаемые, содержащие орбиталь  $a_k$ . Получаем

$$\begin{aligned}
 E_N = & \sum_{\substack{i=1 \\ (i \neq k)}}^{n_a} \langle a_i | \hat{h} | a_i \rangle + \langle a_k | \hat{h} | a_k \rangle + \sum_{i=1}^{n_b} \langle b_i | \hat{h} | b_i \rangle \\
 & + \frac{1}{2} \sum_{\substack{i,j=1 \\ (i,j \neq k)}}^{n_a} ([a_i a_j | a_i a_j] - [a_i a_j | a_j a_i]) + \frac{1}{2} \sum_{j=1}^{n_a} ([a_k a_j | a_k a_j] - [a_k a_j | a_j a_k]) \\
 & + \frac{1}{2} \sum_{i=1}^{n_a} ([a_i a_k | a_i a_k] - [a_i a_k | a_k a_i]) + \frac{1}{2} \sum_{j=1}^{n_a} ([b_i b_j | b_i b_j] - [b_i b_j | b_j b_i]) \\
 & + \sum_{\substack{i=1 \\ (i \neq k)}}^{n_a} \sum_{j=1}^{n_b} [a_i b_j | a_i b_j] + \sum_{j=1}^{n_b} [a_k b_j | a_k b_j] . \quad (6.61)
 \end{aligned}$$

Легко видеть, что слагаемые, не содержащие  $a_k$ , составляют энергию  $E_{N-1}(a_k)$  некоторой  $N-1$ -электронной системы в приближении НХФ. Именно, она соответствует такой волновой функции  $N-1$  электронов, которая получается удалением орбитали  $a_k$  из волновой функции (6.21). Имея в виду тождества  $[ab|cd] \equiv [ba|dc]$ ,  $[a_k a_j | a_k a_j] \equiv \langle a_k | \hat{J}_j^a | a_k \rangle$ ,  $[a_k a_j | a_j a_k] \equiv \langle a_k | \hat{K}_j^a | a_k \rangle$  и т. д., можно преобразовать слагаемые, содержащие  $a_k$ , и записать

$$\begin{aligned}
 E_N = & E_{N-1}(a_k) + \langle a_k | \hat{h} | a_k \rangle + \sum_{j=1}^{n_a} (\langle a_k | \hat{J}_j^a | a_k \rangle - \langle a_k | \hat{K}_j^a | a_k \rangle) \\
 & + \sum_{j=1}^{n_b} \langle a_k | \hat{J}_j^b | a_k \rangle = E_{N-1}(a_k) + \underbrace{\langle a_k | \hat{F}^a | a_k \rangle}_{\varepsilon_{kk}^a} . \quad (6.62)
 \end{aligned}$$

Так мы получаем *теорему Купманса*

$$E_N = E_{N-1}(a_k) + \varepsilon_{kk}^a . \quad (6.63)$$



В соответствии с полученным результатом, диагональные элементы матрицы  $\epsilon^a$  (а также матрицы  $\epsilon^b$ ) равны изменению энергии при удалении электрона с данной орбитали и сохранении остальных занятых орбиталей в неизменном виде. Обычно  $\epsilon_{kk}^a$  отрицательна, что указывает на то, что для ионизации системы нужно затратить энергию; поэтому величину  $-\epsilon_{kk}^a$  можно рассматривать как приближение для соответствующего потенциала ионизации.

Как уже отмечалось, матрица  $\epsilon^a$  задает представление оператора  $\hat{F}^a$  в подпространстве занятых орбиталей. Так как  $\epsilon^a$  эрмитова, ее собственные значения обладают свойствами, обсуждавшимися в гл. 3 в связи с линейной вариационной задачей. Необходимо отметить, что не только наименьшее собственное значение равно наибольшей по абсолютной величине отрицательной  $\epsilon_{kk}^a$ , но также и наивысшее собственное значение дается наименьшим по абсолютной величине из возможных отрицательных значений. (Напомним, что рассматривается только подпространство занятых орбиталей.) Так как энергия ионизованного состояния, в соответствии с формулой (6.63), равна

$$E_{N-1}(a_k) = E_N - \epsilon_{kk}^a, \quad (6.64)$$

то наименьшая энергия иона, которая получается при использовании орбиталей исходной системы (т. е. без рассмотрения «орбитальной релаксации», сопровождающей ионизацию), соответствует наивысшему (наименее отрицательному) возможному значению  $\epsilon_{kk}^a$ , которое как раз и достигается для наивысшей орбитальной энергии в каноническом представлении. Это свойство, помимо удобства вычислений с использованием уравнений псевдозадачи на собственные значения, выявляет особую роль канонических уравнений Хартри—Фока. В дополнение к математическим аспектам, канонические орбитали предпочтительны еще и потому, что они хорошо отражают *глобальные* физические свойства системы, в частности ионизацию и, — правда, в меньшей степени — процессы возбуждения. В то же время канонические орбитали часто совершенно не коррелируют с локальными характеристиками основного состояния, такими, как существование отдельных химических связей. (Для их описания более целесообразно использовать локализованные орбитали.)

В согласии с предшествующим обсуждением, орбитальные энергии дают оценку «вертикальных» потенциалов ионизации, т. е. вычисленных при фиксированной геометрии. Не только наивысшее собственное значение, но часто и более низлежащие собственные значения, хорошо описывают различные процессы ионизации. Получаемые численные значения обычно достаточно точны для того, чтобы различить разные иони-

зационные процессы в одном и том же атоме или молекуле, например, ионизацию с остовных орбиталей,  $\sigma$ - или  $\pi$ -уровней.

Значение потенциала ионизации, оцененное по теореме Купманса, всегда больше (по абсолютной величине), чем так называемое значение « $\Delta\text{SCF}$ », которое получают, проводя независимые вычисления методом ХФ для иона и молекулы и вычисляя разность полных энергий. (Энергия ССП для иона учитывает орбитальную релаксацию и с необходимостью меньше, чем значение, полученное из формулы (6.64), которое соответствует замороженным орбиталям.) Можно ожидать, что эта ошибка частично скомпенсируется эффектами электронной корреляции: понижение энергии, которое получается при учете электронной корреляции, обычно несколько меньше для иона, так как в нем число электронов меньше на единицу.

### 3.1. Орбитальные энергии и полная энергия

*Необходимо сделать важное замечание:* хотя энергия, необходимая для удаления одного электрона из системы, равна соответствующей орбитальной энергии (взятой с обратным знаком), полная электронная энергия системы (т. е. энергия, необходимая для удаления всех электронов при данной зафиксированной ядерной конфигурации) *не* равна сумме орбитальных энергий. Дело в том, что орбитальная энергия  $\varepsilon_k^a$  содержит полную энергию взаимодействия электрона, находящегося на орбитали  $a_k$ , со всеми остальными электронами: энергия же  $E_{N-1}(a_k)$  не включает никаких слагаемых, связанных с орбиталью  $a_k$ . Поэтому если просто просуммировать орбитальные энергии, то энергия электрон-электронного взаимодействия будет учтена *дважды*, тогда как кинетическая энергия и энергия электронно-ядерного притяжения будет учтена только один раз. Тогда можно написать

$$E = \sum_{i=1}^{n_a} \varepsilon_i^a + \sum_{i=1}^{n_b} \varepsilon_i^b - E_{\text{el.-el.}} \quad (6.65)$$

Вместо того, чтобы вычитать лишнюю энергию электрон-электронного взаимодействия из суммы орбитальных энергий, можно добавить еще раз одноэлектронную часть энергии (среднее значение одноэлектронной части гамильтониана) и разделить результат пополам:

$$E = \frac{1}{2} \left[ \sum_{i=1}^{n_a} (\varepsilon_i^a + \langle a_i | \hat{h} | a_i \rangle) + \sum_{i=1}^{n_b} (\varepsilon_i^b + \langle b_i | \hat{h} | b_i \rangle) \right] \quad (6.66)$$

Также можно подставить в выражение (6.66) запись орбитальных энергий через операторы Фока:

$$E = \frac{1}{2} \left[ \sum_{i=1}^{n_a} \langle a_i | \hat{F}^a | a_i \rangle + \sum_{i=1}^{n_b} \langle b_i | \hat{F}^b | b_i \rangle \right]. \quad (6.67)$$

В реальных вычислениях обычно используют какой-либо вариант последнего выражения.

Оператор Фока является эрмитовым, безотносительно к тому, построен ли он с использованием самосогласованных хартри-фоковских или любых других орбиталей. Он имеет также и собственные функции, находящиеся вне подпространства занятых орбиталей, так как собственные векторы эрмитового оператора образуют полный ортонормированный набор. Эти дополнительные собственные векторы («пустые», «незанятые», «вакантные» или «виртуальные» орбитали), разумеется, не влияют на процесс построения оператора Фока. До того, как итерационный процесс не сойдется, собственные векторы оператора Фока не совпадают в точности с занятыми орбиталями, использованными для его построения. Поэтому процесс итерационного решения уравнений ССП можно интерпретировать так, что на каждой итерации происходит некоторое смешивание занятых и вакантных орбиталей, полученных на предыдущем шаге — хотя обычно так явно не делают.

Как правило, занятым орбиталям принадлежат наинизшие орбитальные энергии (наинизшие собственные значения оператора Фока); это правило часто называют принципом «Aufbau» (~ «построения»)<sup>1</sup>, и обычно наблюдается значительная разность энергий (энергетическая щель) между орбитальной энергией ВЗМО (высшей занятой молекулярной орбитали) и НВМО (низшей вакантной молекулярной орбитали)<sup>2</sup>. В большинстве расчетов, проведенных для нейтральных систем, орбитальная энергия ВЗМО отрицательна, а НВМО — положительна. Такое обстоятельство возникает из-за того, что мы добавили к оператору Фока слагаемое «самоотталкивания»  $\hat{J}_k^a a_k - \hat{K}_k^a a_k = 0$ , чтобы привести его к виду (6.25), являющемуся единым уравнением для всех орбиталей  $a_k$ . Однако при действии этого оператора на виртуальную орбиталь  $a_j^v$  он содержит  $n_a$  слагаемых  $\hat{J}_k^a a_j^v - \hat{K}_k^a a_j^v \neq 0$ . Это можно интерпретировать

<sup>1</sup> Можно доказать, что устойчивое решение задачи Хартри—Фока, т. е. отвечающее минимуму, а не седловой точке, соответствующего функционала энергии, должно удовлетворять принципу Aufbau (см. Степанов Н.Ф., Пупышев В.И. Квантовая механика молекул и квантовая химия. — М.: МГУ, 1991). — *Прим. ред.*

<sup>2</sup> В оригинале использованы аббревиатуры НОМО (highest occupied molecular orbital) и LUMO (lowest unoccupied molecular orbital), соответственно. — *Прим. пер.*

так, что электрон, помещенный на виртуальную орбиталь, «видит» на один электрон больше, чем электрон на занятой орбитали, и это «выталкивает вверх» энергии виртуальных орбиталей. Вследствие этого, если начинать процедуру ССП с какой-то заданной схемой орбитального заполнения (например, с заданным числом  $\sigma$ - и  $\pi$ -электронов), то в подавляющем большинстве случаев общий характер этой схемы не меняется при итерациях: у орбитали, которая не заполнена в начальном приближении, будет более высокая орбитальная энергия из-за слагаемого самоотталкивания и, соответственно, она не заполнится и на последующих шагах итерации. Выбирая разные виды начальных приближений, в некоторых случаях процесс ССП можно заставить сходиться к решениям, аппроксимирующим разные точные электронные состояния. (Это обычно имеет место в системах с высокой симметрией.)

В некоторых случаях, однако, процедура ССП расходится — например, может возникнуть такая ситуация, что пара занятых и виртуальных орбиталей меняются местами на каждой итерации. В таких случаях может оказаться полезным ввести так называемый «сдвиг уровня», искусственно увеличивающий орбитальные энергии виртуальных орбиталей. Этого можно достичь, добавляя к оператору Фока слагаемое, равное  $\Delta\epsilon(1 - \hat{P}_{occ})$ , где  $\Delta\epsilon$  есть желаемый сдвиг орбитальной энергии, а  $\hat{P}_{occ}$  — проекционный оператор на подпространство занятых орбиталей. Легко видеть, что в тех случаях, когда сходимость уже достигнута, добавление такого слагаемого ни на что не повлияет, за исключением орбитальных энергий, соответствующих виртуальным орбиталам, не имеющих какого-либо реального значения для основного состояния приближения Хартри—Фока. Качественное рассуждение в духе теории возмущений показывает, что если увеличивать разность орбитальных энергий занятых и виртуальных орбиталей, то это уменьшит их «смешивание» (или воспрепятствует их перестановке) на следующей итерации и, соответственно, погасит осцилляции итерационной процедуры.

Используя достаточно большой, но конечный, сдвиг уровней, можно заставить любую задачу ССП сходиться к какому-то решению уравнений ХФ, которое может быть, а может и не быть, решением с наименьшей полной энергией. Из-за присутствия самоотталкивания в операторе Фока решения после достижения сходимости по большей части будут подчиняться принципу Aufbau, т. е. занятые орбитали будут принадлежать наименьшим орбитальным энергиям даже без использования энергетического сдвига, так что его можно «выключить», когда процесс начинает сходиться. (Это, однако, в общем случае нельзя доказать.) Конечно, если процедура ССП сходится сама по себе без использования сдвига уровней, то последний лишь замедлит итерационный процесс.

Стоит отметить, что существует множество других интерполяционных и экстраполяционных схем для улучшения сходимости в расчетах ССП.

Как замечено ранее, включение самоотталкивания в оператор Фока влечет за собой то, что энергии виртуальных орбиталей содержат взаимодействие еще с одним электроном. Поэтому виртуальные орбитали, скорее всего, соответствуют некоторым состояниям отрицательного иона, а не возбужденным состояниям исходной системы. Таким образом, энергии виртуальных орбиталей можно было бы связать со сродством системы к электрону<sup>1</sup>. Однако корректное описание отрицательных ионов обычно является куда более тонкой задачей, чем описание положительных ионов и нейтральных молекул. В большинстве случаев нельзя получить реалистическое описание связанных состояний отрицательных ионов без учета орбитальной релаксации; кроме того, необходимо также и явное рассмотрение электронной корреляции.

Если учесть самоотталкивание надлежащим образом, то орбитальные энергии можно использовать для получения оценки энергии возбуждения. Как было сказано ранее, эффективный потенциал, определяющий орбитальную энергию виртуальной орбитали (скажем,  $a_j^v$ ), содержит потенциальную энергию взаимодействий с электроном, занимающим орбиталь (скажем,  $a_k^o$ ), с которой происходит возбуждение; поэтому разность  $\varepsilon_j^{av} - \varepsilon_k^{ao}$  не равна энергии возбуждения. Для того чтобы получить лучшую оценку, необходимо вычесть из этой разности вклад самоотталкивания  $\langle a_j^v | \hat{J}_k^a - \hat{K}_k^a | a_j^v \rangle = [a_j^v a_k^o | a_j^v a_k^o] - [a_j^v a_k^o | a_k^o a_j^v]$ . (Эта формула справедлива для однодетерминантного метода НХФ. В случае метода ОХФ ситуация похожа, но формула другая, так как обычно рассматривают двухдетерминантные возбужденные состояния, отвечающие определенному полному спину — синглету или триплету, см. разд. 4.4.)

Из-за того, что канонические виртуальные орбитали не описывают возбужденных состояний, возбужденные состояния обычно не рассматриваются в рамках однодетерминантного приближения. Простейшим методом, который было бы разумно использовать для этой цели, является так называемый метод «КВ в базисе однократных возбуждений», т. е. решение линейной вариационной задачи для линейной комбинации всех детерминантных волновых функций, которые можно получить однократным возбуждением из хартри-фоковского основного состояния<sup>2</sup>.

<sup>1</sup> Формально так оно и есть. Это можно доказать, проведя выкладки, аналогичные тем, которые приводят к теореме Купманса. — *Прим. ред.*

<sup>2</sup> Такой подход (известный также как приближение Тамма—Данкова) занижает энергии возбужденных состояний из-за того, что корреляции оказываются учтены несбалансированно: в основном состоянии, найденном в приближении ХФ, их нет, а в возбужденных, где используется ограниченное КВ, они как-то учтены. В сложных

(Иногда рассматривают лишь часть таких конфигураций, пренебрегая возбуждениями, содержащими низколежащие занятые и/или высоколежащие виртуальные орбитали, т. е. проводят предварительный отбор на основании значений орбитальных энергий.)

В качественных рассуждениях о химических свойствах сравнение отдельных орбитальных энергий часто используется без должной осторожности, которую всегда необходимо проявлять для того, чтобы отличать суммы орбитальных энергий от полной электронной энергии. Подобные рассуждения можно отчасти оправдать ссылкой на результат (6.63), если в рассматриваемом процессе состояния всех электронов, кроме одного, меняются незначительно; в таком случае полная энергия меняется более или менее так же, как и энергия той орбитали, которая подвергается значительной деформации.

В этой связи можно упомянуть и тот факт, что можно построить такие эффективные (ССП) операторы, сумма диагональных матричных элементов которых по подпространствам занятых орбиталей будет равна полной электронной энергии. Тривиально проверяется, что такой оператор можно получить домножая на  $\frac{1}{2}$  каждый кулоновский и обменный оператор в операторе Фока (6.25). Применяя такие операторы, можно представить хартри-фововскую полную электронную энергию в виде суммы одноэлектронных вкладов. Если бы в каноническом расчете ССП использовался такой оператор, то полная энергия была бы равна сумме орбитальных энергий. Однако разумеется, такая процедура сошлась бы к волновой функции с энергией, более высокой, чем энергия ХФ. Тем не менее можно считать, что такой оператор может давать некоторое теоретическое основание многим качественным рассуждениям химиков, проведенным с помощью лишь карандаша и бумаги.

## 4. Метод ОХФ

### 4.1. Схемы ОХФ и НХФ

Наиболее распространенной версией метода Хартри—Фока, применяемой для описания синглетных состояний систем с замкнутой оболочкой является «ограниченный» метод ХФ (ОХФ, restricted HF, RHF), использующий дважды занятые пространственные орбитали. Тогда волновая функция имеет вид:

$$\Psi = \hat{\mathcal{A}}[\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2) \dots \varphi_n(2n-1)\alpha(2n-1)\varphi_n(2n)\beta(2n)] \quad (6.68)$$

случаях (малая энергетическая щель между ВЗМО и НСМО) может оказаться, что энергии низших возбужденных состояний будут отрицательными, т. е. окажется, что они лежат ниже основного состояния приближения ХФ. — *Прим. ред.*

(В этом разделе мы рассматриваем случай только четного числа электронов  $N=2n$  с одинаковым числом  $n$  спинов  $\alpha$  и  $\beta$ .)

Все рассуждения о свойствах и преобразованиях одноэлектронных орбиталей, инвариантности детерминантных функций и т. п., приведенные выше, остаются без изменений и в этом случае. Эквивалентность вариационного принципа ( $\delta E = 0$ ) и теоремы Бриллюэна тоже можно тривиально обосновать для настоящего случая. В самом деле, введение условия двукратной занятости приводит к единственному отличию от вывода, данного в разд. 1 и 2: в случае ОХФ мы должны рассматривать одновременно вариацию обеих спин-орбиталей, получаемых из данной пространственной орбитали  $\varphi_i$ . Тогда вариация орбитали  $\varphi_i$ , как  $\varphi_i \rightarrow \varphi_i + \delta\varphi_i = \varphi_i + \eta\varphi'_i$ , приводит к вариации первого порядка определителя  $\Psi$ , которая представляет собой сумму двух слагаемых:

$$\delta\Psi = \eta\Psi_1(\varphi_i\alpha \rightarrow \varphi'_i\alpha) + \eta\Psi_1(\varphi_i\beta \rightarrow \varphi'_i\beta) . \quad (6.69)$$

Однако матричные элементы между определителем  $\Psi$  и двумя однократно возбужденными определителями в правой части выражения (6.69) равны друг другу, так как они отличаются только перестановкой спинов  $\alpha$  и  $\beta$ . Поэтому присутствие двух слагаемых в выражении (6.69) ведет лишь к появлению множителя 2 в левой части уравнения  $\langle\delta\Psi|\hat{H} - E|\Psi\rangle = 0$ , дающего условие экстремума (ср. гл. 2, разд. 1.3). Таким образом, доказательство теоремы Бриллюэна и вывод уравнений ХФ через это утверждение с помощью пространственных орбиталей остаются точно такими же, как и для случая НХФ. Соответственно, уравнения ОХФ получаются из уравнений НХФ, если везде заменить  $a_i = b_i = \varphi_i$ ;  $n_a = n_b = n$ ;  $\varepsilon_i^a = \varepsilon_i^b = \varepsilon_i$ , и т. д. Таким образом, из уравнений (6.42) получаются канонические уравнения метода ОХФ в виде

$$\hat{F}\varphi_i = \varepsilon_i\varphi_i \quad (i = 1, 2, \dots, n) . \quad (6.70)$$

В этом случае можно также отбросить верхние индексы  $a$  и  $b$  в определениях кулоновского и обменного операторов и записать

$$\hat{F} = \hat{h} + \sum_{j=1}^n (2\hat{J}_j - \hat{K}_j) . \quad (6.71)$$

Очевидно, что мы получаем одни и те же уравнения ОХФ и в том случае, если исходим из уравнений НХФ для орбиталей  $a_i$  и для орбиталей  $b_i$ : число разных занятых пространственных орбиталей и число уравнений вдвое меньше, чем для случая НХФ. Уравнения ОХФ являются частным случаем уравнений НХФ и поэтому решения ОХФ также всегда

являются решениями НХФ. Часто решение НХФ совпадает с решением ОХФ и обычно это имеет место как раз в области равновесных геометрий молекул с замкнутыми оболочками. В таких случаях волновая функция ОХФ представляет собой минимум не только среди детерминантных волновых функций, построенных из дважды занятых орбиталей, но также и среди всех однодетерминантных волновых функций. В таких случаях нельзя добиться понижения энергии, расщепляя орбитали, т. е. используя различные подпространства занятых орбиталей, растянутые орбиталями со спинами  $\alpha$  и  $\beta$  соответственно.

Обратная картина может наблюдаться при больших межатомных расстояниях вдали от равновесия. В этом случае существует решение НХФ с меньшей энергией, чем ОХФ. (Точнее, решения НХФ типа всегда появляются парами, имеющими строго равные энергии и отличающимися друг от друга перестановкой всех проекций спинов  $\alpha$  и  $\beta$ .) В таких случаях решение ОХФ является минимумом только среди определителей с дважды занятыми орбиталями, но для определителей метода НХФ более общего вида, не имеющих такого ограничения, оно является лишь седловой точкой. Значит, волновая функция ОХФ является минимумом для случая дважды занятых орбиталей, но представляет собой максимум по отношению к таким изменениям орбитальных коэффициентов, которые приводят к «расщеплению» орбиталей и поэтому орбитали, занятые электронами с разными проекциями спинов  $\alpha$  и  $\beta$ , будут отличаться. (По этой причине метод НХФ часто также называют схемой «разных орбиталей для разных спинов», или ROPC.) Если существует решение НХФ с энергией, более низкой, чем решение ОХФ с замкнутой оболочкой, то часто говорят о «триплетной нестабильности» последнего. Эта терминология связана с тем фактом, что детерминант ROPC не является чистой синглетной собственной функцией оператора квадрата полного спина  $\hat{S}^2$ , а содержит также и некоторую примесь триплетных и прочих компонент.

Волновая функция приближения НХФ имеет в два раза больше свободно оптимизируемых одноэлектронных пространственных орбиталей, чем имеется для ОХФ, и представляет собой намного более гибкую вариационную пробную функцию. Вопрос о том, получится ли в итоге орбитальное расщепление, определяется относительной важностью двух противодействующих факторов. С одной стороны, расщепление орбиталей в волновой функции НХФ приводит к понижению энергии, так как, если электроны не обязаны занимать попарно одни и те же пространственные орбитали, то среднее расстояние между ними увеличится. (т. е. метод НХФ может учесть *некоторую часть* электронной корреляции.) С другой стороны, имеет место некоторое повышение энергии из-за того,



что средние расстояния между электронами и ядрами тоже несколько увеличиваются, и может также иметь место увеличение кинетической энергии.

## 4.2. Симметрия и метод ОХФ

Для большинства симметричных органических молекул волновая функция ОХФ автоматически отвечает пространственной симметрии системы. Это значит, что операции симметрии (например, отражение в плоскости симметрии или вращение вокруг оси симметрии) оставляют неизменной молекулярную волновую функцию ОХФ. В таких случаях канонические молекулярные орбитали также являются собственными функциями операторов симметрии, т. е. принадлежат неприводимому представлению группы симметрии системы. Невырожденные орбитали при операциях симметрии остаются инвариантными или меняют свои фазы, тогда как вырожденные орбитали преобразуются друг через друга, т. е. результатом действия операции симметрии на вырожденную орбиталь является линейная комбинация вырожденных орбиталей.

Было бы, однако, ошибкой считать, что слэтеровская детерминантная волновая функция, построенная из дважды занятых орбиталей, *всегда* является собственной для операторов симметрии. Это можно гарантировать только в случае, когда исходное приближение для орбиталей выбрано так, что получающийся из них оператор Фока имеет симметрию, совпадающую с симметрией системы, т. е. коммутирует со всеми операциями симметрии рассматриваемой точечной группы. Только тогда его собственные функции — канонические орбитали, — можно также классифицировать по их свойствам симметрии (неприводимым представлениям). Эта особенность сохраняется при дальнейших итерациях, если мы продолжаем использовать гарантирующий сохранение симметрии способ построения оператора Фока. Для этого требуется либо полностью заполнять, либо оставлять совершенно незанятым каждый вакантный при решении набор вырожденных орбиталей.

Может, однако, случиться, что для данной системы существует волновая функция, построенная из дважды занятых орбиталей, которая не обладает полной симметрией системы, но имеет более низкую энергию, чем детерминант, приведенный по симметрии. Такие волновые функции называются, соответственно, неограниченной по симметрии и ограниченной по симметрии функциями ХФ, а о ситуации, когда энергия неограниченной по симметрии функции оказывается ниже, чем для ограниченной, говорят как о «синглетной нестабильности» решений метода ОХФ.

Особое внимание следует уделить тем точкам, в которых симметрия системы выше, чем в их окрестности. Как указал Машер, добиваясь полной симметрии в этих изолированных точках, можно получить точечные разрывы на рассчитанной поверхности потенциальной энергии. (Примером такого поведения является молекула циклобутана в конформации правильного квадрата<sup>1</sup>.)

С помощью теоретико-групповой терминологии предшествующее обсуждение можно кратко подытожить следующим образом. Если оператор Фока соответствует волновой функции, построенной с «полным заполнением неприводимых представлений» (т. е. функции, образующие базис какого-то неприводимого представления, либо все являются дважды занятыми, либо все полностью вакантны), тогда он будет коммутировать с операциями симметрии точечной группы, и его собственные функции можно классифицировать по этим неприводимым представлениям. В таком случае полная волновая функция ОХФ также относится к одномерному полносимметричному представлению точечной группы. Решающее значение имеет правильная симметрия начального приближения, так как его свойства симметрии обычно сохраняются при итерациях<sup>2</sup>.

<sup>1</sup> Если молекула циклобутана имеет форму прямоугольника с неравными сторонами, то волновая функция метода ОХФ соответствует одной из двух возможных «структур Кекуле», а именно той, которая имеет двойные связи по более коротким сторонам прямоугольника. Характер волновой функции остается таким же, даже если совершить предельный переход к структуре правильного квадрата: волновая функция, описывающая одну «структуру Кекуле», соответствует симметрии системы все то время, пока стороны прямоугольника не становятся *строго* равными. Для квадратной геометрии однодетерминантные волновые функции, соответствующие двум «структурам Кекуле», вырождены по энергии. При этом ни та, ни другая не соответствуют полной симметрии квадрата, а их симметризованные линейные комбинации уже не являются однодетерминантными волновыми функциями. Если потребовать, чтобы однодетерминантная волновая функция метода ОХФ с двукратно заполненными орбиталями полностью соответствовала симметрии квадрата, то в этой единственной точке мы должны построить волновую функцию совершенно другого типа, имеющую значительно более высокую энергию.

<sup>2</sup> Это все тесно связано с принципом Aufbau и его возможными нарушениями. Описываемая проблема возникает прежде всего в системах с высокой симметрией, т. е. такой, чья группа симметрии имеет вырожденное (более чем одномерное) неприводимое представление, а именно  $C_{4v}$ ,  $D_{4h}$ ,  $O_h$ ,  $T_d$  и т.п. Содержательные примеры таких систем — прежде всего комплексы переходных металлов, где вырожденные орбитали имеют существенный вклад  $d$ -функций. В данном случае требование «полного заполнения неприводимых представлений» вступает в противоречие с физической картиной электронной структуры этих соединений, которые привычно (и обоснованно) характеризуются как  $d^2$ -,  $d^5$ - и т.п. комплексы. В случае октаэдрической симметрии орбитали  $d$ -оболочки расщепляются на двукратно вырожденное подпространство  $e_g$ -орбиталей и трехкратно вырожденное подпро-

### 4.3. «Диссоциационная катастрофа» и различные варианты метода Хартри—Фока

Метод ОХФ, как правило, хорошо передает основные характеристики электронной структуры молекул в окрестности равновесной геометрии; однако он полностью перестает отвечать действительности, когда рассматривается «гомолитическая» диссоциация молекулы с замкнутой оболочкой на два фрагмента с открытыми оболочками. Причина этого явления заключается в том, что использование дважды занятых орбиталей не позволяет описывать два изолированных атома, каждый из которых несет нечетное число электронов. Анализируя волновую функцию ОХФ, получающуюся в случае больших межатомных расстояний, можно установить, что она представляет собой суперпозицию равных вкладов, один из которых соответствует нейтральным атомам, а другой — паре ионов, в которой один электрон перенесен либо к одному либо к другому атому. (На практике обычно так и происходит — стандартная процедура ОХФ перестает сходиться в случае большого межатомного расстояния, но, по крайней мере в принципе, и в этом случае решение ОХФ существует и его можно найти с помощью соображений, основанных на симметрии, или других средств.)

#### Пример

Рассмотрим молекулу  $H_2$  при большом межатомном расстоянии. Волновая функция ОХФ равна

$$\Psi = \hat{A}[\varphi(1)\alpha(1)\varphi(2)\beta(2)] . \quad (6.72)$$

Орбиталь  $\varphi$  симметрична по перестановкам двух атомов, и ее можно записать как

$$\varphi = \mathcal{N}[\chi_A + \chi_B] , \quad (6.73)$$

где функции  $\chi_a$  и  $\chi_B$  связаны отражением, и каждая из них локализована возле соответствующего ядра. Если они обе нормированы и ортогональны друг другу, т. е.  $\langle \chi_A | \chi_A \rangle = \langle \chi_B | \chi_B \rangle = 1$ , и  $\langle \chi_A | \chi_B \rangle = 0$ , то нормировочный коэффициент  $\mathcal{N}$  равен  $\frac{1}{\sqrt{2}}$ . Таким образом,

$$\Psi = \frac{1}{2} \hat{A}\{[\chi_A(1) + \chi_B(1)]\alpha(1)[\chi_A(2) + \chi_B(2)]\beta(2)\} . \quad (6.74)$$

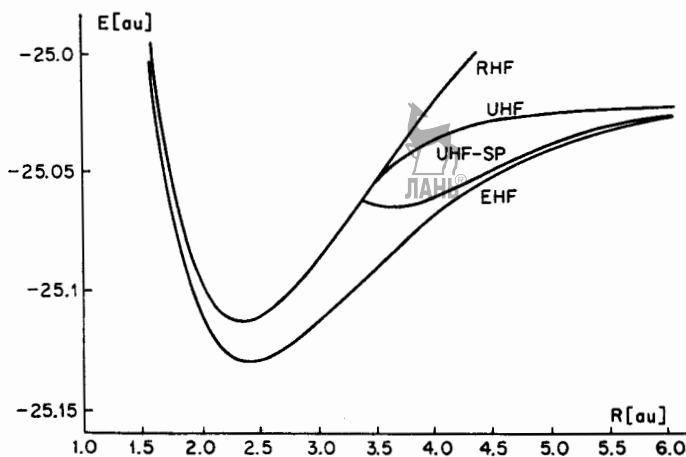
странство  $t_{2g}$ -орбиталей. Понятно, что за исключением случаев конфигураций  $d^0$ ,  $d^4$ ,  $d^6$  и  $d^{10}$  невозможно удовлетворить требованию «полного заполнения». Следовательно для них не подходит метод ХФР и в этих случаях приходится применять метод ОХФ для открытых оболочек (см. разд. 6 этой главы). — *Прим. ред.*

Эта волновая функция представляет собой детерминант, каждый столбец которого является суммой двух слагаемых; поэтому ее можно записать как сумму четырех детерминантов

$$\Psi = \frac{1}{2} \left\{ \hat{A}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{A}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{A}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{A}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right\}. \quad (6.75)$$

Первые два слагаемые этой суммы представляют собой ионные структуры, описывающие ситуацию, в которой оба электрона находятся возле одного и того же ядра. Третье и четвертое слагаемые соответствуют нейтральным атомам; они обычно называются «ковалентными структурами». (В этом контексте термин «ковалентный» означает просто «неионный».)

Так как энергия ионных слагаемых очень велика (два электрона на одном атоме испытывают сильное отталкивание), энергия ОХФ, полученная на больших расстояниях, значительно больше суммы энергий свободных нейтральных атомов. В таких системах метод ОХФ всегда проявляет «триплетную неустойчивость» на расстояниях, больших, чем «критическое», т. е. на достаточно больших межатомных расстояниях существует решение НХФ с энергией, меньшей, чем энергия ОХФ, описывающее нейтральные фрагменты, на которые диссоциирует молекула.



Потенциальные кривые молекулы  $H_2$ , полученные с помощью базисного набора DZ<sup>1</sup>.

<sup>1</sup> Воспроизведено по статье автора «Spin-Projected EHF Method. IV. Comparison of Potential Curves Given by Different One-Electron Methods», I. Mayer, Int. J. Quantum Chem. 14, 29 (1978); Copyright 1978 Wiley Periodicals.

Приведенный рисунок иллюстрирует характерное поведение потенциальных кривых, получающихся разными вариантами метода ССП. Метод ОХФ предсказывает положение минимума кривой (равновесное межатомное расстояние) с приемлемой точностью, однако переоценивает ее кривизну в минимуме примерно на 10-20%. (Колебательные частоты ОХФ, соответствующие растяжению связей, должны быть систематически «скорректированы в сторону уменьшения»<sup>1</sup>.) При больших межатомных расстояниях кривая ОХФ резко возрастает из-за наличия ионных вкладов в волновой функции. Метод НХФ обычно не дает решения, отличного от ОХФ, при равновесных или даже несколько больших расстояниях. Если продолжать увеличивать межатомное расстояние, то обнаружится «критическая точка» (точка бифуркации), в которой внезапно возникает новое, нетривиальное, решение НХФ, и кривая НХФ плавно отходит от кривой ОХФ. В критической точке решения ОХФ и НХФ совпадают. Так как градиент энергии ССП определяется самими орбиталями (ср. гл. 2, разд. 2.1 и разд. 7 в данной главе), первые производные кривых НХФ и ОХФ также совпадают в точке бифуркации. Так как волновая функция НХФ представляет новый тип решения, вторая производная кривой, получающейся в методе НХФ, отличается от аналогичной величины, полученной методом ОХФ. После расхождения с решением ОХФ, кривая НХФ стремится к качественно правильному диссоциационному пределу; обычно можно наблюдать, что на самом деле она приближается к асимптотическому значению энергии слишком быстро.

Можно рассматривать кривую НХФ, как совпадающую с кривой ОХФ на расстояниях меньше критического. Тогда кривая НХФ непрерывна и непрерывно дифференцируема, но в критической точке имеет разрывную вторую производную. (Как указывал Пулай, должен существовать разрыв какой-либо производной либо на ОХФ, либо на НХФ кривой, так как они совпадают при равновесном расстоянии и около него. Если бы обе кривые были бы всюду непрерывны с непрерывными производными всех порядков, тогда они имели бы один и тот же ряд Тейлора и совпадали бы всюду.)

Так как волновая функция ОХФ является частным случаем решения уравнений НХФ, (первая) вариация энергии, как функционала орбиталей, обращается в нуль, т. е. энергия стационарна. Вопрос о существовании нетривиальных решений НХФ можно изучать с помощью второй

<sup>1</sup> Существует метод, введенный Pulay, в котором расчет колебательных частот ведется с применением масштабирующего множителя. Частоты, получаемые в расчете, умножаются на некоторые коэффициенты, меньшие единицы.

вариации. Если она положительно определена, даже и при допущении возможности «орбитального расщепления», то решение ОХФ является истинным минимумом также и на множестве волновых функций НХФ. Если она не является положительно определенной, то можно найти решение НХФ с энергией, более низкой, чем решение ОХФ. Следовательно, критическая точка, в которой появляется нетривиальное решение НХФ, есть точка, в которой вторая вариация энергии становится положительно *полу*определенной. Если, как это обычно бывает на практике, для разложения орбиталей используется конечный базисный набор, то можно вычислить матрицу Гесса (матрицу вторых производных энергии, как функции орбитальных коэффициентов) и диагонализировать ее. Точка, в которой наименьшее собственное значение матрицы Гесса обращается в нуль, есть точка бифуркации, где появляется ССП решение нового типа. Это позволяет, по крайней мере в принципе, локализовать положение критической точки.

Волновая функция метода ОХФ для замкнутой оболочки, построенная из дважды занятых орбиталей, автоматически является чистым синглетом, т. е. собственной функцией операторов  $\hat{S}_z$  и  $\hat{S}^2$  с собственными значениями, равными 0. Это можно легко увидеть для  $\hat{S}_z = \sum_{i=1}^N \hat{s}_z(i)$ , так как спин-функции  $\alpha(i)$  и  $\beta(i)$  являются собственными функциями оператора  $\hat{s}_z(i)$  с собственными значениями  $+\frac{1}{2}$  и  $-\frac{1}{2}$ , соответственно. Так как числа заполнения спин-функций  $\alpha$  и  $\beta$  в детерминанте ОХФ с замкнутой оболочкой равны, последний является собственной функцией  $\hat{S}_z$  с собственным значением 0.

Детерминантная волновая функция ОХФ с замкнутой оболочкой является собственной функцией и для операторов  $\hat{S}_x = \sum_{i=1}^N \hat{s}_x(i)$  и  $\hat{S}_y = \sum_{i=1}^N \hat{s}_y(i)$ , тоже с нулевыми собственными значениями. (Случай  $S_x = S_y = S_z = 0$  исключительный в том смысле, что все три компоненты спинового углового момента можно измерить одновременно.) Чтобы убедиться в том, что детерминант замкнутой оболочки является собственной функцией  $\hat{S}_x$  и  $\hat{S}_y$ , сначала необходимо заметить, что  $\hat{S}_x$  и  $\hat{S}_y$  симметричны по отношению к перестановкам электронов и потому коммутируют с оператором антисимметризации. Так, например, для  $\hat{S}_x$ ,

обозначив число двукратно занятых орбиталей как  $n$  ( $N = 2n$ ), имеем

$$\begin{aligned}
 \hat{S}_x \Psi &= \sum_{i=1}^N \hat{s}_x(i) \hat{A}[\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2)\dots \\
 &\quad \times \psi_n(2n-1)\alpha(2n-1)\varphi_n(2n)\beta(2n)] \\
 &= \sum_{i=1}^N \hat{A}\{\hat{s}_x(i)[\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2)\dots \\
 &\quad \times \psi_n(2n-1)\alpha(2n-1)\varphi_n(2n)\beta(2n)]\}.
 \end{aligned} \tag{6.76}$$

Далее, в соответствии с определениями (5.6) и (5.7), оператор  $\hat{s}_x(i)$  «переворачивает» спин-функцию  $\gamma(i)$ , т. е. если она была  $\alpha$ , то становится  $\frac{1}{2}\beta$ , а если была  $\beta$ , то становится  $\frac{1}{2}\alpha$ . Это действие можно обозначить, как  $\hat{s}_x(i)\gamma(i) = \frac{1}{2}\bar{\gamma}(i)$ , где  $\bar{\gamma}(i)$  является спин-функцией, направленной противоположно  $\gamma(i)$ . Вначале в  $\Psi$  мы имеем пару одинаковых пространственных орбиталей, занятых противоположными спинами  $\gamma(i)$  и  $\bar{\gamma}(i)$ ; после применения  $\hat{s}_x(i)$  одна из них становится дважды занятой спином  $\bar{\gamma}(i)$ . Тогда каждый  $\hat{s}_x(i)$  дает детерминант с двумя идентичными столбцами, который обращается в нуль в соответствии со свойствами определителей. Поэтому детерминант ОХФ с замкнутой оболочкой  $\Psi$  является собственной функцией  $\hat{S}_x$  с нулевым собственным значением. Аналогичные рассуждения верны и для оператора  $\hat{S}_y$ . Таким образом,  $\Psi$  является также собственной функцией  $\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$  с нулевым собственным значением. Ч. т. д.

Однодетерминантная волновая функция НХФ-РОРС является собственной функцией оператора  $\hat{S}_z$ , но не  $\hat{S}^2$ . Это значит, что ей нельзя приписать определенную спиновую мультиплетность и она представляет собой смесь синглетных, триплетных и т. д. компонент. (Часто, когда волновая функция содержит нежелательные спиновые компоненты, пользуются термином «спиновое загрязнение».)

Чтобы решить эту проблему, Лёвдин ввел «оператор спиновой проекции»

$$\hat{O}^S = \prod_{l \neq S} \frac{\hat{S}^2 - l(l+1)}{S(S+1) - l(l+1)}, \tag{6.77}$$

позволяющий получать чистую спиновую компоненту волновой функции, принадлежащую искомому собственному значению  $S(S+1)$  оператора  $\hat{S}^2$ . (В согласии с общими правилами для момента импульса в

квантовой механике, выбранное состояние имеет спиновую мультиплетность  $2S + 1$ .) Оператор спиновой проекции уничтожает все нежелательные спиновые компоненты, оставляя неизменной только искомую. В самом деле, любую волновую функцию  $\Psi$  можно представить в виде суммы  $\Psi = \sum_l {}^l\Psi$  слагаемых  ${}^l\Psi$  с разными спиновыми мультиплетностями  $2l + 1$ . Сомножитель  $\hat{S}^2 - l(l + 1)$  в операторе  $\hat{O}^S$  уничтожает компоненту  ${}^l\Psi$ :

$$[\hat{S}^2 - l(l + 1)] {}^l\Psi = l(l + 1) {}^l\Psi - l(l + 1) {}^l\Psi = 0. \quad (6.78)$$

Под действием произведения таких сомножителей все компоненты  ${}^l\Psi$  с  $l \neq S$  уничтожаются одна за другой. В том же случае, если оператор  $\hat{S}^2 - l(l + 1)$  действует на слагаемое  ${}^S\Psi$  с желаемой мультиплетностью  $2S + 1$ , получим

$$[\hat{S}^2 - l(l + 1)] {}^S\Psi = [S(S + 1) - l(l + 1)] {}^S\Psi. \quad (6.79)$$

Знаменатели введены в определении (6.77) для компенсации этого коэффициента; поэтому компонента  ${}^S\Psi$  остается неизменной, в то время как мы избавляемся от всех других компонент  ${}^l\Psi$ .

Волновая функция, получающаяся спиновым проектированием детерминанта РОРС, больше не является однодетерминантной, что затрудняет использование этого подхода. Подвергая детерминант НХФ спиновому проектированию, не только получают чистую спиновую функцию, но обычно также достигают понижения энергии по сравнению с энергией НХФ. Это происходит, разумеется, только в тех точках, где существует настоящее решение НХФ, отличающееся от волновой функции ОХФ, соответствующей чистому синглету. Представленный выше рис. (см. с. 228) иллюстрирует, что это может приводить к совершенно патологическим потенциальным кривым. Такое поведение является следствием того, что при этом расчете вариационным путем оптимизируется не чистая по спину компонента волновой функции, полученная спиновым проектированием, а однодетерминантная волновая функция, которая была у нас *перед* этим («метод НХФ с последующей спиновой проекцией»). Намного более последовательным было бы рассмотрение самого спин-спроектированного детерминанта в качестве вариационной пробной волновой функции («вариация после проектирования»). Эта процедура приводит к спин-спроектированному «расширенному» методу Хартри—Фока (РХФ, extended-EHF); как показывает рис. на с. 228, метод РХФ способен дать хорошо сбалансированное описание гомолитической диссоциации *одной* химической связи.



Особым свойством спин-спроектированных волновых функций является то, что они хорошо подходят для описания электронной структуры антиферромагнитных и антиароматических систем, которую можно представить в виде совокупности спинов, направленных то вверх, то вниз, и образующих при этом две подрешетки наивысшего возможного спина, которые потом связываются в результирующий синглет.

Волновая функция ОХФ снова представляет собой частное решение уравнений РХФ, но, в отличие от случая НХФ, решение РХФ с энергией, меньшей, чем энергия ОХФ, существует всюду. Метод РХФ все еще можно рассматривать, как обобщенную одноэлектронную модель, так как исходный определитель РОРС содержит столько же спин-орбиталей, сколько имеется электронов в системе. Его существенный недостаток, помимо сложных уравнений и больших вычислительных затрат, состоит в том, что он размерно не согласован и поэтому не может давать столь же хорошее описание системы с несколькими разрывающимися химическими связями, как для одной связи.

Так как схема РХФ оказалась очень сложной, были предложены два ее упрощенных варианта; работы, использующие их, еще и сейчас можно встретить в литературе. Одним из них является метод «простой аннигиляции», который использует только один сомножитель в произведении (6.77), соответствующий наинизшей и, в большинстве случаев, наибольшей по весу нежелательной спиновой компоненте (обычно триплету). В другом, так называемом «полуспроектированном методе Хартри—Фока», волновая функция представляется суммой или разностью всего лишь двух определителей, которые отличаются друг от друга перестановкой всех спин-функций  $\alpha$  и  $\beta$ . Она содержит только каждую вторую спиновую компоненту (например, синглет, квинтет и т. д., но не триплет и септет и т. д.) и обычно представляет собой почти чистое спиновое состояние.

Трудности, связанные со спиновыми собственными функциями, представляют собой практически наиболее важный, но не единственный, пример противоречия, которое Лёвдин назвал «дилеммой симметрии». Он указал на то, что точные решения уравнения Шрёдингера автоматически являются собственными функциями всех операторов, коммутирующих с гамильтонианом, т. е. автоматически приведены по симметрии — например, являются чистыми спиновыми собственными функциями. (Исключениями являются некоторые частные случаи с вырождением, как, например, синглетная и триплетная волновые функции системы  $\text{H}_2$  при бесконечном межъядерном расстоянии. Однако, даже в таком случае собственные функции можно выбрать приведенными по симметрии.) В то же время, свойство «автоматической симметрии» не

распространяется на приближенные волновые функции, определенные с помощью вариационного принципа. Наложение требований симметрии уменьшает гибкость пробной волновой функции, используемой в вариационной процедуре, что может приводить к большим значениям энергии, по сравнению с волновыми функциями без ограничений по симметрии. Вследствие этого, мы сталкиваемся с дилеммой: приходится выбирать между волновой функцией с более низкой энергией и волновой функцией, обладающей правильной симметрией.

#### 4.4. Синглетные и триплетные возбуждения

Рассуждения, касающиеся энергии однократно возбужденного детерминанта, приведенные в разд. 3.1, верны также и в случае метода ОХФ. Однако, так как основное состояние ОХФ является чистым синглетом, имеет смысл обсудить, как можно получить чистые синглеты и триплеты в базисе однократно возбужденных относительно волновой функции ОХФ состояний.

Если рассматривать возбуждение одного электрона с дважды занятой пространственной орбитали  $\varphi_i$  на виртуальную пространственную орбиталь  $\varphi_v$  (без изменения проекции полного спина  $S_z$ ), то можно построить два различных однократно возбужденных детерминанта в зависимости от того, проводится ли возбуждение электрона с проекцией спина  $\alpha$  или  $\beta$ :

$$\begin{aligned}\Psi_1^{iv,\alpha} &= \Psi_1(\varphi_i\alpha \rightarrow \varphi_v\alpha) \\ &= \hat{A}[\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2) \dots \varphi_v(2i-1)\alpha(2i-1)\varphi_i(2i)\beta(2i) \dots]\end{aligned}\quad (6.80)$$

$$\begin{aligned}\Psi_1^{iv,\beta} &= \Psi_1(\varphi_i\beta \rightarrow \varphi_v\beta) \\ &= \hat{A}[\varphi_1(1)\alpha(1)\varphi_1(2)\beta(2) \dots \varphi_i(2i-1)\alpha(2i-1)\varphi_v(2i)\beta(2i) \dots].\end{aligned}$$

Ни тот, ни другой детерминант не являются чистыми спиновыми функциями, но их нормированная сумма

$$^1\Psi = \frac{1}{\sqrt{2}}(\Psi_1^{iv,\alpha} + \Psi_1^{iv,\beta}) \quad (6.81)$$

соответствует синглету, а разность

$$^3\Psi = \frac{1}{\sqrt{2}}(\Psi_1^{iv,\alpha} - \Psi_1^{iv,\beta}) \quad (6.82)$$

— триpletу. Эти формулы очень похожи на уравнение (5.19), так как дважды занятые орбитали образуют синглетные оболочки, которые не

нужно явно рассматривать при обсуждении спиновых функций. Здесь уместны, однако, некоторые небольшие замечания. Несколько необычно то, что в формулах (6.81) и (6.82) присутствуют знаки «+» в синглете и «-» в триплете. Это так, потому что два детерминанта (6.80) имеют одинаковый порядок спин-функций, а не одинаковый порядок пространственных орбиталей, как было в формуле (5.19). Если изменить порядок орбиталей в  $\Psi_1^{iv,\beta}$ , то знак изменится и для  ${}^1\Psi$  получится более привычный вид  $\dots \varphi_v(2i-1)\varphi_i(2i)\frac{1}{\sqrt{2}}[\alpha(2i-1)\beta(2i) - \beta(2i-1)\alpha(2i)]\dots$ , и аналогично спиновый множитель  $\frac{1}{\sqrt{2}}(\alpha\beta + \beta\alpha)$  для  ${}^3\Psi$ . Следует отметить, что компоненты  $S_z = \pm 1$  триплета можно записать в виде одного детерминанта — ср. уравнение (5.18). Напомним также, что множитель  $\frac{1}{2}$  в уравнении (5.19) содержит также множитель  $\frac{1}{\sqrt{2}!}$ , появляющийся из определения оператора антисимметризации.

Теперь вычислим энергию с этими волновыми функциями:

$$E_{S,T} = \langle \frac{1}{\sqrt{2}}(\Psi_1^{iv,\alpha} \pm \Psi_1^{iv,\beta}) | \hat{H} | \frac{1}{\sqrt{2}}(\Psi_1^{iv,\alpha} \pm \Psi_1^{iv,\beta}) \rangle. \quad (6.83)$$

Верхние знаки соответствуют синглету, нижние — триплету.

Расписывая выражение (6.83), получим

$$E_{S,T} = \frac{1}{2} \left( \langle \Psi_1^{iv,\alpha} | \hat{H} | \Psi_1^{iv,\alpha} \rangle + \langle \Psi_1^{iv,\beta} | \hat{H} | \Psi_1^{iv,\beta} \rangle \right. \\ \left. \pm \langle \Psi_1^{iv,\beta} | \hat{H} | \Psi_1^{iv,\alpha} \rangle \pm \langle \Psi_1^{iv,\alpha} | \hat{H} | \Psi_1^{iv,\beta} \rangle \right). \quad (6.84)$$

Первые два слагаемых равны: они отличаются друг от друга только перестановкой спиновых индексов  $\alpha$  и  $\beta$ . Каждое из них представляет собой энергию однократно возбужденного детерминанта  $\Psi_1^{iv,\alpha}$  или  $\Psi_1^{iv,\beta}$ . Согласно обсуждению, приведенному в разд. 3.1, энергия такого детерминанта отличается от энергии хартри-фоковского основного состояния  $E_0$  на разность орбитальных энергий минус поправка на самоотталкивание  $\hat{J}_i - \hat{K}_i$ , присутствующее в операторе Фока:

$$\langle \Psi_1^{iv,\alpha} | \hat{H} | \Psi_1^{iv,\alpha} \rangle = \langle \Psi_1^{iv,\beta} | \hat{H} | \Psi_1^{iv,\beta} \rangle = E_0 + \varepsilon_v - \varepsilon_i - \langle \varphi_v | \hat{J}_i - \hat{K}_i | \varphi_v \rangle \\ = E_0 + \varepsilon_v - \varepsilon_i - J_{iv} + K_{iv}. \quad (6.85)$$

Здесь  $J_{iv} = [\varphi_i \varphi_v | \varphi_i \varphi_v]$  и  $K_{iv} = [\varphi_i \varphi_v | \varphi_v \varphi_i]$  — кулоновский и обменный интегралы для пары орбиталей  $\varphi_i$  и  $\varphi_v$ .

Третье и четвертое слагаемые в выражении (6.84) комплексно сопряжены друг другу и равны между собой, если использовать вещественные

орбитали. Слагаемое  $\langle \Psi_1^{iv,\alpha} | \hat{H} | \Psi_1^{iv,\beta} \rangle$  является матричным элементом гамильтониана между детерминантами, построенными из ортонормированных орбиталей и отличающимися двумя спин-орбиталями: «бра» содержит  $\varphi_v \alpha \varphi_i \beta$ , а «кет» —  $\varphi_i \alpha \varphi_v \beta$ .

Используем правила Слэтера (гл. 5, разд. 6.3.2),

$$\begin{aligned} \langle \Psi_1^{iv,\alpha} | \hat{H} | \Psi_1^{iv,\beta} \rangle = & \langle \varphi_v(1)\alpha(1)\varphi_i(2)\beta(2) | \frac{1}{r_{12}} | \varphi_i(1)\alpha(1)\varphi_v(2)\beta(2) \rangle \\ & - \langle \varphi_v(1)\alpha(1)\varphi_i(2)\beta(2) | \frac{1}{r_{12}} | \varphi_v(1)\beta(1)\varphi_i(2)\alpha(2) \rangle. \end{aligned} \quad (6.86)$$

Второй интеграл равен нулю вследствие ортогональности спиновых функций  $\alpha$  и  $\beta$ . Проводя суммирование по спинам в первом интеграле, получим

$$\langle \Psi_1^{iv,\alpha} | \hat{H} | \Psi_1^{iv,\beta} \rangle = [\varphi_v \varphi_i | \varphi_i \varphi_v] = K_{iv}. \quad (6.87)$$

Подставим результат (6.85) и (6.87) в выражение (6.84) и получим

$$E_{S,T} = \frac{1}{2} [2(E_0 + \varepsilon_v - \varepsilon_i - J_{iv} + K_{iv}) \pm 2K_{iv}], \quad (6.88)$$

т. е.

$$\begin{aligned} E_S &= E_0 + \varepsilon_v - \varepsilon_i - J_{iv} + 2K_{iv} \\ E_T &= E_0 + \varepsilon_v - \varepsilon_i - J_{iv}. \end{aligned} \quad (6.89)$$

В детерминантах  $\Psi_1^{iv,\alpha}$  и  $\Psi_1^{iv,\beta}$ , согласно определениям (6.81) и (6.82), поровну представлены синглетное  $^1\Psi$  и триплетное  $^3\Psi$  состояния. В соответствии с этим, энергия (6.85) представляет собой среднее арифметическое от энергий синглета и триплета (6.89)<sup>1</sup>.

Обменный интеграл  $K_{iv} > 0$ , так как он представляет собой электростатическое самоотталкивание обменной зарядовой плотности  $\varphi_i(\vec{r})\varphi_v(\vec{r})$ . Тогда из результата (6.89) получаем  $E_S > E_T$ , в полном согласии с общим наблюдением, что для молекул с замкнутой оболочкой первое триплетное возбужденное состояние обычно лежит ниже по энергии, чем первое возбужденное синглетное состояние.

Конечно, предпочтительнее не ограничиваться рассмотрением только отдельных возбуждений с какой-то одной занятой орбитали на виртуальную, а провести вариационный расчет (расчет по методу «конфигурационного взаимодействия», КВ) с учетом всех однократно возбужденных детерминантов, как отмечалось в разд. 3.1 в связи со схемой НХФ. Если в качестве функции основного состояния используется

<sup>1</sup> Так получается из-за того, что недиагональный матричный элемент оператора Гамильтона между состояниями разного полного спина равен нулю. — Прим. ред.

волновая функция ОХФ, то решения автоматически будут либо синглетами, либо триплетными (за исключением, возможно, «патологических» вырожденных случаев). В качестве альтернативы можно использовать чистые синглетные и триплетные конфигурации, подобные тем, которые обсуждались выше. Тогда расчет матричных элементов усложняется, но нужно диагонализировать матрицы меньшего размера — одну для синглетных, а другую для триплетных состояний.

## 5. Теория Хартри—Фока в конечном базисе

### 5.1. Уравнения Хартри—Фока—Рутана

Уравнения Хартри—Фока представляют собой систему связанных интегро-дифференциальных уравнений. Их можно решать численными методами в случае свободных атомов, а также, возможно, для систем, подобных  $\text{H}_2$ . Когда рассматриваются молекулы, решения уравнений, т. е. одноэлектронные орбитали, обычно аппроксимируют линейной комбинацией некоторых заранее определенных базисных функций (базисных орбиталей). Этот подход очень похож на тот, который был рассмотрен в гл. 3, но здесь задача является нелинейной. Базисные орбитали обычно привязаны к отдельным атомам молекулы; раньше их часто брали из расчетов свободных атомов, что объясняет, почему сокращение ЛКАО (линейная комбинация атомных орбиталей) часто используется для обозначения этого подхода.

Таким образом, пользуемся разложением одноэлектронных орбиталей  $\varphi_i$  в виде

$$\varphi_i = \sum_{\mu=1}^m c_{\mu}^i \chi_{\mu}; \quad i = 1, 2, \dots, m, \quad (6.90)$$

где  $m$  орбиталей  $\chi_{\mu}$  ( $\mu = 1, 2, \dots, m$ ) образуют базис АО, а  $c_{\mu}^i$  есть коэффициент  $\mu$ -й базисной орбитали в разложении  $i$ -й молекулярной орбитали (МО)  $\varphi_i$ . Коэффициенты каждой орбитали  $\varphi_i$  можно собрать в вектор-столбец  $\mathbf{c}^i$  с элементами  $c_{\mu}^i$ . Базисные функции  $\chi_{\mu}$  обычно (но не обязательно) нормированы на единицу; как правило, они не предполагаются ортогональными друг к другу, но, разумеется, должны быть линейно независимыми. (На практике может получиться так, что базисные орбитали почти линейно зависимы. В таком случае можно использовать каноническую ортогонализацию по Лёвдину, описанную в гл. 3, разд. 3.2, чтобы уменьшить размерность матрицы в задаче диагонализации.)

Если используется ограниченный базис, то интегродифференциальные уравнения задачи на псевдособственные значения метода ХФ заменяются обобщенными матричными уравнениями на псевдособственные значения. Первоначально они были выведены Рутаном (Rutan, Roothaan<sup>1</sup>), а также независимо Холлом (Hall) в 1951 г., и обычно называются уравнениями Хартри—Фока—Рутана (ХФР). Рутан явно выписал полную энергию детерминантной волновой функции, как функцию, зависящую от орбитальных коэффициентов  $c_\mu^i$ , и проделал дифференцирования, необходимые для нахождения условия стационарности энергии при дополнительном условии ортонормировки орбиталей. Вместо этого (несколько громоздкого) вывода мы снова воспользуемся теоремой Бриллюэна (разд. 1), эквивалентной вариационному принципу для однодетерминантных волновых функций:

$$\langle \Psi_1(\varphi_i \rightarrow \varphi_i') | \hat{H} | \Psi \rangle = 0 ; \quad i = 1, 2, \dots, n . \quad (6.91)$$

Здесь  $\langle \varphi_i' | \varphi_j \rangle = 0$ ,  $j = 1, 2, \dots, n$ , т. е.  $\varphi_i'$  ортогональна всем занятым орбиталям, но в *остальном произвольна*. (Для простоты мы рассмотрим случай ОХФ, так как обобщение на случай НХФ тривиально).

Единственное принципиальное отличие от вывода, приведенного в разд. 2.1, заключается в том, что теперь все орбитали имеют вид ЛКАО (6.90), и это свойство должно сохраняться при варьировании. Последнее требование означает, что функция  $\varphi_i'$  также должна представляться в виде разложения по исходным базисным орбиталям  $\chi_\mu$ :

$$\varphi_i' = \sum_{\mu=1}^m q_\mu \chi_\mu . \quad (6.92)$$

Коэффициенты  $q_\mu$  не вполне произвольны, так как на них накладываются условия ортогональности  $\langle \varphi_i' | \varphi_j \rangle = 0$ . Мы можем построить наиболее общую произвольную орбиталь вида (6.92), удовлетворяющую условию ортогональности, следующим образом. Рассмотрим оператор проектирования на подпространство занятых орбиталей:

$$\hat{P} = \sum_{j=1}^m |\varphi_j\rangle \langle \varphi_j| . \quad (6.93)$$

Тогда  $1 - \hat{P}$  является оператором, проектирующим на ортогональное дополнение к этому подпространству, т. е. любая функция  $|\varphi_i'\rangle = (1 -$

<sup>1</sup> См. примечание на с. 216.

$\hat{P})|\sum_{\nu} p_{\nu}\chi_{\nu}\rangle$  с произвольными коэффициентами  $p_{\nu}$  ортогональна ко всем занятым орбиталям. Коэффициенты  $p_{\nu}$  уже совершенно произвольны и независимы друг от друга. Например, мы можем предположить, что для каждой вариации только один из них отличен от нуля. Поэтому необходимо потребовать выполнения теоремы Бриллюэна для каждой функции вида

$$|\varphi'_i\rangle = (1 - \hat{P})|\chi_{\nu}\rangle ; \quad \nu = 1, 2, \dots, m . \quad (6.94)$$

Выражение теоремы Бриллюэна через одноэлектронные орбитали снова приводит к равенству:

$$\langle \varphi'_i | \hat{F} | \varphi_i \rangle = 0 , \quad (6.95)$$

где смысл оператора Фока  $\hat{F}$  тот же, что и ранее, и сейчас он, разумеется, строится из орбиталей, которые можно разложить по данному одноэлектронному базису. Рассмотрев функцию, сопряженную к  $|\varphi'_i\rangle$  (6.94), получим

$$\langle \varphi'_i | = [|\varphi'_i\rangle]^{\dagger} = [(1 - \hat{P})|\chi_{\nu}\rangle]^{\dagger} = \langle \chi_{\nu} | (1 - \hat{P}) . \quad (6.96)$$

Подставим выражение (6.96) в уравнение (6.95):

$$\langle \chi_{\nu} | (1 - \hat{P}) \hat{F} | \varphi_i \rangle = 0 , \quad (6.97)$$

т. е.

$$\langle \chi_{\nu} | \hat{F} | \varphi_i \rangle - \langle \chi_{\nu} | \sum_{j=1}^m |\varphi_j\rangle \underbrace{\langle \varphi_j | \hat{F} | \varphi_i \rangle}_{\varepsilon_{ji}} = 0 . \quad (6.98)$$

Снова можно считать, что орбитали уже были подвергнуты унитарному преобразованию, диагонализующему матрицу  $\varepsilon$ , так что  $\varepsilon_{ji} = \varepsilon_i \delta_{ij}$ . Это приводит к выражению

$$\langle \chi_{\nu} | \hat{F} | \varphi_i \rangle = \varepsilon_i \langle \chi_{\nu} | \varphi_i \rangle . \quad (6.99)$$

Подставляя разложение ЛКАО (6.90) орбитали  $\varphi_i$  и вводя обозначение  $F_{\nu\mu} = \langle \chi_{\nu} | \hat{F} | \chi_{\mu} \rangle$  и  $S_{\nu\mu} = \langle \chi_{\nu} | \chi_{\mu} \rangle$ , получаем уравнения ХФР:

$$\sum_{\mu=1}^m F_{\nu\mu} c_{\mu}^i = \varepsilon_i \sum_{\mu=1}^m S_{\nu\mu} c_{\mu}^i \quad i = 1, 2, \dots, m . \quad (6.100)$$

Если рассматривать  $F_{\nu\mu}$  и  $S_{\nu\mu}$ , как элементы  $m \times m$  матрицы Фока  $\mathbf{F}$  и матрицы перекрывания  $\mathbf{S}$ , соответственно, уравнение (6.100) можно переписать в матричном виде:

$$\mathbf{F}\mathbf{c}^i = \varepsilon_i \mathbf{S}\mathbf{c}^i. \quad (6.101)$$

Если ввести матрицу  $\mathbf{C}$ , столбцами которой являются векторы  $\mathbf{c}^i$ , то уравнение (6.101) можно также переписать как

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}, \quad (6.102)$$

где  $\boldsymbol{\varepsilon} = \text{diag}(\varepsilon_i)$  является диагональной матрицей орбитальных энергий. Обычно уравнение (6.102) записывают, включая в матрицу  $\mathbf{C}$  коэффициенты как занятых, так и виртуальных орбиталей; тогда и  $\mathbf{C}$ , и  $\boldsymbol{\varepsilon}$  оказываются матрицами размера  $m \times m$ . (Можно ограничиться рассмотрением занятых орбиталей; тогда  $\mathbf{C}$  является прямоугольной матрицей размера  $m \times n$ , а  $\boldsymbol{\varepsilon}$  представляет собой диагональную квадратную матрицу размера  $n \times n$ .)

Если бы базис был бесконечным и полным, то переход от интегро-дифференциальной формы к матричной потребовал бы только разложения двух функций (двух частей уравнения  $\hat{F}\varphi_i = \varepsilon_i\varphi_i$ ) по одному и тому же полному базису. *A priori* не очевидно, что в случае конечного базиса условия стационарности энергии получаются просто «обрыванием» этой бесконечной системы уравнений, и сохранением только первых  $m$  уравнений, соответствующих фактически используемому базису. Однако на практике это всегда имеет место, даже при рассмотрении более сложных уравнений ССП, чем уравнения ХФР, хотя автору неизвестно общее доказательство этого утверждения<sup>1</sup>.

Полезно кратко рассмотреть упомянутый подход, так как он позволяет очень просто получить уравнения ХФР, даже если это и *не может рассматриваться, как настоящий вывод*.

Подставим разложение (6.90) в уравнение ХФ  $\hat{F}\varphi_i = \varepsilon_i\varphi_i$ :

$$\hat{F} \sum_{\mu=1}^m c_{\mu}^i \chi_{\mu} = \varepsilon_i \sum_{\mu=1}^m c_{\mu}^i \chi_{\mu}, \quad (6.103)$$

<sup>1</sup> Редактору кажется, что это совсем не так, в том смысле, что доказать это нельзя, поскольку непонятно, куда подеваются матричные элементы оператора Фока между состояниями, включенными в конечный базисный набор, и тем бесконечным набором состояний, которые не войдут в базис. (Очевидно, они не влияют на энергию, полученную с применением конечного базиса. — *Из переписки с авт.*). — Прим. ред.



умножим на  $\chi_\nu^*$  и проинтегрируем:

$$\sum_{\mu=1}^m \langle \chi_\nu | \hat{F} | \chi_\mu \rangle c_\mu^i = \varepsilon_i \sum_{\mu=1}^m \langle \chi_\nu | \chi_\mu \rangle c_\mu^i, \quad (6.104)$$

т. е.

$$\sum_{\mu=1}^m F_{\nu\mu} c_\mu^i = \varepsilon_i \sum_{\mu=1}^m S_{\nu\mu} c_\mu^i, \quad (6.105)$$

или в матричных обозначениях:

$$\mathbf{F} \mathbf{c}^i = \varepsilon_i \mathbf{S} \mathbf{c}^i. \quad (6.106)$$

Когда базисный набор увеличивается и становится все ближе к полному, то решения уравнений ХФР (орбитали, орбитальные энергии и полная энергия) стремятся к так называемому «хартри-фоковскому пределу», который можно было бы получить точным решением интегродифференциальных уравнений ХФ. Расчеты, проводимые с использованием конечных, но весьма больших, базисов, часто называются «почти хартри-фоковскими».

В практических расчетах обычно используют стандартные наборы базисных орбиталей, приписываемых атомам системы; число базисных функций, используемых для каждого атома может меняться в широких пределах. Соответственно, точность расчета (даже в рамках модели Хартри—Фока) разных *ab initio* методов также может очень сильно меняться. Мы подчеркиваем, что термин *ab initio* означает только, что не используются никакие эмпирические параметры<sup>1</sup> — за исключением значений универсальных постоянных, — и все интегралы, встречающиеся в расчете, вычисляются с необходимой точностью<sup>2</sup>. В настоящее время, вероятно, наименьшим и наибольшим «стандартными» базами для молекулы  $\text{H}_2$  являются те, которые обозначают сокращениями «STO-3G» и «aug-ss-pV5Z», состоящими только из двух («минимальный базис») и 160 базисных функций<sup>3</sup>, и дающими при одном и том же межатомном расстоянии 0.74 Å энергию ССП, составляющую  $-1.116759$

<sup>1</sup> В смысле параметров, взятых из эксперимента.

<sup>2</sup> На самом деле параметрами *ab initio* являются параметры базисных функций, наборы которых приписываются каждому атому; неудобочитаемые сокращения (типа «aug-ss...») представляют собой указания на развернутые наборы таких параметров для большего или меньшего числа атомов. — *Прим. ред.*

<sup>3</sup> На самом деле это представлялось недооцененным даже во время написания английского оригинала книги. Теперь используются намного бóльшие базисные наборы, что имеет смысл при учете электронной корреляции.

и  $-1.133619$  а. е., соответственно. (Последнее значение можно рассматривать, как хартри-фоковский предел; численные расчеты по методу Хартри—Фока вряд ли смогут достичь лучшей точности.)

## 5.2. Матрица $\mathbf{P}$

Однодетерминантная волновая функция однозначно определяется (с точностью до физически несущественного постоянного множителя) подпространством занятых орбиталей, которое можно задать, определив оператор проекции на него. Здесь мы рассмотрим этот момент подробно для случая, когда орбитали представлены в виде разложения по конечному базису.

Возьмем занятые орбитали  $\varphi_i$ ,  $i = 1, 2, \dots, n$  задачи ОХФ и выразим оператор  $\hat{P}$  проекции на занятое подпространство через коэффициенты ЛКАО, использованные в разложении (6.90):

$$\begin{aligned}\hat{P} &= \sum_{i=1}^n |\varphi_i\rangle\langle\varphi_i| = \sum_{i=1}^n \left| \sum_{\mu=1}^m c_{\mu}^i \chi_{\mu} \right\rangle \left\langle \sum_{\nu=1}^m c_{\nu}^i \chi_{\nu} \right| \\ &= \sum_{\mu,\nu=1}^m \sum_{i=1}^n c_{\mu}^i c_{\nu}^{i*} |\chi_{\mu}\rangle\langle\chi_{\nu}| = \sum_{\mu,\nu=1}^m P_{\mu\nu} |\chi_{\mu}\rangle\langle\chi_{\nu}|,\end{aligned}\tag{6.107}$$

где мы ввели « $P$ -матрицу» с элементами

$$P_{\mu\nu} = \sum_{i=1}^n c_{\mu}^i c_{\nu}^{i*}.\tag{6.108}$$

В матричных обозначениях матрицу  $\mathbf{P}$  можно выразить в виде суммы диадных произведений:

$$\mathbf{P} = \sum_{i=1}^n \mathbf{c}^i \mathbf{c}^{i\dagger}.\tag{6.109}$$

Эрмитова матрица  $\mathbf{P}$  играет важную роль в теории. Обычно ее также называют «матрицей плотности», хотя эту терминологию нельзя признать вполне удачной, если используется неортогональный базис.

В случае теории ОХФ для замкнутой оболочки часто используют матрицу  $\mathbf{P}$ , умноженную на двойку; обозначим ее как

$$\mathbf{D} = 2\mathbf{P} = 2 \sum_{i=1}^n \mathbf{c}^i \mathbf{c}^{i\dagger}.\tag{6.110}$$

Хотя, как мы видели, матрица  $\mathbf{P}$  непосредственно связана с оператором проекции на подпространство занятых орбиталей, она является истинной проектирующей матрицей только в том случае, если базис  $\{\chi_{\mu}\}$

ортонормирован. В общем случае, когда  $S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle \neq \delta_{\mu\nu}$ , матрица  $\mathbf{P}$  эрмитова, но не идемпотентна, и тогда правильным свойством идемпотентности обладает матрица  $\mathbf{PS}$ , т. е.  $(\mathbf{PS})^2 = \mathbf{PS}$ . В самом деле,

$$\begin{aligned} (\mathbf{PS})^2 &= \mathbf{PSPS} = \sum_{i=1}^n \mathbf{c}^i \mathbf{c}^{i\dagger} \mathbf{S} \sum_{j=1}^n \mathbf{c}^j \mathbf{c}^{j\dagger} \mathbf{S} \\ &= \sum_{i,j=1}^n \mathbf{c}^i \underbrace{\mathbf{c}^{i\dagger} \mathbf{S} \mathbf{c}^j}_{\delta_{ij}} \mathbf{c}^{j\dagger} \mathbf{S} = \sum_{i=1}^n \mathbf{c}^i \mathbf{c}^{i\dagger} \mathbf{S} = \mathbf{PS}. \quad \text{Ч. т. д.} \end{aligned} \quad (6.111)$$

В цепочке равенств (6.111) мы воспользовались тем, что занятые орбитали ортонормированы, т. е.  $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$ :

$$\langle \varphi_i | \varphi_j \rangle = \left\langle \sum_{\mu=1}^m c_\mu^i \chi_\mu \middle| \sum_{\nu=1}^m c_\nu^j \chi_\nu \right\rangle = \sum_{\mu,\nu=1}^m c_\mu^{i*} S_{\mu\nu} c_\nu^j = \mathbf{c}^{i\dagger} \mathbf{S} \mathbf{c}^j = \delta_{ij}. \quad (6.112)$$

Часто свойство идемпотентности (6.111) выражают<sup>1</sup> в форме  $\mathbf{PSP} = \mathbf{P}$ .

Если применяется конечный базис, то одноэлектронные орбитали представляются вектор-столбцами, составленными из коэффициентов их разложения по этому базису. В этой схеме матрица  $\mathbf{PS}$  есть матрица проекции на подпространство занятых орбиталей, в согласии с результатом (6.111). В самом деле, пусть  $|\psi\rangle = \sum_{\varrho=1}^m q_\varrho |\chi_\varrho\rangle$  — произвольная одноэлектронная функция: найдем коэффициенты ЛКАО  $p_\mu$  орбитали  $\hat{P}|\psi\rangle = \sum_{\mu=1}^m p_\mu |\chi_\mu\rangle$  с помощью разложения (6.107) ( $\hat{P}$  — оператор проекции на подпространство):

$$\begin{aligned} \sum_{\mu=1}^m p_\mu |\chi_\mu\rangle &\equiv \hat{P}|\psi\rangle = \sum_{\mu,\nu=1}^m P_{\mu\nu} |\chi_\mu\rangle \langle \chi_\nu | \sum_{\varrho=1}^m q_\varrho |\chi_\varrho\rangle \\ &= \sum_{\mu,\nu,\varrho=1}^m |\chi_\mu\rangle P_{\mu\nu} S_{\nu\varrho} q_\varrho = \sum_{\mu,\varrho=1}^m (\mathbf{PS})_{\mu\varrho} q_\varrho |\chi_\mu\rangle, \end{aligned} \quad (6.113)$$

т. е. в матричном представлении оказывается, что

$$\mathbf{p} = \mathbf{PSq}. \quad \text{Ч. т. д.} \quad (6.114)$$

<sup>1</sup> Получается умножением выражения  $\mathbf{PSPS} = \mathbf{PS}$  справа на  $\mathbf{S}^{-1}$ , которая существует, так как  $\mathbf{S}$  невырождена. — Прим. ред.

В ортонормированном базисе матрица **S** равна единичной, а матрица **P** сама является идемпотентной проектирующей матрицей.

В случае НХФ необходимо вводить две *P*-матрицы: одну для орбиталей спина  $\alpha$ , а другую для орбиталей спина  $\beta$ . Предполагая, что коэффициенты разложения орбиталей  $|a_i\rangle$  и  $|b_i\rangle$  заданы векторами  $\mathbf{a}^i$  и  $\mathbf{b}^i$  соответственно, т. е.

$$|a_i\rangle = \sum_{\mu=1}^m a_{\mu}^i |\chi_{\mu}\rangle ; \quad |b_i\rangle = \sum_{\mu=1}^m b_{\mu}^i |\chi_{\mu}\rangle , \quad (6.115)$$

имеем

$$\mathbf{P}^a = \sum_{i=1}^{n_a} \mathbf{a}^i \mathbf{a}^{i\dagger} ; \quad \mathbf{P}^b = \sum_{i=1}^{n_b} \mathbf{b}^i \mathbf{b}^{i\dagger} . \quad (6.116)$$

(Покомпонентно имеем, например,  $P_{\mu\nu}^a = \sum_{i=1}^{n_a} a_{\mu}^i a_{\nu}^{i*}$ .)

Здесь может быть уместным следующее предостережение. Если имеется выражение для какого-то сложного оператора, содержащее некоторый проекционный оператор  $\hat{Q}$  и мы хотим получить эквивалент этого выражения для случая ЛКАО, то необходимо построить соответствующую матрицу **Q**, аналогично тому, как мы сделали раньше для матрицы **P**. Однако, вообще говоря, неверно, что во всех случаях необходимо использовать матрицу **QS** в качестве эквивалента оператора  $\hat{Q}$  для случая ЛКАО. Нельзя делать никакие подстановки механически, а всегда необходимо рассматривать явно разложения ЛКАО всех выражений, так как правильный вид матрицы, соответствующей оператору  $\hat{Q}$  в неортogonalном базисе, может быть как **Q**, так и **QS**, **SQ**, или даже **SQS**, в зависимости от того, каким образом оператор  $\hat{Q}$  входит в выражение. В некоторых случаях представлением тождественного оператора, также представляющего проекционный оператор, в ЛКАО является не единичная матрица, а матрица перекрывания («метрика») **S**. Так происходит, например, в методе сдвига уровня, упомянутом в разд. 3.1. В этом случае к матрице Фока **F** требуется добавить матрицу  $\Delta\epsilon(\mathbf{S} - \mathbf{SPS}) = \Delta\epsilon(\mathbf{1} - \mathbf{PS})$ , которая и является представлением оператора сдвига  $\Delta\epsilon(1 - \hat{P})$  для случая ЛКАО.

### 5.3. Пример использования проекционных операторов в методе ЛКАО. Уравнения НХФ для последовательной оптимизации орбиталей

Выводом уравнений НХФ, приведенным в разделе 2.1, можно воспользоваться для получения такого варианта уравнений НХФ, который можно

применить для разработки безусловно сходящегося алгоритма оптимизации орбиталей, аналогичного тому, который обсуждался в разд. 1.3.

Напомним, что выражение (6.26) дает условие стационарности энергии при вариациях орбитали  $|a_i\rangle$ . При выводе этого уравнения не был использован тот факт, что другие орбитали уже являются самосогласованными, поэтому, решая уравнение для  $|a_i\rangle$  мы можем получить эту орбиталь, доставляющую минимум полной энергии, при условии, что остальные орбитали фиксированы. Предполагая, что слагаемое самоотталкивания  $\hat{J}_i^a - \hat{K}_i^a$  не было добавлено к фокиану (таким образом, вместо обычного оператора Фока  $\hat{F}^a$  мы используем « $i$ -зависимый»,  $\hat{F}_i^a$ ), задачу поиска оптимальной орбитали  $a_i$  для заданных (фиксированных) остальных орбиталей можно решить без использования итераций. Это позволяет нам сформулировать безусловно сходящуюся процедуру, заключающуюся в циклическом решении этих уравнений для всех орбиталей.

Получаем уравнение, которое необходимо решить для орбитали  $|a_i\rangle$ ,

$$\hat{F}_i^a |a_i\rangle = \sum_{j=1}^{n_a} \varepsilon_{ji}^a |a_j\rangle, \quad (6.117)$$

где орбитали  $|a_j\rangle$ ,  $j \neq i$  являются некоторыми фиксированными ортонормированными орбиталями, а искомая орбиталь  $|a_i\rangle$  тоже ортогональна им. Умножая уравнение (6.117) на  $\langle a_l|$  и используя ортогональность орбиталей, получим, аналогично формуле (6.37),

$$\varepsilon_{li}^a = \langle a_l | \hat{F}_i^a | a_i \rangle. \quad (6.118)$$

Подставим матричный элемент (6.118) в равенство (6.117) и перенесем недиагональные слагаемые в левую часть:

$$\hat{F}_i^a |a_i\rangle - \sum_{\substack{j=1 \\ j \neq i}}^{n_a} |a_j\rangle \langle a_j | \hat{F}_i^a | a_i \rangle = \varepsilon_{ii}^a |a_i\rangle \quad (6.119)$$

или же

$$\left[ 1 - \sum_{\substack{j=1 \\ (j \neq i)}}^{n_a} |a_j\rangle \langle a_j| \right] \hat{F}_i^a |a_i\rangle = \varepsilon_{ii}^a |a_i\rangle. \quad (6.120)$$

Теорема Купманса (разд. 3) указывает, что наименьшая полная энергия системы будет соответствовать такому решению  $|a_i\rangle$ , которое является собственным вектором, отвечающим наименьшему собственному

значению этого уравнения. (Здесь предполагается, что все остальные орбитали  $|a_j\rangle$ ,  $j \neq i$  фиксированы.) Введем обозначение

$$\hat{P}_{(i)}^a = \sum_{\substack{j=1 \\ (j \neq i)}}^{n_a} |a_j\rangle\langle a_j|, \quad (6.121)$$

и тогда можем записать

$$(1 - \hat{P}_{(i)}^a) \hat{F}_i^a |a_i\rangle = \varepsilon_{ii}^a |a_i\rangle. \quad (6.122)$$

Как и в случае, рассмотренном в гл. 2, разд. 1.5, все решения этого уравнения, для которых  $\varepsilon_{ii}^a \neq 0$ , автоматически являются ортогональными всем векторам  $|a_j\rangle$ , использованным при построении проекционного оператора  $\hat{P}_{(i)}^a$ . Поэтому можно вставить дополнительный проекционный оператор  $1 - \hat{P}_{(i)}^a$  без нарушения равенства и получить уравнение на собственные значения эрмитова оператора:

$$(1 - \hat{P}_{(i)}^a) \hat{F}_i^a (1 - \hat{P}_{(i)}^a) |a_i\rangle = \varepsilon_{ii}^a |a_i\rangle. \quad (6.123)$$

Чтобы переписать это уравнение в конечном базисе  $\{\chi_\mu\}$ , будем для краткости использовать упрощенную процедуру, рассмотренную в конце разд. 5.1. Введем разложение ЛКАО для орбиталей

$$|a_l\rangle = \sum_{\mu=1}^m a_\mu^l |\chi_\mu\rangle \quad (6.124)$$

и умножим уравнение (6.123) на  $\langle \chi_\nu |$ . Получим

$$\begin{aligned} \langle \chi_\nu | \left[ 1 - \sum_{\substack{j=1 \\ (j \neq i)}}^{n_a} \left| \sum_{\varrho=1}^m a_\varrho^j \chi_\varrho \right\rangle \left\langle \sum_{\tau=1}^m a_\tau^j \chi_\tau \right| \right] \hat{F}_i^a \left[ 1 - \sum_{\substack{k=1 \\ (k \neq i)}}^{n_a} \left| \sum_{\eta=1}^m a_\eta^k \chi_\eta \right\rangle \left\langle \sum_{\lambda=1}^m a_\lambda^k \chi_\lambda \right| \right] \\ \times \left| \sum_{\mu=1}^m a_\mu^i \chi_\mu \right\rangle = \varepsilon_{ii}^a \langle \chi_\nu | \sum_{\mu=1}^m a_\mu^i \chi_\mu \rangle. \end{aligned} \quad (6.125)$$

Распишем это уравнение, введя матрицу:

$$\mathbf{P}_{(i)}^a = \sum_{\substack{j=1 \\ (j \neq i)}}^{n_a} \mathbf{a}^j \mathbf{a}^{j\dagger}, \quad (6.126)$$

и тривиально получим обобщенное уравнение на собственные значения для эрмитовой матрицы:

$$(1 - \mathbf{S}\mathbf{P}_{(i)}^a)\mathbf{F}_i^a(1 - \mathbf{P}_{(i)}^a\mathbf{S})\mathbf{a}^i = \varepsilon_{ii}^a\mathbf{S}\mathbf{a}^i, \quad (6.127)$$

в котором ЛКАО представлением оператора  $\hat{P}_{(i)}^a$  в базисе неортогональных АО в одной скобке является матрица  $\mathbf{S}\mathbf{P}_{(i)}^a$ , а в другой — матрица  $\mathbf{P}_{(i)}^a\mathbf{S}$ .

#### 5.4. Матрица Фока и энергия

Согласно результату (6.100), нам понадобятся элементы матрицы Фока (рассматривается случай ОХФ):

$$\begin{aligned} F_{\nu\mu} &= \langle \chi_\nu | \hat{F} | \chi_\mu \rangle = \langle \chi_\nu | \hat{h} + \sum_{i=1}^n (2\hat{J}_i - \hat{K}_i) | \chi_\mu \rangle \\ &= h_{\nu\mu} + \sum_{i=1}^n (2\langle \chi_\nu | \hat{J}_i | \chi_\mu \rangle - \langle \chi_\nu | \hat{K}_i | \chi_\mu \rangle). \end{aligned} \quad (6.128)$$

Согласно определению (6.22) кулоновского оператора, можно написать

$$\begin{aligned} \langle \chi_\nu | \hat{J}_i | \chi_\mu \rangle &= \int \chi_\nu^*(1) \int \frac{\varphi_i^*(2)\varphi_i(2)}{r_{12}} dv_2 \chi_\mu(1) dv_1 \\ &= \iint \frac{\chi_\nu^*(1) \sum_{\varrho=1}^m c_\varrho^{i*} \chi_\varrho^*(2) \sum_{\tau=1}^m c_\tau^i \chi_\tau(2) \chi_\mu(1)}{r_{12}} dv_1 dv_2 \\ &= \sum_{\varrho, \tau=1}^m c_\tau^i c_\varrho^{i*} \iint \frac{\chi_\nu^*(1) \chi_\varrho^*(2) \chi_\mu(1) \chi_\tau(2)}{r_{12}} dv_1 dv_2 \\ &= \sum_{\varrho, \tau=1}^m c_\tau^i c_\varrho^{i*} [\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau]. \end{aligned} \quad (6.129)$$

Совершенно аналогично получаем матричный элемент обменного оператора (6.23):

$$\langle \chi_\nu | \hat{K}_i | \chi_\mu \rangle = \sum_{\varrho, \tau=1}^m c_\tau^i c_\varrho^{i*} [\chi_\nu \chi_\varrho | \chi_\tau \chi_\mu]. \quad (6.130)$$

Подставляя выражения (6.129) и (6.130) в уравнение (6.128) и воспользовавшись определением (6.108) элементов  $P$ -матрицы, получим

$$F_{\nu\mu} = h_{\nu\mu} + \sum_{i=1}^n \left( 2 \sum_{\varrho, \tau=1}^m c_\tau^i c_\varrho^{i*} [\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau] - \sum_{\varrho, \tau=1}^m c_\tau^i c_\varrho^{i*} [\chi_\nu \chi_\varrho | \chi_\tau \chi_\mu] \right)$$

$$= h_{\nu\mu} + \sum_{\varrho, \tau=1}^m P_{\tau\varrho} (2[\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau] - [\chi_\nu \chi_\varrho | \chi_\tau \chi_\mu]) . \quad (6.131)$$

Этот результат показывает, что для вычисления матрицы Фока необходима матрица **P**, а также одно- и двухэлектронные интегралы  $h_{\nu\mu}$  и  $[\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau]$ .

Матрицы Фока метода НХФ можно получить, проводя очень похожие выкладки. Получаем

$$F_{\nu\mu}^a = h_{\nu\mu} + \sum_{\varrho, \tau=1}^m \left[ (P_{\tau\varrho}^a + P_{\tau\varrho}^b) [\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau] - P_{\tau\varrho}^a [\chi_\nu \chi_\varrho | \chi_\tau \chi_\mu] \right] . \quad (6.132)$$

Элементы матрицы **F**<sup>b</sup> можно получить, заменяя в последнем слагаемом  $P_{\tau\varrho}^a$  на  $P_{\tau\varrho}^b$ .

Проводя теоретический анализ, удобно использовать «запись [12|12] для интегралов», т. е. применить такое обозначение двухэлектронных интегралов  $[\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau] \equiv \langle \chi_\mu(1) \chi_\nu(2) | \frac{1}{r_{12}} | \chi_\varrho(1) \chi_\tau(2) \rangle$ , в котором можно различить сомножители, происходящие от «бра» и от «кет»-векторов. По мнению автора, в абстрактно-теоретических рассуждениях предпочтительно сохранять возможность использования комплексных волновых функций. Однако, когда идет речь о программировании полученных формул, нужно пользоваться преимуществами того, что базисные функции, применяемые на практике, вещественны. Поэтому нет необходимости различать функции с комплексным сопряжением и без него, и тогда более удобно применять обозначения двухэлектронных интегралов в «записи (11|22)»:

$$(\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau) = \iint \frac{\chi_\mu(1) \chi_\nu(1) \chi_\varrho(2) \chi_\tau(2)}{r_{12}} dv_1 dv_2 . \quad (6.133)$$

(Легко видеть, что для вещественных орбиталей  $(\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau) \equiv [\chi_\mu \chi_\varrho | \chi_\nu \chi_\tau]$ .)

Анализ определения (6.133) показывает его инвариантность по отношению к перестановкам орбиталей  $\chi_\mu$  с  $\chi_\nu$  и  $\chi_\varrho$  с  $\chi_\tau$ , а также двух наборов орбиталей, соответствующих переменным интегрирования «1» и «2.» Если для краткости обозначить интеграл  $(\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau)$ , как  $(\mu\nu | \varrho\tau)$ , то эта инвариантность приводит к равенствам

$$\begin{aligned} (\mu\nu | \varrho\tau) &= (\nu\mu | \varrho\tau) = (\mu\nu | \tau\varrho) = (\nu\mu | \tau\varrho) \\ &= (\varrho\tau | \mu\nu) = (\varrho\tau | \nu\mu) = (\tau\varrho | \mu\nu) = (\tau\varrho | \nu\mu) . \end{aligned} \quad (6.134)$$

Эту симметрию легко понять, если рассмотреть двухэлектронный интеграл (6.133) как электростатическое взаимодействие между двумя зарядовыми плотностями  $\chi_\mu(\vec{r}_1) \chi_\nu(\vec{r}_1)$  и  $\chi_\varrho(\vec{r}_2) \chi_\tau(\vec{r}_2)$ .



В соответствии с (6.134), если все четыре индекса  $\mu, \nu, \varrho$ , и  $\tau$  различны, то существует восемь интегралов, у которых порядок индексов в определении (6.133) различен, но значения равны. (Это число уменьшается до четырех, если имеется одна пара совпадающих индексов, двух, если две пары совпадают, и единицы, если все индексы одинаковы.)

Воспользовавшись обозначениями (6.133) и используя элементы матрицы  $\mathbf{D}$ , определенной выражением (6.110), элементы матрицы Фока в случае метода ОХФ можно записать как

$$F_{\nu\mu} = h_{\nu\mu} + \sum_{\varrho, \tau=1}^m D_{\tau\varrho} [(\chi_\nu \chi_\mu | \chi_\varrho \chi_\tau) - \frac{1}{2} (\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau)] . \quad (6.135)$$

(Это уравнение можно получить из его аналога (6.132) для метода НХФ, с учетом того, что в случае ОХФ  $\mathbf{P}^a = \mathbf{P}^b = \frac{1}{2}\mathbf{D}$ . В более ранней литературе разности  $(\chi_\nu \chi_\mu | \chi_\varrho \chi_\tau) - \frac{1}{2} (\chi_\nu \chi_\varrho | \chi_\mu \chi_\tau)$  назывались элементами некоторой «суперматрицы».)

Согласно результату (6.67), электронную энергию однодетерминантной функции в случае НХФ можно записать как

$$E = \frac{1}{2} \left( \sum_{i=1}^{n_a} \langle a_i | \hat{h} + \hat{F}^a | a_i \rangle + \sum_{i=1}^{n_b} \langle b_i | \hat{h} + \hat{F}^b | b_i \rangle \right) . \quad (6.136)$$

Подставляя разложение ЛКАО орбиталей  $a_i$ ,  $b_i$ , мы получим:

$$\begin{aligned} E &= \frac{1}{2} \left( \sum_{i=1}^{n_a} \langle \sum_{\mu=1}^m a_\mu^i \chi_\mu | \hat{h} + \hat{F}^a | \sum_{\nu=1}^m a_\nu^i \chi_\nu \rangle \right. \\ &\quad \left. + \sum_{i=1}^{n_b} \langle \sum_{\mu=1}^m b_\mu^i \chi_\mu | \hat{h} + \hat{F}^b | \sum_{\nu=1}^m b_\nu^i \chi_\nu \rangle \right) \\ &= \frac{1}{2} \left( \sum_{i=1}^{n_a} \sum_{\mu, \nu=1}^m a_\nu^i a_\mu^{i*} \langle \chi_\mu | \hat{h} + \hat{F}^a | \chi_\nu \rangle \right. \\ &\quad \left. + \sum_{i=1}^{n_b} \sum_{\mu, \nu=1}^m b_\nu^i b_\mu^{i*} \langle \chi_\mu | \hat{h} + \hat{F}^b | \chi_\nu \rangle \right) \\ &= \frac{1}{2} \sum_{\mu, \nu=1}^m \left[ P_{\nu\mu}^a (h_{\mu\nu} + F_{\mu\nu}^a) + P_{\nu\mu}^b (h_{\mu\nu} + F_{\mu\nu}^b) \right] \\ &= \frac{1}{2} \left\{ \text{Tr} [\mathbf{P}^a (\mathbf{h} + \mathbf{F}^a)] + \text{Tr} [\mathbf{P}^b (\mathbf{h} + \mathbf{F}^b)] \right\} . \end{aligned} \quad (6.137)$$

В случае ОХФ это выражение сводится просто к

$$E = \frac{1}{2} \text{Tr}[\mathbf{D}(\mathbf{h} + \mathbf{F})] . \quad (6.138)$$

Конечно, подстановкой явного выражения элементов матрицы Фока энергию можно выразить непосредственно через одно- и двухэлектронные интегралы. Так, например, подставляя матричные элементы (6.135) в формулу (6.138), получаем электронную энергию ОХФ как

$$E = \sum_{\mu, \nu=1}^m D_{\mu\nu} h_{\nu\mu} + \frac{1}{2} \sum_{\mu, \nu, \varrho, \tau=1}^m (D_{\mu\nu} D_{\tau\varrho} - \frac{1}{2} D_{\varrho\nu} D_{\tau\mu}) (\chi_\nu \chi_\mu | \chi_\varrho \chi_\tau) . \quad (6.139)$$

(Последнее слагаемое, появляющееся из обменной части матрицы Фока, было получено перестановкой индексов суммирования  $\mu$  и  $\varrho$ .)

## 6. Ограниченный метод Хартри—Фока для открытых оболочек

Системы с открытыми оболочками (радикалы, триплетные состояния и т. д.) часто описываются однодетерминантным методом НХФ. Как обсуждалось ранее, его недостаток состоит в том, что волновая функция в этом случае не является собственной функцией оператора полного спина. Другая проблема появляется тогда, когда требуется сравнить энергии и другие свойства системы с замкнутой оболочкой и родственных ей систем с открытыми оболочками. Уже отмечалось, что в методе НХФ часто не удается прийти к решению, отличному от решения ОХФ около равновесной геометрии молекулы. Это означает, что мы описываем систему с замкнутой оболочкой, применяя дважды занятые орбитали. В то же время, система с открытой оболочкой, получаемая при ионизации/возбуждении исходной молекулы, либо представляющая какой-то ее фрагмент, содержит разное число электронов со спином  $\alpha$  и  $\beta$ . Следовательно, из-за того, что обменные слагаемые фокианов для спинов  $\alpha$  и  $\beta$  различны, НХФ описание системы с открытой оболочкой не может привести к строго дважды занятым орбиталам. (Некоторое орбитальное расщепление появится даже в тех частях системы, которые в самом деле можно считать неизменными при переходе от замкнутой оболочки к открытой.) Использование разных орбиталей для разных спинов дает более гибкую пробную волновую функцию, что может привести к учету некоторых эффектов электронной корреляции. С другой стороны, можно исследовать некоторые из этих задач, основываясь на расчетах исходной системы с замкнутой оболочкой, как мы делали, рассматривая процесс ионизации с помощью теоремы Купманса (разд. 3),

или же синглетных и триплетных возбуждений относительно основного состояния метода ОХФ (разд. 4.4). В таких расчетах, однако, невозможно учесть эффекты орбитальной релаксации при возбуждении или ионизации.

Эти и сходные проблемы рассматриваются с помощью «ограниченного метода Хартри—Фока для открытой оболочки» (ROHF, ОХФО), в котором замкнутые и открытые оболочки описываются по-разному. Замкнутые оболочки состоят из дважды занятых орбиталей; их расщепление не предполагается. В зависимости от природы открытой оболочки, волновая функция может быть одним детерминантом, как в случае дублета или высокоспиновой компоненты ( $S_z = 1$ ) триплета, или же может быть линейной комбинацией двух (или более) детерминантов, как в случае синглета с открытой оболочкой (синглетного бирадикала) или компоненты триплета с  $S_z = 0$ . Многодетерминантная волновая функция также может возникать, если требуется воспроизвести высокую пространственную симметрию, например, в случае свободных атомов или же комплексов переходных металлов с открытыми  $d$ -оболочками.

Нельзя не отметить, что существуют задачи, в которых насильственное получение состояний с правильной спиновой симметрией при использовании наибольшего возможного числа дважды занятых орбиталей может также иметь свои недостатки: так, волновая функция ОХФО не учитывает явление спиновой поляризации, что совершенно необходимо для понимания сверхтонкого расщепления в спектрах ЭПР и т. д.

В схеме ОХФО разные системы открытых оболочек описываются по единой схеме на основе достаточно общего выражения для энергии, содержащего численные коэффициенты, значения которых отличаются для разных задач. Во всех случаях предполагается, что мы имеем дело с некоторым числом  $N$  ортонормированных (пространственных) орбиталей  $\varphi_i$ , каждой из которых можно сопоставить свое число заполнения  $\omega_i$ , и тогда энергию можно записать как

$$E = \sum_{i=1}^N \omega_i h_{ii} + \sum_{i,j=1}^N (\alpha_{ij} J_{ij} - \beta_{ij} K_{ij}), \quad (6.140)$$

где «коэффициенты связи»  $\alpha_{ij}$  и  $\beta_{ij}$  определяются типом состояния с открытой оболочкой. В выражении (6.140)  $h_{ii} = \langle \varphi_i | \hat{h} | \varphi_i \rangle$  есть матричный элемент одноэлектронной части гамильтониана, а  $J_{ij} = [\varphi_i \varphi_j | \varphi_i \varphi_j]$  и  $K_{ij} = [\varphi_i \varphi_j | \varphi_j \varphi_i]$  являются кулоновским и обменным интегралами, соответственно. Заметим, что в выражении (6.140) суммирование по  $i$  и  $j$  идет независимо.

Разумеется, выражение (6.140) достаточно общее, и включает как случай обычной волновой функции ОХФ замкнутой оболочки (при значениях параметров  $N = n$ ;  $\omega_i = 2$ ;  $\alpha_{ij} = 2$ ;  $\beta_{ij} = 1$ ), так и волновой функции НХФ ( $N = n_a + n_b$ ;  $\omega_i = 1$ ;  $\alpha_{ij} = \frac{1}{2}$ ;  $\beta_{ij} = \frac{1}{2}$ , если  $i$  и  $j$  оба принадлежат к одному набору  $\{a_i\}$  или  $\{b_i\}$ , и  $\beta_{ij} = 0$  в противном случае). В конце этого раздела мы получим значения коэффициентов связи для наиболее важных случаев открытых оболочек, упомянутых ранее.

## 6.1. Уравнения ОХФО

Схема ОХФО по существу определяется формулой общего вида (6.140) для энергии и условием ортонормировки всех орбиталей. Поэтому наиболее подходящим будет изучение вариационной задачи с использованием метода множителей Лагранжа. (Процедура практически такая же, как и та, которую мы использовали в разд. 2.2; главное отличие в том, что в формуле для энергии (6.140) присутствуют коэффициенты  $\omega_i$ ,  $\alpha_{ij}$ , и  $\beta_{ij}$ .) Определим вспомогательный функционал:

$$F = E - \sum_{i,j=1}^N 2\varepsilon_{ji} (\langle \varphi_i | \varphi_j \rangle - \delta_{ij}) \quad (6.141)$$

и потребуем обращения в нуль его вариации  $\delta F = 0$  при вариациях  $\delta \varphi_i$  орбиталей  $\varphi_i$ ,  $i = 1, 2, \dots, N$ . Здесь  $E$  означает энергию (6.140) и, согласно детальному обсуждению в разд. 2.2, матрица множителей Лагранжа  $\varepsilon$  должна быть эрмитовой, а решения  $\varphi_i$  должны быть ортонормированы. (Для дальнейшего удобства мы определили множители Лагранжа как  $-2\varepsilon_{ji}$ .)

Вариация  $\delta F$ , имеющая место при варьировании орбитали  $\varphi_k$ , имеет вид

$$\begin{aligned} \delta F = & \omega_k \langle \delta \varphi_k | \hat{h} | \varphi_k \rangle + \sum_{j=1}^N (\alpha_{kj} \langle \delta \varphi_k | \hat{J}_j | \varphi_k \rangle - \beta_{kj} \langle \delta \varphi_k | \hat{K}_j | \varphi_k \rangle) \\ & + \sum_{i=1}^N (\alpha_{ik} \langle \delta \varphi_k | \hat{J}_i | \varphi_k \rangle - \beta_{ik} \langle \delta \varphi_k | \hat{K}_i | \varphi_k \rangle) - \sum_{j=1}^N 2\varepsilon_{jk} \langle \delta \varphi_k | \varphi_j \rangle + \text{с.с.} = 0. \end{aligned} \quad (6.142)$$

Замечая, что суммы по  $i$  и  $j$  равны, и вводя « $k$ -зависимый» оператор Фока ОХФО

$$\hat{F}_k = \frac{\omega_k}{2} \hat{h} + \sum_{j=1}^N (\alpha_{kj} \hat{J}_j - \beta_{kj} \hat{K}_j), \quad (6.143)$$

перепишем условие (6.142) в виде уравнений ОХФО:

$$\hat{F}_k |\varphi_k\rangle = \sum_{j=1}^N \varepsilon_{jk} |\varphi_j\rangle \quad k = 1, 2, \dots, N. \quad (6.144)$$

$k$ -зависимость оператора Фока означает, что для каждой орбитали или, по крайней мере, для каждого класса орбиталей имеется свой оператор Фока. В случае ОХФО необходимо также явно потребовать эрмитовости матрицы множителей Лагранжа:

$$\varepsilon_{jk} = \varepsilon_{kj}^*, \quad (6.145)$$

потому что в общем случае оно не выполняется автоматически из-за  $k$ -зависимости оператора Фока (6.143). В отличие от методов ОХФ или НХФ, в методе ОХФО унитарные преобразования допустимы только внутри данного класса орбиталей (только между однократно занятыми, либо между двукратно занятыми) и, следовательно, нельзя полностью освободиться от внедиагональных множителей Лагранжа, связывающих разные наборы орбиталей.

Часто желательно иметь, хотя бы формально, общее уравнение на собственные значения, для которого все орбитали  $\varphi_i$  являются собственными функциями. Мы кратко рассмотрим два возможных уравнения такого типа.

В первом случае начнем с выкладок, сходных с теми, которые уже использовались в разд. 5.3. Умножим уравнение (6.144) на  $\langle \varphi_l |$  и получим значение множителя Лагранжа  $\varepsilon_{lk}$ :

$$\langle \varphi_l | \hat{F}_k | \varphi_k \rangle = \sum_{j=1}^N \varepsilon_{jk} \langle \varphi_l | \varphi_j \rangle = \sum_{j=1}^N \varepsilon_{jk} \delta_{lj} = \varepsilon_{lk}. \quad (6.146)$$

Подставим матричные элементы (6.146) в уравнение (6.144) и, перенося все слагаемые, кроме тех, для которых  $j = k$ , в левую часть, получим

$$\hat{F}_k |\varphi_k\rangle - \sum_{\substack{j=1 \\ (j \neq k)}}^N |\varphi_j\rangle \langle \varphi_j | \hat{F}_k | \varphi_k \rangle = \varepsilon_{kk} |\varphi_k\rangle, \quad (6.147)$$

т. е.

$$(1 - \hat{P}_{(k)}) \hat{F}_k |\varphi_k\rangle = \varepsilon_{kk} |\varphi_k\rangle, \quad (6.148)$$

где, как и в случае, рассмотренном в разд. 5.3, используется обозначение  $\hat{P}_{(k)}$  для проекционного оператора:

$$\hat{P}_{(k)} = \sum_{\substack{j=1 \\ (j \neq k)}}^N |\varphi_j\rangle\langle\varphi_j|. \quad (6.149)$$

Пользуясь равенством  $(1 - \hat{P}_{(k)})|\varphi_k\rangle = |\varphi_k\rangle$ , мы можем свести уравнение (6.148) к уравнению с эрмитовым оператором:

$$(1 - \hat{P}_{(k)})\hat{F}_k(1 - \hat{P}_{(k)})|\varphi_k\rangle = \varepsilon_{kk}|\varphi_k\rangle. \quad (6.150)$$

Каждая занятая орбиталь  $|\varphi_i\rangle$ ,  $i \neq k$ , является собственным вектором этого оператора с собственным значением 0. Поэтому мы можем составить сумму таких операторов и получить уравнение на собственные значения некоторого эрмитова оператора, которое должно выполняться одновременно для всех орбиталей  $\varphi_i$ :

$$\hat{G}|\varphi_i\rangle = \varepsilon_{ii}|\varphi_i\rangle, \quad (6.151)$$

где

$$\hat{G} = \sum_{k=1}^N (1 - \hat{P}_{(k)})\hat{F}_k(1 - \hat{P}_{(k)}). \quad (6.152)$$

К сожалению, для решений этого уравнения не гарантировано выполнение условий (6.145) эрмитовости матрицы множителей Лагранжа. Эти условия можно переформулировать следующим образом: орбитали ОХФО должны быть *одновременно* собственными векторами и некоторого другого эрмитова оператора

$$\hat{Q} = \sum_{i,j=1}^N \hat{P}_i(\hat{F}_i - \hat{F}_j)\hat{P}_j. \quad (6.153)$$

В выражении (6.153) мы использовали обозначение проекционного оператора

$$\hat{P}_i = |\varphi_i\rangle\langle\varphi_i|, \quad (6.154)$$

которое необходимо отличать от оператора  $\hat{P}_{(i)}$ , определенного формулой (6.149). В самом деле, пусть  $|\varphi_l\rangle$  будет собственным вектором оператора  $\hat{Q}$  с собственным значением  $\kappa_l$ :

$$\hat{Q}|\varphi_l\rangle = \sum_{i,j=1}^N \hat{P}_i(\hat{F}_i - \hat{F}_j)\hat{P}_j|\varphi_l\rangle = \sum_{i=1}^N \hat{P}_i(\hat{F}_i - \hat{F}_l)|\varphi_l\rangle = \kappa_l|\varphi_l\rangle. \quad (6.155)$$

Умножая соотношение (6.155) на некоторую  $\langle \varphi_k |$  ( $k \neq l$ ), и воспользовавшись ортогональностью орбиталей  $|\varphi_k\rangle$  и  $|\varphi_l\rangle$ , получим

$$\langle \varphi_k | \sum_{i=1}^N \hat{P}_i (\hat{F}_i - \hat{F}_l) | \varphi_l \rangle = \langle \varphi_k | \hat{F}_k | \varphi_l \rangle - \langle \varphi_k | \hat{F}_l | \varphi_l \rangle = \kappa_l \langle \varphi_k | \varphi_l \rangle = 0. \quad (6.156)$$

Согласно результату (6.146), имеем  $\varepsilon_{lk} = \langle \varphi_l | \hat{F}_k | \varphi_k \rangle$  и аналогично  $\varepsilon_{kl} = \langle \varphi_k | \hat{F}_l | \varphi_l \rangle$ . Поэтому равенство (6.156) показывает, что требование  $\hat{Q}|\varphi_l\rangle = \kappa_l|\varphi_l\rangle$  ( $l = 1, 2, \dots, N$ ) эквивалентно условию эрмитовости (6.145) матрицы множителей Лагранжа.

Обеспечить одновременную диагональность двух эрмитовых операторов  $\hat{G}$  и  $\hat{Q}$  не так уж тривиально. Поэтому может показаться интересным отыскать такое уравнение на собственные значения, которое автоматически обеспечивает выполнение условий (6.145). Для этого необходимо ввести их явно в уравнения ОХФО (6.144) заменой  $\varepsilon_{jk}$  на  $\varepsilon_{kj}^* = \langle \varphi_k | \hat{F}_j | \varphi_j \rangle^* = \langle \varphi_j | \hat{F}_j | \varphi_k \rangle$ . Получаем

$$\hat{F}_k |\varphi_k\rangle = \sum_{\substack{j=1 \\ (j \neq k)}}^N |\varphi_j\rangle \langle \varphi_j | \hat{F}_j | \varphi_k \rangle + \varepsilon_{kk} |\varphi_k\rangle, \quad (6.157)$$

т. е.

$$\left( \hat{F}_k - \sum_{\substack{j=1 \\ (j \neq k)}}^N |\varphi_j\rangle \langle \varphi_j | \hat{F}_j \right) |\varphi_k\rangle = \varepsilon_{kk} |\varphi_k\rangle. \quad (6.158)$$

Это уравнение отличается от уравнения (6.147) появлением в сумме операторов  $\hat{F}_j$ , вместо  $\hat{F}_k$ . Умножим каждую из частей уравнения (6.158) на соответствующие левую и правую части условия ортонормировки

$$\langle \varphi_k | \varphi_i \rangle = \delta_{ki} \quad (6.159)$$

и просуммируем по  $k$

$$\sum_{k=1}^N \left[ \left( \hat{F}_k - \sum_{\substack{j=1 \\ (j \neq k)}}^N \hat{P}_j \hat{F}_j \right) \hat{P}_k \right] |\varphi_i\rangle = \varepsilon_{ii} |\varphi_i\rangle. \quad (6.160)$$

Умножим это уравнение на  $\langle \varphi_l |$  ( $l \neq i$ ) и получим, учитывая, что  $\hat{P}_k |\varphi_i\rangle = \delta_{ki} |\varphi_i\rangle$  и  $\langle \varphi_l | \hat{P}_j = \delta_{jl} \langle \varphi_l |$ ,



$$\begin{aligned}
 & \langle \varphi_l | \sum_{k=1}^N \left[ \left( \hat{F}_k - \sum_{\substack{j=1 \\ (j \neq k)}}^N \hat{P}_j \hat{F}_j \right) \hat{P}_k \right] |\varphi_i\rangle \\
 &= \langle \varphi_l | \sum_{k=1}^N \hat{F}_k \hat{P}_k |\varphi_i\rangle - \langle \varphi_l | \sum_{k=1}^N \sum_{\substack{j=1 \\ (j \neq k)}}^N \hat{P}_j \hat{F}_j \hat{P}_k |\varphi_i\rangle \\
 &= \langle \varphi_l | \hat{F}_i |\varphi_i\rangle - \langle \varphi_l | \hat{F}_l |\varphi_i\rangle = 0.
 \end{aligned} \tag{6.161}$$

Это значит, что решения уравнений (6.160) автоматически удовлетворяют условию эрмитовости (6.145). К сожалению, однако, оператор в уравнении (6.160) не является эрмитовым; если просто привести его к эрмитовому виду, то автоматическое выполнение условия (6.145) может быть утрачено.

Мы не будем углубляться в рассмотрение технических проблем, связанных со схемой ОХФО; заметим только, что разные общедоступные квантовохимические программы дают одни и те же орбитали ОХФО, при этом присваивая им совершенно разные «орбитальные энергии» (собственные значения).



## 6.2. Коэффициенты связи для некоторых систем с открытой оболочкой

Простейший случай единственной (высокоспиновой) открытой оболочки (дублета, компоненты триплета с  $S_z = 1$  и т. д.) можно описать с помощью однодетерминантной волновой функции. В этом случае мы рассматриваем  $n_c$  дважды занятых орбиталей (замкнутую оболочку) и  $n_o$  орбиталей, однократно занятых электронами с проекцией спина  $\alpha$  (открытую оболочку). Тогда энергию можно легко получить с помощью общей формулы (6.44) для энергии однодетерминантной волновой функции, подставляя в нее  $n_b = n_c$ ,  $n_a = n_c + n_o$ ,  $a_p = b_p = \varphi_p$  ( $p \leq n_c$ ) и  $a_p = \varphi_p$  ( $n_c + 1 \leq p \leq n_c + n_o$ ):

$$\begin{aligned}
 E = & \sum_{p=1}^{n_c} 2h_{pp} + \sum_{q=n_c+1}^{n_c+n_o} h_{qq} + \sum_{p,q=1}^{n_c} (2J_{pq} - K_{pq}) + \sum_{p=1}^{n_c} \sum_{q=n_c+1}^{n_c+n_o} (2J_{pq} - K_{pq}) \\
 & + \frac{1}{2} \sum_{p,q=n_c+1}^{n_c+n_o} (J_{pq} - K_{pq})
 \end{aligned} \tag{6.162}$$



Соответственно, значения коэффициентов, введенных в (6.140), суть  $\omega_c = 2$ ,  $\omega_o = 1$ ,  $\alpha_{cc} = 2$ ,  $\alpha_{co} = \alpha_{oc} = 1$ ,  $\alpha_{oo} = 1/2$ ;  $\beta_{cc} = 1$ , и  $\beta_{co} = \beta_{oc} = \beta_{oo} = 1/2$ , где индексы  $c$  или  $o$  указывают на оболочку,  $\{C\}$  или  $\{O\}$ , к которой принадлежат орбитали  $\varphi_i$ ,  $\varphi_j$ . (C=замкнутая, closed; O=открытая, open.) При определении этих коэффициентов мы учли, что суммирование по  $i$  и  $j$  в выражении (6.140) идет независимо, так что существует случай, когда  $i \in \{C\}$  и  $j \in \{O\}$ , а также, когда  $j \in \{C\}$  и  $i \in \{O\}$ . В случае дублета ( $n_o = 1$ ) в открытую оболочку входит только одна орбиталь и последняя сумма в выражении (6.162) исчезает ( $J_{pp} \equiv K_{pp}$ ).

В случае открытой оболочки, содержащей две функции и два электрона в синглетном или триплетном состоянии с  $S_z = 0$ , можно действовать точно так же, как мы делали ранее при рассмотрении синглетных и триплетных возбуждений волновой функции ОХФ (разд. 4.4). Имеем

$$^{1,3}\Psi = \frac{1}{\sqrt{2}}[\Psi_1 \pm \Psi_2], \quad (6.163)$$

где два детерминанта  $\Psi_1$  и  $\Psi_2$  определены как

$$\Psi_1 = \mathcal{A}[\varphi_1^2, \varphi_1^2 \dots \varphi_{n_c}^2 \varphi_i^\alpha \varphi_j^\beta] \quad (6.164)$$

и

$$\Psi_2 = \mathcal{A}[\varphi_1^2, \varphi_1^2 \dots \varphi_{n_c}^2 \varphi_j^\alpha \varphi_i^\beta] \quad (6.165)$$

и верхний знак снова соответствует синглету, а нижний — триплету. В выражениях (6.164) и (6.165) мы ввели упрощенное обозначение  $\varphi_p^2$  для дважды занятой орбитали и опустили аргументы (координаты отдельных электронов).

Необходимо вычислить

$$^{1,3}E = \frac{1}{2} \langle \Psi_1 \pm \Psi_2 | \hat{H} | \Psi_1 \pm \Psi_2 \rangle. \quad (6.166)$$

Здесь  $\langle \Psi_1 | \hat{H} | \Psi_1 \rangle = \langle \Psi_2 | \hat{H} | \Psi_2 \rangle$  — энергии однодетерминантных волновых функций  $\Psi_1$  и  $\Psi_2$ , которые можно вычислить с помощью общей формулы (6.44), подставив в нее  $n_a = n_b = n_c + 1$ ,  $a_p = b_p = \varphi_p$  ( $p \leq n_c$ ),  $a_{n_c+1} = \varphi_i$ , и  $b_{n_c+1} = \varphi_j$ . Получим

$$\begin{aligned} \langle \Psi_1 | \hat{H} | \Psi_1 \rangle &= \sum_{p=1}^{n_c} 2h_{pp} + h_{ii} + h_{jj} + \sum_{p,q=1}^{n_c} (2J_{pq} - K_{pq}) \\ &+ \sum_{p=1}^{n_c} [2(J_{pi} + J_{pj}) - (K_{pi} + K_{pj})] + J_{ij}. \end{aligned} \quad (6.167)$$

Предполагая, что орбитали вещественны, мы получим, как и в случае, рассмотренном в разд. 4.4, что

$$\langle \Psi_1 | \hat{H} | \Psi_2 \rangle = \langle \Psi_2 | \hat{H} | \Psi_1 \rangle = K_{ij} . \quad (6.168)$$

Введем обозначение  $i = n_c + 1$ ,  $j = n_c + 2$  для упрощения формулы:

$$\begin{aligned} {}^{1,3}E = & \sum_{p=1}^{n_c} 2h_{pp} + \sum_{q=n_c+1}^{n_c+2} h_{qq} + \sum_{p,q=1}^{n_c} (2J_{pq} - K_{pq}) \\ & + \sum_{p=1}^{n_c} \sum_{q=n_c+1}^{n_c+2} (2J_{pq} - K_{pq}) + J_{ij} \pm K_{ji} . \end{aligned} \quad (6.169)$$

Таким образом, мы получаем следующие значения коэффициентов: общие значения —  $\omega_c = 2$ ,  $\omega_o = 1$ ,  $\alpha_{cc} = 2$ ,  $\alpha_{co} = \alpha_{oc} = 1$ ,  $\alpha_{oo} = 1/2$ ;  $\beta_{cc} = 1$ ;  $\beta_{co} = 1/2$ ; частные значения —  $\beta_{oo} = -1/2$  для синглета и  $\beta_{oo} = 1/2$  для триплета.

Как и ожидалось, компоненты триплета с  $S_z = 1$  и  $S_z = 0$  имеют одну и ту же энергию, несмотря на то, что первая является однодетерминантной волновой функцией, а вторая — линейной комбинацией двух детерминантов.

## 7. Градиент энергии

Одной из основных задач квантовой химии является поиск минимумов и других стационарных точек на поверхности потенциальной энергии Борна—Оппенгеймера молекулярной системы. Знание минимумов необходимо для изучения геометрических параметров стабильных молекул, их различных конформаций и т. д.; седловые точки важны для исследования процессов изомеризации и химических реакций. Несмотря на то, что исследование потенциальных поверхностей, в принципе, возможно путем вычисления энергии для достаточно большого числа разных геометрических конфигураций ядер, задача оптимизации геометрии может решаться намного более эффективно, если можно вычислить градиент энергии, т. е. ее производные по различным ядерным координатам.

Как уже обсуждалось в гл. 2, разд. 2.1, теорема Гельмана—Фейнмана не выполняется для волновой функции метода ХФР, если, как обычно, базисные орбитали «следуют за ядрами», и поэтому изменение геометрии порождает также изменение используемых базисных функций. Другими словами, в случае конечного базиса вариации базисных орбиталей

дают определенный вклад в вариацию энергии  $\delta E$ , возникающую при бесконечно малом изменении геометрии  $\delta \mathbf{R}$ . (Здесь  $\mathbf{R}$  обозначает набор всех ядерных координат.) Эти вклады приводят к так называемым «силам волновой функции», которые отличаются от «сил Гельмана—Фейнмана», характерных для случая, когда теорема Гельмана—Фейнмана выполняется.

Мы рассмотрим<sup>1</sup> волновую функцию ССП  $\Psi$ , которая может быть одним детерминантом или линейной комбинацией детерминантов; в последнем случае предполагается, что разные детерминанты либо имеют фиксированные коэффициенты (как в ОХФО или теориях спроектированного ХФ), либо коэффициенты при отдельных детерминантах оптимизируются вариационно (как в различных теориях МК ССП). Волновая функция  $\Psi$  зависит от  $N$  орбиталей  $\psi_i$ , которые разлагаются по базису  $\{\chi_\mu\}$  согласно

$$\psi_i = \sum_{\mu=1}^m c_\mu^i \chi_\mu. \quad (6.170)$$

Перепишем уравнение (2.55) в виде

$$\delta E = \frac{\langle \Psi | \delta \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} + \frac{1}{\langle \Psi | \Psi \rangle} \{ \langle \delta \Psi | \hat{H} - E | \Psi \rangle + \text{с.с.} \}. \quad (6.171)$$

Здесь как изменение гамильтониана  $\delta \hat{H}$ , так и изменение волновой функции  $\delta \Psi$  являются следствием изменения молекулярной геометрии  $\delta \mathbf{R}$ . Очевидно, первое слагаемое в правой части равенства (6.171) соответствует силе Гельмана—Фейнмана, а второе — «силе волновой функции».

Изменение молекулярной геометрии  $\delta \mathbf{R}$  вызывает изменение орбитали  $\psi_i$ , которое состоит из двух частей:

$$\delta \psi_i = \delta \psi_i' + \delta \psi_i'' = \sum_{\mu=1}^m \delta c_\mu^i \chi_\mu + \sum_{\mu=1}^m c_\mu^i \delta \chi_\mu. \quad (6.172)$$

Первое слагаемое ( $\delta \psi_i'$ ) отвечает изменению орбитальных коэффициентов  $c_\mu^i$ , второе ( $\delta \psi_i''$ ) — базисных орбиталей  $\chi_\mu$ .

<sup>1</sup> Последующий вывод воспроизводится с разрешения Elsevier Science из нашей статьи М. В. Ruiz, I. Mayer, «Deriving Gradient Formulae for SCF Methods by Using Brillouin-type Theorems. Gradients in the HPHF Method», Chemical Physics Letters, **236**, pages 217–228, Copyright 1995 Elsevier Science.

В силу линейности вариаций первого порядка и известных свойств определителей, вариацию первого порядка функции  $\Psi$  можно представить в виде суммы

$$\delta\Psi = \sum_{i=1}^N (\delta\Psi'_i + \delta\Psi''_i) + \delta\Psi_C, \quad (6.173)$$

где  $\delta\Psi'_i$  и  $\delta\Psi''_i$  являются вариациями многоэлектронной волновой функции  $\Psi$ , возникающими, когда орбиталь  $\psi_i$  меняется на  $\delta\psi'_i$  и  $\delta\psi''_i$ , соответственно. Если  $\Psi$  является многодетерминантной (МК ССП) волновой функцией с переменными коэффициентами детерминантов, то  $\delta\Psi_C$  задает вариацию  $\Psi$ , связанную с изменением этих коэффициентов. Если  $\Psi$  — это один детерминант или сумма детерминантов с фиксированными коэффициентами (случаи ОХФ, НХФ, ОХФО, спроектированного ХФ), то слагаемое  $\delta\Psi_C$  отсутствует.

Слагаемые  $\delta\Psi'_i$  и  $\delta\Psi_C$  (если таковые имеются) являются вариациями тех типов, которые допустимы при оптимизации волновой функции при фиксированной геометрии, так что для них выполняется равенство  $\langle \delta\Psi | \hat{H} - E | \Psi \rangle = 0$  (см. гл. 2, разд. 1.3). Соответственно, для вычисления градиента *вариационной* волновой функции достаточно рассмотреть в выражении (6.171) только вариации  $\delta\Psi''_i$ , связанные с изменениями базисных орбиталей, но не надо заниматься теми вариациями, которые относятся к изменениям коэффициентов разложения орбиталей (или коэффициентов при разных детерминантах).

Так как  $\delta\Psi''_i$  есть ничто иное, как однократно возбужденная конфигурация  $\Psi_1(\psi_i \rightarrow \delta\psi''_i)$ , общую формулу для изменения энергии в первом порядке по волновой функции  $\Psi$ , возникающем при изменении геометрии  $\delta\mathbf{R}$ , можно записать как

$$\delta E = \frac{\langle \Psi | \delta \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} + \frac{1}{\langle \Psi | \Psi \rangle} \sum_{i=1}^N \{ \langle \Psi_1(\psi_i \rightarrow \delta\psi''_i) | \hat{H} - E | \Psi \rangle + \text{с.с.} \}. \quad (6.174)$$

При вычислении выражения (6.174) необходимо помнить, что в обычном гамильтониане Борна—Оппенгеймера только слагаемое электронно-ядерного отталкивания

$$\hat{V}_{eN} = - \sum_{\alpha=1}^{N_N} \sum_{i=1}^{N_e} \frac{Z_{\alpha}}{r_{\alpha i}} \quad (6.175)$$

зависит от молекулярной геометрии, так что  $\delta \hat{H} \equiv \delta \hat{V}_{eN}$ .

Теперь мы конкретизируем полученные результаты для случая НХФ, когда  $\Psi$  является однодетерминантной волновой функцией, по-

строенной из  $n_a$  орбиталей  $a_i$  с проекцией спина  $\alpha$  и  $n_b$  орбиталей  $b_i$  с проекцией спина  $\beta$ , которые разлагаются по ограниченному базису как

$$a_i = \sum_{\mu=1}^m a_{\mu}^i \chi_{\mu} ; \quad b_i = \sum_{\mu=1}^m b_{\mu}^i \chi_{\mu} \quad (6.176)$$

Уравнение (6.174) можно переписать для этого случая как

$$\begin{aligned} \delta E = \langle \Psi | \delta \hat{V}_{Ne} | \Psi \rangle + \sum_{i=1}^{n_a} \{ \langle \Psi | (a_i \rightarrow \delta a_i'') | \hat{H} - E | \Psi \rangle + \text{с.с.} \} \\ + \sum_{i=1}^{n_b} \{ \langle \Psi | (b_i \rightarrow \delta b_i'') | \hat{H} - E | \Psi \rangle + \text{с.с.} \} . \end{aligned} \quad (6.177)$$

(Орбитали предполагаются ортонормированными, так что  $\langle \Psi | \Psi \rangle = 1$ .)

Так как  $\delta \hat{V}_{eN}$  представляет собой сумму одноэлектронных операторов, правила Слэтера дают

$$\langle \Psi | \delta \hat{H} | \Psi \rangle = \langle \Psi | \delta \hat{V}_{Ne} | \Psi \rangle = \sum_{i=1}^{n_a} \langle a_i | \delta \hat{h} | a_i \rangle + \sum_{i=1}^{n_b} \langle b_i | \delta \hat{h} | b_i \rangle , \quad (6.178)$$

где

$$\delta \hat{h} = \delta \sum_{\alpha=1}^{N_N} \frac{-Z_{\alpha}}{r_{\alpha}} . \quad (6.179)$$

Теперь подробно рассмотрим вариацию  $\delta a_i'' = \sum_{\mu} a_{\mu}^i \delta \chi_{\mu}$ . Представим

единичный оператор в виде  $1 \equiv \hat{P}^a + (1 - \hat{P}^a)$  (разложение единицы), где  $\hat{P}^a$  есть проектор на подпространство молекулярных орбиталей  $\{a_j\}$ , занятых спином  $\alpha$ . Тогда вариацию  $\delta a_i''$  можно представить в виде суммы двух слагаемых:

$$\delta a_i'' = \delta a_{i,1}'' + \delta a_{i,2}'' = \hat{P}^a \delta a_i'' + (1 - \hat{P}^a) \delta a_i'' , \quad (6.180)$$

первое из которых полностью лежит в подпространстве занятых орбиталей  $a_j$ , а второе ортогонально к этому подпространству. Имеем также

$$\Psi_1(a_i \rightarrow \delta a_i'') = \Psi_1(a_i \rightarrow \delta a_{i,1}'') + \Psi_1(a_i \rightarrow \delta a_{i,2}'') , \quad (6.181)$$

согласно свойствам определителей. Так как  $\delta a_{i,1}'' = \hat{P}^a \delta a_i''$  является линейной комбинацией занятых орбиталей  $a_j$ , вариация  $\delta a_{i,1}''$  не изменит детерминантной волновой функции  $\Psi$ , за исключением, возможно, ее фазы или нормировки. Следовательно,  $\Psi_1(a_i \rightarrow \delta a_{i,1}'')$  либо обращается в

нуль, либо пропорционально  $\Psi$ , а  $\langle \Psi_1(a_i \rightarrow \delta a''_{i,1}) | \hat{H} - E | \Psi \rangle = 0$ . Таким образом, на самом деле нужно найти только слагаемое  $\langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \hat{H} - E | \Psi \rangle$ . Однако  $\delta a''_{i,2} = (1 - \hat{P}^a) \delta a''_i$  является орбиталью, ортогональной ко всем орбиталям, занятым в  $\Psi$ , и поэтому  $\langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \Psi \rangle = 0$ , согласно правилам Слэтера, и те слагаемые в выражении (6.177), которые содержат полную энергию, также пропадают.

Подобные рассуждения применимы также и к орбиталям со спином  $\beta$ , так что вариацию энергии, вызванную вариацией  $\delta \mathbf{R}$ , в случае НХФ можно записать как

$$\begin{aligned} \delta E = & \sum_{i=1}^{n_a} \left[ \langle a_i | \delta \hat{h} | a_i \rangle + \{ \langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \hat{H} | \Psi \rangle + \text{с.с.} \} \right] \\ & + \sum_{i=1}^{n_b} \left[ \langle b_i | \delta \hat{h} | b_i \rangle + \{ \langle \Psi_1(b_i \rightarrow \delta b''_{i,2}) | \hat{H} | \Psi \rangle + \text{с.с.} \} \right] . \end{aligned} \quad (6.182)$$

Мы видим тесную аналогию с методом, которым мы пользовались при выводе уравнений НХФ через теорему Бриллюэна (разд. 2.1).

Теперь обсудим несколько подробнее слагаемое  $\langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \hat{H} | \Psi \rangle$  в правой части выражения (6.182). Расписывая его с помощью правил Слэтера, получаем

$$\langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \hat{H} | \Psi \rangle = \langle \delta a''_{i,2} | \hat{F}^a | a_i \rangle , \quad (6.183)$$

с учетом обычного определения оператора Фока (6.25), включающего слагаемые самоотталкивания.

Подставляя определение  $\delta a''_{i,2}$  в «бра»-часть, имеем

$$\langle \delta a''_{i,2} | = \langle \delta a''_i | (1 - \hat{P}^a) = \langle \delta a''_i | \left( 1 - \sum_{k=1}^{n_a} |a_k\rangle \langle a_k| \right) , \quad (6.184)$$

так что правую часть выражения (6.183) можно преобразовать к виду

$$\langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \hat{H} | \Psi \rangle = \langle \delta a''_i | \hat{F}^a | a_i \rangle - \sum_{k=1}^{n_a} \langle \delta a''_i | a_k \rangle \langle a_k | \hat{F}^a | a_i \rangle . \quad (6.185)$$

Теперь, предположив, что мы имеем дело с каноническими орбиталями НХФ, для которых  $\langle a_k | \hat{F}^a | a_i \rangle = \varepsilon_i^a \delta_{ik}$ , получим

$$\langle \Psi_1(a_i \rightarrow \delta a''_{i,2}) | \hat{H} | \Psi \rangle = \langle \delta a''_i | \hat{F}^a | a_i \rangle - \varepsilon_i^a \langle \delta a''_i | a_i \rangle , \quad (6.186)$$

и аналогичное выражение верно для орбиталей со спином  $\beta$ . Используя этот результат, вводя  $P$ -матрицы  $\mathbf{P}^a, \mathbf{P}^b$ , а также определяя «матрицы плотности энергии»  $\mathbf{W}^a, \mathbf{W}^b$  с элементами

$$W_{\mu\nu}^a = \sum_{i=1}^{n_a} \varepsilon_i^a a_{\mu}^i a_{\nu}^{i*}, \quad W_{\mu\nu}^b = \sum_{i=1}^{n_b} \varepsilon_i^b b_{\mu}^i b_{\nu}^{i*}, \quad (6.187)$$

и их бесспиновые аналоги  $\mathbf{D} = \mathbf{P}^a + \mathbf{P}^b$ ;  $\mathbf{W} = \mathbf{W}^a + \mathbf{W}^b$ , мы получим элементарными выкладками вариацию энергии, вызванную вариацией  $\delta\mathbf{R}$ , в виде

$$\begin{aligned} \delta E = & \sum_{\mu,\nu} \left[ P_{\nu\mu}^a \langle \chi_{\mu} | \delta \hat{h} | \chi_{\nu} \rangle + \{ P_{\nu\mu}^a \langle \delta \chi_{\mu} | \hat{F}^a | \chi_{\nu} \rangle - W_{\nu\mu}^a \langle \delta \chi_{\mu} | \chi_{\nu} \rangle + \text{c.c.} \} \right] \\ & + \sum_{\mu,\nu} \left[ P_{\nu\mu}^b \langle \chi_{\mu} | \delta \hat{h} | \chi_{\nu} \rangle + \{ P_{\nu\mu}^b \langle \delta \chi_{\mu} | \hat{F}^b | \chi_{\nu} \rangle - W_{\nu\mu}^b \langle \delta \chi_{\mu} | \chi_{\nu} \rangle + \text{c.c.} \} \right]. \end{aligned} \quad (6.188)$$

Из определения оператора Фока следует, что

$$\langle \delta \chi_{\mu} | \hat{F}^a | \chi_{\nu} \rangle = \langle \delta \chi_{\mu} | \hat{h} | \chi_{\nu} \rangle + \sum_{\rho,\tau} \{ D_{\tau\rho} [\delta \chi_{\mu} \chi_{\rho} | \chi_{\nu} \chi_{\tau}] - P_{\tau\rho}^a [\delta \chi_{\mu} \chi_{\rho} | \chi_{\tau} \chi_{\nu}] \}, \quad (6.189)$$

где использована «запись [12|12]» для двухэлектронных интегралов. Аналогичное выражение верно и для  $\langle \delta \chi_{\mu} | \hat{F}^b | \chi_{\nu} \rangle$ . Подставляя результат (6.189) в (6.188) и воспользовавшись симметрией двухэлектронных интегралов  $[\chi_{\mu} \chi_{\rho} | \chi_{\nu} \chi_{\tau}] = [\chi_{\rho} \chi_{\mu} | \chi_{\tau} \chi_{\nu}]$ , мы получаем после простых выкладок (таких, как разложение суммы на две идентичных половины и перестановка некоторых индексов) выражение

$$\begin{aligned} \delta E = & \sum_{\mu,\nu} \left[ D_{\nu\mu} \langle \chi_{\mu} | \delta \hat{h} | \chi_{\nu} \rangle + (D_{\nu\mu} \langle \delta \chi_{\mu} | \hat{h} | \chi_{\nu} \rangle - W_{\nu\mu} \langle \delta \chi_{\mu} | \chi_{\nu} \rangle + \text{c.c.}) \right] \\ & + \frac{1}{2} \sum_{\mu,\nu,\rho,\tau} \{ ([\delta \chi_{\mu} \chi_{\rho} | \chi_{\nu} \chi_{\tau}] + [\chi_{\mu} \delta \chi_{\rho} | \chi_{\nu} \chi_{\tau}]) (D_{\nu\mu} D_{\tau\rho} \\ & - P_{\nu\rho}^a P_{\tau\mu}^a - P_{\nu\rho}^b P_{\tau\mu}^b) + \text{c.c.} \}. \end{aligned} \quad (6.190)$$

Теперь выпишем явно комплексно сопряженные выражения, замечая, что

$$\begin{aligned} \delta S_{\mu\nu} &= \delta \langle \chi_{\mu} | \chi_{\nu} \rangle = \langle \delta \chi_{\mu} | \chi_{\nu} \rangle + \langle \chi_{\mu} | \delta \chi_{\nu} \rangle; \\ \delta h_{\mu\nu} &= \delta \langle \chi_{\mu} | \hat{h} | \chi_{\nu} \rangle = \langle \delta \chi_{\mu} | \hat{h} | \chi_{\nu} \rangle + \langle \chi_{\mu} | \delta \hat{h} | \chi_{\nu} \rangle + \langle \chi_{\mu} | \hat{h} | \delta \chi_{\nu} \rangle; \\ \delta [\chi_{\mu} \chi_{\rho} | \chi_{\nu} \chi_{\tau}] &= [\delta \chi_{\mu} \chi_{\rho} | \chi_{\nu} \chi_{\tau}] + [\chi_{\mu} \delta \chi_{\rho} | \chi_{\nu} \chi_{\tau}] \\ &+ [\chi_{\mu} \chi_{\rho} | \delta \chi_{\nu} \chi_{\tau}] + [\chi_{\mu} \chi_{\rho} | \chi_{\nu} \delta \chi_{\tau}], \end{aligned} \quad (6.191)$$

переставим некоторые индексы и получим окончательную формулу для изменения энергии НХФ  $\delta E$ , вызванного изменением геометрии  $\delta \mathbf{R}$ :

$$\begin{aligned} \delta E = & \sum_{\mu, \nu} (D_{\nu\mu} \delta h_{\mu\nu} - W_{\nu\mu} \delta S_{\mu\nu}) \\ & + \frac{1}{2} \sum_{\mu, \nu, \rho, \tau} (D_{\nu\mu} D_{\tau\rho} - P_{\nu\rho}^a P_{\tau\mu}^a - P_{\nu\rho}^b P_{\tau\mu}^b) \delta[\chi_\mu \chi_\rho | \chi_\nu \chi_\tau]. \end{aligned} \quad (6.192)$$

В случае ОХФ это выражение сводится просто к

$$\begin{aligned} \delta E = & \sum_{\mu, \nu} (D_{\nu\mu} \delta h_{\mu\nu} - W_{\nu\mu} \delta S_{\mu\nu}) \\ & + \frac{1}{2} \sum_{\mu, \nu, \rho, \tau} (D_{\nu\mu} D_{\tau\rho} - \frac{1}{2} D_{\nu\rho} D_{\tau\mu}) \delta[\chi_\mu \chi_\rho | \chi_\nu \chi_\tau]. \end{aligned} \quad (6.193)$$

*Альтернативный вывод с помощью множителей Лагранжа*

Введенный в выражении (6.46) вспомогательный функционал  $F$ , который использовался при выводе уравнений НХФ по методу множителей Лагранжа, имеет то свойство, что его значения всегда равны однодетерминантной энергии, если выполняются дополнительные условия ортонормировки. Поэтому градиенты энергии и вспомогательного функционала  $F$  для этого класса волновых функций также должны совпадать. При рассмотрении «физической» вариации орбиталей, вызванной вариацией геометрии  $\delta \mathbf{R}$ , мы работаем с орбиталями, меняющимися при переходе от одного набора ортонормированных орбиталей Хартри—Фока, соответствующих начальной геометрии, к другому такому набору, получаемому после изменения геометрии  $\delta \mathbf{R}$ . Это означает, что при такой вариации сохраняется ортонормировка орбиталей.

Отметим, что величина  $E$  (6.46) представляет собой выражение (6.44) для энергии для случая ортонормированных орбиталей. Если орбиталь  $a_i$  получает приращение  $\delta a_i$ , то вариация энергии, в соответствии с выражением (6.50), равна  $\langle \delta a_i | \hat{F}^a | a_i \rangle + \text{с.с.}$  В данном случае изменение геометрии  $\delta \mathbf{R}$  приведет к вариации всех орбиталей, так что мы должны просуммировать по всем ним. Необходимо также учесть, что одноэлектронная часть гамильтониана тоже меняется с геометрией. Тогда, варьируя также слагаемые, содержащие множители Лагранжа, мы получаем, что

$$\begin{aligned} \delta E = \delta F = & \sum_{i=1}^{n_a} \left[ \langle \delta a_i | \hat{F}^a | a_i \rangle - \sum_{j=1}^{n_a} \varepsilon_{ji}^a \langle \delta a_i | a_j \rangle + \text{с.с.} \right] \\ & + \sum_{i=1}^{n_b} \left[ \langle \delta b_i | \hat{F}^b | b_i \rangle - \sum_{j=1}^{n_b} \varepsilon_{ji}^b \langle \delta b_i | b_j \rangle + \text{с.с.} \right] \\ & + \sum_{i=1}^{n_a} \langle a_i | \delta \hat{h} | a_i \rangle + \sum_{i=1}^{n_b} \langle b_i | \delta \hat{h} | b_i \rangle. \end{aligned} \quad (6.194)$$



Теперь можно снова представить вариацию орбиталей в виде двух слагаемых, одно из которых содержит вариации коэффициентов разложения молекулярных орбиталей, а другое — вариации самих базисных орбиталей:

$$\delta a_i = \delta a'_i + \delta a''_i = \sum_{\mu=1}^m \delta a_{\mu}^i \chi_{\mu} + \sum_{\mu=1}^m a_{\mu}^i \delta \chi_{\mu}, \quad (6.195)$$

и аналогично для  $\delta b_i$ . Вариации  $\delta a'_i$  и  $\delta b'_i$  являются «математическими» вариациями, допустимыми при выводе уравнений ХФ, в котором потребовалось выполнение условий  $\delta F = 0$  при данной геометрической конфигурации, так что они не дают вклада в градиент. Поэтому в (6.194) должны рассматриваться только вариации  $\delta a''_i$  и  $\delta b''_i$ . Мы снова можем предположить, что используются канонические орбитали, т. е.  $\varepsilon_{ji}^a = \varepsilon_i^a \delta_{ji}$ ;  $\varepsilon_{ji}^b = \varepsilon_i^b \delta_{ij}$ . Тогда получим

$$\begin{aligned} \delta E = & \sum_{i=1}^{n_a} \left( \langle \delta a''_i | \hat{F}^a | a_i \rangle - \varepsilon_i^a \langle \delta a''_i | a_i \rangle + \text{с.с.} \right) \\ & + \sum_{i=1}^{n_b} \left( \langle \delta b''_i | \hat{F}^b | b_i \rangle - \varepsilon_i^b \langle \delta b''_i | b_i \rangle + \text{с.с.} \right) \\ & + \sum_{i=1}^{n_a} \langle a_i | \delta \hat{h} | a_i \rangle + \sum_{i=1}^{n_b} \langle b_i | \delta \hat{h} | b_i \rangle. \end{aligned} \quad (6.196)$$

Подставляя сюда разложение орбиталей и их вариаций, и вводя матрицы  $\mathbf{P}^a, \mathbf{P}^b, \mathbf{W}^a, \mathbf{W}^b$ , приходим к (6.188) и, повторив все последующие шаги, к (6.192) или (6.193).

## Библиографические заметки

### Раздел 1.

Оба доказательства теоремы Бриллюэна были опубликованы в [1]. (Также см. [2].) Алгоритм, напоминающий тот, который описан в разделе 1.3, использовался Доди [3] для получения локализованных орбиталей непосредственно из вариационного принципа, а не путем *a posteriori* локализации канонических орбиталей. (Похожий алгоритм, основанный на обобщенной теореме Бриллюэна теории спин-спроектированного «расширенного» метода Хартри—Фока, был использован в [4].)

Теорема Бриллюэна первоначально [5] рассматривалась, как следствие уравнений Хартри—Фока, и все еще рассматривается именно так многими авторами.

### Раздел 2.1.

Вывод в основном следует [1]. Возможность вывода уравнений Хартри—Фока из теоремы Бриллюэна обсуждалась также в [6, 7], но наш вывод более общий.

### Раздел 2.2.

Вывод с помощью множителей Лагранжа является наиболее стандартным (например, [8]). Проблема, связанная с использованием множителей Лагранжа, обсуждалась в [2] и [9].

### Раздел 2.3.

Неопубликованные результаты автора. (Техника такого типа была использована автором для альтернативного вывода [10] спин-спроектированных расширенных уравнений Хартри—Фока.)

### Раздел 3.

Теорема Купманса имеет исключительную важность для понимания электронных спектров любого типа, а также ионизационных процессов. Была предложена Купмансом [11] (Нобелевская премия по экономике<sup>1</sup>, 1975). Согласно автобиографии Купманса [12], доказательство теоремы было дано Крамерсом.

### Раздел 4.

В первые десятилетия существования метода Хартри—Фока его понимали в смысле ОХФ. (Для атомов обычно использовался сферически усредненный эффективный потенциал.) Схема НХФ (РОРС) была введена Поплом и Несбетом [13] и Бертье [14]. (См. также важные работы Слэтера [15, 16], который был, вероятно, первым, кто изучал проблему «диссоциационной катастрофы» и появление особого решения НХФ на больших межатомных расстояниях. Для исследования математических аспектов «критической точки», где имеет место бифуркация решений ОХФ и НХФ, упомянем [17]; см. также [18]. Машер указал на возможность точечных разрывов в случае принудительного поддержания полной симметрии [19].

Спин-спроектированный оператор, служащий для решения «дилеммы симметрии», был введен Лёвдином [20]. Впервые он был широко применен в  $\pi$ -электронном «Метод Альтернативных Молекулярных Орбиталей» [21]. Спин-спроектированный расширенный метод Хартри—Фока [22] можно рассматривать, как обобщение одноэлектронного приближения, в котором волновая функция является *спин-спроектированным* детерминантом Слэтера. Обзор см. в [2]. (Существование особых решений РХФ для любых геометрий можно вывести из результатов [23].)

### Раздел 5.1.

Уравнения ХФР были независимо выведены Ротаном (Рутаном) [24] и Холлом [25]. Приведенный вывод принадлежит автору [26].

### Раздел 5.2.

$P$ -матрица (точнее «матрица зарядов и порядков связей»  $\mathbf{D} = 2\mathbf{P}$ ) была введена Коулсоном [27] в рамках простой  $\pi$ -электронной теории Хюккеля. Относительно обсуждения соотношения между матрицей  $\mathbf{P}$  и «одноэлектронной матрицей плотности»  $\rho_1(x, x')$  в случае неортогонального базиса упомянем [28].

### Раздел 5.3.

Эти уравнения впервые были выведены другим методом [29]. Метод, позво-

<sup>1</sup> Совместно с Л. В. Канторовичем. — Прим. ред.

ляющий перейти от неэрмитового уравнения (6.122) к эрмитовому уравнению (6.123), был предложен Куприевичем [30].

#### Раздел 6.

Уравнения РОНФ впервые были систематически изучены Рутаном [31]. Наше изложение в основном следует [32], за исключением преобразования, ведущего от уравнения (6.158) к уравнению (6.160), которое похоже на использованное в [33] и основано на неопубликованном результате Скленера [34].

#### Раздел 7.

Вывод градиента, описанный вначале, взят из [35], второй (стандартный) основан на [36, 37].



### Для дальнейшего изучения

Метод Хартри—Фока имеет фундаментальное значение в квантовой химии; все учебники рассматривают его детально и содержат полезную информацию о нем. Весьма субъективно можем указать классическую работу Слэтера [38], исчерпывающую монографию Фудзинаги [39] и замечательные учебники МакВини [40] и Сабо и Остлунда [41]. Для сравнения разных методов Хартри—Фока можно порекомендовать обзор автора [2]. Таблицы почти хартри-фовских волновых функций свободных атомов и ионов, вычисленные Клементи и Роегги [42] с помощью орбиталей слэтеровского типа, являются очень полезным источником справочных данных.

### Литература

1. Mayer I. *Acta Phys. Hung.* **30**, 373 (1971).
2. Mayer I. *Adv. Quantum Chem.* **12**, 189 (1980).
3. Daudey J. P. *Chem. Phys. Letters* **24**, 574 (1974).
4. Lefebvre R., Smeyers Y. G. *Int. J. Quantum Chem.* **1**, 403 (1967).
5. Brillouin L. *Actual. Sci. Ind.*, **71** (1933), **159** (1934).
6. Dahl J. P., Johansen H., Truax D. R., Ziegler T. *Chem. Phys. Letters* **6**, 64 (1970).
7. Lefebvre R. *Cahiers Phys. (Paris)* **13**, 369 (1959).
8. Bethe H. A. *Intermediate Quantum Mechanics*, Benjamin, New York 1964.
9. Goddard W. A. III, Dunning T. H. Jr., Hunt W. J. *Chem. Phys. Letters* **4**, 231 (1969).
10. Mayer I. *Acta Phys. Hung.* **37**, 39 (1974).
11. Koopmans T. C. *Physica* **1**, 104 (1933).
12. См. веб-страницу  
<http://www.nobelprize.org/economics/laureates/1975/koopmans-autobio.html>
13. Pople J. A., Nesbet R. K. J. *Chem. Phys.* **22**, 571 (1954).
14. Berthier G. J. *Chim. Phys.* **51**, 363 (1954).
15. Slater J. C. *Phys. Rev.* **35**, 509 (1930).
16. Slater J. C. *Phys. Rev.* **82**, 538 (1951).
17. Mayer I. *Acta Phys. Hung.* **54**, 249 (1983).
18. Löwdin P.-O., Mayer I. *Adv. Quantum Chem.* **24**, 79 (1992).
19. Musher J. I. *Chem. Phys. Lett.* **7**, 397 (1970).
20. Löwdin P.-O. *Phys. Rev.* **97**, 1510 (1955).
21. Pauncz R. *Alternant Molecular Orbital Method* Saunders, Philadelphia 1967.

22. Löwdin P.-O., in *Quantum Theory of Atoms, Molecules, and the Solid State*, ed. P.-O. Löwdin, Academic Press, New York 1966.
23. Karadakov P., Castaño O. *Int. J. Quant. Chem.* **24**, 453 (1983).
24. Roothaan C. C. J. *Rev. Mod. Phys.* **23**, 69 (1951).
25. Hall G. G. *Proc. Roy. Soc. (London)* **A205**, 541 (1951).
26. Mayer I. *Acta Phys. Hung.* **36**, 11 (1974).
27. Coulson C. A. *Proc. Roy. Soc. (London)* **A169**, 413 (1939).
28. Mayer I. *J. Mol. Struct. (Theochem)* **255**, 1 (1992).
29. Mayer I. *Acta Phys. Hung.* **34**, 83 (1973).
30. Куприевич В. А., частное сообщение, Киев, 1971.
31. Roothaan C. C. J. *Rev. Mod. Phys.* **32**, 179 (1960).
32. Carbo R., Gropen O. *Adv. Quantum Chem.* **12**, 159 (1980).
33. Mayer I., Ladik J. and Biczó G., *Int. J. Quantum Chem.* **7**, 583 (1973).
34. Sklenar H., частное сообщение, Будапешт, 1972.
35. Ruiz M. B., Mayer I. *Chem. Phys. Letters* **236**, 217 (1995).
36. Pulay P. in *Modern Theoretical Chemistry* vol. 4., ed. Schaefer III, Plenum Press, New York 1977.
37. Pulay P. *Adv. Chem. Phys.* **69**, 241 (1987).
38. Slater J. C. *Quantum Theory of Molecules and Solids vol. 1. Electronic Structure of Molecules*, McGraw-Hill, New York 1963. (Имеется русский перевод: Слэтер Дж. *Электронная структура молекул.* — М.: Мир, 1965).
39. Фудзинага С. *Метод молекулярных орбиталей.* — М.: Мир, 1983.
40. McWeeny R. *Methods of Molecular Quantum Mechanics*, Academic Press, London 1992. (Имеется русский перевод первого издания: МакВини Р., Сатклиф С. *Квантовая механика молекул.* — М.: Мир, 1973).
41. Szabo A., Ostlund N. S. *Modern Quantum Chemistry*, McGraw Hill, New York 1989.
42. Clementi E., Roetti C. *Roothaan–Hartree–Fock Atomic Wavefunctions*, Academic Press, New York 1974. (*Atomic Data and Nuclear Data Tables* **14**, 177, 1974)

# Анализ заселенностей, порядки связей и валентности

При проведении квантовохимических исследований обычно основное внимание уделяется значениям полной энергии и различным величинам, связанным с ней, таким, как геометрические параметры, отвечающие минимуму энергии, колебательные частоты, определяемые формой поверхности потенциальной энергии в окрестности минимума энергии, и т. д. Часто, однако, необходимо добиться лучшего понимания изучаемой системы, используя для этого не только энергетические характеристики, но и информацию, содержащуюся в волновой функции. Сама волновая функция обычно содержит слишком много чисел для того, чтобы непосредственно служить этой цели, и необходимо выполнить некоторое «сжатие данных», чтобы какая-либо интерпретация полученных результатов стала возможной.

Очень важной *измеримой физической величиной* является электронная плотность, т. е. распределение электронов в физическом (трехмерном) пространстве. Ее детальный анализ, проводимый, скажем, в теории «атомов-в-молекулах» Бэйдера, представляется интересным, но выходит за рамки нашей книги. С другой стороны, имеет смысл рассмотреть распределение электронов между базисными орбиталями, и в особенности между разными атомами, и определить результирующие заряды, которые можно приписать отдельным атомам молекулы. Можно также искать различные другие параметры, характеризующие валентное состояние атома в молекуле и химическое связывание между отдельными атомами. Мы опишем в этой главе некоторые такие параметры. Они часто бывают полезны для интерпретации результатов вычислений и могут иметь даже определенную предсказательную силу. Тем не менее необходимо помнить, что ни одна из величин, обсуждаемых в этой главе, не является истинной *измеримой физической величиной*, за исключением пространственных электронной и спиновой плотностей; эти величины всего лишь служат для представления результатов расчетов в относительно простой и сжатой форме через понятия, которые могут быть соотнесены с интуитивными химическими представлениями. (Ни

базисная орбиталь, ни ее заселенность не являются измеримыми физическими величинами.)

В отличие от исследования электронного распределения в физическом пространстве, методы, рассматриваемые в этой главе, описывают анализ результатов, проводимый в гильбертовом пространстве, растянутом базисными орбиталями.

## 1. Анализ заселенностей по Малликену

### 1.1. Электронная плотность

В гл. 2, разд. 2.1 мы уже приводили выражение для электронной плотности  $\varrho(\vec{r})$ , которое получается при интегрировании по координатам всех электронов, кроме одного, и суммировании по всем электронам. Электронную плотность в точке  $\vec{r}$  можно рассматривать, как среднее значение оператора  $\hat{\varrho}(\vec{r})$ , представляющего собой сумму одноэлектронных операторов, симметричную по перестановкам координат отдельных электронов:

$$\hat{\varrho}(\vec{r}) = \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) . \quad (7.1)$$

Так как заряд электрона отрицателен, электронная плотность  $\varrho(\vec{r})$  соответствует электростатической плотности заряда  $-\varrho(\vec{r})$ , если использовать атомные единицы. (В иных случаях необходимо писать  $-e_0\varrho(\vec{r})$ , где  $e_0$  — элементарный заряд.)

Если многоэлектронная волновая функция является детерминантом Слэтера  $\Psi = \hat{A}[\psi_1(1)\psi_2(2)\dots\psi_N(N)]$ , построенным из ортонормированных спин-орбиталей, то можно воспользоваться правилом Слэтера (5.100), и, аналогично случаю других одноэлектронных величин, представить электронную плотность  $\varrho(\vec{r})$  в виде суммы вкладов от отдельных занятых спин-орбиталей  $\psi_i(\vec{r}, \sigma) = \varphi_i(\vec{r})\gamma(\sigma)$  в виде

$$\begin{aligned} \varrho(\vec{r}) &= \langle \Psi | \hat{\varrho}(\vec{r}) | \Psi \rangle = \sum_{j=1}^N \langle \psi_j(\vec{r}_1, \sigma_1) | \delta(\vec{r} - \vec{r}_1) | \psi_j(\vec{r}_1, \sigma_1) \rangle \\ &= \sum_{j=1}^N \langle \varphi_j(\vec{r}_1) | \delta(\vec{r} - \vec{r}_1) | \varphi_j(\vec{r}_1) \rangle = \sum_{j=1}^N \varphi_j^*(\vec{r}) \varphi_j(\vec{r}) = \sum_{j=1}^N |\varphi_j(\vec{r})|^2 . \end{aligned} \quad (7.2)$$

В частности, если  $\Psi$  содержит  $n_a$  орбиталей  $a_i$ , занятых электронами с проекцией спина  $\alpha$ , и  $n_b$  орбиталей  $b_i$ , занятых электронами с

проекцией спинами  $\beta$ , то выражение (7.2) примет вид

$$\varrho(\vec{r}) = \sum_{j=1}^{n_a} |a_j(\vec{r})|^2 + \sum_{j=1}^{n_b} |b_j(\vec{r})|^2. \quad (7.3)$$

Конечно, электронная плотность инвариантна относительно унитарных преобразований наборов орбиталей  $\{a_i\}$  или  $\{b_i\}$ , так как унитарные преобразования оставляют инвариантной однодетерминантную волновую функцию  $\Psi$ . Например, если

$$a_j(\vec{r}) = \sum_{k=1}^{n_a} U_{kj} a'_k(\vec{r}), \quad (7.4)$$

и матрица  $\mathbf{U}$  унитарна ( $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ ), то

$$\begin{aligned} \sum_{j=1}^{n_a} |a_j(\vec{r})|^2 &= \sum_{j=1}^{n_a} a_j^*(\vec{r}) a_j(\vec{r}) = \sum_{j=1}^{n_a} \sum_{k=1}^{n_a} U_{kj}^* a'_k(\vec{r}) \sum_{l=1}^{n_a} U_{lj} a'_l(\vec{r}) \\ &= \sum_{j,k,l=1}^{n_a} a'_k(\vec{r}) a'_l(\vec{r}) U_{lj} (\mathbf{U}^\dagger)_{jk} = \sum_{k,l=1}^{n_a} a'_k(\vec{r}) a'_l(\vec{r}) \delta_{kl} \\ &= \sum_{k=1}^{n_a} a'_k(\vec{r}) a'_k(\vec{r}) = \sum_{k=1}^{n_a} |a'_k(\vec{r})|^2 \quad \text{Ч.т.д.} \end{aligned} \quad (7.5)$$

Согласно предыдущим результатам, электронная плотность в случае однодетерминантной волновой функции ведет себя так, как будто мы имеем дело с совокупностью независимых электронов, занимающих ортонормированные орбитали. Однако мы снова подчеркнем, что нельзя приписывать слишком большой физический смысл этим *отдельным* орбиталиям, так как любой набор ортонормированных орбиталей, который растягивает данное подпространство занятых орбиталей, будет вести себя точно так же и давать ту же самую электронную плотность. (То же верно, конечно, и для любого другого одноэлектронного свойства.)

Если использовать разложения ЛКАО вида (6.90), (6.115) то электронную плотность можно также выразить через базисные орбитали  $\chi_\mu$ :

$$\begin{aligned} \varrho(\vec{r}) &= \sum_{j=1}^{n_a} a_j^*(\vec{r}) a_j(\vec{r}) + \sum_{j=1}^{n_b} b_j^*(\vec{r}) b_j(\vec{r}) \\ &= \sum_{j=1}^{n_a} \sum_{\nu=1}^m a_\nu^{j*} \chi_\nu^*(\vec{r}) \sum_{\mu=1}^m a_\mu^j \chi_\mu(\vec{r}) + \sum_{j=1}^{n_b} \sum_{\nu=1}^m b_\nu^{j*} \chi_\nu^*(\vec{r}) \sum_{\mu=1}^m b_\mu^j \chi_\mu(\vec{r}). \end{aligned} \quad (7.6)$$

С помощью элементов  $P$ -матрицы, определенных в выражении (6.116), это выражение можно переписать как

$$\varrho(\vec{r}) = \sum_{\mu, \nu=1}^m (P_{\mu\nu}^a + P_{\mu\nu}^b) \chi_{\nu}^*(\vec{r}) \chi_{\mu}(\vec{r}) = \sum_{\mu, \nu=1}^m D_{\mu\nu} \chi_{\nu}^*(\vec{r}) \chi_{\mu}(\vec{r}). \quad (7.7)$$

Здесь мы ввели «бесспиновую матрицу плотности»  $\mathbf{D}$ :

$$\mathbf{D} = \mathbf{P}^a + \mathbf{P}^b. \quad (7.8)$$

В случае метода ОХФ это определение совпадает с определением (6.110). Уравнение (7.7), очевидно, содержит сумму двух плотностей, которые можно сопоставить с электронами, имеющими проекции спинов  $\alpha$  и  $\beta$ . Можно также определить спиновую плотность  $\varrho_s(\vec{r})$  как разность этих плотностей

$$\varrho_s(\vec{r}) = \sum_{\mu, \nu=1}^m (P_{\mu\nu}^a - P_{\mu\nu}^b) \chi_{\nu}^*(\vec{r}) \chi_{\mu}(\vec{r}) = \sum_{\mu, \nu=1}^m P_{\mu\nu}^s \chi_{\nu}^*(\vec{r}) \chi_{\mu}(\vec{r}), \quad (7.9)$$

где

$$\mathbf{P}^s = \mathbf{P}^a - \mathbf{P}^b \quad (7.10)$$

является «матрицей спиновой плотности».

Если проинтегрировать электронную плотность  $\varrho(\vec{r})$  по всему пространству, то должно, очевидно, получиться число электронов  $N = n_a + n_b$ . Это следует из того, что орбитали  $a_i$  и  $b_i$  нормированы, так что каждое слагаемое в суммах (7.3) вносит единичный вклад в интеграл  $\int \varrho(\vec{r}) dv$ . С помощью разложения (7.7) интегралы от электронной плотности можно выразить через следы матриц  $\mathbf{P}^a \mathbf{S}$ ,  $\mathbf{P}^b \mathbf{S}$ , и  $\mathbf{D} \mathbf{S}$ :

$$\begin{aligned} \sum_{\mu, \nu=1}^m P_{\mu\nu}^a S_{\nu\mu} &= \sum_{\mu=1}^m (\mathbf{P}^a \mathbf{S})_{\mu\mu} = \text{Tr}(\mathbf{P}^a \mathbf{S}) = n_a \\ \sum_{\mu, \nu=1}^m P_{\mu\nu}^b S_{\nu\mu} &= \sum_{\mu=1}^m (\mathbf{P}^b \mathbf{S})_{\mu\mu} = \text{Tr}(\mathbf{P}^b \mathbf{S}) = n_b \\ \sum_{\mu, \nu=1}^m D_{\mu\nu} S_{\nu\mu} &= \sum_{\mu=1}^m (\mathbf{D} \mathbf{S})_{\mu\mu} = \text{Tr}(\mathbf{D} \mathbf{S}) = N. \end{aligned} \quad (7.11)$$

Мы должны заметить, что разложение (7.7) и равенства (7.11) верны для любой  $N$ -электронной волновой функции с  $n_a$  электронов с проекцией спина  $\alpha$  и  $n_b$  электронов с проекцией спина  $\beta$ , т. е. имеющей проекцию полного спина  $S_z = \frac{1}{2}(n_a - n_b)$ . Это так, потому что разложение (7.7)



имеет наиболее общий возможный вид для волновых функций, построенных из  $m$  базисных орбиталей  $\chi_\mu$ . Однако свойства идемпотентности  $(\mathbf{P}^a \mathbf{S})^2 = \mathbf{P}^a \mathbf{S}$ ;  $(\mathbf{P}^b \mathbf{S})^2 = \mathbf{P}^b \mathbf{S}$  выполняются только в том случае, если волновая функция является однодетерминатной.

## 1.2. Анализ заселенностей

Согласно формулам (7.11), след матрицы  $\mathbf{DS}$  равен числу электронов в системе. Эти выражения также определяют некоторое распределение электронов по базисным орбиталям  $\chi_\mu$ . Поэтому величина

$$q_\mu = (\mathbf{DS})_{\mu\mu} = \sum_{\nu=1}^m D_{\mu\nu} S_{\nu\mu} \quad (7.12)$$

называется «полной орбитальной заселенностью по Малликену» базисной орбитали  $\chi_\mu$ , а сумма величин  $q_\mu$  по  $\mu$  дает полное число электронов.

Можно формально разбить величину  $q_\mu$  на два слагаемых:

$$q_\mu = D_{\mu\mu} + \sum_{\substack{\nu=1 \\ (\nu \neq \mu)}}^m D_{\mu\nu} S_{\nu\mu} . \quad (7.13)$$

(Базисные функции предполагаются нормированными,  $S_{\mu\mu} = 1$ .)

Первое слагаемое называется «чистой орбитальной заселенностью по Малликену» орбитали  $\chi_\mu$ ; оно получается при интегрировании компоненты  $D_{\mu\mu} \chi_\mu^*(\vec{r}) \chi_\mu(\vec{r})$  в разложении (7.7). Второе слагаемое получается при интегрировании компонент  $D_{\mu\nu} \chi_\nu^*(\vec{r}) \chi_\mu(\vec{r})$  с  $\nu \neq \mu$ . Его появление связано с существованием «плотности перекрывания»  $\chi_\nu^*(\vec{r}) \chi_\mu(\vec{r})$ , которая при интегрировании давала бы нуль только в случае ортогонального базиса.

Каждая пара орбиталей  $\chi_\mu$  и  $\chi_\nu$  вносит вклад  $\chi_\nu^*(\vec{r}) \chi_\mu(\vec{r}) + \chi_\mu^*(\vec{r}) \chi_\nu(\vec{r})$  в полную электронную плотность (7.7). Первое из этих слагаемых учитывается при вычислении полной орбитальной заселенности  $q_\mu$ , а второе дает вклад в полную орбитальную заселенность  $q_\nu$ . В случае вещественных орбиталей эти два слагаемых равны, что объясняет, почему часто говорят, что полная электронная плотность перекрывания  $2\chi_\mu(\vec{r})\chi_\nu(\vec{r})$  и соответствующая «заселенность перекрывания»  $2D_{\mu\nu}S_{\mu\nu}$  «делятся пополам» при вычислении полных заселенностей по Малликену  $q_\mu$  и  $q_\nu$ . (Если орбитали вещественны, то и  $D_{\mu\nu} = D_{\nu\mu}$ .) Автор не может согласиться с тем мнением, что это деление пополам является «произвольным»; напротив, оно является следствием соотношений (7.11), и поэтому является единственной возможностью, согласующейся

с математической структурой формализма, в котором перекрывания базисных орбиталей отличны от нуля (матрицы  $\mathbf{P}^a\mathbf{S}$  и  $\mathbf{P}^b\mathbf{S}$  суть матрицы идемпотентных проекционных операторов).

Если использовать такой базис, в котором каждая орбиталь центрирована на одном из атомов (что является общепринятым), то можно просуммировать полные орбитальные заселенности  $q_\mu$  всех орбиталей, принадлежащих данному атому  $A$  (эту принадлежность будем обозначать, как  $\mu \in A$ ), и получить полную атомную заселенность по Малликену для рассматриваемого атома:

$$Q_A = \sum_{\mu \in A} q_\mu = \sum_{\mu \in A} (\mathbf{DS})_{\mu\mu} = \sum_{\mu \in A} \sum_{\nu=1}^m D_{\mu\nu} S_{\nu\mu}. \quad (7.14)$$

Полную атомную заселенность  $Q_A$  снова можно естественным образом представить, как

$$Q_A = d_A + \sum_{\substack{B \\ (B \neq A)}} d_{AB}, \quad (7.15)$$

где величина  $d_A$  содержит слагаемые, в которых обе орбитали принадлежат атому  $A$ :

$$d_A = \sum_{\mu, \nu \in A} D_{\mu\nu} S_{\nu\mu}, \quad (7.16)$$

и она называется чистой атомной заселенностью по Малликену, тогда как величина, имеющая двухатомный характер:

$$d_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} D_{\mu\nu} S_{\nu\mu}, \quad (7.17)$$

называется «заселенностью перекрывания по Малликену». (Часто это название используется для величины  $d_{AB} + d_{BA}$ , которая равна  $2d_{AB}$ , если применяются вещественные орбитали.) Стоит указать, что чистая атомная заселенность  $d_A$  не есть просто сумма чистых орбитальных заселенностей, но содержит также заселенности перекрывания для всех пар орбиталей, которые центрированы на том же самом атоме.

Заселенность перекрывания по Малликену является существенно положительным числом (скажем, порядка  $\sim 0.5$ ) для пар химически связанных атомов и отражает накопление электронного заряда в области связывания. Ее значение обычно хорошо коррелирует с силой (прочностью) связи между атомами. С другой стороны, отрицательное значение заселенности перекрывания между химически несвязанными атомами

указывает на то, что вероятно отталкивательное взаимодействие между данными атомами.

### Инвариантность

На практике обычно используют базисные орбитали, которые центрированы на данном атоме и ориентированы относительно «лабораторной» системы координат — например, орбитали  $p_x, p_y, p_z$  ориентированы по осям  $x, y$ , и  $z$  именно лабораторной системы координат. Очевидно, если повернуть молекулу относительно лабораторной системы, то коэффициенты разложения каждой молекулярной орбитали по атомным орбиталям будут меняться, хотя молекулярная волновая функция останется той же самой. Это особенно очевидно, если представить такой процесс в виде вращения лабораторной системы вокруг молекулы. Изменение орбитальных коэффициентов приводит также и к изменению матрицы  $\mathbf{D}$ , хотя физическое состояние молекулы остается неизменным. (В частности, в  $D$ -матрице только элементы между орбиталями  $s$ -типа остаются прежними, а все другие меняются при вращениях.)<sup>1</sup>

Часто используют также представление о «гибридных атомных орбиталях», представляющих собой линейные комбинации базисных орбиталей  $s$  и  $p$  типа (иногда и  $d$ ) для обсуждения разных химических задач, в особенности направленного характера химических связей. Очевидно, переход к такому гибриднему базису не изменит окончательных результатов расчетов.

Согласно предшествующему обсуждению, для интерпретации можно использовать только величины, которые имеют «вращательно-гибридизационную инвариантность», т. е. остаются неизменными при любом невырожденном линейном преобразовании орбиталей, центрированных на данном атоме. (По этой причине ни отдельным матричным элементам  $D$ -матрицы, ни их сумме для данной пары атомов нельзя приписать какой-либо определенный смысл, в отличие от старых  $\pi$ -электронных теорий, в которых недиагональные элементы  $D_{\mu\nu}$  рассматривались, как «порядки связей».)

Мы докажем здесь вращательно-гибридизационную инвариантность малликеновской чистой и полной атомных заселенностей данного атома и малликеновской заселенности перекрывания пар атомов. Для этого рассмотрим невырожденные линейные преобразования базисных орбиталей, центрированных на одном и том же атоме:

$$\chi'_\mu = \sum_{\nu \in A} T_{\nu\mu}^A \chi_\nu \quad (\mu \in A). \quad (7.18)$$

<sup>1</sup> Аналогично и в разложении МО ЛКАО только коэффициенты при  $s$ -функциях вращательно инвариантны. — Прим. ред.

Вращательно-гибридизационные преобразования (7.18) базисных орбиталей всех атомов одновременно можно описать блочно-диагональной матрицей преобразования  $\mathbf{T}$ , диагональные блоки которой являются матрицами преобразования  $\mathbf{T}^A$  базисных функций, принадлежащих атому  $A$ :

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^2 & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{T}^M \end{pmatrix}, \quad (7.19)$$

где  $M$  есть число атомов в системе. С помощью матрицы  $\mathbf{T}$  преобразование всего базиса можно компактно записать как

$$\chi'_\mu = \sum_{\nu=1}^m T_{\nu\mu} \chi_\nu \quad \mu = 1, 2, \dots, m. \quad (7.20)$$

Так как каждое преобразование  $\mathbf{T}^A$  предполагается невырожденным, то же самое верно и для матрицы  $\mathbf{T}$ ; обратная матрица  $\mathbf{T}^{-1}$  снова является блочно-диагональной и ее диагональные блоки суть матрицы  $(\mathbf{T}^A)^{-1}$  соответствующих обратных матриц преобразований базисных функций отдельных атомов. Обратная матрица  $\mathbf{T}^{-1}$  позволяет определить преобразование, обратное к преобразованию (7.20):

$$\chi_\mu = \sum_{\nu=1}^m (\mathbf{T}^{-1})_{\nu\mu} \chi'_\nu. \quad (7.21)$$

Пользуясь этими матрицами, мы можем выразить каждую молекулярную орбиталь  $a_i = \sum_{\mu=1}^m a_\mu^i \chi_\mu$  через новые базисные орбитали:

$$\begin{aligned} a_i &= \sum_{\mu=1}^m a_\mu^i \chi_\mu = \sum_{\mu=1}^m a_\mu^i \sum_{\nu=1}^m (\mathbf{T}^{-1})_{\nu\mu} \chi'_\nu \\ &= \sum_{\nu=1}^m \left( \sum_{\mu=1}^m (\mathbf{T}^{-1})_{\nu\mu} a_\mu^i \right) \chi'_\nu = \sum_{\nu=1}^m a_{\nu}^{i'} \chi'_\nu \end{aligned} \quad (7.22)$$

т. е. векторы орбитальных коэффициентов преобразуются с помощью обратной матрицы  $\mathbf{T}$ :

$$\mathbf{a}^{i'} = \mathbf{T}^{-1} \mathbf{a}^i \quad (7.23)$$

и аналогично для орбиталей с проекцией спина  $\beta$ .

Согласно этим результатам, матрица  $\mathbf{D}$  преобразуется как

$$\begin{aligned}\mathbf{D}' &= \mathbf{P}^{a'} + \mathbf{P}^{b'} = \sum_{i=1}^{n_a} \mathbf{a}^{i'} (\mathbf{a}^{i'})^\dagger + \sum_{i=1}^{n_b} \mathbf{b}^{i'} (\mathbf{b}^{i'})^\dagger \\ &= \sum_{i=1}^{n_a} \mathbf{T}^{-1} \mathbf{a}^i \mathbf{a}^{i\dagger} \mathbf{T}^{-1\dagger} + \sum_{i=1}^{n_b} \mathbf{T}^{-1} \mathbf{b}^i \mathbf{b}^{i\dagger} \mathbf{T}^{-1\dagger} \\ &= \mathbf{T}^{-1} \mathbf{D} \mathbf{T}^{-1\dagger},\end{aligned}\quad (7.24)$$

где было использовано тождество  $(\mathbf{A}\mathbf{b})^\dagger = \mathbf{b}^\dagger \mathbf{A}^\dagger$ .

Матрица перекрывания преобразуется как

$$\begin{aligned}S'_{\mu\nu} &= \langle \chi'_\mu | \chi'_\nu \rangle = \left\langle \sum_{\varrho=1}^m T_{\varrho\mu} \chi_\varrho \middle| \sum_{\tau=1}^m T_{\tau\nu} \chi_\tau \right\rangle \\ &= \sum_{\varrho, \tau=1}^m T_{\varrho\mu}^* S_{\varrho\tau} T_{\tau\nu} = \sum_{\varrho, \tau=1}^m (\mathbf{T}^\dagger)_{\mu\varrho} S_{\varrho\tau} T_{\tau\nu}\end{aligned}\quad (7.25)$$

т. е.

$$\mathbf{S}' = \mathbf{T}^\dagger \mathbf{S} \mathbf{T} . \quad (7.26)$$

Чистая атомная заселенность  $d_A$  и межатомные заселенности перекрывания  $d_{AB}$  инвариантны по отношению к вращательно-гибридизационным преобразованиям, заданным матрицей  $\mathbf{T}$ ; следовательно, малликеновская полная атомная заселенность, которая, согласно определению (7.15), является их суммой, также инвариантна.

#### Доказательство

Для чистой атомной заселенности в преобразованном базисе имеем

$$\begin{aligned}d'_A &= \sum_{\mu, \nu \in A} D'_{\mu\nu} S'_{\nu\mu} = \sum_{\mu, \nu \in A} (\mathbf{T}^{-1} \mathbf{D} \mathbf{T}^{-1\dagger})_{\mu\nu} (\mathbf{T}^\dagger \mathbf{S} \mathbf{T})_{\nu\mu} \\ &= \sum_{\kappa, \lambda, \varrho, \tau=1}^m \sum_{\mu, \nu \in A} (\mathbf{T}^{-1})_{\mu\kappa} D_{\kappa\lambda} (\mathbf{T}^{-1\dagger})_{\lambda\nu} (\mathbf{T}^\dagger)_{\nu\varrho} S_{\varrho\tau} T_{\tau\mu} .\end{aligned}\quad (7.27)$$

Благодаря блочно-диагональной структуре матриц  $\mathbf{T}$ ,  $\mathbf{T}^{-1}$  и  $\mathbf{T}^\dagger$ ,  $\mathbf{T}^{-1\dagger}$ , при суммировании по  $\mu$  и  $\nu$  имеем

$$\sum_{\mu \in A} T_{\tau\mu} (\mathbf{T}^{-1})_{\mu\kappa} = \begin{cases} \delta_{\tau\kappa} & \text{если } \tau, \kappa \in A \\ 0 & \text{иначе} \end{cases}, \quad (7.28)$$

и

$$\sum_{\nu \in A} (\mathbf{T}^{-1\dagger})_{\lambda\nu} (\mathbf{T}^\dagger)_{\nu\varrho} = \begin{cases} \delta_{\lambda\varrho} & \text{если } \lambda, \varrho \in A \\ 0 & \text{иначе} \end{cases}, \quad (7.29)$$

соответственно. Поэтому при суммировании по  $\kappa, \lambda, \varrho$  и  $\tau$  можно ограничиться индексами, относящимися к атому  $A$ . Получаем

$$d'_A = \sum_{\kappa, \lambda, \varrho, \tau \in A} \delta_{\tau\kappa} D_{\kappa\lambda} \delta_{\lambda\varrho} S_{\varrho\tau} = \sum_{\tau, \varrho \in A} D_{\tau\varrho} S_{\varrho\tau} = d_A, \quad (7.30)$$

т. е. чистая атомная заселенность инвариантна. Заселенность перекрытия между атомами  $A$  и  $B$  также инвариантна. Имеем, как и в предыдущем случае:

$$\begin{aligned} d'_{AB} &= \sum_{\mu \in A} \sum_{\nu \in B} D'_{\mu\nu} S'_{\nu\mu} \\ &= \sum_{\kappa, \lambda, \varrho, \tau=1}^m \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{T}^{-1})_{\mu\kappa} D_{\kappa\lambda} (\mathbf{T}^{-1\dagger})_{\lambda\nu} (\mathbf{T}^\dagger)_{\nu\varrho} S_{\varrho\tau} T_{\tau\mu}. \end{aligned} \quad (7.31)$$

Для суммы по  $\mu$  снова имеем выражение (7.28), тогда как для суммы по  $\nu$  получаем выражение (7.29), но с заменой  $A$  на  $B$ . Таким образом, суммирование по  $\kappa$  и  $\tau$  должно быть ограничено атомом  $A$ , тогда как суммы по  $\lambda$  и  $\varrho$  ограничены атомом  $B$ :

$$d'_{AB} = \sum_{\kappa, \tau \in A} \sum_{\lambda, \varrho \in B} \delta_{\tau\kappa} D_{\kappa\lambda} \delta_{\lambda\varrho} S_{\varrho\tau} = \sum_{\tau \in A} \sum_{\varrho \in B} D_{\tau\varrho} S_{\varrho\tau} = d_{AB}. \quad \text{Ч. т. д.} \quad (7.32)$$

Можно заключить, что малликеновская заселенность перекрытия является простейшей величиной, линейной по межатомным элементам  $D$ -матрицы и обладающей необходимой вращательно-гибридизационной инвариантностью.

## 2. Порядки связей и валентности

### 2.1. Индекс порядка связи

Несмотря на то, что малликеновская заселенность перекрытия является параметром, который позволяет идентифицировать химически связанные атомы, ее нельзя поставить в соответствие такому понятию, как порядок (кратность) химической связи, так как она не принимает значений, близких к 1, 2 и 3 для одинарной, двойной и тройной связей, соответственно. Классический порядок связи между атомами в двухатомной молекуле можно определить, как  $\frac{1}{2}(N_{\text{связ}} - N_{\text{антисвяз}})$ , где  $N_{\text{связ}}$

и  $N_{\text{антисвяз}}$  равны числам электронов, занимающих связывающую и антисвязывающую орбитали. Таким образом, молекулярный ион  $\text{H}_2^+$  характеризуется порядком связи  $\frac{1}{2}$  (один электрон на связывающей орбитали), а молекула  $\text{H}_2$  имеет порядок связи 1 (2 электрона на связывающей МО), тогда как димер гелия имеет нулевой порядок связи. Можно описать  $\text{He}_2$  либо с помощью двух дважды занятых локализованных атомных  $1s$ -орбиталей, либо с помощью их суммы и разности, соответствующих делокализованным МО связывающего и антисвязывающего характера. Так как обе они дважды заняты,  $N_{\text{связ}} = N_{\text{антисвяз}} = 2$ , то получается порядок связи 0 между двумя атомами, в полном согласии с химической интуицией. Аналогично, получаем, например, порядок связи 3 для  $\text{N}_2$  и 1 для  $\text{F}_2$ . Очевидно, такой подсчет электронов в строгом смысле возможен только для двухатомных молекул.

В рамках простого метода Хюккеля (и вообще в  $\pi$ -электронных методах) недиагональные элементы матрицы **D** назывались порядками связей по Коулсону. Это определение, однако, нельзя обобщить на теории, в которых каждый атом может нести больше одной базисной орбитали. Простейшей квантовохимической величиной, которую можно поставить в соответствие классическому химическому представлению о кратности связи для многоатомных молекул, и вычисляемой из волновой функции, являлся так называемый индекс Вайберга. Первоначально он был введен в рамках полуэмпирической теории ППДП (CNDO) и пригоден только в случае, если базисные орбитали ортонормированы.

Индекс Вайберга для пары атомов  $A$  и  $B$  системы с замкнутой оболочкой определяется как

$$W_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} |D_{\mu\nu}|^2. \quad (7.33)$$

Адекватное обобщение этой формулы на системы с открытой оболочкой, описываемой в рамках однодетерминантного НХФ, есть

$$W_{AB} = 2 \sum_{\mu \in A} \sum_{\nu \in B} (|P_{\mu\nu}^a|^2 + |P_{\mu\nu}^b|^2) = \sum_{\mu \in A} \sum_{\nu \in B} (|D_{\mu\nu}|^2 + |P_{\mu\nu}^s|^2), \quad (7.34)$$

где матрицы **D**,  $\mathbf{P}^a$ ,  $\mathbf{P}^b$ , и  $\mathbf{P}^s$  определены соотношениями (7.8), (6.116) и (7.10). Эквивалентность двух представлений (7.34), можно тривиально проверить с помощью тождеств  $\mathbf{D} = \mathbf{P}^a + \mathbf{P}^b$ ;  $\mathbf{P}^s = \mathbf{P}^a - \mathbf{P}^b$ .

Записывая  $|D_{\mu\nu}|^2$ , как  $D_{\mu\nu}D_{\nu\mu}$ , и т. д., и используя соотношения (7.24), можно легко увидеть, что индекс Вайберга инвариантен по отношению к унитарным преобразованиям орбиталей, центрированных на

одном и том же атоме, в том числе и тем, которые вызваны вращениями системы координат (тогда  $\mathbf{T}^{-1\dagger} = \mathbf{T}$ ), но не по отношению к общему (невыврожденному, но не унитарному) линейному преобразованию. Поэтому индекс Вайберга, определенный выше, неприменим в *ab initio* теории, в которой требуется инвариантность также относительно таких преобразований (базисные орбитали могут быть неортогональными). Обобщение индекса Вайберга, подходящее для *ab initio* теории, было предложено автором. (Аналогичное определение было предложено Джьямбьяджи в рамках полуквантового «расширенного метода Хюккеля», в котором тоже применяют перекрывающиеся базисы.)

Соответствующие определения индекса порядка связи суть

$$B_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} \quad (7.35)$$

для случая замкнутой оболочки (ОХФ) и

$$\begin{aligned} B_{AB} &= 2 \sum_{\mu \in A} \sum_{\nu \in B} [(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu}] \\ &= \sum_{\mu \in A} \sum_{\nu \in B} [(\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} + (\mathbf{P}^s \mathbf{S})_{\mu\nu} (\mathbf{P}^s \mathbf{S})_{\nu\mu}] \end{aligned} \quad (7.36)$$

для случая открытой оболочки (НХФ). Опять-таки, эквивалентность двух форм записи (7.36) можно проверить с помощью определений  $\mathbf{D} = \mathbf{P}^a + \mathbf{P}^b$ ,  $\mathbf{P}^s = \mathbf{P}^a - \mathbf{P}^b$ . Очевидно, определение (7.36) сводится к предыдущему, если использовать его для волновой функции ОХФ замкнутой оболочки ( $\mathbf{P}^s = \mathbf{0}$ ). Далее, если базис ортонормирован ( $\mathbf{S} = \mathbf{1}$ ), то имеем  $B_{AB} = W_{AB}$ . Отметим также, что эти определения остаются в силе и для коррелированных волновых функций. В этом случае определения матриц  $\mathbf{P}^a$  и  $\mathbf{P}^b$  или  $\mathbf{D}$  и  $\mathbf{P}^s$  через орбитальные коэффициенты, разумеется, неприменимы, но эти матрицы можно определить как матрицы коэффициентов разложения электронной и спиновой плотностей (7.7) и (7.9).

### Инвариантность

Индекс порядка связи, определенный в (7.35) для случая замкнутой оболочки в методе ОХФ, представляет собой простейшую величина, которая квадратична по межатомным элементам матрицы  $\mathbf{D}$  и обладает вращательно-гибридизационной инвариантностью. Похожие рассуждения верны и для определения (7.36), применимого в случае НХФ, включающем две матрицы  $\mathbf{P}^a$  и  $\mathbf{P}^b$  или  $\mathbf{D}$  и  $\mathbf{P}^s$ . Мы докажем инвариантность



определения (7.35); матрицы, входящие в определение (7.36), ведут себя таким же образом.

Рассмотрим теперь индекс порядка связи (7.35) и покажем его инвариантность при преобразовании (7.18), имеющем блочно-диагональную матрицу (7.19). Имеем после преобразования

$$B'_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{D}'\mathbf{S}')_{\mu\nu} (\mathbf{D}'\mathbf{S}')_{\nu\mu}, \quad (7.37)$$

где преобразованные матрицы  $\mathbf{D}'$  и  $\mathbf{S}'$  заданы выражениями (7.24) и (7.26), соответственно. Для их произведения имеем

$$\mathbf{D}'\mathbf{S}' = \mathbf{T}^{-1}\mathbf{D}\mathbf{T}^{-1\dagger}\mathbf{T}^{\dagger}\mathbf{S}\mathbf{T} = \mathbf{T}^{-1}\mathbf{D}\mathbf{S}\mathbf{T}, \quad (7.38)$$

и

$$B'_{AB} = \sum_{\kappa, \lambda, \varrho, \tau=1}^m \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{T}^{-1})_{\mu\kappa} (\mathbf{D}\mathbf{S})_{\kappa\lambda} T_{\lambda\nu} (\mathbf{T}^{-1})_{\nu\varrho} (\mathbf{D}\mathbf{S})_{\varrho\tau} T_{\tau\mu}. \quad (7.39)$$

Как и в случае заселенности перекрывания  $d_{AB}$ , мы должны пользоваться соотношениями (7.28) и (7.29), заменяя в последнем  $A$  на  $B$ . Суммы по  $\kappa$  и  $\tau$  должны быть ограничены атомом  $A$ , а по  $\lambda$  и  $\varrho$  — атомом  $B$ :

$$\begin{aligned} B'_{AB} &= \sum_{\kappa, \tau \in A} \sum_{\lambda, \varrho \in B} \delta_{\tau\kappa} (\mathbf{D}\mathbf{S})_{\kappa\lambda} \delta_{\lambda\varrho} (\mathbf{D}\mathbf{S})_{\varrho\tau} \\ &= \sum_{\tau \in A} \sum_{\varrho \in B} (\mathbf{D}\mathbf{S})_{\tau\varrho} (\mathbf{D}\mathbf{S})_{\varrho\tau} = B_{AB}. \quad \text{Ч. т. д.} \end{aligned} \quad (7.40)$$

## 2.2. Индексы валентности

Следуя Вайбергу, целесообразно определить величину  $b_{\mu} = 2q_{\mu} - q_{\mu}^2$ , где  $q_{\mu}$  есть маллиkenовская полная орбитальная заселенность базисной орбитали  $\chi_{\mu}$ . Эту величину можно рассматривать, как меру способности к связыванию орбитали  $\chi_{\mu}$  в данной молекуле: она имеет максимум при  $q_{\mu} = 1$ , характерном для случая, когда орбиталь участвует в чисто ковалентной связи, и обращается в нуль как для пустой орбитали, не играющей никакой роли в молекуле ( $q_{\mu} = 0$ ), так и для дважды занятой ( $q_{\mu} = 2$ ), также не имеющей химического значения (как, например, остовные орбитали). Если просуммировать величины  $b_{\mu}$  для всех базисных орбиталей, центрированных на данном атоме, но вычесть внутри-

атомные частные порядки связей, не имеющие химического значения, то получим определение эффективной *валентности* атома в молекуле:

$$V_A = 2Q_A - \sum_{\mu, \nu \in A} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} , \quad (7.41)$$

где  $Q_A$  есть маллиkenовская полная атомная заселенность атома  $A$ .

В случае замкнутой оболочки, описываемой методом ОХФ, для матрицы  $\mathbf{P}$  выполнено свойство идемпотентности  $(\mathbf{PS})^2 = \mathbf{PS}$ , что приводит к равенству  $(\mathbf{DS})^2 = 2\mathbf{DS}$  для матрицы  $\mathbf{D} = 2\mathbf{P}$ . Как следствие, получаем равенство

$$V_A = \sum_{\substack{B \\ (B \neq A)}} B_{AB} , \quad (7.42)$$

т. е. если используется однодетерминантная волновая функция замкнутой оболочки, то валентность атома равна сумме его порядков связей со всеми другими атомами.

В самом деле, подставляя  $Q_A = \sum_{\mu \in A} q_\mu = \sum_{\mu \in A} (\mathbf{DS})_{\mu\mu}$ , получим

$$V_A = \sum_{\mu \in A} 2(\mathbf{DS})_{\mu\mu} - \sum_{\mu, \nu \in A} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} . \quad (7.43)$$

Из указанного выше свойства идемпотентности следует, что

$$2(\mathbf{DS})_{\mu\mu} = [(\mathbf{DS})^2]_{\mu\mu} = \sum_{\nu=1}^m (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} . \quad (7.44)$$

Подставляя это выражение в формулу (7.43), выделяя слагаемые с  $\nu \in A$  и группируя остальные по соответствующим атомам, имеем

$$\begin{aligned} V_A &= \sum_{\mu \in A} \left[ \sum_{\nu \in A} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} + \sum_{\substack{B \\ B \neq A}} \sum_{\nu \in B} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} \right] \\ &\quad - \sum_{\mu, \nu \in A} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} = \sum_{\substack{B \\ B \neq A}} \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{DS})_{\mu\nu} (\mathbf{DS})_{\nu\mu} \\ &= \sum_{\substack{B \\ (B \neq A)}} B_{AB} . \quad \text{Ч. т. д.} \end{aligned} \quad (7.45)$$

Равенство (7.45) не выполняется для волновой функции НХФ открытой оболочки, или если используется коррелированная волновая функция. В этих случаях разность между полной валентностью атома и суммой его порядков связей определяется как *свободная валентность* атома

$$F_A = V_A - \sum_{\substack{B \\ (B \neq A)}} B_{AB} . \quad (7.46)$$

Индекс валентности  $V_A$  можно рассматривать, как квантовохимический аналог химического понятия ковалентности атома, и обычно его величина близка к целому числу, соответствующему классической картине химических связей в молекуле. В случае свободных радикалов, бирадикалов и т. д. индекс свободной валентности можно поставить в близкое соответствие «точкам», используемым химиками для обозначения радикальных центров в химических формулах. (Вообще говоря, заметные значения индекса свободной валентности указывают на атомы, по которым можно ожидать свободнорадикальной атаки.)

Предыдущая интерпретация индекса свободной валентности подтверждается тем фактом, что в случае метода НХФ этот индекс можно выразить через матрицу спиновой плотности  $\mathbf{P}^s$  как

$$F_A = \sum_{\mu, \nu \in A} (\mathbf{P}^s \mathbf{S})_{\mu\nu} (\mathbf{P}^s \mathbf{S})_{\nu\mu} . \quad (7.47)$$

Чтобы вывести формулу (7.47), сначала преобразуем выражение для величины  $2Q_A$  из определения (7.41) для  $V_A$ , с помощью равенства  $\mathbf{D} = \mathbf{P}^a + \mathbf{P}^b$  и свойств идемпотентности (6.111) для матриц  $\mathbf{P}^a \mathbf{S}$  и  $\mathbf{P}^b \mathbf{S}$ :

$$\begin{aligned} 2Q_A &= 2 \sum_{\mu \in A} (\mathbf{D} \mathbf{S})_{\mu\mu} = 2 \sum_{\mu \in A} [(\mathbf{P}^a \mathbf{S})_{\mu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\mu}] \\ &= 2 \sum_{\mu \in A} \left\{ [(\mathbf{P}^a \mathbf{S})^2]_{\mu\mu} + [(\mathbf{P}^b \mathbf{S})^2]_{\mu\mu} \right\} \\ &= 2 \sum_{\mu \in A} \sum_{\nu=1}^m [(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu}] \\ &= 2 \sum_{\mu, \nu \in A} [(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu}] \\ &\quad + 2 \sum_{\mu \in A} \sum_B \sum_{\substack{\nu \in B \\ (B \neq A)}} [(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu}] = \end{aligned} \quad (7.48)$$

$$= 2 \sum_{\mu, \nu \in A} [(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu}] + \sum_{\substack{B \\ (B \neq A)}} B_{AB},$$

где также использовано определение (7.36) индекса порядка связи  $B_{AB}$ . Подставляя определение (7.41) в определение индекса свободной валентности (7.46), получаем

$$\begin{aligned} F_A - \sum_{\substack{B \\ (B \neq A)}} B_{AB} &= 2Q_A - \sum_{\mu, \nu \in A} (\mathbf{D} \mathbf{S})_{\mu\nu} (\mathbf{D} \mathbf{S})_{\nu\mu} - \sum_{\substack{B \\ (B \neq A)}} B_{AB} \\ &= 2Q_A - \sum_{\mu, \nu \in A} [(\mathbf{P}^a + \mathbf{P}^b) \mathbf{S}]_{\mu\nu} [(\mathbf{P}^a + \mathbf{P}^b) \mathbf{S}]_{\nu\mu} - \sum_{\substack{B \\ (B \neq A)}} B_{AB}. \end{aligned} \quad (7.49)$$

Подставляя сюда выражение (7.48), получим (слагаемые, содержащие  $B_{AB}$ , уничтожаются)

$$\begin{aligned} F_A &= \sum_{\mu, \nu \in A} \{2(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + 2(\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu} \\ &\quad - [(\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} + (\mathbf{P}^a \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu} + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^a \mathbf{S})_{\nu\mu} \\ &\quad + (\mathbf{P}^b \mathbf{S})_{\mu\nu} (\mathbf{P}^b \mathbf{S})_{\nu\mu}]\} = \sum_{\mu, \nu \in A} [(\mathbf{P}^a - \mathbf{P}^b) \mathbf{S}]_{\mu\nu} [(\mathbf{P}^a - \mathbf{P}^b) \mathbf{S}]_{\nu\mu} \\ &= \sum_{\mu, \nu \in A} (\mathbf{P}^s \mathbf{S})_{\mu\nu} (\mathbf{P}^s \mathbf{S})_{\nu\mu}. \quad \text{Ч. т. д.} \end{aligned}$$

### 2.3. Обменная плотность и индекс порядка связи

Среднее значение  $\varrho(\vec{r}) = \langle \Psi | \sum_{i=1}^N \delta(\vec{r}_i - \vec{r}) | \Psi \rangle$  оператора  $\hat{\varrho} = \sum_{i=1}^N \delta(\vec{r}_i - \vec{r})$  дает электронную плотность, т. е. плотность вероятности найти электрон в окрестности точки с радиус-вектором  $\vec{r}$ . Аналогично мы можем определить парную плотность  $\varrho_2(\vec{r}, \vec{r}')$ , дающую плотность вероятности найти один электрон около точки  $\vec{r}$ , а другой электрон около точки  $\vec{r}'$ . Очевидно, она определяется как среднее значение оператора

$$\hat{\varrho}_2(\vec{r}, \vec{r}') = \sum_{\substack{i, j \\ (i \neq j)}} \delta(\vec{r}_i - \vec{r}) \delta(\vec{r}_j - \vec{r}'). \quad (7.50)$$

Этот оператор можно преобразовать к виду

$$\begin{aligned}\hat{\varrho}_2(\vec{r}, \vec{r}') &= \sum_{i < j} \delta(\vec{r}_i - \vec{r}) \delta(\vec{r}_j - \vec{r}') + \sum_{j < i} \delta(\vec{r}_i - \vec{r}) \delta(\vec{r}_j - \vec{r}') \\ &= \sum_{i < j} [\delta(\vec{r}_i - \vec{r}) \delta(\vec{r}_j - \vec{r}') + \delta(\vec{r}_j - \vec{r}) \delta(\vec{r}_i - \vec{r}')] .\end{aligned}\quad (7.51)$$

(Мы поменяли индексы суммирования  $i$  и  $j$  во втором слагаемом.)

Теперь оператор  $\hat{\varrho}_2(\vec{r}, \vec{r}')$  является симметричным двухэлектронным оператором типа, рассмотренного в гл. 5, разд. 6.3, в котором  $\hat{g}(i, j)$  определено, как:

$$\hat{g}(i, j) = \delta(\vec{r}_i - \vec{r}) \delta(\vec{r}_j - \vec{r}') + \delta(\vec{r}_i - \vec{r}') \delta(\vec{r}_j - \vec{r}) . \quad (7.52)$$

Если  $\Psi$  является однодетерминантной волновой функцией, построенной из ортонормированных спин-орбиталей  $\psi_i(\vec{r}, \sigma) = \varphi_i(\vec{r}) \gamma_i(\sigma)$ , то среднее значение  $\hat{\varrho}_2(\vec{r}, \vec{r}')$  можно вычислить с помощью формулы (5.116):

$$\begin{aligned}\varrho_2(\vec{r}, \vec{r}') &= \\ \frac{1}{2} \sum_{i,j=1}^N &\left[ \langle \varphi_i(1) \varphi_j(2) | \delta(\vec{r}_1 - \vec{r}) \delta(\vec{r}_2 - \vec{r}') + \delta(\vec{r}_1 - \vec{r}') \delta(\vec{r}_2 - \vec{r}) | \varphi_i(1) \varphi_j(2) \rangle \right. \\ &- \left. \langle \varphi_i(1) \varphi_j(2) | \delta(\vec{r}_1 - \vec{r}) \delta(\vec{r}_2 - \vec{r}') + \delta(\vec{r}_1 - \vec{r}') \delta(\vec{r}_2 - \vec{r}) | \varphi_j(1) \varphi_i(2) \rangle \delta_{\gamma_i \gamma_j} \right] = \\ \frac{1}{2} \sum_{i,j=1}^N &\left[ |\varphi_i(\vec{r})|^2 |\varphi_j(\vec{r}')|^2 + |\varphi_i(\vec{r}')|^2 |\varphi_j(\vec{r})|^2 \right. \\ &- \left. (\varphi_i^*(\vec{r}) \varphi_j(\vec{r}) \varphi_j^*(\vec{r}') \varphi_i(\vec{r}') + \varphi_i^*(\vec{r}') \varphi_j(\vec{r}') \varphi_j^*(\vec{r}) \varphi_i(\vec{r})) \delta_{\gamma_i \gamma_j} \right] .\end{aligned}\quad (7.53)$$

После перестановки индексов суммирования  $i$  и  $j$  во втором и четвертом слагаемых, получим

$$\varrho_2(\vec{r}, \vec{r}') = \sum_{i,j=1}^N (|\varphi_i(\vec{r})|^2 |\varphi_j(\vec{r}')|^2 - \varphi_i^*(\vec{r}) \varphi_j(\vec{r}) \varphi_j^*(\vec{r}') \varphi_i(\vec{r}') \delta_{\gamma_i \gamma_j}) . \quad (7.54)$$

Так как  $\sum_{i=1}^N |\varphi_i(\vec{r})|^2 = \varrho(\vec{r})$ , можно записать

$$\varrho_2(\vec{r}, \vec{r}') = \varrho(\vec{r}) \varrho(\vec{r}') - \varrho_2^x(\vec{r}, \vec{r}') , \quad (7.55)$$

где первое слагаемое (можно его назвать «прямым» или «кулоновского типа») есть произведение электронных плотностей в точках  $\vec{r}$  и  $\vec{r}'$ , а

второе есть двухэлектронная «обменная парная плотность». Обменная парная плотность является следствием антисимметричности волновой функции; в том числе, она отвечает и за компенсацию «самовзаимодействия» электронов (см. ниже).

Согласно результату (7.54), имеем

$$\varrho_2^x(\vec{r}, \vec{r}') = \sum_{i,j=1}^N \varphi_i^*(\vec{r}) \varphi_j(\vec{r}) \varphi_j^*(\vec{r}') \varphi_i(\vec{r}') \delta_{\gamma_i \gamma_j}. \quad (7.56)$$

Если  $\Psi$  — детерминант, построенный из  $n_a$  орбиталей  $a_i$ , занятых спином  $\alpha$  и  $n_b$  орбиталей  $b_i$ , занятых спином  $\beta$ , то

$$\varrho_2^x(\vec{r}, \vec{r}') = \sum_{i,j=1}^{n_a} a_i^*(\vec{r}) a_j(\vec{r}) a_j^*(\vec{r}') a_i(\vec{r}') + \sum_{i,j=1}^{n_b} b_i^*(\vec{r}) b_j(\vec{r}) b_j^*(\vec{r}') b_i(\vec{r}'). \quad (7.57)$$

Теперь проинтегрируем  $\varrho_2^x(\vec{r}, \vec{r}')$  по обоим переменным  $\vec{r}$  и  $\vec{r}'$ , и, учитывая, что оба набора орбиталей  $\{a_i\}$  и  $\{b_i\}$  ортонормированы, получим:

$$\begin{aligned} \int \varrho_2^x(\vec{r}, \vec{r}') dv dv' &= \sum_{i,j=1}^{n_a} \langle a_i | a_j \rangle \langle a_j | a_i \rangle + \sum_{i,j=1}^{n_b} \langle b_i | b_j \rangle \langle b_j | b_i \rangle \\ &= \sum_{i,j=1}^{n_a} \delta_{ij} + \sum_{i,j=1}^{n_b} \delta_{ij} = n_a + n_b = N. \end{aligned} \quad (7.58)$$

Подставим разложение орбиталей  $a_i$ ,  $b_i$  (6.115) в формулу (7.58):

$$\begin{aligned} &\sum_{i,j=1}^{n_a} \langle \sum_{\mu=1}^m a_{\mu}^i \chi_{\mu} | \sum_{\nu=1}^m a_{\nu}^j \chi_{\nu} \rangle \langle \sum_{\varrho=1}^m a_{\varrho}^j \chi_{\varrho} | \sum_{\tau=1}^m a_{\tau}^i \chi_{\tau} \rangle \\ &+ \sum_{i,j=1}^{n_b} \langle \sum_{\mu=1}^m b_{\mu}^i \chi_{\mu} | \sum_{\nu=1}^m b_{\nu}^j \chi_{\nu} \rangle \langle \sum_{\varrho=1}^m b_{\varrho}^j \chi_{\varrho} | \sum_{\tau=1}^m b_{\tau}^i \chi_{\tau} \rangle = N \end{aligned} \quad (7.59)$$

т. е.

$$\sum_{i,j=1}^{n_a} \sum_{\mu,\nu,\varrho,\tau=1}^m a_{\tau}^i a_{\mu}^{i*} S_{\mu\nu} a_{\nu}^j a_{\varrho}^{j*} S_{\varrho\tau} + \sum_{i,j=1}^{n_b} \sum_{\mu,\nu,\varrho,\tau=1}^m b_{\tau}^i b_{\mu}^{i*} S_{\mu\nu} b_{\nu}^j b_{\varrho}^{j*} S_{\varrho\tau} = N. \quad (7.60)$$

С помощью определений (6.116) это выражение можно переписать как

$$\sum_{\nu,\tau=1}^m [(\mathbf{P}^a \mathbf{S})_{\tau\nu} (\mathbf{P}^a \mathbf{S})_{\nu\tau} + (\mathbf{P}^b \mathbf{S})_{\tau\nu} (\mathbf{P}^b \mathbf{S})_{\nu\tau}] = N. \quad (7.61)$$

Мы можем сгруппировать слагаемые в левой части по атомам, на которых центрированы базисные орбитали. Получим

$$\sum_{A,B} \sum_{\nu \in A} \sum_{\tau \in B} [(\mathbf{P}^a \mathbf{S})_{\tau\nu} (\mathbf{P}^a \mathbf{S})_{\nu\tau} + (\mathbf{P}^b \mathbf{S})_{\tau\nu} (\mathbf{P}^b \mathbf{S})_{\nu\tau}] = N. \quad (7.62)$$

Сравнивая это выражение с определением (7.36), видим, что индекс порядка связи  $B_{AB}$  между атомами  $A$  и  $B$  является двухатомным вкладом в интеграл обменной парной плотности  $\varrho^x(\vec{r}, \vec{r}')$ . (Множитель 2, присутствовавший в определении (7.36), также воспроизводится, так как суммирование по  $A$  и  $B$  в разложении (7.62) идет по всем атомам независимо.)

Заметим, что равенство (7.61) можно также получить, как следствие идемпотентности матриц  $\mathbf{P}^a \mathbf{S}$  и  $\mathbf{P}^b \mathbf{S}$ , используя равенства  $\text{Tr}(\mathbf{P}^a \mathbf{S}) = n_a$  и  $\text{Tr}(\mathbf{P}^b \mathbf{S}) = n_b$  (см. приложение П8). С помощью соотношений идемпотентности можно также ввести индексы связи высших порядков, например, рассматривая трехцентровые вклады, присутствующие в величине  $\text{Tr}(\mathbf{P}^a \mathbf{S} \mathbf{P}^a \mathbf{S} \mathbf{P}^a \mathbf{S} + \mathbf{P}^b \mathbf{S} \mathbf{P}^b \mathbf{S} \mathbf{P}^b \mathbf{S}) = N$ , и т. д.

Сделаем несколько замечаний относительно тесной связи между обменом и химическим связыванием. Можно заметить, что энергия кулоновского отталкивания электронов  $\langle \Psi | \sum_{i < j} \frac{1}{r_{ij}} | \Psi \rangle$  выражается через парную плотность  $\varrho_2(\vec{r}, \vec{r}')$  как

$$E_{el-el} = \langle \Psi | \sum_{i < j} \frac{1}{r_{ij}} | \Psi \rangle = \iint \frac{\varrho_2(\vec{r}, \vec{r}')}{|\vec{r} - \vec{r}'|} dv dv'. \quad (7.63)$$

Это следует из возможности записи  $\frac{1}{r_{ij}}$  в виде интеграла

$$\frac{1}{r_{ij}} = \frac{1}{|\vec{r}_i - \vec{r}_j|} = \frac{1}{2} \iint \frac{\delta(\vec{r}_i - \vec{r}) \delta(\vec{r}_j - \vec{r}') + \delta(\vec{r}_i - \vec{r}') \delta(\vec{r}_j - \vec{r})}{|\vec{r} - \vec{r}'|} dv dv'. \quad (7.64)$$

Тогда при вычислении суммы  $\langle \Psi | \sum_{i < j} \frac{1}{r_{ij}} | \Psi \rangle$  можно поменять порядок интегрирования и выполнить все из них, кроме интегралов по  $\vec{r}$  и  $\vec{r}'$ . Для однодетерминантной волновой функции, можно далее переписать энергию (7.63) с помощью разбиения (7.55) как

$$E_{el-el} = \iint \frac{\varrho(\vec{r}) \varrho(\vec{r}')}{|\vec{r} - \vec{r}'|} dv dv' - \iint \frac{\varrho_2^x(\vec{r}, \vec{r}')}{|\vec{r} - \vec{r}'|} dv dv'. \quad (7.65)$$

Так как  $\varrho(\vec{r})$  есть плотность, полученная суммированием по всем электронам, первое слагаемое в правой части содержит также и электростатическое взаимодействие плотности, созданной данным электроном в точке  $\vec{r}$ , с его же собственной зарядовой плотностью в точке  $\vec{r}'$ . Это имеет место, потому что мы включили слагаемые самоотталкивания  $\hat{J}_i - \hat{K}_i$  в оператор Фока и в хартри-фоковскую энергию. Необходимая компенсация этого самовзаимодействия содержится во втором слагаемом в правой части (обменной энергии).

Очевидно, аналогичное представление  $E_{el-el}$  возможно также для коррелированных волновых функций, однако в этом случае обменную часть  $\varrho_2^x(\vec{r}, \vec{r}')$  парной плотности необходимо заменить «обменно-корреляционной», которую можно (формально) определить, как разность  $\varrho_2^{xc}(\vec{r}, \vec{r}') = \varrho_2(\vec{r}, \vec{r}') - \varrho(\vec{r})\varrho(\vec{r}')$ . На практике ее вычисляют с помощью применяемой в конкретном расчетном методе коррелированной волновой функции.

В однодетерминантном случае, подставляя представление (7.54) в выражение (7.65), немедленно получаем двухэлектронную часть выражения для энергии Хартри—Фока (6.44). Ее обменная составляющая равна:

$$\begin{aligned} E^x &= - \iint \frac{\varrho_2^x(\vec{r}, \vec{r}')}{\vec{r} - \vec{r}'} dv dv' = - \sum_{i,j=1}^N [\varphi_i \varphi_j | \varphi_j \varphi_i] \delta_{\gamma_i \gamma_j} \\ &= - \sum_{i,j=1}^{n_a} [a_i a_j | a_j a_i] - \sum_{i,j=1}^{n_b} [b_i b_j | b_j b_i]. \end{aligned} \quad (7.66)$$

Это равенство показывает тесную связь между обменной парной плотностью  $\varrho_2^x(\vec{r}, \vec{r}')$  и обменной частью  $E^x$  полной энергии Хартри—Фока. Стоит отметить, что обменная энергия отрицательна. В самом деле, каждый интеграл  $[\varphi_i \varphi_j | \varphi_j \varphi_i]$  положителен — по крайней мере, если используются вещественные орбитали,<sup>1</sup> — поскольку он представляет собой электростатическое самоотталкивание соответствующей плотности  $\varphi_i(\vec{r})\varphi_j(\vec{r})$ . (Это находится в согласии с тем, что обменная энергия компенсирует самовзаимодействие, включенное в кулоновское слагаемое<sup>2</sup>.)

<sup>1</sup> На самом деле это всегда так, если произведение функций  $\varphi_i \varphi_j \neq 0$  (см. Степанов Н.Ф., Пупышев В.И. Квантовая механика молекул и квантовая химия. — М.: Изд-во МГУ, 1991). — *Прим. ред.*

<sup>2</sup> Стоит отметить, что в отличие от кулоновских и обменных вкладов в двухэлектронную энергию, сумма  $\sum_{i=1}^N [\varphi_i \varphi_i | \varphi_i \varphi_i]$ , описывающая самовзаимодействие, не является инвариантной при унитарных преобразованиях занятых орбиталей одно-



Как мы видели раньше, если используются базисные функции, центрированные на атомах, то интеграл  $\int \int \varrho_2^x(\vec{r}, \vec{r}') dv dv'$  обменной парной плотности можно разложить на одно- и двухцентровые составляющие. Этого нельзя проделать с обменной энергией (7.66), так как разложение двухэлектронных интегралов по МО будет содержать также, вообще говоря, трех- и четырехцентровые интегралы от базисных орбиталей.

Однако такое разложение возможно и весьма поучительно в простейшем случае модельных теорий типа ППДП (CNDO), использующих приближение «нулевого дифференциального перекрывания»:

$$[\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau] = \gamma_{AB} \delta_{\mu\varrho} \delta_{\nu\tau} \quad \mu \in A; \nu \in B. \quad (7.67)$$

Здесь  $A$  и  $B$  суть атомы, на которых центрированы базисные орбитали  $\chi_\mu$  и  $\chi_\nu$ , соответственно. С помощью приближения (7.67) получаем, подставив разложение орбиталей  $a_i$  и  $b_i$  (6.115) в выражение (7.66) и используя определения (6.116) матриц  $\mathbf{P}^a, \mathbf{P}^b$ :

$$\begin{aligned} E^x(CNDO) &= - \sum_{i,j=1}^{n_a} \sum_{\mu,\nu,\varrho,\tau=1}^m a_\mu^{i*} a_\nu^{j*} a_\varrho^j a_\tau^i [\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau] \\ &\quad - \sum_{i,j=1}^{n_b} \sum_{\mu,\nu,\varrho,\tau=1}^m b_\mu^{i*} b_\nu^{j*} b_\varrho^j b_\tau^i [\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau] \\ &= - \sum_{A,B} \sum_{\mu \in A} \sum_{\nu \in B} \sum_{\varrho,\tau=1}^m (P_{\tau\mu}^a P_{\varrho\nu}^a + P_{\tau\mu}^b P_{\varrho\nu}^b) \gamma_{AB} \delta_{\mu\varrho} \delta_{\nu\tau} \\ &= - \sum_{A,B} \sum_{\mu \in A} \sum_{\nu \in B} (|P_{\mu\nu}^a|^2 + |P_{\mu\nu}^b|^2) \gamma_{AB}. \end{aligned} \quad (7.68)$$

Сравнивая это выражение с (7.34) и учитывая, что сумма по  $A$  и  $B$  ведется по всем атомам независимо, приходим к заключению, что двухатомный вклад в обменную энергию пропорционален соответствующему индексу Вайберга:

$$E_{AB}^x = -\frac{1}{2} \gamma_{AB} W_{AB}. \quad (7.69)$$

Автор ввел индекс порядка связи (7.36) в контексте *ab initio* теорий, рассмотрев приближенное обобщение этого важного соотношения.

детерминантной волновой функции. Соответственно, самоотталкиванию электронов нельзя приписать какое-то строго определенное значение.

## 2.4. Порядки связей в трехцентровых системах

Большинство химических связей можно рассматривать, как образованные парой электронов, занимающих двухцентровую локализованную молекулярную орбиталь. (Для многократно связанных атомов необходимо предположить наличие нескольких таких связывающих электронных пар.) Можно даже на основе этих соображений обсуждать сопряжение и ароматичность, предполагая, что волновая функция является линейной комбинацией нескольких слагаемых, описывающих «резонанс» между различными «структурами Кекуле», каждая из которых описывается несколькими двухэлектронными двухцентровыми связями. Имеются, однако, системы, в которых необходимо предположить существование *трехцентровых связей*. Вероятно, наиболее известным примером такого рода является молекула диборана  $B_2H_6$ , в которой два атома бора связаны мостиками из двух симметрично расположенных атомов водорода. Электронная структура этой молекулы обычно описывается в терминах двух трехцентровых двухэлектронных связей, каждая из которых в первом приближении построена из соответствующих гибридных атомных орбиталей двух атомов бора и  $1s$ -орбитали «мостикового» водорода. Это означает, что водород соединен связью дробного порядка симметрично с обоими атомами бора. (Имеется принципиальное отличие от обычной водородной связи, в которой водород связан с одним из партнеров лишь добавочной связью.) Значительное расстояние между атомами бора, как будто, указывает на то, что между ними не должно быть заметного взаимодействия. Это, однако, не так: как мы увидим, образование трехцентровой двухэлектронной связи приводит также к появлению ненулевого порядка связи между атомами бора. Притягивающее обменное взаимодействие, соответствующее этим добавочным порядкам связи, может иметь существенное значение с точки зрения энергетики (стабильности) молекулы диборана (или похожих систем).

Рассмотрим простую модель, в которой два электрона занимают молекулярную орбиталь, образованную из линейной комбинации трех нормированных базисных орбиталей  $\chi_1 \div \chi_3$ . Орбитали  $\chi_1$  и  $\chi_3$  являются орбиталями атомов бора, направленными к водородной орбитали  $\chi_2$ . Мы предполагаем, что перекрывание орбиталей бора и водорода равно  $\langle \chi_1 | \chi_2 \rangle = \langle \chi_2 | \chi_3 \rangle = S$ , тогда как перекрывание орбиталей двух атомов бора пренебрежимо мало. Матрица перекрывания  $\mathbf{S}$ , таким образом, имеет вид

$$\mathbf{S} = \begin{pmatrix} 1 & S & 0 \\ S & 1 & S \\ 0 & S & 1 \end{pmatrix}. \quad (7.70)$$

Построим простейшую симметричную нормированную трехцентровую орбиталь:

$$\psi = \frac{1}{\sqrt{2(1 + \sqrt{2}S)}} \left[ \chi_2 + \frac{1}{\sqrt{2}}(\chi_1 + \chi_3) \right]. \quad (7.71)$$

Ее вектор коэффициентов ЛКАО равен

$$\mathbf{c} = \frac{1}{\sqrt{2(1 + \sqrt{2}S)}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \quad (7.72)$$

Тривиальное вычисление матриц  $\mathbf{D} = 2\mathbf{c}\mathbf{c}^\dagger$  и  $\mathbf{D}\mathbf{S}$  дает

$$\mathbf{D} = \frac{1}{2(1 + \sqrt{2}S)} \begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}; \quad (7.73)$$

$$\mathbf{D}\mathbf{S} = \frac{1}{2} \begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}. \quad (7.74)$$

Это значит, что индексы порядков связей  $B_{12} = B_{23} = \frac{1}{2}$ ;  $B_{13} = \frac{1}{4}$ . (При использовании определения (7.35) необходимо учесть, что в нашей упрощенной модели каждый центр несет только одну базисную орбиталь.)

Эти результаты означают, что в молекуле диборана должны существовать порядки связей примерно  $\frac{1}{2}$  между мостиковыми атомами водорода и каждым из атомов бора, а также между двумя атомами бора: так как молекула диборана содержит две независимые двухэлектронные трехцентровые связи, порядок связи бор—бор должен быть в два раза больше, чем значение  $\frac{1}{4}$ , полученное выше. Эти значения находятся в очень хорошем согласии с теми, которые получены в реальных *ab initio* расчетах.

Похожий вывод можно провести для случая трехцентровой четырехэлектронной связи, в которой несвязывающая орбиталь тоже дважды занята. Типичным примером является система аксиальных связей, образованных гипервалентным атомом серы. В этом случае необходимо предположить, что центральная орбиталь  $\chi_2$ , находящаяся на атоме

серы, является орбиталью  $p$ -типа; тогда  $\langle \chi_1 | \chi_2 \rangle = S$  и  $\langle \chi_2 | \chi_3 \rangle = -S$ . Связывающая и несвязывающая орбитали в этом случае суть

$$\psi_b = \frac{1}{\sqrt{2(1 + \sqrt{2}S)}} \left[ \chi_2 + \frac{1}{\sqrt{2}}(\chi_1 - \chi_3) \right] \quad (7.75)$$

и

$$\psi_n = \frac{1}{\sqrt{2}}(\chi_1 + \chi_3). \quad (7.76)$$

Легко проверить, однако, что это отличие не влияет на величины порядков связей:  $B_{12} = B_{23} = \frac{1}{2}$  и  $B_{13} = \frac{1}{4}$ . (Неудивительно, что заполнение несвязывающей орбитали  $\psi_n$  не повлияет на значения индексов порядков связей.) Относительно малые величины этих индексов порядков связей могут послужить для объяснения того, почему аксиальные связи в гипервалентных соединениях серы, подобно  $\text{SF}_4$ , удлинены, по сравнению с обычными двухэлектронными двухцентровыми экваториальными связями.

### Библиографические заметки

Важность различения между анализом заселенностей, проводимом в трехмерном физическом пространстве и в гильбертовом пространстве базисных орбиталей, была указана еще Холлом [1].

#### Раздел 1.

Анализ заселенности по Малликену [2] обсуждается в большинстве учебников по квантовой химии (например, в [3]). Требование вращательно-гибридизационной инвариантности было впервые систематически исследовано в связи с параметризацией полуэмпирического метода ППДП (CNDO) [4]. Несмотря на фундаментальную важность вращательно-гибридизационной инвариантности малликеновских чистой заселенности и заселенности перекрытия, она не обсуждается в явном виде ни в каком из источников, известных автору.

Недостатком малликеновских заселенностей является то, что полная орбитальная заселенность  $q_\mu$  орбитали  $\chi_\mu$  не обязана быть заключена в интервале от 0 до 2, как это было бы в случае ортонормированного базиса. Это может вызывать проблемы, если использовать большие базисы, содержащие несколько «диффузных» функций, которые медленно убывают с расстоянием, и, следовательно, фактически не имеют атомного характера [5]. (Их заселенность может быть существенной в отрицательных ионах.) В таком случае необходимо прибегать к базису, ортогонализованному по Лёвдину (гл. 3, разд. 2.3), и проводить анализ заселенностей и порядков связей относительно этого вспомогательного ортонормированного базиса<sup>1</sup>.

<sup>1</sup> Уже после выхода английского оригинала этой книги автор обнаружил [Mayer I. Chem. Phys. Letters **393**, 209 (2004).], что анализ заселенностей, проведенный в

### Раздел 2.1.

Вайберг ввел свой индекс в [6] для систем с замкнутой оболочкой. Корректное обобщение на системы с открытыми оболочками было дано Борисовой и Семеновым [7]. Они также доказали [8], что индекс Вайберга равен  $\frac{1}{2}(N_{\text{связ}} - N_{\text{антисвяз}})$  для гомоядерных двухатомных молекул первого ряда — за исключением  $\text{C}_2$ , которая имеет весьма своеобразную электронную структуру [9].

Общий индекс порядка связи (7.35) был предложен на интуитивной основе Giambiagi *et al.* [10] в рамках полуэмпирического расширенного метода Хюккеля. Будучи неосведомленным об этом, автор ввел его для случая *ab initio* [11] на основе анализа обменной части энергии, полученной в схеме разложения энергии на компоненты [12]. Точный вид определения (7.36) для случая открытой оболочки был приведен в [13], как *ab initio* обобщение результатов Борисовой и Семенова [7]. (См. также [14, 15].)

Инвариантность индекса порядка связи по отношению к вращательно-гибридизационным преобразованиям была доказана в [16].

### Раздел 2.2.

Понятие индекса валентности введено в рамках метода CNDO независимо Борисовой и Семеновым [7] и Армстронгом с соавторами [17]; для случая *ab initio* его обобщил автор [11], который также ввел индекс свободной валентности (см. также [13, 14, 15]).

### Раздел 2.3.

Представленный вывод тесно связан с анализом соотношений между индексом порядка связи и нормировкой обменной части «матрицы плотности второго порядка»  $\rho_2(x_1, x_2; x'_1, x'_2)$  [13, 14, 15].

### Раздел 2.3.

Материал основан на работе [18].

базисе, ортогонализованном по Лёвдину, не имеет вращательно-гибридизационной инвариантности. Гибридизационной инвариантности просто нет, а вращательная инвариантность тоже нарушается, если в атомных базисах применяются 6 «декартовых»  $d$ -орбиталей (10  $f$ -орбиталей и т. д.), а не 5 «чистых»  $d$ -орбиталей (7 «чистых»  $f$ -орбиталей и т. д.), пропорциональных соответствующим сферическим гармоникам. Таким образом, применяя, например, популярные стандартные базисные наборы 6-31G\* или 6-31G\*\*, можно получать атомные заряды в базисе, ортогонализованном по Лёвдину, зависящие от ориентации молекулы относительно лабораторной системы координат; кроме того, эквивалентным атомам могут быть приписаны (несколько) различные заряды. Этот результат указывает на ограниченную применимость анализа заселенностей (и порядков связи), проведенного в лёвдиновском базисе. Указанная проблема не возникает в том варианте ортогонализации по Лёвдину, который использовал Дэвидсон: он сначала проводит ортогонализацию базисных орбиталей на каждом из атомов по отдельности, и проводит ортогонализацию по Лёвдину только после этого, т. е. для учета межатомных перекрытий. Этот подход, конечно, представляет собой вполне корректный выход из положения, нужно только иметь в виду, что он приведет к совершенно другим численным результатам, чем те, которые получаются при применении обычного метода ортогонализации по Лёвдину.

## Для дальнейшего изучения

Рекомендуются обзорные статьи [19, 20].

## Литература

1. Холл Г. Г. (Hall G. G.) *Речь Председателя*, 5-й Международный Конгресс по Квантовой Химии, Монреаль, 1985.
2. Mulliken R. S. J. Chem. Phys. **23**, 1833, 1841, 2338, 2343 (1955).
3. Veszprémi T., Fehér M. *Quantum Chemistry: Fundamentals to Applications*, Kluwer Academic/Plenum New York, 1999.
4. Pople J. A., Beveridge D. L. *Approximate Molecular Orbital Theory*, McGraw-Hill, New York 1970.
5. Baker J. Theor. Chim. Acta **68**, 221 (1985).
6. Wiberg K. A. Tetrahedron **24**, 1083 (1968).
7. Борисова Н. П., Семенов С. Г. Вестн. ЛГУ, No. 16, 98 (1976).
8. Борисова Н. П., Семенов С. Г. Вестн. ЛГУ, No. 16, 119 (1973).
9. Юг К. (Jug K.), частное сообщение, 1984.
10. Giambiagi M., de Giambiagi M. D., Grepel D. R., Heynemann C. D. J. Chim. Phys. **72**, 15 (1975).
11. Mayer I. Chem. Phys. Letters, **97**, 270 (1983); Addendum: **117**, 396 (1985).
12. Mayer I. Int. J. Quantum Chem. **23**, 341<sup>®</sup> (1983).
13. Mayer I. Int. J. Quantum Chem. **29**, 73 (1986).
14. Mayer I. Int. J. Quantum Chem. **29**, 477 (1986).
15. Mayer I. Theor. Chim. Acta **67**, 315 (1985).
16. Mayer I. DSc. Thesis, Hungarian Academy of Sciences, Budapest 1986 (Диссертация на соискание степени доктора химических наук, Венгерская Академия Наук, Будапешт, 1986, неопубликовано).
17. Armstrong D. R., Perkins P. G., Stewart J. J. P. J. Chem. Soc. Dalton Trans. **1973**, 838, 2273 (1973).
18. Mayer I. J. Mol. Struct. (Theochem) **186**, 43 (1989).
19. Sannigrahi A. B. Adv. Quantum Chem. **23**, 301 (1992).
20. Bridgeman A. J., Cavigliasso G., Ireland L. R., Rothery J. J. Chem. Soc Dalton Trans. No. 14, 2095 (2001).



# Электронная корреляция



## 1. Разложение КВ

Рассмотрим произвольную нормируемую  $N$ -электронную волновую функцию  $\Psi = \Psi(1, 2, \dots, N)$  при фиксированных значениях пространственных и спиновых координат первых  $N - 1$  электронов:  $x_k \equiv \{\vec{r}_k, \sigma_k\} = \text{const}$ ;  $k = 1, 2, \dots, N - 1$ . При этих условиях  $\Psi$  есть нормируемая функция пространственных и спиновых координат  $N$ -го электрона и как таковая может быть разложена в ряд по некоторому полному набору  $\psi_i = \psi_i(x)$ ,  $i = 1, 2, \dots, \infty$  одноэлектронных функций

$$\Psi(1, 2, \dots, N) = \sum_{i_N=1}^{\infty} c_{i_N}(x_1, x_2, \dots, x_{N-1})\psi_{i_N}(x_N). \quad (8.1)$$

Коэффициенты этого разложения являются функциями координат остальных электронов  $i = 1, 2, \dots, N - 1$ . Очевидно, они также должны быть нормируемыми функциями их аргументов. Поэтому для фиксированных значений  $x_i$ ,  $i = 1, 2, \dots, N - 2$  мы можем снова разложить  $c_{i_N}(x_1, x_2, \dots, x_{N-1})$  в ряд как функцию  $x_{N-1}$ ; тогда получим

$$\Psi = \sum_{i_{N-1}, i_N=1}^{\infty} c_{i_{N-1}i_N}(x_1, x_2, \dots, x_{N-2})\psi_{i_{N-1}}(x_{N-1})\psi_{i_N}(x_N) \quad (8.2)$$

и т. д. В конце концов получаем

$$\Psi = \sum_{i_1, i_2, \dots, i_N=1}^{\infty} c_{i_1 i_2 \dots i_N} \psi_{i_1}(x_1) \psi_{i_2}(x_2) \dots \psi_{i_N}(x_N), \quad (8.3)$$

где коэффициенты  $c_{i_1 i_2 \dots i_N}$  больше не зависят ни от каких электронных координат.

Функция  $\Psi$  должна быть антисимметричной по перестановкам электронов; это приводит к определенным ограничениям на значения коэффициентов  $c_{i_1 i_2 \dots i_N}$ . Например, из антисимметрии  $\Psi$  по перестановке электронов «1» и «2» следует, что

$$c_{i_1 i_2 i_3 \dots i_N} = -c_{i_2 i_1 i_3 \dots i_N} \quad \text{и т. д.} \quad (8.4)$$

Можно собрать слагаемые, которые содержат одни и те же спин-орбитали, и отличаются только порядком электронных координат и порядком индексов  $i_1, i_2, \dots, i_N$  при коэффициентах  $c_{i_1 i_2 \dots i_N}$ . Тогда получим детерминант Слэтера, построенный из рассматриваемых спин-орбиталей, умноженный на некоторый общий коэффициент<sup>1</sup>. При условии, что индексы  $i_j$  образуют возрастающую последовательность ( $i_1 < i_2 < \dots < i_N$ ), набор  $N$  индексов  $\{i_j\}$  называется  $I$ -й «упорядоченной конфигурацией». Тогда волновую функцию  $\Psi$  можно представить в виде разложения

$$\Psi = \sum_I c_I \Psi_I \quad (8.5)$$

по детерминантам Слэтера, соответствующим разным упорядоченным конфигурациям

$$\Psi_I = \hat{\mathcal{A}}[\psi_{i_1}(1)\psi_{i_2}(2)\dots\psi_{i_N}(N)] . \quad (8.6)$$

Если одноэлектронные функции  $\psi_i$  образуют полный набор в одноэлектронном гильбертовом пространстве, то детерминанты, соответствующие всем построенным из этих функций упорядоченным конфигурациям, образуют полный набор в линейном пространстве антисимметричных  $N$ -электронных функций. Если одноэлектронные функции ортонормированы, то детерминанты, построенные из них, также ортонормированы.

Точно так же, если начать с конечного числа спин-орбиталей  $\psi_i$ , то детерминанты, соответствующие упорядоченным конфигурациям этих орбиталей, образуют полный базис в пространстве антисимметричных волновых функций, которые можно построить из данного набора одноэлектронных функций.

Разложение волновых функций по детерминантам, описанное выше, обычно называется разложением «конфигурационного взаимодействия» (КВ). Это традиционное название; более удачными были бы термины «наложение, суперпозиция или линейная комбинация конфигураций». Метод КВ, используемый на практике, состоит в вариационном определении коэффициентов  $c_I$ , что требует решения векового уравнения линейной вариационной задачи (ср. гл. 3, разд. 1).

Характерной особенностью гамильтониана Борна—Оппенгеймера, используемого в квантовой химии, является то, что он представляет собой сингулярную функцию межэлектронных расстояний  $\left(\frac{1}{r_{ij}} \rightarrow \infty \text{ если}$

<sup>1</sup> Равный, очевидно,  $\pm\sqrt{N!}|c_{i_1 i_2 \dots i_N}|$ . — *Прим. ред.*



$r_{ij} \rightarrow 0$ ). В свою очередь, разложение КВ выписано через одноэлектронные спин-орбитали, так что оно не содержит межэлектронных расстояний явно. Следовательно, поведение волновой функции в области пространства координат всех электронов, отвечающей сближению двух из них, описывается только как суммарный эффект большого числа слабых в разложении КВ, или, можно сказать, как их интерференция. Вследствие этого, разложение КВ сходится весьма медленно, особенно когда это касается локальных свойств (ср. гл. 9, разд. 2.2) волновой функции.

## 2. Корреляционная энергия и теорема Несбета

Волновые функции Хартри—Фока не являются точными решениями уравнения Шрёдингера, поскольку они учитывают межэлектронное взаимодействие только в усредненном виде. Это означает, что волновая функция Хартри—Фока не учитывает тот факт, что разные электроны не движутся независимо друг от друга, а напротив, их движения *коррелированы*. (Эта корреляция приводит, среди прочего, к зависимости волновой функции от межэлектронного расстояния, упомянутой выше.)

Электронную корреляцию можно охарактеризовать следующим образом. Вероятность найти электрон в данной точке зависит также от положения всех электронов. Пусть  $p_1(\vec{r}_1)$  есть плотность вероятности обнаружить первый электрон в точке  $\vec{r}_1$ , независимо от положений остальных электронов (т. е. при произвольных положениях других электронов). Точно так же мы можем определить плотность вероятности  $p_2(\vec{r}_2)$  для второго электрона. Тогда корреляция движений двух электронов означает, что для плотности вероятности  $p_{12}(\vec{r}_1, \vec{r}_2)$  того, что первый электрон будет найден в точке  $\vec{r}_1$  и одновременно второй электрон в точке  $\vec{r}_2$ , имеет место неравенство

$$p_{12}(\vec{r}_1, \vec{r}_2) \neq p_1(\vec{r}_1)p_2(\vec{r}_2), \quad (8.7)$$

что отвечает условной вероятности событий, которые статистически не являются независимыми.

Строго говоря, движение всех электронов описывалось как полностью независимое друг от друга только в рамках (совершенно устаревшего) метода Хартри, который пренебрегал даже требованием антисимметрии и использовал волновую функцию в виде простого произведения спин-орбиталей. Введение антисимметризации исключает возможность найти два электрона, имеющие одинаковые спины, в одной и той же точке пространства. На примере детерминантов Слэтера это проявля-

ется следующим образом: если два электрона — скажем,  $i$ -й и  $j$ -й, — обладают одними и теми же координатами и проекцией спина, то  $\vec{r}_i = \vec{r}_j$ ,  $\sigma_i = \sigma_j$ , и  $i$ -й и  $j$ -й столбцы детерминантной волновой функции (5.15) станут равными и, таким образом, детерминант (определитель) обратится в нуль. Ввиду непрерывности волновой функции, вероятность найти два электрона одного спина близко друг к другу также мала. Эту характерную особенность электронного распределения называют «ферми-дыркой»; она появляется вследствие антисимметрии волновой функции и существует даже в случае невзаимодействующих электронов. Таким образом, ферми-дырку можно рассматривать, как чисто кинематический эффект.

Из-за существования ферми-дырки, для волновой функции Хартри—Фока совершенно некоррелировано только движение электронов с противоположными проекциями спинов. Тем не менее, в силу практической важности стандартного ограниченного метода Хартри—Фока, корреляционную энергию определяют, согласно общепринятой терминологии, введенной Лёвдиным, как ошибку в энергии, допущенную методом ОХФ, т. е. как отклонение энергии ОХФ от точного значения нерелятивистской энергии. В соответствии с этим, представление об электронной корреляции включает только те эффекты, которые не являются следствием одной лишь антисимметрии и поэтому не учитываются в методе ОХФ. Иначе можно сказать, что под электронной корреляцией понимают лишь те эффекты, которые появляются вследствие кулоновского отталкивания между электронами, уменьшающего вероятность найти их вблизи друг друга («кулоновская дырка»). Поэтому электронная корреляция является динамическим эффектом, который в сущности не зависит от спина, в отличие от ферми-дырки, являющейся кинематическим эффектом, зависящим от спина.

Имеется и другая, тоже широко распространенная, классификация, согласно которой различают «статическую» и «динамическую» корреляции. О статической корреляции говорят тогда, когда описание системы с помощью однодетерминантной волновой функции ОХФ, построенной из дважды занятых орбиталей, совершенно не отражает физическую ситуацию. (Типичным примером является гомолитическая диссоциация, обсуждавшаяся в гл. 6, разд. 4.3, но могут встречаться и более сложные случаи.) Динамическая корреляция преобладает вблизи равновесных конфигураций обычных молекул, как прямое следствие межэлектронного отталкивания. Хотя такая классификация часто бывает полезной и в некоторых случаях даже очевидной, терминам «статическая» и «динамическая корреляция» трудно дать четкое определение.

Во многих задачах, интересных с химической точки зрения, можно предположить с некоторой степенью уверенности, что корреляционная энергия почти постоянна. (Хотя это, конечно, не так, когда в системе меняется число или характер химических связей.) Поэтому результаты, получающиеся при проведении расчетов методом Хартри—Фока—Рутана с хорошим базисным набором, часто достаточны для полуколичественных оценок. Однако, если нужны точные числа, особенно в случаях, когда рассматриваются малые разности энергий, то невозможно избежать явного учета электронной корреляции. С другой стороны, существуют такие проблемы, как, например, исследование путей реакций, в которых нельзя ожидать даже качественно правильных результатов без учета электронной корреляции.

Наиболее полное описание электронной корреляции можно получить, если явно ввести зависимость волновой функции от межэлектронных расстояний  $r_{ij}$ . Традиционно это делалось только для самых малых систем, таких, как атом He или молекула  $H_2$ , но зато это позволяло получать для них практически точные решения уравнения Шрёдингера. В настоящее время существуют (приближенные) методы, позволяющие включать некоторую зависимость от  $r_{ij}$  также и в волновые функции больших систем. Однако, на практике абсолютное большинство расчетов электронной корреляции проводится на основе волновой функции разложения КВ, обсуждавшегося в предыдущем разделе.

Почти все методы, учитывающие корреляцию, исходят из волновой функции (ограниченного или неограниченного) метода Хартри—Фока, как «состояния отсчета». Тогда упорядоченные конфигурации  $I$  соответствуют детерминантам, построенным из занятых и виртуальных орбиталей ХФ. (Это обычно — но не обязательно, — канонические орбитали ХФ.) Обозначая волновую функцию ХФ, как  $\Psi_0$  и используя промежуточную (или «корреляционную») нормировку, которая уже применялась в гл. 4, можем записать

$$\Psi = \Psi_0 + \varphi = \Psi_0 + \sum_{I=1} c_I \Psi_I, \quad (8.8)$$

где

$$\langle \Psi_0 | \Psi_0 \rangle = 1; \quad \langle \Psi_0 | \varphi \rangle = 0. \quad (8.9)$$

(Последние равенства следуют из ортонормированности занятых и виртуальных орбиталей ХФ, приводящей к  $\langle \Psi_0 | \Psi_I \rangle = \delta_{0I}$ .)

Подставляя разложение (8.8) в уравнение Шрёдингера

$$\hat{H}\Psi = E\Psi, \quad (8.10)$$

имеем

$$\hat{H} \left( \Psi_0 + \sum_{I=1} c_I \Psi_I \right) = E \left( \Psi_0 + \sum_{I=1} c_I \Psi_I \right). \quad (8.11)$$

Умножая равенство (8.11) на  $\Psi_0^*$  и интегрируя, получаем с учетом того, что  $\langle \Psi_0 | \Psi_I \rangle = 0$ , если  $I \neq 0$ :

$$E = \underbrace{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}_{E_{HF}} + \sum_{I=1} c_I \underbrace{\langle \Psi_0 | \hat{H} | \Psi_I \rangle}_{H_{0I}} \quad (8.12)$$

т. е.

$$\Delta E = E - E_{HF} = \sum_I c_I H_{0I}. \quad (8.13)$$

Так как детерминанты  $\Psi_I$  построены из занятых и виртуальных орбиталей ХФ, их можно классифицировать по числу спин-орбиталей, которые заменены виртуальными относительно состояния отсчета — волновой функции ХФ  $\Psi_0$ .

Таким образом, имеем одно-, дву- и трехкратно замещенные (или «возбужденные») детерминанты, и т. д. до  $N$ -кратно замещенных, где  $N$  есть число электронов в системе. Если  $\Psi_I$  — однократно возбужденный детерминант, то матричный элемент  $\langle \Psi_0 | \hat{H} | \Psi_I \rangle = 0$  по теореме Бриллюэна (гл. 6, разд. 1). Для детерминантов с тремя и более замещениями по отношению к  $\Psi_0$ , матричный элемент обращается в нуль по правилам Слэтера (гл. 5, разд. 6). Поэтому единственными слагаемыми, дающими вклад в правую часть равенства (8.13), являются двукратно возбужденные детерминанты. Вводя для них обозначение  $\Psi_{ij}^{uv} = \Psi_2(\psi_i \rightarrow \psi_u; \psi_j \rightarrow \psi_v)$  и обозначая соответствующие им коэффициенты в разложении КВ (8.8) точной волновой функции, как  $c_{ij}^{uv}$ , получаем выражение для корреляционной энергии в виде

$$\begin{aligned} \Delta E &= \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} c_{ij}^{uv} \langle \Psi_0 | \hat{H} | \Psi_{ij}^{uv} \rangle \\ &= \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} c_{ij}^{uv} \left( \langle \psi_i(1) \psi_j(2) | \frac{1}{r_{12}} | \psi_u(1) \psi_v(2) \rangle \right. \\ &\quad \left. - \langle \psi_i(1) \psi_j(2) | \frac{1}{r_{12}} | \psi_v(1) \psi_u(2) \rangle \right), \end{aligned} \quad (8.14)$$

где интегралы относятся к спин-орбиталям, т. е. интегрирование включает суммирование по спинам.

Результаты (8.14) показывают, что для вычисления точной корреляционной энергии нужно знать только коэффициенты дважды возбужденных конфигураций в точном разложении КВ. Этот результат обычно называется теоремой Несбета.

Важным следствием (8.14) является то, что после введения некоторых очевидных обозначений энергию корреляции можно представить в виде суммы

$$\Delta E = \sum_{i < j}^{(\text{зан.})} \varepsilon_{ij}, \quad (8.15)$$

где суммирование ведется по занятым *спин-орбиталям* ХФ. Этот результат показывает, что точную корреляционную энергию всегда можно представить суммой вкладов

$$\varepsilon_{ij} = \sum_{u < v}^{(\text{вирт.})} c_{ij}^{uv} \left( \langle \psi_i(1) \psi_j(2) | \frac{1}{r_{12}} | \psi_u(1) \psi_v(2) \rangle - \langle \psi_i(1) \psi_j(2) | \frac{1}{r_{12}} | \psi_v(1) \psi_u(2) \rangle \right), \quad (8.16)$$

которые обычно называют «парными корреляционными энергиями» отдельных пар занятых спин-орбиталей. Строго говоря, каждая такая парная корреляционная энергия зависит также от всех других орбиталей, так как на значения коэффициентов  $c_{ij}^{uv}$  прямо или косвенно влияют все конфигурации разложения КВ.

Если пространственная часть двух спин-орбиталей одинакова, т. е. они соответствуют той же самой дважды занятой пространственной орбитали, заполненной спинами  $\alpha$  и  $\beta$ , то говорят о «внутрипарной корреляции», в противном случае — о «межпарной корреляции». Обычно вклады внутрипарной корреляции наибольшие, так как они соответствуют паре электронов, которые в методе ХФ принуждены занимать одну и ту же пространственную орбиталь, что влечет за собой относительно большую вероятность того, что эти два электрона окажутся близко друг к другу. Аналогично различию между внутри- и межпарными корреляциями, можно также определить «внутриоболочечные» и «межоболочечные» корреляции для атомов и т. д.

Рассуждения, приводящие к равенствам (8.14) и (8.15), одинаково верны, независимо от того, являются ли хартри-фоковские спин-орбитали каноническими или нет. В то время как точная корреляционная энергия  $\Delta E$  инвариантна относительно унитарных преобразований занятых орбиталей, отдельные парные корреляционные энергии,

разумеется, не инвариантны. Однако, как  $\Delta E$ , так и каждая  $\varepsilon_{ij}$  инвариантны по отношению к унитарным преобразованиям виртуальных орбиталей. (Инвариантность  $\Delta E$  следует из того факта, что решение уравнения Шрёдингера не может зависеть от выбора базиса в линейном пространстве нормируемых антисимметричных  $N$ -электронных волновых функций. Из свойств определителей следует, что *подпространство*, растянутое детерминантами, получающимися двукратными возбуждениями с *данной пары* занятых спин-орбиталей, остается инвариантным при унитарных преобразованиях виртуальных орбиталей, что приводит к инвариантности парных корреляционных энергий.)

Можно ожидать, что доминирование внутривпарной корреляции будет даже более выраженным, если использовать локализованные орбитали, и что заметные значения будут иметь только парные корреляционные энергии, которые соответствуют орбиталям, находящимся не очень далеко друг от друга в физическом пространстве. (Как интегралы, так и коэффициенты в формуле (8.16) быстро спадают с увеличением расстояния.) Для похожих орбиталей (например, атомных  $1s$ -орбиталей) в разных молекулах получаются очень похожие значения. Это часто дает возможность исключить из рассмотрения корреляцию остовных орбиталей без заметного изменения интересующих нас разностей энергии.

С точки зрения внутривпарной корреляции, соответствующей данной дважды занятой орбитали, наибольшее значение имеют виртуальные орбитали, которые находятся преимущественно в той же области пространства, что и рассматриваемая занятая орбиталь, но обладают большим количеством узловых поверхностей. Например, в случае химической связи, описываемой в методе ССП локализованными связывающими орбиталями, существенная часть внутривпарной корреляции будет получаться при учете возбуждений на соответствующую антисвязывающую орбиталь. Таким образом описывается так называемая «право-левая» корреляция, что следует понимать так, что если один электрон пары находится около одного ядра («справа»), то другой электрон будет найден с большей вероятностью около другого («слева»), и наоборот. (Таким образом, учет электронной корреляции уменьшает относительный вес ионных составляемых, который переоценивается в методе ССП; см. гл. 6, разд. 4.3) Аналогично, особенно в случае свободных атомов, говорят о радиальной (или «внутри-наружу») корреляции, для описания которой необходимо использовать радиальные функции с разными показателями орбитальных экспонент и разным числом узловых поверхностей. (Например, чтобы описать радиальную корреляцию в  $1s$ -оболочке необходимы орбитали типа  $2s$ ,  $3s$  и т. д.) Другим важным классом корреляции является «угловая корреляция», которая означает, что

вероятность найти два электрона на разных направлениях относительно ядра, больше, чем найти их на том же самом направлении. Этот эффект описывается с помощью орбиталей с высшими азимутальными квантовыми числами. (Необходимо, например, использовать  $p, d, f$  и т. д. орбитали для описания угловой корреляции пары  $1s$ -электронов.) Вообще, при построении базиса разложения КВ необходимо использовать такие орбитали, которые сильно осциллируют и тем самым обеспечивают наличие нескольких изолированных областей в пространстве, где разные орбитали существенно отличаются от нуля. Важно помнить, что для адекватного описания кулоновской дырки нужно очень большое число разных одноэлектронных функций: как уже было отмечено, правильная  $r_{ij}$ -зависимость волновой функции воспроизводится в разложении КВ только как результат интерференции многих слагаемых.

Нужно отметить, что существует интересный подход, который сводится к попытке «срезать угол» при учете электронной корреляции, рассматривая все интересные величины, как функционалы единственной переменной, а именно электронной плотности. Этот подход называется «теорией функционала плотности» (Density functional theory, DFT). В его рамках *формально* доказывается, что существует «универсальный функционал» электронной плотности, задающий сумму обменной и корреляционной энергии. Точный вид этого обменно-корреляционного функционала, однако, неизвестен, и в нашем распоряжении нет никаких методов, которые позволили бы, начав с приближенного вида функционала, улучшать его каким-либо систематическим образом. По этой причине мнение автора таково, что DFT должен рассматриваться, как полезная полуэмпирическая схема расчетов, позволяющая приближенно учесть определенную часть электронной корреляции. Мы не будем здесь детально рассматривать методы, основанные на DFT.



### Пример

Рассмотрим простейший пример молекулы  $H_2$ , описываемой минимальным базисным набором, состоящим из двух атомных орбиталей  $\chi_A$  и  $\chi_B$ . Пространственная симметрия системы однозначно определяет МО. Имеем симметричную связывающую орбиталь

$$\varphi = \mathcal{N}(\chi_A + \chi_B) \quad (8.17)$$

и антисимметричную антисвязывающую орбиталь

$$\psi = \mathcal{N}'(\chi_A - \chi_B) . \quad (8.18)$$

Нормировочные коэффициенты  $\mathcal{N}$  и  $\mathcal{N}'$ , как легко проверить, равны

$$\mathcal{N} = \frac{1}{\sqrt{2(1+S)}}; \quad \mathcal{N}' = \frac{1}{\sqrt{2(1-S)}}, \quad (8.19)$$

где  $S = \langle \chi_1 | \chi_2 \rangle$ . Вследствие их разной симметрии, орбитали  $\varphi$  и  $\psi$  автоматически ортогональны.

В этой простой модельной задаче существуют только две детерминантные волновые функции  $\Psi_0 = \hat{\mathcal{A}}[\varphi(1)\alpha(1)\varphi(2)\beta(2)]$  и  $\Psi_d = \hat{\mathcal{A}}[\psi(1)\alpha(1)\psi(2)\beta(2)]$ , которые симметричны по отношению к инверсии относительно центра молекулы; обе они построены из дважды занятых орбиталей. Волновая функция основного состояния, записанная с использованием корреляционной нормировки, равна

$$\Psi = \Psi_0 + c\Psi_d. \quad (8.20)$$

(Заметим, что однократно возбужденные детерминанты  $\hat{\mathcal{A}}[\varphi(1)\alpha(1)\psi(2)\beta(2)]$  и  $\hat{\mathcal{A}}[\psi(1)\alpha(1)\varphi(2)\beta(2)]$  антисимметричны по отношению к инверсии и потому отсутствуют в разложении КВ волновой функции основного состояния.)

Подставив явные выражения для МО через АО в определение  $\Psi$  (8.20), получим

$$\begin{aligned} \Psi &= \hat{\mathcal{A}}[\varphi(1)\alpha(1)\varphi(2)\beta(2)] + c\hat{\mathcal{A}}[\psi(1)\alpha(1)\psi(2)\beta(2)] \\ &= \mathcal{N}^2 \hat{\mathcal{A}}\{[\chi_A(1) + \chi_B(1)]\alpha(1)[\chi_A(2) + \chi_B(2)]\beta(2)\} \\ &\quad + c\mathcal{N}'^2 \hat{\mathcal{A}}\{[\chi_A(1) - \chi_B(1)]\alpha(1)[\chi_A(2) - \chi_B(2)]\beta(2)\} \\ &= \mathcal{N}^2 \left\{ \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right. \\ &\quad \left. + \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] \right\} \\ &\quad + c\mathcal{N}'^2 \left\{ \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] - \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right. \\ &\quad \left. - \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] \right\} \\ &= (\mathcal{N}^2 + \mathcal{N}'^2 c) \underbrace{\left\{ \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] \right\}}_{\text{ионные слагаемые (H}^-\text{H}^+ \text{ и H}^+\text{H}^-)} \\ &\quad + (\mathcal{N}^2 - \mathcal{N}'^2 c) \underbrace{\left\{ \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right\}}_{\text{ковалентные слагаемые}}. \end{aligned} \quad (8.21)$$

Ковалентные слагаемые представляют собой так называемую волновую функцию Гайтлера—Лондона. Рождение квантовой химии, как



науки, можно датировать работой Гайтлера и Лондона (1927), которая являлась первым примером, в котором удалось по-настоящему объяснить существование химической связи на основе законов физики (квантовой механики) и дать ее полук количественное описание.

Однодетерминантная волновая функция МО  $\Psi_0$  содержит избыток ионных слагаемых — она не учитывает электронную корреляцию «право-лево». (Это, собственно, и приводит к «диссоциационной катастрофе», обсуждавшейся в гл. 6, разд. 4.3) В то же время, волновую функцию Гайтлера—Лондона, содержащую только ковалентные слагаемые, можно рассматривать, как «слишком коррелированную». Оптимальная, т. е. истинная, величина корреляции «право-лево» дается решением задачи КВ (оптимизацией коэффициента  $c$ ); это эквивалентно оптимизации веса ионных слагаемых, что было впервые сделано Вайнбаумом (1933).



### 3. Методы учета электронной корреляции

Полное описание методов, используемых для расчета электронной корреляции находится вне рамок нашей книги. Мы обсудим лишь очень кратко наиболее широко используемые методы, указывая на те их аспекты, которые кажутся важными с общетеоретической точки зрения.

#### 3.1. Метод КВ

Метод КВ концептуально (но, конечно, не практически) очень прост. Запишем разложение КВ (8.5), соответствующее данному набору одноэлектронных базисных функций и оптимизируем коэффициенты  $c_I$ , решая вековое уравнение (см. гл. 3, разд. 1). Часто многоэлектронные функции, используемые как базисные функции в линейной вариационной задаче, являются не отдельными детерминантами, а их комбинациями, представляющими собой чистые спиновые функции (собственные векторы операторов  $\hat{S}_z$  и  $\hat{S}^2$ ) и, возможно, соответствуют пространственной симметрии системы. Использование таких приведенных по спину и симметрии конфигураций может существенно снизить размерность векового уравнения.

В том случае, если решают уравнение на собственные значения, учитывая все детерминанты (или приведенные по спину конфигурации), которые можно построить для данного одноэлектронного базиса, то говорят о методе «полного КВ» (ПКВ, FCI, full CI) — он является точным решением вариационной задачи в данном базисе одноэлектронных функций. Можно рассматривать это так, что используемый базис определяет

некоторую модельную задачу с соответствующим модельным гамильтонианом, и метод полного КВ дает *точное решение* уравнения Шрёдингера, записанного для такого *модельного гамильтониана*. Можно сформулировать способ решения следующим образом: вместо решения уравнения Шрёдингера  $\hat{H}\Psi = E\Psi$  решают уравнение Шрёдингера со спроектированным гамильтонианом

$$\hat{P}\hat{H}\hat{P}\Psi = E\Psi, \quad (8.22)$$

где  $\hat{P}$  является оператором проекции на подпространство всех антисимметричных  $N$ -электронных волновых функций, возможных в данном одноэлектронном базисе. (По мере того, как базис увеличивается и становится полным ( $\hat{P} \rightarrow \hat{1}$ ), решения задачи полного КВ стремятся к точным решениям исходного уравнения Шрёдингера.) Очевидно, при работе в данном конечном базисе нет необходимости использовать явно проектор  $\hat{P}$ , так как рассматривают только те волновые функции, для которых  $\hat{P}\Psi \equiv \Psi$ .

Согласно сказанному, результаты ПКВ зависят только от используемого подпространства одноэлектронных функций, но не зависят от конкретного базиса одноэлектронных функций, выбираемого в этом подпространстве для построения детерминантов (или их комбинаций, приведенных по спину и симметрии), служащих базисными функциями при решении  $N$ -электронной линейной вариационной задачи. Работать с детерминантами (конфигурациями), построенными прямо из неортогональных одноэлектронных базисных орбиталей, обычно применяемых в квантовой химии, очевидно, было бы непрактично. Приемлем любой ортогонализованный набор орбиталей, растягивающий то же самое одноэлектронное подпространство. Обычно используют занятые и виртуальные хартри-фоковские орбитали, которые позволяют достичь быстрой сходимости при итерационном решении уравнения на собственные значения.

Как отмечалось в предыдущем разделе, число детерминантов растет очень быстро с увеличением размера базиса и числа электронов. Поэтому на практике возможно проводить расчеты ПКВ только для совсем малых систем, описываемых очень небольшими базисами, даже если используются самые мощные компьютеры. Тем не менее, эти расчеты очень важны с методологической точки зрения, так как дают нам эталоны, с которыми должны сравниваться все другие (приближенные) результаты.

Расчеты полным методом КВ за пределами дороги практически для всех задач, представляющих химический интерес, но нет и реальной необходимости проводить их (за исключением упомянутых эталонных

целей), так как абсолютное большинство конфигураций (детерминантов) не вносит какого-либо ощутимого вклада в волновую функцию и практически не влияет на результаты, особенно если речь идет об энергии (которая обычно наиболее важна для приложений).

Если использовать разложение (8.5), применяя хартри-фоковские орбитали, то из теоремы Несбета, обсуждавшейся в предыдущем разделе, следует, что — кроме основного состояния ХФ волновой функции  $\Psi_0$ , — наиболее важными детерминантами являются дважды возбужденные: если бы мы знали их точные коэффициенты в разложении КВ, то могли бы вычислить и точную энергию ПКВ.

Согласно предыдущему замечанию, простейшим методом КВ для описания основных состояний молекул с замкнутыми оболочками был бы такой, в котором рассматриваются только двукратные возбуждения. Этот метод, однако, не используется, так как, хотя однократно возбужденные детерминанты (конфигурации) и не взаимодействуют с хартри-фоковским основным состоянием согласно теореме Бриллюэна, тем не менее каждый однократно возбужденный детерминант (конфигурация) взаимодействует с несколькими двукратно возбужденными, и, таким образом, их косвенным влиянием на рассчитанную корреляционную энергию нельзя пренебречь. В то же время, число однократно возбужденных конфигураций много меньше, чем число двукратно возбужденных, поэтому их включение в расчет не приводит к заметному увеличению вычислительных затрат. (Число однократно возбужденных состояний пропорционально  $n_{\text{зан}} \times n_{\text{вирт}}$ , где  $n_{\text{зан}}$  и  $n_{\text{вирт}}$  — число занятых и виртуальных ХФ орбиталей, тогда как число двукратно возбужденных орбиталей, очевидно, пропорционально  $n_{\text{зан}}^2 \times n_{\text{вирт}}^2$ ).

Включение всех однократно и двукратно возбужденных детерминантов в расчет КВ (так называемый метод CISD, configuration interaction singles and doubles) позволяет учесть значительную часть корреляционной энергии полного КВ в малых молекулах. Серьезным недостатком данного метода является отсутствие у него размерной согласованности (ср. гл. 4, разд. 4.1). Если мы рассмотрим две невзаимодействующие системы, и опишем каждую из них методом CISD, как подсистемы формальной объединенной системы, то волновые функции должны быть антисимметризованным произведением волновых функций двух подсистем. Однако, такое произведение двух волновых функций CISD не является волновой функцией типа CISD: оно содержит также слагаемые, которые формально соответствуют трех- и четырехкратным возбуждениям и не включаются в CISD-описание объединенной системы. По этой причине метод CISD больше уже не используется широко, исключая, возможно, случаи, когда вводятся приближенные поправки на размер-

ную согласованность. Необходимо помнить, что никакой «оборванный» ряд КВ (т. е. КВ, не включающее все возбуждения) никогда не будет размерно согласованным. Схема КВ, включающая возбуждения вплоть до четырехкратных, вероятно давала бы ответ с достаточно малой ошибкой за счет размерной несогласованности. Такие расчеты КВ, однако, были бы чрезмерно затратны для химически интересных проблем, вследствие очень большого числа различных трех- и четырехкратных возбуждений. В то же время, строго размерно-согласованные результаты подобного качества можно получить с помощью методов возмущений или «связанных кластеров» (разд. 3.3 и 3.4) с намного меньшими вычислительными усилиями.

При выполнении практических расчетов КВ матрицу гамильтониана (КВ)  $\mathbf{H}$  с элементами  $H_{IJ} = \langle \Psi_I | \hat{H} | \Psi_J \rangle$ , где  $\Psi_I$  и  $\Psi_J$  суть  $N$ -электронные детерминанты (конфигурации) явно не строят, а используют так называемый «прямой» метод, который позволяет непосредственно рассчитывать векторы  $\mathbf{y} = \mathbf{H}\mathbf{x}$  по векторам  $\mathbf{x}$ , содержащим коэффициенты разложения КВ, используя одно- и двухэлектронные интегралы по ортонормированным орбиталям  $\varphi_i$ . Для этой цели необходимо перебирать интегралы и находить те компоненты произведения матрицы на вектор  $\mathbf{H}\mathbf{x}$ , в которые может дать вклад данный интеграл, и вычислять эти вклады (без явного расчета матричных элементов  $H_{IJ}$ ).

### 3.2. Методы МК ССП

При проведении расчетов КВ с очень большим числом конфигураций большинство из них в конечном счете могут оказаться несущественными. Вместо этого можно попытаться использовать ограниченное число подходящим образом подобранных конфигураций, которые содержат лишь часть доступных одноэлектронных орбиталей, и провести оптимизацию не только коэффициентов разложения КВ, но и выбранных орбиталей. Этот подход называется методом «многоконфигурационного ССП» (МК ССП, MCSCF), объединяющим подходы схем ССП и КВ. Расчет МК ССП состоит из двух отдельных частей, которые должны повторяться циклически до достижения сходимости: оптимизации коэффициентов КВ при фиксированных текущих одноэлектронных орбиталях, что является стандартной линейной вариационной задачей, и оптимизации одноэлектронных орбиталей с фиксированными текущими коэффициентами КВ, что представляет собой процедуру, аналогичную методу ОХФО, обсуждавшемуся в гл. 6, разд. 6.

Приведенное определение охватывает огромное многообразие возможных схем МК ССП. Одним из предельных случаев тут является

использование только нескольких — возможно, лишь двух — детерминантов, которые необходимы для учета статической корреляции (разд. 2), характерной для рассматриваемой задачи. Другим пределом является так называемый метод «полного активного пространства ССП» (CASSCF), в котором выбирают несколько электронов и небольшое число орбиталей и для них по существу решается задача полного КВ, тогда как остающиеся электроны распределяются по дважды занятым орбиталям.

Так называемый метод АПСГ («антисимметризованное произведение строго ортогональных геминалей», APSG) тоже можно рассматривать в качестве специального случая метода МК ССП. В методе АПСГ используется тот факт, что многие химические системы можно рассматривать, как состоящие из синглетно связанных пар электронов. Тогда пробную волновую функцию соответствующего класса («Ansatz») можно записать, как:

$$\Psi(1, 2, \dots, N) = \hat{A} \prod_{i=1}^{N/2} g_i(2i-1, 2i) \frac{1}{\sqrt{2}} [\alpha(2i-1)\beta(2i) - \beta(2i-1)\alpha(2i)] . \quad (8.23)$$

При этом предполагается, что *двухэлектронные функции* («геминали»)  $g_i$ ,  $i = 1, 2, \dots, N/2$  подлежат оптимизации. (Каждая геминаль является симметричной функцией своих пространственных аргументов).

С волновыми функциями такого типа можно работать практически, только если ввести так называемое условие строгой ортогональности:

$$\int g_i(1, 2)g_j(1, 2)dv_1 = \int g_i(1, 2)g_j(1, 2)dv_2 = 0 \text{ если } i \neq j, \quad (8.24)$$

т. е. интеграл произведения разных геминалей, взятый по координатам даже только одного электрона, дает нуль. Геминали обычно разлагают по произведениям одноэлектронных функций  $\varphi_k$ :

$$g_i(1, 2) = \sum_{k,l} c_{kl}^i \varphi_k(1) \varphi_l(2), \quad (8.25)$$

где  $c_{kl}^i = c_{lk}^i$ . Таким образом, функция (8.23) по существу является частным случаем разложения (8.5). Строгую ортогональность (8.24) можно гарантировать тривиально, если приписать отдельным геминалям ортогональные поднаборы ортонормированного базиса одноэлектронных функций  $\{\varphi_k\}$ , не имеющие общих элементов (непересекающиеся подмножества).

Возможность рассматривать волновую функцию АПСГ как частный случай МК ССП не означает, что на практике ее оптимизация проводится разложением в сумму детерминантов. Но в случае АПСГ также необходимо оптимизировать как орбитали  $\varphi_k, \varphi_l$ , так и коэффициенты линейных комбинаций  $c_{kl}^i$ .

Заметим здесь, что метод «обобщенных валентных связей» (generalized valence bond, GVB) Годдарда является простым случаем АПСГ: для каждой геминали используется только одна пара одноэлектронных орбиталей, но здесь эти орбитали не обязаны быть ортогональны друг другу. (Однако они должны быть ортогональными к каждой орбитали, используемой в других геминалях). После ортогонализации этих орбиталей можно выразить волновую функцию таким образом, что для каждой пары электронов появляются четыре слагаемых, обсуждавшихся выше в разд. 2 для молекулы  $\text{H}_2$ . (Такая волновая функция, выписанная для одной электронной пары молекулы  $\text{H}_2$  является классической волновой функцией Коулсона—Фишер (1949). В случае минимального базиса она эквивалентна волновой функции КВ, приведенной в разд. 2.)

### 3.3. Методы многочастичной теории возмущений или теория возмущений Меллера—Плессета

Электронную корреляцию можно учесть (частично), применяя теорию возмущений (гл. 4). Наиболее широко используемый вариант по причинам исторического характера часто называется «многочастичной теорией возмущений» (many body perturbation theory, МВРТ), что указывает на диаграммную технику выкладок, перенятую из теории поля. Многочастичная теория возмущений является ничем иным, как приложением детально обсуждавшейся в гл. 4 теории возмущений Рэля—Шрёдингера к проблеме электронной корреляции, с помощью так называемого разбиения гамильтониана Борна—Оппенгеймера по Меллеру—Плессету. Соответственно,  $n$ -й порядок такой теории обычно сокращенно обозначают  $\text{MP}n$  ( $\text{MP}2$ ,  $\text{MP}3$  и т. д.), среди них наиболее употребимыми вариантами являются  $\text{MP}2$  и  $\text{MP}4$ .

При разбиении по Меллеру—Плессету невозмущенный гамильтониан выбирается в виде суммы операторов Фока всех электронов системы:

$$\hat{H}^0 = \sum_{i=1}^N \hat{F}(i). \quad (8.26)$$

Любая детерминантная волновая функция, построенная из (занятых или виртуальных) канонических орбиталей Хартри—Фока является соб-

ственным вектором этого невозмущенного гамильтониана с собственным значением, равным сумме орбитальных энергий тех орбиталей, из которых построен рассматриваемый детерминант. (Заметим, что  $\hat{H}^0$  коммутирует с оператором антисимметризации; ср. гл. 5, разд. 4.2.)

В силу свойств инвариантности оператора Фока (гл. 6, разд. 2.1), можно построить *тот же*  $\hat{H}^0$  не только с помощью канонических орбиталей, но и с помощью любого набора ортогональных орбиталей (например, локализованных), растягивающих занятое подпространство. Однако только использование канонических орбиталей обеспечивает то, что все «упорядоченные конфигурации», построенные из них (ср. разд. 1), являются собственными функциями *этого* невозмущенного гамильтониана. Тем не менее, при описании действительно больших систем может иметь смысл применять локализованные орбитали и воспользоваться сильно выраженным локальным характером парных корреляционных энергий (ср. разд. 2), даже идя на то, что для расчета поправок теории возмущений придется решать несколько более сложные уравнения.

Строго говоря, выражение (8.26) верно только в ОХФ случае, когда используются дважды занятые орбитали и необходимо работать с одним оператором Фока, и оно относится к ограниченной теории Меллера-Плессета «RMPn» («restricted» Møller-Plesset), или теории MPn в узком смысле. Если для однопредетерминантной волновой функции использовать теорию НХФ, как это обычно делается в случае нечетного числа электронов или других систем с открытой оболочкой, то для орбиталей спина  $\alpha$  и  $\beta$  будем иметь различные фокианы. Тогда невозмущенный гамильтониан можно записать как

$$\hat{H}^0 = \sum_{i=1}^N \left[ \hat{F}^{\alpha}(i) \delta_{\alpha\gamma_i} + \hat{F}^{\beta}(i) \delta_{\beta\gamma_i} \right], \quad (8.27)$$

где символы Кронекера показывают, что операторы  $\hat{F}^{\alpha}(i)$  и  $\hat{F}^{\beta}(i)$  действуют только на спин-орбитали соответствующего спина. (Определение (8.27) применимо и в случае ОХФ, для которого  $\hat{F}^{\alpha} = \hat{F}^{\beta} = \hat{F}$ ).

По определению (8.27), энергию нулевого порядка можно записать как

$$E_0 = \sum_i^{(\text{зан.})} \varepsilon_i^{\gamma_i}, \quad (8.28)$$

где «(зан.)» показывает, что суммирование должно вестись по всем орбиталиям, занятым в детерминанте основного состояния Хартри—Фока.

(Несколько позже аналогичное обозначение «(вирт.)» будет использовано для суммирования по всем виртуальным орбиталям.)

Согласно обычной практике теории возмущений (ср. разбиение (4.2)), запишем  $\hat{H} = \hat{H}^0 + \hat{V}$ , т. е. оператор возмущения будет равен

$$\hat{V} = \hat{H} - \hat{H}^0. \quad (8.29)$$

(Мы не используем здесь формальный параметр интенсивности возмущения  $\lambda$ .)

Поправка первого порядка к энергии равна среднему значению оператора возмущения по невозмущенной волновой функции — ср. уравнение (4.23):

$$E^{(1)} = \langle \Psi_0 | \hat{V} | \Psi_0 \rangle = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle - \sum_i^{(\text{зан.})} \varepsilon_i^{\gamma_i}. \quad (8.30)$$

(Хартри-фоковские орбитали предполагаются ортонормированными, так что  $\langle \Psi_0 | \Psi_0 \rangle = 1$ .)

Согласно результатам (8.28) и (8.30), сумма энергий нулевого и первого порядков равна энергии Хартри—Фока, т. е. среднему значению полного гамильтониана по невозмущенной (хартри-фоковской) волновой функции.

Поправку второго порядка к энергии (4.30) можно записать через упорядоченные конфигурации как

$$E^{(2)} = \sum_{\substack{K \\ (K \neq 0)}} \frac{-|\langle \Psi_K | \hat{H} | \Psi_0 \rangle|^2}{E_K - E_0} = \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} \frac{-| [uv || ij]' |^2}{\varepsilon_u^{\gamma_u} + \varepsilon_v^{\gamma_v} - \varepsilon_i^{\gamma_i} - \varepsilon_j^{\gamma_j}} \quad (8.31)$$

где сумма пробегает по двукратно возбужденным конфигурациям, ибо, в соответствии с обсуждением в разд. 2, только двукратно возбужденные детерминанты взаимодействуют с хартри-фоковским основным состоянием. (Мы воспользовались тем фактом, что  $\langle \Psi_K | \hat{H}^0 | \Psi_0 \rangle = 0$ , если  $K \neq 0$ , так что  $\langle \Psi_K | \hat{V} | \Psi_0 \rangle = \langle \Psi_K | \hat{H} | \Psi_0 \rangle$ ).

В выражении (8.31) мы использовали упрощенное обозначение «с двойными черточками» для разности интегралов:

$$[ab || cd]' = \langle \psi_a(1) \psi_b(2) | \frac{1}{r_{12}} | \psi_c(1) \psi_d(2) \rangle - \langle \psi_a(1) \psi_b(2) | \frac{1}{r_{12}} | \psi_d(1) \psi_c(2) \rangle. \quad (8.32)$$

(Штрих указывает, что интегрирование включает суммирование по проекциям спина, что отличает его от аналогичных интегралов, записанных для чисто пространственных орбиталей).



Результат (8.31) полностью согласуется с теоремой Несбета, так как волновая функция первого порядка теории возмущений, отвечающая приближению (8.31), представляет собой линейную комбинацию двукратно возбужденных детерминантов с коэффициентами, равными, в соответствии с результатом (4.28):

$$c_{ij}^{uv} = - \frac{[uv||ij]'}{\varepsilon_u^{\gamma_u} + \varepsilon_v^{\gamma_v} - \varepsilon_i^{\gamma_i} - \varepsilon_j^{\gamma_j}}. \quad (8.33)$$

Чтобы получить выражение для энергии второго порядка в виде, более удобном для практических расчетов, необходимо переписать выражение (8.31) через пространственные орбитали. Чтобы сделать это, предположим, что волновая функция Хартри—Фока построена из  $n_\alpha$  орбиталей  $a_i$ , занятых спином  $\alpha$ , и  $n_\beta$  орбиталей  $b_i$ , занятых спином  $\beta$ . (Здесь мы рассматриваем случай НХФ, который содержит в себе ОХФ, как частный случай, с  $n_\alpha = n_\beta = n$ ;  $a_i = b_i = \varphi_i$ ). Допустим, что спин-орбитали упорядочены так, что первые  $n_\alpha$  орбиталей заполнены электронами с проекцией спина  $\alpha$ , а затем следуют спин-орбитали, заполненные электронами с проекцией спина  $\beta$ . Аналогичную нумерацию можно ввести также и для виртуальных орбиталей. В этом случае сумма по спин-орбиталям  $\psi_i$  и  $\psi_j$  с ограничением  $i < j$ , введенным в формуле (8.31), распадется на три суммы — две для случая, когда  $\psi_i$  и  $\psi_j$  имеют одинаковые проекции спина, и одна для случая, когда  $\psi_i$  имеет проекцию спина  $\alpha$ , а  $\psi_j$  — проекцию спина  $\beta$ . Первые два суммирования должны идти по парам орбиталей  $a_i, a_j$  и  $b_i, b_j$  с ограничением  $i < j$  для обоих, тогда как третье суммирование идет по парам орбиталей  $a_i, b_j$  без ограничения на значения индексов  $i$  и  $j$ , так как, согласно принятому порядку орбиталей, каждая орбиталь, занятая электроном с проекцией спина  $\alpha$ , имеет номер, меньший, чем любая орбиталь, занятая электроном с проекцией спина  $\beta$ . Кроме того, необходимо учитывать, что если обе занятые орбитали, встречающиеся в данном слагаемом, имеют проекцию спина  $\alpha$  ( $\beta$ ), то виртуальные орбитали также обе должны иметь те же проекции спинов, иначе при суммировании по проекциям спина интегралы исчезают. Аналогично, если одна орбиталь занята электроном с проекцией спина  $\alpha$ , а другая — с проекцией спина  $\beta$ , то же самое должно иметь место и для виртуальных орбиталей. В этом случае второе слагаемое в правой части выражения (8.32) обращается в нуль вследствие ортогональности спиновых функций. Первый интеграл в правой части выражения (8.32) обозначим как

$$[ab|cd]' = \langle \psi_a(1)\psi_b(2) | \frac{1}{r_{12}} | \psi_c(1)\psi_d(2) \rangle, \quad (8.34)$$

— без двойной вертикальной черточки.

Далее опустим штрихи, так как в выражении остаются только слагаемые, для которых суммирование по спину дает множитель 1, и будем считать, что интегралы берутся по пространственным орбиталям. Тогда мы получим общую формулу для корреляционной энергии МР2 (обычно обозначаемой UMP2), соответствующую невозмущенной волновой функции НХФ:

$$\begin{aligned} \Delta E_{\text{UMP2}} = & - \left\{ \sum_{i < j}^{(\text{зан.}\alpha)} \sum_{u < v}^{(\text{вирт.}\alpha)} \frac{|[a_i a_j][a_u a_v]|^2}{\varepsilon_u^a + \varepsilon_v^a - \varepsilon_i^a - \varepsilon_j^a} \right. \\ & + \sum_{i < j}^{(\text{зан.}\beta)} \sum_{u < v}^{(\text{вирт.}\beta)} \frac{|[b_i b_j][b_u b_v]|^2}{\varepsilon_u^b + \varepsilon_v^b - \varepsilon_i^b - \varepsilon_j^b} \\ & \left. + \sum_i^{(\text{зан.}\alpha)} \sum_j^{(\text{зан.}\beta)} \sum_u^{(\text{вирт.}\alpha)} \sum_v^{(\text{вирт.}\beta)} \frac{|[a_i b_j][a_u b_v]|^2}{\varepsilon_u^a + \varepsilon_v^b - \varepsilon_i^a - \varepsilon_j^b} \right\}. \end{aligned} \quad (8.35)$$

Так как волновая функция ОХФ есть частный случай НХФ, эта формула может использоваться и для случая ОХФ (часто так и делается), если подставить в нее  $a_i = b_i = \varphi_i$ . Однако в последнем случае возможны и некоторые дальнейшие упрощения.

Так, для случая ОХФ первые две суммы в формуле (8.35) становятся одинаковыми. Можно объединить их в одну сумму по  $i, j$ ;  $i \neq j$ , переставляя  $i$  и  $j$  во второй сумме и используя тождество  $[\varphi_j \varphi_i | \varphi_u \varphi_v]^2 \equiv (-[\varphi_i \varphi_j | \varphi_u \varphi_v])^2 = [\varphi_i \varphi_j | \varphi_u \varphi_v]^2$ . (Теперь мы предполагаем, что орбитали вещественны и нет нужды указывать абсолютные значения). Третью сумму в формуле (8.35) можно разбить на три, соответствующие случаям  $u < v$ ,  $u = v$ , и  $u > v$ . Переставим индексы  $u$  и  $v$  в последнем случае, и выделим слагаемые с  $i = j$ . ( $[\varphi_i \varphi_i | \varphi_v \varphi_u] \equiv [\varphi_i \varphi_i | \varphi_u \varphi_v]$ .) Затем распишем квадрат  $[\varphi_i \varphi_j | \varphi_u \varphi_v]^2 = ([\varphi_i \varphi_j | \varphi_u \varphi_v] - [\varphi_i \varphi_j | \varphi_v \varphi_u])^2$  и объединим суммы по  $i, j$ ;  $i \neq j$  в одну, учитывая тождество  $[\varphi_i \varphi_j | \varphi_v \varphi_u] \equiv [\varphi_j \varphi_i | \varphi_u \varphi_v]$  и тот факт, что  $i$  и  $j$  являются независимыми индексами суммирования, которые можно переставлять. Получим

$$\begin{aligned} \Delta E_{\text{RMP2}} = & \left\{ \sum_{\substack{i, j \\ (i \neq j)}}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} \frac{2[\varphi_i \varphi_j | \varphi_u \varphi_v] (2[\varphi_i \varphi_j | \varphi_u \varphi_v] - [\varphi_i \varphi_j | \varphi_v \varphi_u])}{\varepsilon_u + \varepsilon_v - \varepsilon_i - \varepsilon_j} \right. \\ & \left. + \sum_i^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} \frac{2[\varphi_i \varphi_i | \varphi_u \varphi_v]^2}{\varepsilon_u + \varepsilon_v - 2\varepsilon_i} + \sum_{i, j}^{(\text{зан.})} \sum_u^{(\text{вирт.})} \frac{[\varphi_i \varphi_j | \varphi_u \varphi_u]^2}{2\varepsilon_u - \varepsilon_i - \varepsilon_j} \right\}. \end{aligned} \quad (8.36)$$

Это выражение является одной из эквивалентных форм, удобной для программирования корреляционной энергии в приближении RMP2.

С помощью выражений, полученных в рамках теории возмущений Рэлея—Шрёдингера (гл. 4), можно, в принципе, получить выражения и для поправок более высоких порядков теории возмущений МР к энергии. Так как число слагаемых растет очень быстро, эти выкладки обычно проводят с использованием диаграммной техники, позволяющей вести «учет» нужных слагаемых. Мы не будем детально обсуждать поправки высших порядков, а заметим лишь, что формула для поправки третьего порядка к энергии содержит только двукратно возбужденные конфигурации, как и поправка второго порядка к энергии. (Обычно считается, что результаты МР3 плохо сбалансированы).

Для того, чтобы найти поправку четвертого порядка к энергии, требуется знать, по крайней мере (ср. гл. 4, разд. 1.4.2), поправку второго порядка к волновой функции, которая включает все конфигурации, взаимодействующие с основным состоянием Хартри—Фока и двукратно возбужденными конфигурациями, входящими в поправку первого порядка к волновой функции, т. е. все конфигурации до четырехкратно возбужденных включительно. Хотя роль четырехкратно возбужденных конфигураций («четверных»), главным образом, состоит в обеспечении размерной согласованности, и работа с ними не представляет слишком большой проблемы на практике, описание большого числа трехкратно возбужденных конфигураций («тройных») оказывается более трудной задачей (требует больше компьютерного времени). Так как влияние трехкратных возбуждений на энергию корреляции обычно весьма умеренное, иногда ими пренебрегают и получают приближенную схему МР4, часто обозначаемую как МР4(SPQ), по сравнению с «полной» МР4, которую можно обозначить как МР4(SPTQ), чтобы подчеркнуть, что все возбуждения включены.

Отметим, что и  $1s$ -оболочки атомов часто оставляют «замороженными», т. е. не включают их в теорию МР и другие расчеты корреляционной энергии. Хотя корреляционная энергия, которую можно сопоставить  $1s$ -орбиталям, вовсе не мала<sup>1</sup>, она практически не зависит от химического окружения атомов и не влияет на *разности* энергий, которые интересны с практической точки зрения.

Вследствие присутствия слагаемых самоотталкивания в операторе Фока, энергии виртуальных орбиталей сдвинуты вверх (гл. 6, разд. 3.1) и это приводит к сравнительно большим энергетическим знаменателям

<sup>1</sup> Учет зависимости от  $r_{12}$  существенен именно для сильно сжатых дважды занятых остовных состояний. — *Прим. ред.*

в формулах теории МР. В результате, в обычных МР-расчетах молекул в окрестности их равновесных конфигураций получается только часть корреляционной энергии. (Напротив, так называемое разбиение Эпштейна—Несбета, при котором возмущение представляют собой внедиагональные элементы матрицы КВ, иногда может привести к значениям энергии, более низким, чем точные).

Еще раз подчеркнем, что методы теории возмущений не являются вариационными и поэтому не дают оценок сверху для энергии. В методах МР это может привести к совершенно неправильным потенциальным кривым вдали от равновесных конформаций, например, когда рассматривается гомолитическая диссоциация, т. е. методы МР плохо подходят для описания статической корреляции (ср. разд. 2).

### 3.4. Метод связанных кластеров

Как мы уже неоднократно говорили, если имеется несколько невзаимодействующих подсистем, то полная волновая функция является (антисимметризованным) произведением волновых функций отдельных подсистем. Это означает, что для обеспечения размерной согласованности необходимо включать в волновую функцию такие возбуждения, которые — по крайней мере, формально, — имеют более высокий порядок, чем используемые для описания отдельных подсистем. С точки зрения объединенной системы, появление высших возбуждений, являющихся произведениями независимых возбуждений низшего порядка, локализованных в подсистемах, можно рассматривать, как кинематический эффект; их необходимо отличать от тех же возбуждений высших порядков, которые отражают истинные динамические эффекты в системе. Однако, приведенная формулировка «или—или» описывает лишь два предельных случая, а на самом деле речь идет об одних и тех же возбуждениях. Если две подсистемы сближаются, то между ними возникают все более интенсивные взаимодействия, вследствие чего коэффициенты слагаемых, соответствующих возбуждениям высших порядков, просто будут все больше отклоняться от произведений коэффициентов возбуждений низших порядков внутри подсистем. В некоторых случаях, тем не менее, можно ожидать, что эти коэффициенты останутся близки к указанным произведениям даже на малых расстояниях, так как истинные (динамические) эффекты высших порядков малы; в других случаях могут иметь место значительные изменения. Например, однократно возбужденные конфигурации вносят вклад в волновые функции полного КВ изолированных подсистем только в малой степени, так как они проявляются лишь косвенным образом, вследствие их взаимо-

действий с двукратно (и трехкратно) возбужденными конфигурациями. Волновая функция объединенной системы, состоящей из двух невзаимодействующих подсистем, будет содержать двукратные возбуждения, соответствующие одновременным однократным возбуждениям двух отдельных подсистем, с коэффициентами, равными произведению коэффициентов этих однократных возбуждений. (Тогда эти коэффициенты будут очень малы.) Согласно требованию размерной согласованности, эти двукратные возбуждения не внесут никакого вклада в получаемую корреляционную энергию, так как соответствующие интегралы в выражении (8.14) обращаются в нуль из-за того, что подсистемы не взаимодействуют. На большом, но конечном расстоянии между подсистемами упомянутые интегралы больше не исчезают и это приводит к тому, что коэффициенты при двукратно возбужденных конфигурациях начинают быстро отклоняться от простого произведения, также появится и ненулевой вклад в корреляционную энергию. (Компонента корреляционной энергии, связанная с этими двукратными возбуждениями, описывает важное — дисперсионное — взаимодействие). При еще меньших расстояниях рассуждения о возбуждениях отдельных подсистем становятся бессмысленными, и мы имеем постепенный переход к обычным двукратным возбуждениям, вызванным электронной корреляцией.

Желательно иметь формализм, в рамках которого можно было бы отделить друг от друга кинематические и динамические вклады высших порядков возбуждений в данную волновую функцию. Это позволяет проводить расчеты, куда автоматически включены все кинематические возбуждения высшего порядка (тем самым обеспечивается размерная согласованность) и вычислять настоящие динамические эффекты высших порядков только в той степени, которая требуется физикой задачи (и доступна в вычислительном плане).

Чтобы представить общие черты соответствующего формализма, рассмотрим сначала простую модель  $M$  невзаимодействующих подсистем. Допустим, что волновая функция отдельной подсистемы может быть задана с помощью промежуточной нормировки как

$$\Psi^{(i)} = \Psi_0^{(i)} + \varphi^{(i)} = \Psi_0^{(i)} + q_i \Psi_1^{(i)} = (1 + \hat{X}_i) \Psi_0^{(i)}, \quad (8.37)$$

где оператор возбуждения  $\hat{X}_i$  проще всего задать с помощью формализма «бра-кет»

$$\hat{X}_i = q_i |\Psi_1^{(i)}\rangle \langle \Psi_0^{(i)}|. \quad (8.38)$$

( $\Psi_0^{(i)}$  и  $\Psi_1^{(i)}$  предполагаются ортонормированными,  $q_i = \text{const.}$ )

Волновая функция объединенной системы имеет вид

$$\begin{aligned}\Psi &= \Psi^{(1)}\Psi^{(2)} \dots \Psi^{(M)} = (1 + \hat{X}_1)\Psi_0^{(1)}(1 + \hat{X}_2)\Psi_0^{(2)} \dots (1 + \hat{X}_M)\Psi_0^{(M)} \\ &= (1 + \hat{X}_1)(1 + \hat{X}_2) \dots (1 + \hat{X}_M)\Psi_0 = \prod_{i=1}^M (1 + \hat{X}_i)\Psi_0,\end{aligned}\quad (8.39)$$

где

$$\Psi_0 = \Psi_0^{(1)}\Psi_0^{(2)} \dots \Psi_0^{(M)} . \quad (8.40)$$

В формуле (8.39) мы использовали тот факт, что  $\hat{X}_i\Psi_0^{(j)} = \Psi_0^{(j)}\hat{X}_i$  и  $\hat{X}_i\hat{X}_j = \hat{X}_j\hat{X}_i$ , если  $i \neq j$ , поскольку они относятся к разным переменным. (В этой связи напомним, что антисимметризация между невзаимодействующими системами не имеет практического значения— ср. гл. 5, разд. 4.3).

Расписывая произведение (8.39), легко получаем

$$\Psi = \left(1 + \sum_{i=1}^M \hat{X}_i + \sum_{i<j} \hat{X}_i\hat{X}_j + \sum_{i<j<k} \hat{X}_i\hat{X}_j\hat{X}_k + \dots\right)\Psi_0 . \quad (8.41)$$

Это выражение можно преобразовать с учетом коммутативности операторов возбуждения  $\hat{X}_i$  как

$$\begin{aligned}\Psi &= \left(1 + \sum_{i=1}^M \hat{X}_i + \frac{1}{2!} \sum_{i,j=1} \hat{X}_i\hat{X}_j + \frac{1}{3!} \sum_{i,j,k=1} \hat{X}_i\hat{X}_j\hat{X}_k + \dots\right)\Psi_0 \\ &= \left(1 + \sum_{i=1}^M \hat{X}_i + \frac{1}{2!} \sum_{i=1} \hat{X}_i \sum_{j=1} \hat{X}_j + \frac{1}{3!} \sum_{i=1} \hat{X}_i \sum_{j=1} \hat{X}_j \sum_{k=1} \hat{X}_k + \dots\right)\Psi_0 .\end{aligned}\quad (8.42)$$

Такое преобразование допустимо, так как слагаемые с  $i = j$  и т. д. не нужно исключать из этих сумм, ибо они все равно обращаются в нуль:

$$\hat{X}_i\hat{X}_i = q_i^2|\Psi_1^{(i)}\rangle\langle\Psi_0^{(i)}|\Psi_1^{(i)}\rangle\langle\Psi_0^{(i)}| = 0, \quad (8.43)$$

согласно определению (8.38) и в силу ортогональности состояний  $\Psi_0^{(i)}$  и  $\Psi_1^{(i)}$ .

Известно, что функции операторов определяются их степенными рядами; таким образом, для любого «приличного» линейного оператора  $\hat{L}$  имеем экспоненциальную функцию от него (его операторную экспоненту):

$$e^{\hat{L}} = 1 + \hat{L} + \frac{1}{2!}\hat{L}^2 + \frac{1}{3!}\hat{L}^3 + \dots . \quad (8.44)$$

Тогда уравнение (8.42) можно переписать как

$$\Psi = e^{\hat{X}} \Psi_0, \quad (8.45)$$

где, по определению,

$$\hat{X} = \sum_{i=1}^M \hat{X}_i. \quad (8.46)$$

Степени  $\hat{X}^k$  с  $k > M$  обращаются в нуль, вследствие условия нильпотентности (8.43); поэтому степенное разложение  $e^{\hat{T}}$  обрывается после слагаемого  $\hat{X}^M$ . В частности, если имеются только одна или две подсистемы ( $M = 1$  или  $M = 2$ ), то остаются, соответственно, слагаемые

$$e^{\hat{X}} = 1 + \hat{X}_1 \quad (8.47)$$

или

$$\begin{aligned} e^{\hat{X}} &= 1 + \hat{X} + \frac{1}{2} \hat{X}^2 = 1 + \hat{X}_1 + \hat{X}_2 + \frac{1}{2} (\hat{X}_1 + \hat{X}_2)^2 \\ &= 1 + \hat{X}_1 + \hat{X}_2 + \hat{X}_1 \hat{X}_2 = (1 + \hat{X}_1)(1 + \hat{X}_2) \end{aligned} \quad (8.48)$$

как и ожидалось. (При разложении квадрата в выражении (8.48) мы воспользовались тем фактом, что  $\hat{X}_i^2 = 0$ .)

В то время как волновая функция объединенной системы, состоящей из независимых подсистем, является «мультипликативно сепарабельной» величиной, использование экспоненциальной формулы («кластерного разложения») позволяет работать с операторами возбуждений  $\hat{T}_i$ , которые представляют собой «аддитивно сепарабельные» объекты. Важно еще раз подчеркнуть, что все эти преобразования возможны потому, что разные операторы  $\hat{X}_i$  коммутируют друг с другом.

Метод «связанных кластеров» (coupled cluster approach, CCA) состоит в применении только что описанной схемы к проблеме электронной корреляции. В CCA используется кластерное разложение волновой функции в виде

$$\Psi = e^{\hat{T}} \Psi_0. \quad (8.49)$$

Здесь  $\Psi_0$  является волновой функцией «состояния отсчета» — в простейшем и самом распространенном случае волновой функцией ОХФ для основного состояния системы, — а полный оператор возбуждения  $\hat{T}$  равен сумме:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots \quad (8.50)$$

Разные слагаемые этого разложения определяются следующим образом. Оператор  $\hat{T}_1$  порождает всевозможные однократно возбужденные конфигурации. Аналогично,  $\hat{T}_2$  порождает двукратно возбужденные конфигурации, но за вычетом вкладов, которые происходят из произведений двух однократных возбуждений<sup>1</sup> и, таким образом, уже включены в  $e^{\hat{T}_1}$ . (Разные операторы  $\hat{T}_i$  коммутируют — см. ниже, — и тогда экспоненту в формуле (8.49) можно переписать, как  $e^{\hat{T}_1}e^{\hat{T}_2}e^{\hat{T}_3}\dots$ ). Аналогично,  $\hat{T}_3$  порождает собственные трехкратные возбуждения, т. е. те, которые не являются произведениями независимых одно- и двукратных возбуждений. Разложение можно в принципе продолжать, но на практике обычно не идут дальше  $\hat{T}_3$ . Детальный вид этих операторов следующий:

$$\begin{aligned}\hat{T}_1 &= \sum_i^{(\text{зан.})} \sum_u^{(\text{вирт.})} t_i^u \hat{T}_i^u \\ \hat{T}_2 &= \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} t_{ij}^{uv} \hat{T}_{ij}^{uv} \\ \hat{T}_3 &= \sum_{i < j < k}^{(\text{зан.})} \sum_{u < v < w}^{(\text{вирт.})} t_{ijk}^{uvw} \hat{T}_{ijk}^{uvw}.\end{aligned}\tag{8.51}$$

В выражениях (8.51) оператор  $\hat{T}_i^u$  есть оператор возбуждения одного электрона с занятой спин-орбитали  $\psi_i$  на виртуальную спин-орбиталь  $\psi_u$ . Он проводит такое замещение для любой детерминантной волновой функции, в которой  $\psi_i$  занята, а  $\psi_u$  — нет; если эти условия не выполнены, то результат действия  $\hat{T}_i^u$  на детерминант равен нулю. Обычно операторы такого типа выписывают в терминах операторов «рождения» и «уничтожения» в формализме так называемого «вторичного квантования»; но здесь их явный вид нам не понадобится. Аналогично,  $\hat{T}_{ij}^{uv}$  ( $i \neq j; u \neq v$ ) есть оператор возбуждения, замещающий занятые спин-орбитали  $\psi_i$  и  $\psi_j$  на виртуальные спин-орбитали  $\psi_u$  и  $\psi_v$  соответственно. Очевидно, при действии на детерминант общего вида он снова приводит

<sup>1</sup> Здесь речь идет сразу о двух вещах: собственно операторах кратных возбуждений — они всегда будут произведениями операторов однократных возбуждений, и о коэффициентах при этих операторах. Понятно, что произведение одноэлектронных операторов  $\hat{T}_i^u \hat{T}_j^v$  и  $\hat{T}_j^u \hat{T}_i^v$  порождает ту же возбужденную конфигурацию, что и операторы двукратного возбуждения  $\hat{T}_{ij}^{uv}$ . Смысл происходящего в том, чтобы ввести в качестве независимой переменной «избыток» (или «недостаток») амплитуды (коэффициента) при этой возбужденной конфигурации по сравнению с произведением амплитуд (коэффициентов) при однократных возбуждениях. — *Прим. ред.*



к ненулевому результату, только если и  $\psi_i$ , и  $\psi_j$  в нем присутствуют, а  $\psi_u$  и  $\psi_v$  обе отсутствуют. Оператор  $\hat{T}_{ijk}^{uvw}$  определяется аналогично. Коэффициенты  $t_i^u$ ,  $t_{ij}^{uv}$  и  $t_{ijk}^{uvw}$  называются «кластерными амплитудами».

Произведение двух (или более) операторов возбуждения  $\hat{T}_i^u$ ,  $\hat{T}_{ij}^{uv}$  и т. д. указывает на то, что соответствующие замещения должны проводиться последовательно, и поэтому является оператором замещения высшего порядка, или будет равно нулю, если замещения невозможны (например,  $\hat{T}_i^u \hat{T}_i^v = \hat{T}_i^u \hat{T}_j^u = 0$ , так как нельзя возбудить два электрона с одной и той же спин-орбитали  $\psi_i$ , или на одну и ту же спин-орбиталь  $\psi_u$ ). Таким образом,  $\hat{T}_1^2$  состоит из двухэлектронных возбуждений,  $\hat{T}_1 \hat{T}_2$  и  $\hat{T}_1^3$  — из трехэлектронных и т. д. Очевидно, операторы возбуждений коммутируют, так как порядок, в котором проводятся разные замещения, несуществен. Следовательно, операторы  $\hat{T}_i$ , представляющие собой линейные комбинации операторов возбуждения, также коммутируют.

Разлагая экспоненту (8.49) в ряд, получим

$$\Psi = \left( 1 + \hat{T} + \frac{1}{2!} \hat{T}^2 + \frac{1}{3!} \hat{T}^3 + \dots \right) \Psi_0. \quad (8.52)$$

Разные степени в (8.52) должны быть рассчитаны с помощью разложения (8.50), т. е.

$$\hat{T}^2 = \hat{T}_1^2 + 2\hat{T}_1 \hat{T}_2 + \hat{T}_2^2 + 2\hat{T}_1 \hat{T}_3 + 2\hat{T}_2 \hat{T}_3 + \hat{T}_3^2 + \dots \quad (8.53)$$

и т. д. Можно сгруппировать слагаемые разложения (8.52) по кратности входящих в них электронных возбуждений:

$$\Psi = \left( 1 + \hat{T}_1 + \frac{1}{2!} \hat{T}_1^2 + \hat{T}_2 + \frac{1}{3!} \hat{T}_1^3 + \hat{T}_1 \hat{T}_2 + \hat{T}_3 + \dots \right) \Psi_0. \quad (8.54)$$

Тогда, подставляя это разложение в (8.51), можно собрать вместе слагаемые, содержащие возбуждения одинаковой кратности. Например, опе-

раторы двухэлектронных возбуждений будут иметь вид

$$\begin{aligned}
 \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 &= \frac{1}{2} \sum_i^{(\text{зан.})} \sum_u^{(\text{вирт.})} t_i^u \hat{T}_i^u \sum_j^{(\text{зан.})} \sum_v^{(\text{вирт.})} t_j^v \hat{T}_j^v + \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} t_{ij}^{uv} \hat{T}_{ij}^{uv} \\
 &= \frac{1}{2} \sum_{i,j}^{(\text{зан.})} \sum_{u,v}^{(\text{вирт.})} t_i^u t_j^v \hat{T}_i^u \hat{T}_j^v + \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} t_{ij}^{uv} \hat{T}_{ij}^{uv} \\
 &= \frac{1}{2} \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} t_i^u t_j^v \hat{T}_i^u \hat{T}_j^v + \frac{1}{2} \sum_{i < j}^{(\text{зан.})} \sum_{u > v}^{(\text{вирт.})} t_i^u t_j^v \hat{T}_i^u \hat{T}_j^v \quad (8.55) \\
 &\quad + \frac{1}{2} \sum_{i > j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} t_i^u t_j^v \hat{T}_i^u \hat{T}_j^v + \frac{1}{2} \sum_{i > j}^{(\text{зан.})} \sum_{u > v}^{(\text{вирт.})} t_i^u t_j^v \hat{T}_i^u \hat{T}_j^v \\
 &\quad + \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} t_{ij}^{uv} \hat{T}_{ij}^{uv} \\
 &= \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} (t_i^u t_j^v - t_i^v t_j^u + t_{ij}^{uv}) \hat{T}_{ij}^{uv},
 \end{aligned}$$

где мы переставили некоторые индексы суммирования, чтобы получить  $i < j$  и  $u < v$  в каждом слагаемом, и учли, что операторное произведение  $\hat{T}_i^u \hat{T}_j^v$  дает то же замещение, что и  $\hat{T}_{ij}^{uv}$ , но произведение  $\hat{T}_i^v \hat{T}_j^u$  приводит к детерминанту, в котором спин-орбитали  $\psi_u$  и  $\psi_v$  переставлены — отсюда знак «минус» при втором слагаемом в скобках.

Сравнивая эти результаты с разложением КВ (8.8), которое можно переписать как

$$\Psi = \Psi_0 + \sum_i^{(\text{зан.})} \sum_u^{(\text{вирт.})} c_i^u \Psi_i^u + \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} c_{ij}^{uv} \Psi_{ij}^{uv} + \sum_{i < j < k}^{(\text{зан.})} \sum_{u < v < w}^{(\text{вирт.})} c_{ijk}^{uvw} \Psi_{ijk}^{uvw} + \dots, \quad (8.56)$$

получаем

$$c_i^u = t_i^u \quad (8.57)$$

и

$$c_{ij}^{uv} = t_i^u t_j^v - t_i^v t_j^u + t_{ij}^{uv}. \quad (8.58)$$

Точно так же, коэффициенты  $c_{ijk}^{uvw}$  можно выразить через амплитуды  $t_i^u$ ,  $t_{ij}^{uv}$ , и  $t_{ijk}^{uvw}$ , и т. д. для коэффициентов КВ высших возбужденных состояний. Каждый коэффициент КВ, таким образом, содержит различные сочетания кластерных амплитуд низших порядков, описывающих

кинематические эффекты возбуждений низшего порядка, и единственную собственную кластерную амплитуду, описывающую вклад динамических эффектов в амплитуду данного возбуждения.

Рассмотрим пример, в котором спин-орбитали  $\psi_i$  и  $\psi_u$  находятся в одной и той же подсистеме, а спин-орбитали  $\psi_j$  и  $\psi_v$  находятся в другой, и эти две подсистемы не взаимодействуют. Тогда представление (8.58) дает  $c_{ij}^{uv} = t_i^u t_j^v$  для кинематически связанного двукратного возбуждения. (Другие слагаемые в правой части выражения (8.58) исчезают, так как подсистемы предполагаются невзаимодействующими.) Если нумерация орбиталей оказывается такой, что две невзаимодействующие системы содержат спин-орбитали  $\psi_i, \psi_v$  и  $\psi_j, \psi_u$ , соответственно, то мы получим  $c_{ij}^{uv} = -t_i^v t_j^u$ . (При этой нумерации орбиталей детерминант  $\Psi_{ij}^{vu}$ , содержащий внутривзаимодействующие возбуждения  $\psi_i \rightarrow \psi_v, \psi_j \rightarrow \psi_u$ , равен упорядоченной конфигурационной функции  $\Psi_{ij}^{uv}$  со знаком минус.) Если все четыре орбитали находятся в области взаимодействия, то оба предшествующих слагаемых появятся в результате кинематической связи однократных возбуждений. Динамические эффекты вызовут отклонение коэффициента  $c_{ij}^{uv}$  от его чисто кинематического значения  $t_i^u t_j^v - t_i^v t_j^u$ , что будет учтено собственной амплитудой двукратного возбуждения  $t_{ij}^{uv}$ . (Напомним еще раз, что в задачах электронной корреляции собственные амплитуды двукратных возбуждений являются наиболее важными слагаемыми; роль однократно возбужденных конфигураций весьма невелика.)

Подставляя формулу (8.58) в выражение (8.14) для теоремы Несбета, получаем выражение для энергии корреляции через кластерные амплитуды:

$$\Delta E = \sum_{i < j}^{(\text{зан.})} \sum_{u < v}^{(\text{вирт.})} (t_i^u t_j^v - t_i^v t_j^u + t_{ij}^{uv}) \times \left( \langle \psi_i(1) \psi_j(2) | \frac{1}{r_{12}} | \psi_u(1) \psi_v(2) \rangle - \langle \psi_i(1) \psi_j(2) | \frac{1}{r_{12}} | \psi_v(1) \psi_u(2) \rangle \right) \quad (8.59)$$

Уравнение (8.59) представляет собой точную корреляционную энергию для данного базиса, если используются точные кластерные амплитуды (т. е. те, которые соответствуют расчету полного КВ). В этом случае волновая функция удовлетворяет (спроектированному) уравнению Шрёдингера (8.22). Его можно переписать как

$$\langle \Phi | \hat{H} | \Psi \rangle = E \langle \Phi | \Psi \rangle \quad (8.60)$$

для произвольной  $N$ -электронной волновой функции  $\Phi$ , которую можно выразить с помощью используемого одноэлектронного базиса. (Так как

поиск  $\Psi$  также ведется в этом базисе, имеем  $\langle \Phi | \hat{\mathcal{P}} \equiv \langle \Phi | ; \hat{\mathcal{P}} | \Psi \rangle \equiv | \Psi \rangle$ . Обычно говорят о «проекции уравнения Шрёдингера на разные функции  $\Phi$ ».

Уравнение (8.60) можно также переписать как

$$\langle \Phi | \hat{H}' | \Psi \rangle = \Delta E \langle \Phi | \Psi \rangle, \quad (8.61)$$

где  $\hat{H}'$  есть гамильтониан с энергетическим сдвигом, определенный как

$$\hat{H}' = \hat{H} - E_0, \quad (8.62)$$

где  $E_0$  есть энергия Хартри—Фока:  $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ .

При выборе  $\Phi = \Psi_0$  воспроизводится результат (8.59), а выбирая  $\Phi$  в виде различных возбужденных конфигураций  $\Psi_i^u$ ,  $\Psi_{ij}^{uv}$ , и т. д., получаем столько уравнений, сколько имеется упорядоченных конфигураций (ср. разд. 1). Это число равно числу независимых коэффициентов КВ  $c_i^u$ ,  $c_{ij}^{uv}$  и т. д., или собственных кластерных амплитуд  $t_i^u$ ,  $t_{ij}^{uv}$  и т. д. Уравнения, получаемые таким образом, являются ничем иным, как уравнениями соответствующей задачи полного КВ. Каждое уравнение, полученное проектированием на  $\Psi_i^u$ ,  $\Psi_{ij}^{uv}$  и т. д., содержит коэффициенты КВ всех детерминантов, которые взаимодействуют с данным (т. е. тех, которые отличаются от него не более, чем двумя спин-орбиталями). Как уже обсуждалось выше, коэффициенты КВ являются сложными функциями кластерных амплитуд и поэтому уравнения полного КВ становятся нелинейными, если выражать их через кластерные амплитуды.

В практических расчетах ССА невозможно использовать точное кластерное разложение, содержащее операторы возбуждений вплоть до  $\hat{T}_N$  для  $N$ -электронной системы, и необходимо обрывать его после слагаемых  $\hat{T}_2$  или  $\hat{T}_3$  (в исключительных случаях возможен и учет слагаемого  $\hat{T}_4$ ). При этом понадобится столько уравнений, сколько есть собственных амплитуд, содержащихся в оборванном кластерном разложении. По этой причине необходимо спроектировать уравнение Шрёдингера на все функции  $\Psi_i^u$  и  $\Psi_{ij}^{uv}$ . (Если  $\hat{T}_3$  также сохраняется в разложении, то необходима и проекция на функции  $\Psi_{ijk}^{uvw}$ ). Проектируя на  $\Psi_0$ , получаем корреляционную энергию

$$\langle \Psi_0 | \hat{H}' | \Psi \rangle = \Delta E, \quad (8.63)$$

что равносильно результату (8.59).

Мы не будем обсуждать подробно уравнения метода связанных кластеров, полученные описанным образом. Это — сложные нелинейные уравнения, содержащие все собственные амплитуды. Кластерные амплитуды высших порядков являются малыми числами; пренебрежение

ими в уравнениях не поменяет слишком сильно значения тех коэффициентов, которые интересуют нас в связи с корреляционной энергией, т. е. выражением (8.59). Тем не менее, если решать усеченную систему уравнений ССА, то получающаяся корреляционная энергия не будет строго вариационной. Это не вызывает серьезных проблем, пока рассматриваются значения самой корреляционной энергии, но существенно усложняет расчет градиентов энергии и вычисление значений различных физических величин. (В принципе можно было бы воспользоваться пробной волновой функцией экспоненциального типа («Ansatz»)  $\Psi = e^T \Psi_0$  и в вариационной процедуре, но это было бы слишком сложно.)

Схема ССА, однако, имеет то очень важное достоинство, что она размерно согласована, независимо от того, на возбуждениях какой кратности обрывать кластерное разложение. Чтобы доказать это, рассмотрим две не взаимодействующие системы  $A$  и  $B$  с волновыми функциями  $\Psi^A = e^{\hat{T}^A} \Psi_0^A$  и  $\Psi^B = e^{\hat{T}^B} \Psi_0^B$ , соответственно. Мы можем составить волновую функцию объединенной системы, как  $\Psi^{AB} = \Psi^A \Psi^B = e^{\hat{T}^A + \hat{T}^B} \Psi_0^A \Psi_0^B = e^{\hat{T}^{AB}} \Psi_0^{AB}$ . (Снова напомним, что антисимметризация между не взаимодействующими системами не имеет практического значения; ср. гл. 5, разд. 4.3. Между тем в ССА обычно используется формализм вторичного квантования, автоматически учитывающий антисимметрию.)

Допустим, что  $\Psi^A$  и  $\Psi^B$  представляют собой решения уравнений ССА для соответствующих подсистем и покажем, что волновая функция  $\Psi^{AB} = \Psi^A \Psi^B$  удовлетворяет уравнениям ССА, записанным для объединенной системы, и оборванным на возбуждениях той же кратности, что и для подсистем.

Уравнения ССА (8.61) и (8.63) для подсистемы  $A$  можно записать как

$$\begin{aligned} \langle \Phi^A | \hat{H}^{A'} | \Psi^A \rangle &= \Delta E^A \langle \Phi^A | \Psi^A \rangle ; \\ \langle \Psi_0^A | \hat{H}^{A'} | \Psi^A \rangle &= \Delta E^A , \end{aligned} \quad (8.64)$$

где  $\hat{H}^{A'}$  есть гамильтониан с энергетическим сдвигом для подсистемы  $A$ ,  $\Phi^A$  есть произвольный детерминант подсистемы  $A$  в рассматриваемом диапазоне возбуждений, а  $\Delta E^A$  есть соответствующая корреляционная энергия. Имеется аналогичный набор уравнений и для подсистемы  $B$ .

Волновой функцией объединенной системы, которая содержит те же возбуждения, что и  $\Phi^A$ , является произведение  $\Phi^A \Psi_0^B$ . Тогда мы должны показать, что выполнения уравнений (8.64) для подсистем  $A$  и

$B$  по отдельности достаточно для выполнения соответствующих уравнений для объединенной системы:

$$\langle \Phi^A \Psi_0^B | \hat{H}^{A'} + \hat{H}^{B'} | \Psi^A \Psi^B \rangle = (\Delta E^A + \Delta E^B) \langle \Phi^A \Psi_0^B | \Psi^A \Psi^B \rangle, \quad (8.65)$$

а также тех уравнений, которые получаются перестановкой  $A$  и  $B$ . В самом деле, уравнение (8.65) можно переписать как

$$\begin{aligned} \langle \Phi^A | \hat{H}^{A'} | \Psi^A \rangle \langle \Psi_0^B | \Psi^B \rangle + \langle \Psi_0^B | \hat{H}^{B'} | \Psi^B \rangle \langle \Phi^A | \Psi^A \rangle \\ = (\Delta E^A + \Delta E^B) \langle \Phi^A | \Psi^A \rangle \langle \Psi_0^B | \Psi^B \rangle \end{aligned} \quad (8.66)$$

Так как  $\langle \Psi_0^B | \Psi^B \rangle = 1$ , согласно условию промежуточной нормировки, и  $\langle \Psi_0^B | \hat{H}^{B'} | \Psi^B \rangle = \Delta E^B$ , согласно второму уравнению (8.64), выписанному для подсистемы  $B$ , то уравнение (8.66) сводится к первому из уравнений (8.64).

Далее, для объединенной системы уравнение (8.63) выглядит как

$$\langle \Psi_0^{AB} | \hat{H}^{A'} + \hat{H}^{B'} | \Psi^{AB} \rangle = \Delta E^{AB}. \quad (8.67)$$

Подставляя определения  $\Psi^{AB} = \Psi^A \Psi^B$ ,  $\Psi_0^{AB} = \Psi_0^A \Psi_0^B$ , получаем

$$\langle \Psi_0^A | \hat{H}^{A'} | \Psi^A \rangle \langle \Psi_0^B | \Psi^B \rangle + \langle \Psi_0^B | \hat{H}^{B'} | \Psi^B \rangle \langle \Psi_0^A | \Psi^A \rangle = \Delta E^{AB}. \quad (8.68)$$

С помощью второго из уравнений (8.64), записанного и для  $A$ , и для  $B$  с учетом промежуточной нормировки, получим

$$\Delta E^{AB} = \Delta E^A + \Delta E^B, \quad (8.69)$$

как и ожидалось.

Чтобы закончить доказательство размерной согласованности, необходимо записать эквивалент выражения (8.61) для объединенной системы с возбужденным детерминантом, содержащим одно или более межсистемных возбуждений. Легко видеть, что это уравнение выполняется: все интегралы исчезают в обеих частях равенства, так как число электронов в подсистемах  $A$  и  $B$  отличается в «бра» и «кет»-функциях, тогда как подсистемы предполагаются невзаимодействующими. Ч. т. д.

Напомним, что произведение двух усеченных волновых функций КВ удовлетворяло бы аналогичным уравнениям; однако, оно содержало бы также и возбуждения высшей кратности, не соответствующие усечению волновой функции объединенной системы на том же уровне возбуждения, который использовался для отдельных подсистем. Усеченная волновая функция КВ является вариационной, но не размерно согласованной. (В формуле (8.66) невозможно исключить слагаемые, содержащие

$\Delta E^B$ , если в волновую функцию не включены кинематически связанные четырехкратные возбуждения. Таким образом, оказывается невозможным расщепление уравнений, описывающих возбуждения в не взаимодействующих подсистемах).

В зависимости от того, какие возбуждения включены в волновую функцию, говорят о схемах «связанных кластеров с двукратным возбуждением» (coupled cluster doubles, CCD), «связанных кластеров с одно- и двукратными возбуждениями» (coupled cluster singles and doubles, CCSD), или «связанных кластеров с одно-, дву- и трехкратными возбуждениями» (coupled cluster singles, doubles and triples, CCSDT) и т. д. Простейшая схема CCD на практике не применяется. В то же время, полный учет трехкратных возбуждений, число которых велико, часто недостижим, но на самом деле и не очень-то и нужен из-за их малой важности. По этой причине амплитуды  $t_{ijk}^{uvw}$  часто вычисляют лишь приближенно (например, по теории возмущений). В зависимости от применяемых приближений, получаются несколько разные схемы; наиболее широко используется та, которую обычно обозначают CCSD(T), где скобки указывают на приближенный характер учета трехкратных возбуждений.

## Библиографические заметки

### Раздел 1.

Рассуждения, представленные здесь, с формально математической точки зрения не являются строгими. Согласно [1], Фадеев [2] строго доказал возможность представления точного решения уравнения Шрёдингера в виде разложения (8.5), если имеется не более трех электронов, но для общего  $N$ -электронного случая строгого доказательства не существует<sup>1</sup>. (МакВини [3] называет эту возможность «верой», основу для которой составляют описанные рассуждения). Этот математический недостаток — не очень серьезный, по мнению автора, — обычно игнорируется, и все методы учета электронной корреляции основаны на разложении КВ волновой функции, за исключением «явно коррелированных» методов, в которых межэлектронные расстояния присутствуют, как аргументы волновой функции. На практике используют конечный набор одноэлектронных базисных функций и в этом случае разложение (8.5) охватывает, разумеется, любые возможные антисимметричные волновые функции, которые можно построить с помощью данного базиса.

<sup>1</sup> В книге Степанов Н. Ф., Пупышев В. И. Квантовая механика молекул и квантовая химия. — М.: Изд-во МГУ, 1991, содержится доказательство следующего утверждения: пусть  $\phi_i(x)$  — полный ортонормированный набор в  $L^2(\mathbb{R}^3)$ , тогда всевозможные детерминанты Слэтера порядка  $N$  образуют полный набор в пространстве антисимметричных интегрируемых с квадратом функций координат  $N$  частиц. — Прим. ред.

## Раздел 2.

Общепринятое определение корреляционной энергии было дано Лёвдиным [4]. Теорема Несбета была получена в [5]; ее формулировка, представленная здесь, была приведена в [6], где также была сформулирована концепция парных корреляционных энергий.

Классические расчеты молекулы  $H_2$ , упоминавшиеся здесь: Гайтлер и Лондон [7, 8], МО [9], Вайнбаум [10].

## Раздел 3.1.

Об очень эффективной реализации схемы полного КВ см. [11, 12]. Крупномасштабные расчеты методом КВ стали возможны благодаря введенному Росом (Roos) «прямому» методу расчета результата действия гамильтониана на вектор коэффициентов [13]. Также см. его обзорную статью [14].

Наиболее важными методами учета приближенной поправки на размерную согласованность для результатов усеченного КВ являются методы Давидсона и Попла. Поправка Давидсона [15] оценивает вклад исключенных высших (трех- и особенно четырехкратных) возбуждений в расчете CISD, как  $\Delta E = (1 - C_0^2)E_{CISD}^{corr.}$ , где  $E_{CISD}^{corr.}$  есть корреляционная энергия, полученная в расчете CISD, а  $C_0^2$  есть коэффициент детерминанта ОХФ в волновой функции CISD. Метод «квадратичного КВ» (quadratic CI, QCI), предложенный Поплом [16], можно рассматривать [17, 18] как (приближенный) вариант «метода связанных кластеров» (разд. 3.4).

Относительно способов решения вековых уравнений большой размерности в методе КВ можно сослаться на [19].

## Раздел 3.2.

Метод МК ССП был переоткрыт несколько раз (например, [20-22]); по поводу технических аспектов см. [23].

Некоторые классические работы по упомянутым методам CAS-SCF: [24]; АПСГ: [25-27]; GVB: [30]; Коулсон—Фишер: [31].

## Раздел 3.3.

Как отмечено в тексте, теория возмущений Меллера—Плессета является частным случаем общей теории возмущений Рэля—Шрёдингера и обычно применяется в сочетании с использованием диаграммных методов. Некоторые уместные ссылки: [32-35].

## Раздел 3.4.

Автору было интересно наблюдать за тем, как в последние десятилетия теория связанных кластеров выросла из предмета, развиваемого немногими коллегами с развитыми математическими интересами, в рутинный метод расчетов, вначале использовавшийся для получения эталонных данных, а затем и для многих прикладных исследований. Очевидно, это произошло потому, что ССА оказался подходом, во многих отношениях очень близким к вариационному методу КВ, но «по построению» обеспечивающим размерную согласованность.

После нескольких пионерских работ по кластерному разложению волновой функции и матриц плотности [36, 37], теория ССА была перенесена из ядерной физики [38, 39] в квантовую химию Чижеком (Čížek) [40]. Затем, в результате постепенного развития, главным образом, благодаря усилиям групп



во Флориде и Чехословакии (например, [41, 42]) и их сотрудничеству [43], (а также развитию Поплом метода «квадратично сходящегося КВ» [16], который, как уже отмечалось, также является вариантом ССА), в последнее десятилетие метод ССА вытеснил так называемые методы парной корреляции [44, 45], например «приближение связанных электронных пар» (coupled electron pair approximation, СЕРА), которые, по существу, вышли из употребления.

Теория связанных кластеров обычно излагается с обильным применением формализма вторичного квантования и диаграммных методов, не используемых в нашей книге. По этой причине наше обсуждение было сосредоточено на экспоненциальном представлении (exponential Ansatz) пробной волновой функции и свойстве размерной согласованности метода ССА.

Недавно метод ССА был объединен (например, в [46, 47]) с явным (приближенным) включением межэлектронных расстояний в волновую функцию, что привело к чрезвычайно многообещающему методу.

### Для дальнейшего изучения

Такие учебники, как [3, 34, 48] содержат подробные обсуждения многих проблем электронной корреляции. Существует огромное число обзоров по отдельным методам, некоторые из них уже цитировались. Также рекомендуется монография [49].

### Литература

1. Biczó G., Ladik J. Számológép **1**, No. 3, 5 (1971).
2. Фадеев Л. Д. Труды МИАН им. В.А. Стеклова, **69**, 1 (1963).
3. McWeeny R. *Methods of Molecular Quantum Mechanics*, Academic Press, London 1992. (Имеется русский перевод первого издания: МакВини Р., Сатклиф С. *Квантовая механика молекул*. — М.: Мир, 1973).
4. Löwdin P.-O. Adv. Chem. Phys. **2**, 207 (1959).
5. Nesbet R. K. Phys. Rev. **109**, 1632 (1958).
6. Nesbet R. K. Adv. Chem. Phys. **9**, 321 (1965).
7. Heitler W., London F. Z. Physik **44**, 455 (1927).
8. Sugiura Y. Z. Physik **45**, 484 (1927).
9. Mulliken R. Phys. Rev. **32** 186, 761 (1928), **33**, 730 (1929).
10. Weinbaum S. J. Chem. Phys. **1**, 593 (1933).
11. Knowles P. J., Handy N. C. Chem. Phys. Letters. **111**, 315 (1984).
12. Knowles P. J., Handy N. C. Computer Physics Communications **54**, 75 (1989).
13. Roos B. O. Chem. Phys. Letters, **15**, 153 (1972).
14. Roos B. *The Configuration Interaction Method in Computational Techniques in Quantum Chemistry and Molecular Physics* (ed. G.H.F. Dierksen, B.T. Sutcliffe and A. Veillard) Reidel, Dordrecht 1974.
15. Langhoff S. R., Davidson E. R. Int. J. Quantum Chem. **8**, 61 (1974).
16. Pople J. A., Head-Gordon M., Raghavachari K. J. Chem. Phys. **87**, 5968 (1987).
17. Paldus J., Čížek J., Jeziorski B. J. Chem. Phys. **90**, 4356 (1989).
18. Noga J., неопубликовано; см. также Noga J. Acta Univ. Debreceniensis Ser. Physica et Chimica XXX/2, 67 (1995).
19. Davidson E. R. *Matrix Eigenvector Methods in Methods in Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.

20. McWeeny R. Proc. Roy. Soc (London) **A232**, 114 (1955).
21. Gilbert T. L. J. Chem. Phys. **43**, S245, (1965).
22. Clementi E. J. Chem. Phys. **46**, 3842 (1967).
23. Roos B. O. *The Multiconfigurational (MC) SCF Method in Methods in Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.
24. Roos B. O., Taylor P. R. Chem. Phys. **48** 157 (1980).
25. Hurley A. C., Lenard-Jones J., Pople J. A. Proc Roy. Soc. (London) **A220**, 446 (1953).
26. Kapuy E. Acta Phys. Hung. **11**, 409 (1960); **12**, 185 (1960).
27. Ahlrichs R., Kutzelnigg W. J. Chem. Phys. **48**, 1819 (1967).
28. Kutzelnigg W. Theor. Chim. Acta **1**, 327 (1963).
29. Kutzelnigg W. J. Chem. Phys. **40**, 3640 (1964).
30. Goddard III W. A., Dunning T. H. Jr., Hunt W. J., Hay P. J. Acc. Chem. Res. **6**, 368 (1973).
31. Coulson C. A., Fischer I. Philos. Mag. **40**, 386 (1949).
32. Møller C., Plesset M. S. Phys. Rev. **46**, 618 (1934).
33. Paldus J., Čížek J. Adv. Quantum Chem. **9**, 10 (1975).
34. Jørgensen P., Simons J. *Second Quantization-Based Methods in Quantum Chemistry*, Academic Press, New York 1981.
35. Wilson S. *Diagrammatic Many-Body Perturbation Theory in Methods in Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.
36. Sinanoğlu O. Adv. Chem. Phys. **6**, 315 (1964).
37. Primas H. in *Modern Quantum Chemistry (Istanbul Lectures)*, ed. O. Sinanoğlu, Academic Press, New York 1965. (Имеется русский перевод: Примас Х., в кн. *Современная квантовая химия* (ред. О. Синаноглу). — М.: Мир, 1968).
38. Coester F. Nucl. Phys. **7**, 421 (1958).
39. Coester F., Kümmel H. Nucl. Phys. **17**, 477 (1960).
40. Čížek J. Adv. Chem. Phys. **14**, 35 (1969).
41. Urban M., Černušák I., Kellö V., Noga J. *Electron correlation in molecules*, in *Computational Methods in Chemistry*, ed. S. Wilson, Plenum Press, New York 1987.
42. Bartlett R. J. *Coupled-Cluster Theory: an Overview and Recent Developments in Modern Electronic Structure Theory* ed. D.R. Yarkony, World Scientific, Singapore 1995.
43. Bartlett R. J., Kucharski S. A., Noga J., Watts J. D., Trucks G. W. *Some Consideration of Alternative Ansatz in Coupled-Cluster Theory*, in *Lecture Notes in Chemistry*, Vol. **52** Springer Verlag, Heidelberg 1989.
44. Robb M. A. *Pair Functions and Diagrammatic Perturbation Theory in Computational Techniques in Quantum Chemistry and Molecular Physics*, (ed. G.H.F. Dierksen, B.T. Sutcliffe and A. Veillard) Reidel, Dordrecht 1974.
45. Ahlrichs R. *Pair Correlation theories in Methods in Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.
46. Noga J., Klopper W., Kutzelnigg W. *CC-R12: an explicitly correlated coupled-cluster theory*, in *Recent Advances in Computational Chemistry - Vol 3.: Recent Advances in Coupled-Cluster Methods*, ed. R. J. Bartlett, World Scientific, Singapore 1997.
47. Noga J., Valiron P. *Explicitly correlated coupled cluster R12 calculations*, in *Computational Chemistry: Reviews of Current Trends*. Vol. 7 ed. J. Leszczynski, World Scientific, Singapore 2002.
48. Szabo A., Ostlund N. S. *Modern Quantum Chemistry*, McGraw Hill, New York 1989.
49. Hurley A. C. *Electron Correlation in Small Molecules* Academic Press, New York 1976.

# Разное



В этой главе мы рассматриваем разнообразные результаты, которые часто упоминаются, но которые не так уж легко найти в литературе, и которые не удалось логично включить в состав предшествующих глав.

## 1. Разложение по степеням $1/Z$

Здесь мы кратко обсудим схему, позволяющую описывать разные атомы (ионы) в едином энергетическом масштабе, что может быть полезно для получения лучшего качественного представления о Периодической системе элементов. Существует формальное сходство между этим выводом и процедурой масштабирования, обсуждавшейся в гл. 3, разд. 3.3, но здесь мы не рассматриваем преобразование волновой функции, а переписываем гамильтониан в новых координатах.

Гамильтониан Борна—Оппенгеймера для  $N$ -электронного атома (иона) с зарядом ядра  $Z$  равен (в атомных единицах)

$$\hat{H} = \sum_{i=1}^N \left( -\frac{1}{2} \Delta_i \right) - \sum_{i=1}^N \frac{Z}{r_i} + \sum_{i < j} \frac{1}{r_{ij}} . \quad (9.1)$$

Введем новые координаты  $\vec{\varrho}_i$ :

$$\vec{\varrho}_i = Z \vec{r}_i , \quad (9.2)$$

что означает  $\vec{r}_i = \frac{1}{Z} \vec{\varrho}_i$ ;  $\frac{1}{r_i} = \frac{Z}{\varrho_i}$ ;  $\frac{1}{r_{ij}} = \frac{Z}{\varrho_{ij}}$ . Далее в соответствии с правилом дифференцирования сложных функций, имеем

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial \varrho_{x_i}} \frac{\partial \varrho_{x_i}}{\partial x_i} = Z \frac{\partial}{\partial \varrho_{x_i}} ; \quad \frac{\partial^2}{\partial x_i^2} = Z^2 \frac{\partial^2}{\partial \varrho_{x_i}^2} \quad (9.3)$$

т. е.

$$\Delta_i = Z^2 \Delta_{\varrho_i} . \quad (9.4)$$

Тогда гамильтониан можно переписать в новых координатах как

$$\begin{aligned}
 \hat{H} &= \sum_{i=1}^N \left( -\frac{1}{2} Z^2 \Delta_{\varrho_i} \right) - \sum_{i=1}^N \frac{Z^2}{\varrho_i} + \sum_{i < j} \frac{Z}{\varrho_{ij}} \\
 &= Z^2 \left[ \sum_{i=1}^N \left( -\frac{1}{2} \Delta_{\varrho_i} \right) - \sum_{i=1}^N \frac{1}{\varrho_i} + \frac{1}{Z} \sum_{i < j} \frac{1}{\varrho_{ij}} \right] \\
 &= Z^2 \left[ \sum_{i=1}^N \hat{h}_H(i_{\varrho}) + \frac{1}{Z} \sum_{i < j} \frac{1}{\varrho_{ij}} \right],
 \end{aligned} \tag{9.5}$$

где  $\hat{h}_H$  есть гамильтониан свободного атома водорода,  $i_{\varrho}$  обозначает все координаты  $i$ -го электрона в преобразованной системе координат, а множитель  $Z^2$  явно указывает на изменение масштаба энергий.

Первая сумма в выражении (9.5) является суммой гамильтонианов  $N$  невзаимодействующих атомов водорода, а электрон-электронное взаимодействие можно рассматривать, как возмущение, умноженное на  $\frac{1}{Z}$ , уменьшающееся с увеличением заряда ядра  $Z$ .

Формула (9.5) выражает очевидный факт, что с увеличением заряда ядра  $Z$  роль межэлектронного отталкивания уменьшается по сравнению с электронно-ядерным притяжением. Это должно особенно четко проявляться в случае внутренних оболочек, для которых влияние внешних (валентных) электронов незначительно. Таким образом, гамильтониан (9.5) предсказывает, что характерная энергия рентгеновского кванта (потенциал ионизации внутренней оболочки) разных элементов с хорошей точностью должна быть пропорциональна  $Z^2$ , что, в самом деле, можно наблюдать в виде линейной зависимости между  $\sqrt{\frac{1}{\lambda}}$  и  $Z$ , где  $\lambda$  есть длина волны рентгеновской линии (закон Мозли). В 1913 г. это послужило первым экспериментальным доказательством предположения ван ден Брока о том, что атомные номера элементов в Периодической системе равны зарядам их ядер (и числам электронов в нейтральных атомах).

В частном случае гелиеподобных ионов (ионов с двумя электронами) точный потенциал ионизации как функция  $Z$  хорошо аппроксимируется рядом

$$I(Z) = Z^2 \left( \frac{1}{2} - \frac{5}{8} \cdot \frac{1}{Z} + 0.1576556 \frac{1}{Z^2} + 0.008535 \frac{1}{Z^3} + 0.00034 \frac{1}{Z^4} + \dots \right). \tag{9.6}$$

Первое слагаемое есть энергия соответствующего «атома Н» с зарядом ядра  $Z$ . Второе слагаемое есть поправка первого порядка к

энергии, которая равна (ср. гл. 4) среднему значению оператора кулоновского взаимодействия электронов, рассматриваемого как возмущение по невозмущенной волновой функции. [Значение  $\frac{5}{8}$  равно интегралу  $\iint |\varphi_{1s}(\vec{r}_1)|^2 \frac{1}{r_{12}} |\varphi_{1s}(\vec{r}_2)|^2 dv_1 dv_2$  для водородной орбитали  $1s$   $\varphi_{1s} = \sqrt{\frac{1}{\pi}} e^{-r}$ ; множитель  $\frac{1}{Z}$  происходит от коэффициента  $\frac{1}{Z}$  в последнем слагаемом в правой части формулы (9.5)].

## 2. Граничные условия для волновой функции в точках сингулярности потенциала

### 2.1. Сингулярность потенциала взаимодействия электрона с ядром

Из предыдущего раздела видно, что поведение внутренних электронных оболочек атома в наибольшей степени определяется зарядом ядра  $Z$ . Это еще более наглядно видно, если рассматривать поведение электронной волновой функции в *непосредственной близости* от ядра, т. е. в предельном случае, когда одно из электронно-ядерных расстояний стремится к нулю.

В пределе, когда электрон подходит к ядру очень близко ( $r \rightarrow 0$ ), можно пренебречь всеми остальными электронами и ядрами системы и рассматривать движение данного электрона в поле только этого ядра. Таким образом, очень близко к ядру, по существу, необходимо рассматривать задачу для водородоподобного иона, но только в области  $r \ll 1$ . Соответствующее уравнение Шрёдингера тогда выглядит как уравнение для атома водорода с зарядом ядра  $Z$

$$-\frac{1}{2}\Delta\psi - \frac{Z}{r}\psi = \varepsilon\psi. \quad (9.7)$$

Оператор Лапласа  $\Delta \equiv \nabla^2$  в сферических координатах  $r, \vartheta, \varphi$  имеет вид

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \left[ \frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left( \sin \vartheta \frac{\partial}{\partial \vartheta} \right) + \frac{1}{\sin^2 \vartheta} \frac{\partial^2}{\partial \varphi^2} \right]. \quad (9.8)$$

Из квантовой механики известно, что для угловой части этого оператора имеет место равенство

$$\frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left( \sin \vartheta \frac{\partial}{\partial \vartheta} \right) + \frac{1}{\sin^2 \vartheta} \frac{\partial^2}{\partial \varphi^2} = -\hat{l}^2, \quad (9.9)$$

т. е. она равна оператору квадрата углового момента со знаком минус. Тогда можем переписать уравнение (9.7), как:

$$-\frac{1}{2} \left[ \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) - \frac{\hat{l}^2}{r^2} \psi \right] - \frac{Z}{r} \psi = E \psi. \quad (9.10)$$

Волновую функцию можно записать в виде  $\psi(r, \vartheta, \varphi) = R(r)Y_l^m(\vartheta, \varphi)$ , где  $R(r)$  — радиальная часть, а  $Y_l^m(\vartheta, \varphi)$  — сферическая гармоника.  $Y_l^m(\vartheta, \varphi)$  является собственной функцией оператора  $\hat{l}^2$  с собственным значением  $l(l+1)$ :  $\hat{l}^2 Y_l^m(\vartheta, \varphi) = l(l+1)Y_l^m(\vartheta, \varphi)$ . Если мы интересуемся сферически симметричным  $s$ -состоянием, то в этом случае  $l = 0$  и производные по угловым переменным исчезают. Строго говоря, если рассматривается многоэлектронная система, то состояние  $\psi$  электрона вблизи ядра не обязательно является собственной функцией углового момента. Однако ее всегда можно представить как линейную комбинацию таковых, и последующее рассуждение будет верно для каждой ее компоненты в отдельности. (Очевидно, наиболее интересны все же состояния  $1s$ , так как речь идет об окрестности ядра.)

Таким образом, получаем для радиальной функции  $R = R(r)$

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) - \frac{l(l+1)}{r^2} R + \frac{2Z}{r} R = -2ER. \quad (9.11)$$

Очевидно, в пределе  $r \rightarrow 0$  можно пренебречь правой частью, так как она содержит конечный коэффициент, а коэффициенты при всех слагаемых в левой части расходятся.

Положим, что ведущее слагаемое  $R(r)$  пропорционально  $r^q$  ( $q > 0$ ), и учтем, что в пределе  $r \rightarrow 0$  любое слагаемое, пропорциональное  $r^p$  с  $p > q$  пренебрежимо мало, по сравнению с  $r^q$ . Тогда  $\frac{dR}{dr} \sim qr^{q-1}$ ,  $r^2 \frac{dR}{dr} \sim qr^{q+1}$ , и  $\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) \sim q(q+1)r^{q-2}$ . Второе слагаемое в левой части имеет тот же порядок  $r^{q-2}$ , тогда как третье слагаемое ведет себя, как  $r^{q-1}$ , т. е. при  $r \rightarrow 0$  им также можно пренебречь, по сравнению с первыми двумя слагаемыми. Следовательно, получаем  $q = l$ .

Этот результат показывает, что для любого ненулевого значения  $l$  имеем  $\lim_{r \rightarrow 0} R(r) = 0$ , в то время как для  $l = 0$  ( $s$ -состояний)  $\lim_{r \rightarrow 0} R(r) \neq 0$  и вероятность обнаружить электрон на самом ядре отлична от нуля.

Чтобы получить более детальное представление о поведении волновой функции вблизи ядра, рассмотрим еще одно слагаемое ряда разложения  $R(r)$  и запишем

$$R(r) \cong Cr^l(1 + \alpha r) = C(r^l + \alpha r^{l+1}), \quad (9.12)$$

где  $C = \text{const}$ . Тогда имеем

$$\frac{dR}{dr} \cong C[lr^{l-1} + \alpha(l+1)r^l] \quad (9.13)$$

и

$$\frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) \cong C[l(l+1)r^l + \alpha(l+1)(l+2)r^{l+1}] . \quad (9.14)$$

Подставляя оценку (9.14) в уравнение (9.11) и снова пренебрегая слагаемым в правой части, получим

$$C[l(l+1)r^{l-2} + \alpha(l+1)(l+2)r^{l-1}] - l(l+1)C(r^{l-2} + \alpha r^{l-1}) + 2ZC(r^{l-1} + \alpha r^l) \cong 0 . \quad (9.15)$$

Слагаемые, пропорциональные  $r^{l-2}$ , взаимно уничтожаются, а слагаемым, пропорциональным  $r^l$ , можно пренебречь, по сравнению со слагаемыми  $\sim r^{l-1}$ , когда  $r \rightarrow 0$ . Тогда получаем

$$C\alpha(l+1)(l+2) - \alpha Cl(l+1) + 2ZC = 0 , \quad (9.16)$$

т. е.

$$\alpha = -\frac{Z}{l+1} . \quad (9.17)$$

Подставляя этот результат в выражение (9.12) и (9.13), имеем

$$R(r) \cong C \left( r^l - \frac{Z}{l+1} r^{l+1} \right) ; \quad (9.18)$$

$$\frac{dR(r)}{dr} \cong C(lr^{l-1} - Zr^l) .$$

Это показывает, что только в  $s$ -состояниях ( $l = 0$ ) имеет место  $\lim_{r \rightarrow 0} R(r) \neq 0$ , и что они удовлетворяют условию излома («cusp condition»)

$$\lim_{r \rightarrow 0} \frac{\frac{dR}{dr}}{R(r)} = -Z, \quad (9.19)$$

или

$$\lim_{r \rightarrow 0} \frac{dR}{dr} = -ZR(0) . \quad (9.20)$$

Если не предполагать наличия сферической симметрии вокруг ядра, то необходимо рассматривать более общее разложение вида

$$\psi(r, \vartheta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l C_{lm} R_l(r) Y_l^m(\vartheta, \varphi) . \quad (9.21)$$



В этом случае, согласно результатам (9.18), только компонента с  $l = 0$  вносит вклад в значение  $\psi$  на ядре, и только компоненты с  $l = 0$  и  $1$  имеют там ненулевые производные. Состояние  $l = 1$  нечетно, т. е. соответствующая компонента функции имеет разный знак при приближении к ядру с двух противоположных направлений  $\vec{r}$  и  $-\vec{r}$ . (Угловые координаты равны  $\vartheta$  и  $\varphi$  в одном случае и  $\pi - \vartheta$  и  $\varphi + \pi$  в другом). Это означает, что инверсия приводит к  $\psi \rightarrow -\psi$ , если рассматривается состояние  $l = 1$ , что также влечет за собой  $\frac{d\psi}{dr} \rightarrow -\frac{d\psi}{dr}$  для этого случая.

Поэтому, если усреднить  $\frac{d\psi}{dr}$  по всем углам, то останется только вклад от  $s$ -состояния. Следовательно, условие излома для разложения общего вида (9.21) можно записать как

$$\lim_{r \rightarrow 0} \left\langle \frac{d\psi}{dr} \right\rangle_{\text{среднее}} = -Z\psi(0) . \quad (9.22)$$

## 2.2. Сингулярность потенциала электрон-электронного взаимодействия

Когда два электрона очень близко подходят друг к другу, то можно снова пренебречь всеми остальными частицами системы, и исследовать поведение одной пары электронов. Уравнения, которые необходимо рассмотреть в этом случае, аналогичны уравнениям, описывающим излом электронной волновой функции на ядре, но с некоторыми важными отличиями.

Во-первых, в рамках приближения Борна—Оппенгеймера ядро рассматривается, как фиксированное; таким образом, масса, входящая в оператор кинетической энергии  $\frac{\hbar^2}{2m} \Delta$ , фигурирующий в гамильтониане, равна массе электрона  $m_e$ . (Заметьте, что уравнение (9.7) выписано в атомных единицах, в которых  $\hbar = m_e = 1$ ). Когда рассматривается относительное движение двух электронов, мы должны использовать приведенную массу (ср. гл. 1, разд. 1.1)  $\mu = \frac{1}{2}m_e$ . Поэтому множитель  $\frac{1}{2}$  в слагаемом, содержащем оператор Лапласа, сокращается. Далее, два электрона отталкиваются, а не притягиваются, поэтому необходимо заменить заряд ядра  $+Z$  на заряд электрона  $-1$ . Соответственно, уравнение Шрёдингера для радиальной функции, которое надо рассмотреть, если два электрона находятся очень близко друг к другу, приобретает вид

$$\frac{1}{r_{12}^2} \frac{d}{dr_{12}} \left( r_{12}^2 \frac{dR}{dr_{12}} \right) - \frac{l(l+1)}{r_{12}^2} R - \frac{1}{r_{12}} R = -ER . \quad (9.23)$$



Все рассуждения, приведенные для излома на ядре, можно применить и здесь, если заменить  $Z$  на  $-\frac{1}{2}$ . Таким образом, вместо условия (9.22) получаем условие излома двухэлектронной волновой функции:

$$\lim_{r_{12} \rightarrow 0} \left\langle \frac{d\psi}{dr_{12}} \right\rangle_{\text{среднее}} = \frac{1}{2} \psi \Big|_{r_{12}=0} . \quad (9.24)$$

Эта формула верна как для пары электронов, имеющих одинаковый спин, так и для электронов с разным спином. Однако, если два электрона имеют один и тот же спин, то волновая функция должна быть антисимметричной, а это означает, что допустимы только нечетные значения  $l$ . В этом случае условие (9.24) на самом деле неинформативно, так как обе части равенства просто обращаются в нуль. Для электронов противоположного спина условие излома (9.24) содержит тот важный результат, что если они подходят друг к другу очень близко, то волновая функция, в среднем, ведет себя как  $\sim C(1 + \frac{1}{2}r_{12} + \dots)$ . Этим можно воспользоваться при конструировании явно коррелированных волновых функций.

### 3. Асимптотическое поведение волновой функции на больших расстояниях

Если один из электронов находится очень далеко от остальной части системы, то снова надо решать эффективную одноэлектронную задачу, описываемую уравнением (9.10), но мы должны заменить в нем  $-\frac{Z}{r}$  на эффективный потенциал  $U(r)$ . В этом случае, по существу, рассматривается асимптотическое поведение решения уравнения Хартри—Фока для  $i$ -й орбитали. Поэтому мы будем писать  $\psi_i$  и  $\varepsilon_i$  вместо  $\psi$  и  $E$ . Выписывая производную произведения

$$\frac{d}{dr} \left( r^2 \frac{d\psi_i}{dr} \right) = 2r \frac{d\psi_i}{dr} + r^2 \frac{d^2\psi_i}{dr^2}, \quad (9.25)$$

получим

$$\frac{2}{r} \frac{d\psi_i}{dr} + \frac{d^2\psi_i}{dr^2} - \frac{\hat{l}^2}{r^2} \psi_i - 2U(r)\psi_i = -2\varepsilon_i \psi_i . \quad (9.26)$$

Так как мы рассматриваем случай  $r \rightarrow \infty$ , любое слагаемое  $\sim \frac{1}{r}$  или  $\frac{1}{r^2}$  становится пренебрежимо малым. Вдали от системы общее поведение

потенциала  $U(r)$  также является кулоновским  $\left( \sim \frac{1}{r} \right)$ . Тогда остается выражение

$$\frac{d^2\psi_i}{dr^2} = -2\varepsilon_i\psi_i. \quad (9.27)$$

Общее решение дифференциального уравнения (9.27) имеет вид  $\psi_i = Ae^{\zeta r} + Be^{-\zeta r}$ , где  $\zeta = \sqrt{-2\varepsilon_i}$ . (Заметим, что  $\varepsilon_i < 0$ ). Однако, первое слагаемое расходитсся при  $r \rightarrow \infty$ , и его нужно отбросить. Таким образом, получаем асимптотическое выражение

$$\psi \sim Be^{\sqrt{-2\varepsilon_i}r}. \quad (9.28)$$

Если нас интересует, как полная волновая функция молекулы ведет себя на очень больших расстояниях, то необходимо рассмотреть орбиталь, которая убывает медленнее всего. Тогда, учитывая теорему Купманса (гл. 6, разд. 3), устанавливающую связь между орбитальными энергиями и потенциалами ионизации, можно записать для наиболее медленно убывающей компоненты волновой функции асимптотическое выражение

$$\psi \sim Be^{-\sqrt{2I_{\min}}r}, \quad (9.29)$$

где  $I_{\min} = \min |\varepsilon_i|$  равен наименьшему (первому) потенциалу ионизации молекулы (атома).

#### 4. Базисные функции: теорема о произведении гауссовых функций

Мы видели в предыдущем разделе, что адекватная базисная орбиталь для расчетов электронной структуры должна экспоненциально убывать на больших расстояниях. Орбитали такого типа обычно записываются как

$$\chi_{nlm}(\vec{r}_a) = \mathcal{N} r_a^{n-1} e^{-\zeta r_a} Y_l^m(\vartheta_a, \varphi_a), \quad (9.30)$$

где  $\vec{r}_a = \vec{r} - \vec{R}_A$  — радиус-вектор электрона относительно центра (атома)  $A$ ;  $r_a, \vartheta_a$ , и  $\varphi_a$  являются сферическими координатами, относящимися к центру  $A$ ;  $Y_l^m(\vartheta_a, \varphi_a)$  — сферические гармоники, а  $\mathcal{N}$  — нормировочный коэффициент.

Орбитали вида (9.30) называются орбиталями слэтеровского типа (Slater-type orbital, STO). Точные решения задачи об атоме водорода можно также представить, как линейные комбинации нескольких STO (с одинаковым показателем экспоненты), так что STO можно рассматривать, как адекватное обобщение (и в то же время упрощение) водородоподобных орбиталей для использования в многоэлектронных задачах.

В соответствии с этим, использование STO было стандартной практикой на раннем этапе развития квантовой химии.

Применение STO позволило проводить замечательные расчеты атомов методом Хартри—Фока, а также и некоторые весьма полезные расчеты двухатомных молекул. Однако, при использовании STO «узким местом» оказалось то, что не удавалось вывести поддающиеся вычислению выражения для трех- и четырехцентровых интегралов, которые необходимы для расчетов многоатомных молекул. Хотя попытки развить практические методы расчета интегралов по STO продолжают, в абсолютном большинстве реальных квантовохимических программ используются «орбитали гауссова типа» (Gaussian-type orbitals, GTOs). Первоначально Бойс (1950) предложил вычислять интегралы с STO, аппроксимируя их линейными комбинациями гауссовых функций, т. е. таких, в которых радиальная зависимость равна  $\sim e^{-\alpha r^2}$ , вместо  $e^{-\zeta r}$ . По существу, эта идея до сих пор применяется при использовании простейших базисов «STO-nG», в котором орбиталь слэтеровского типа приближена суммой  $n$  гауссовых функций ( $n$  обычно заключено между 3 и 6). Однако, в большинстве случаев в качестве базисных орбиталей используют гауссовы функции сами по себе; базисная функция может быть либо одной гауссовой функцией, либо линейной комбинацией («группировкой» или «контрактированием») нескольких гауссовых функций с фиксированными коэффициентами.

По функциональному виду, гауссовы функции не очень хорошо подходят для задач квантовой химии<sup>1</sup>; их преимущество заключается только в том, что для них можно сравнительно просто вычислить все необходимые одно- и двухэлектронные интегралы. Возможность расчета интегралов по GTO предоставляется так называемой «теоремой о произведении гауссовых функций», позволяющей представить произведение двух гауссовых функций, центрированных в разных точках пространства, в виде третьей гауссовой функции, центрированной в точке, лежащей на прямой, соединяющей центры двух первых функций. Тогда трех- и четырехцентровые интегралы сводятся к двухцентровым, которые можно относительно легко рассчитать.

Обычно элементарную гауссову функцию записывают в «декартовой» форме

$$\chi_{uvw}(\vec{r}_a) = \mathcal{N} x_a^u y_a^v z_a^w e^{-\alpha r_a^2}, \quad (9.31)$$

<sup>1</sup> Асимптотика не равна (9.28) и условие излома на ядре не может быть в точности удовлетворено никакой конечной линейной комбинацией гауссовых функций. — Прим. ред.

где  $u, v$ , и  $w$  являются неотрицательными целыми числами, и

$$\vec{r}_a = \vec{i}x_a + \vec{j}y_a + \vec{k}z_a = \vec{i}(x - X_A) + \vec{j}(y - Y_A) + \vec{k}(z - Z_A) \quad (9.32)$$

есть радиус-вектор относительно центра  $A$ , а  $X_A, Y_A$ , и  $Z_A$  являются координатами центра  $A$ . («Чистые»  $d, f$ , и т. д. функции, угловая зависимость которых описывается сферическими гармониками  $Y_l^m(\vartheta_a, \varphi_a)$ , можно представить как линейные комбинации декартовых слагаемых.)

Нас интересует произведение:

$$e^{-\alpha r_a^2} e^{-\beta r_b^2} = \exp \left\{ -\alpha [(x - X_A)^2 + (y - Y_A)^2 + (z - Z_A)^2] - \beta [(x - X_B)^2 + (y - Y_B)^2 + (z - Z_B)^2] \right\}. \quad (9.33)$$

Это выражение можно представить в виде произведения трех сомножителей, относящихся соответственно к трем декартовым осям. Один из них равен

$$\begin{aligned} & \exp \left\{ -\alpha (x - X_A)^2 - \beta (x - X_B)^2 \right\} \\ &= \exp \left\{ -\alpha (x^2 - 2xX_A + X_A^2) - \beta (x^2 - 2xX_B + X_B^2) \right\} \\ &= \exp \left\{ -(\alpha + \beta)x^2 + 2x(\alpha X_A + \beta X_B) - (\alpha X_A^2 + \beta X_B^2) \right\} \\ &= \exp \left\{ -(\alpha + \beta) \left[ x^2 - 2x \frac{\alpha X_A + \beta X_B}{\alpha + \beta} + \frac{\alpha X_A^2 + \beta X_B^2}{\alpha + \beta} \right] \right\} \quad (9.34) \\ &= \exp \left\{ -(\alpha + \beta) \left[ \left( x - \frac{\alpha X_A + \beta X_B}{\alpha + \beta} \right)^2 - \left( \frac{\alpha X_A + \beta X_B}{\alpha + \beta} \right)^2 + \frac{\alpha X_A^2 + \beta X_B^2}{\alpha + \beta} \right] \right\} \\ &= \exp \left\{ -(\alpha + \beta) \left[ \left( x - \frac{\alpha X_A + \beta X_B}{\alpha + \beta} \right)^2 + \frac{\alpha \beta}{(\alpha + \beta)^2} (X_A - X_B)^2 \right] \right\} \\ &= \text{const} \times \exp[-(\alpha + \beta)(x - X_0)^2], \end{aligned}$$

где

$$X_0 = \frac{\alpha X_A + \beta X_B}{\alpha + \beta}. \quad (9.35)$$

Поэтому

$$e^{-\alpha r_a^2} e^{-\beta r_b^2} = \text{const} \times e^{-(\alpha + \beta)(\vec{r} - \vec{R}_0)^2}, \quad (9.36)$$

где

$$\vec{R}_0 = \frac{\alpha \vec{R}_A + \beta \vec{R}_B}{\alpha + \beta}. \quad \text{Ч. т. д.} \quad (9.37)$$

## 5. Проблема преобразования интегралов

Энергия корреляции почти всегда определяется выражениями, содержащими интегралы по молекулярным орбиталям; см., например, выражение (8.14), представляющее теорему Несбета или выражения (8.35) и (8.36) для корреляционной энергии UMP2 и RMP2, соответственно. Сами МО, в свою очередь, обычно являются линейными комбинациями базисных (атомных) орбиталей (ср. гл. 6, разд. 5). Поэтому интегралы по МО являются суммами интегралов по базисным функциям. Если

$$\varphi_i = \sum_{\mu=1}^m c_{\mu}^i \chi_{\mu}, \quad (9.38)$$

то, как легко видеть, имеем

$$\langle \varphi_i | \hat{h} | \varphi_j \rangle = \sum_{\mu, \nu=1}^m c_{\mu}^i c_{\nu}^j \langle \chi_{\mu} | \hat{h} | \chi_{\nu} \rangle, \quad (9.39)$$

и

$$[\varphi_i \varphi_j | \varphi_k \varphi_l] = \sum_{\mu, \nu, \varrho, \tau=1}^m c_{\mu}^i c_{\nu}^j c_{\varrho}^k c_{\tau}^l [\chi_{\mu} \chi_{\nu} | \chi_{\varrho} \chi_{\tau}]. \quad (9.40)$$

(Мы предположили для простоты, что орбитали вещественны, как это обычно и бывает на практике).

Преобразование одноэлектронных интегралов (9.39), если оно вообще требуется, не вызывает каких-либо проблем. Оно соответствует переходу от представления одноэлектронного гамильтониана  $\hat{h}$  в одном конечном базисе к другому в том же самом  $m$ -мерном одноэлектронном подпространстве. Тогда, обозначив интегралы  $\langle \varphi_i | \hat{h} | \varphi_j \rangle = h_{ij}^{MO}$  и  $\langle \chi_{\mu} | \hat{h} | \chi_{\nu} \rangle = h_{\mu\nu}^{AO}$ , и рассматривая их как элементы  $m \times m$  матриц  $\mathbf{h}^{MO}$  и  $\mathbf{h}^{AO}$  соответственно, а также обозначая матрицу, образованную коэффициентами ЛКАО, как  $\mathbf{C}$ , имеем

$$\begin{aligned} h_{ij}^{MO} &= \sum_{\mu, \nu=1}^m c_{\mu}^i h_{\mu\nu} c_{\nu}^j = \sum_{\mu, \nu=1}^m C_{\mu i} h_{\mu\nu}^{AO} C_{\nu j} \\ &= \sum_{\mu, \nu=1}^m (\mathbf{C}^{\dagger})_{i\mu} h_{\mu\nu}^{AO} C_{\nu j} = (\mathbf{C}^{\dagger} \mathbf{h}^{AO} \mathbf{C})_{ij} \end{aligned} \quad (9.41)$$

т. е.  $\mathbf{h}^{MO} = \mathbf{C}^{\dagger} \mathbf{h}^{AO} \mathbf{C}$ . Это означает, что  $\mathbf{h}^{MO}$  можно вычислять двумя последовательными матричными умножениями, которые требуют вычислительных затрат, зависящих от размерности  $m$  базиса одноэлектронных состояний, как  $\sim m^3$ .

Ситуация с двухэлектронными интегралами (9.40) намного более тяжелая. Число двухэлектронных интегралов по МО пропорционально четвертой степени размерности системы, так как необходимо вычислить  $\sim n_{\text{зан.}}^2 \times n_{\text{вирт.}}^2$  разных интегралов, если используется теория возмущений MP2, и  $\sim m^4$  для применения теорий высшего порядка, которые требуют знания всех возможных интегралов по МО. (Здесь  $n_{\text{зан.}}$  и  $n_{\text{вирт.}}$  — числа занятых и виртуальных орбиталей, соответственно;  $n_{\text{зан.}} + n_{\text{вирт.}} = m$ . Так как для достаточно хорошего базиса  $n_{\text{зан.}} \ll m$ , это сравнение объясняет, почему MP2 намного дешевле любого другого метода расчета электронной корреляции.) Каждый интеграл в правой части выражения (9.40) требует  $\sim m^4$  арифметических операций, и тогда «лобовой» расчет всех интегралов требовал бы вычислительных ресурсов, растущих как восьмая степень размера системы (числа орбиталей), что неприемлемо в любом практически интересном случае. Однако эту трудность можно обойти, если проводить преобразование пошагово. (Это вполне аналогично сведению преобразования (9.39), которое требовало бы  $\sim m^4$  операций, к двум матричным умножениям, каждое с  $\sim m^3$  операциями.)

Как первый шаг необходимо вычислить промежуточные величины

$$[\chi_\mu \chi_\nu | \chi_\varrho \varphi_l] = \sum_{\tau=1}^m c_\tau^l [\chi_\mu \chi_\nu | \chi_\varrho \chi_\tau], \quad (9.42)$$

а затем величины

$$[\chi_\mu \chi_\nu | \varphi_k \varphi_l] = \sum_{\varrho=1}^m c_\varrho^k [\chi_\mu \chi_\nu | \chi_\varrho \varphi_l], \quad (9.43)$$

и т. д. Каждый такой шаг требует «только»  $\sim m^5$  операций, если проводится полное преобразование интегралов. Эта маленькая хитрость весьма существенна: без нее нельзя было бы провести расчеты электронной корреляции для систем, представляющих интерес для химиков. (Конечно, в практических расчетах можно ввести дальнейшие упрощения, но они не повлияют на концептуальные аспекты проблемы.)

## Библиографические заметки

### Раздел 1.

Разложение по степеням  $1/Z$  восходит к Хиллераасу [1, 2] (также см., например, [3, 4]). Ссылки на Мозли и ван ден Брока: [5, 6]

### Раздел 2.

Поведение  $R(r) \sim r^l$  около ядра рассматривается в учебниках по квантовой механике (например, [7]). Условия излома, определенные уравнениями (9.22), (9.24), были получены Като [8].



### Раздел 3.

Вывод представляет собой незначительное обобщение вывода, приведенного в [7].

### Раздел 4.

Слэтер первоначально ввел [9] свои орбитали для приближенного описания свободных атомов; он также дал эмпирическое правило для определения показателей экспонент и дал оценки полных атомных энергий, как сумм одно-электронных вкладов — ср. гл. 6, разд. 3.1. (Эти оценки являются неожиданно хорошими). Хартри—фоковские расчеты свободных атомов с помощью STO описаны в [10]. В книге Слэтера [11] обсуждаются некоторые классические расчеты двухатомных молекул, также проведенные с применением STO.

GTO были впервые введены в квантовую химию Бойсом [12], которому также принадлежит «теорема о произведении гауссовых функций». Шавитт [13], вероятно, первым предложил использовать гауссовы функции, как независимые базисные функции, а не только как средства для оценки интегралов по STO.

### Раздел 5.

Метод преобразования интегралов подробно обсуждается, например, в [14].

## Для дальнейшего изучения

Сводка классических формул для расчета интегралов по STO и GTO приведена в [15]. Относительно подробного обсуждения математических методов мы сошлемся на [16-18]. Статья [17] содержит также обсуждение метода полиномов Риса (Rys), введенного в [19] и теперь повсеместно используемого в расчетах двухэлектронных интегралов в базисах гауссовых функций.

## Литература

1. Hylleraas E. A. Z. Phys. **65**, 209 (1930).
2. Hylleraas E. A. Phys. Rev. **103**, 829 (1956).
3. Löwdin P.-O. J. Mol. Spectry, **3**, 46 (1959).
4. Silverman J. N. Phys. Rev. **A23**, 441 (1981).
5. Moseley H. Phil. Mag. **26**, 1024 (1913); **27**, 703 (1914).
6. van den Broek A. Nature **92**, 372 (1913).
7. Ландау Л. Д., Лифшиц И. М. *Квантовая механика*. — М.: Наука, 1974.
8. Kato T. Commun. Pure Appl. Math. **10**, 15 (1957).
9. Slater J. C. Phys. Rev. **36**, 57 (1930).
10. Clementi E., Roetti C. *Roothaan-Hartree-Fock Atomic Wavefunctions*, Academic Press, New York 1974. (Atomic Data and Nuclear Data Tables **14**, 177, 1974).
11. Slater J. C. *Quantum Theory of Molecules and Solids vol. 1. Electronic Structure of Molecules*, McGraw-Hill, New York 1963. (Имеется перевод: Слэтер Дж. *Электронная структура молекул*. — М.: Мир, 1965).
12. Boys S. F. Proc. Roy. Soc. (London) **A200**, 542 (1950).
13. Shavitt I., in *Methods in Computational Physics*, Academic Press, New York 1963.
14. Roos B. *The Configuration Interaction Method in Computational Techniques in Quantum Chemistry and Molecular Physics* (ed. G.H.F. Dierksen, B.T. Suttcliffe and A. Veillard) Reidel, Dordrecht 1974.
15. Жоголев Д. А., Волков В. Б. *Методы, алгоритмы и программы для квантово-химических расчетов молекул*. — Киев, Наукова Думка, 1976.

16. Saunders V. R. *An Introduction to Molecular Integral Evaluation*, in *Computational Techniques in Quantum Chemistry and Molecular Physics* (ed. G.H.F. Dierksen, B.T. Sutcliffe and A. Veillard) Reidel, Dordrecht 1974.
17. Saunders V. R. in *Methods in Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.
18. Steinborn R. O. in *Methods in Computational Molecular Physics* (ed. G.H.F. Dierksen and S. Wilson) Reidel, Dordrecht 1983.
19. Dupuis M., Rys J., King H. F. J. Chem. Phys. **65**, 111 (1976).





# Дополнение редактора перевода



Хотя задача об электронной структуре молекул в рамках приближения Хартри—Фока в общем случае менее ресурсоемка, чем «точное» решение полного метода КВ, она оказывается очень богатой в силу своей нелинейности, что приводит к исключительному разнообразию возможных решений. Пренебрежение этими существенными качественными особенностями приводит к колоссальной путанице в сложных случаях, которые, однако, и являются самыми интересными с точки зрения химии: в случаях разрыва старых и образования новых связей. В этих ситуациях оказывается, что наиболее распространенный вариант метода Хартри—Фока — ограниченный метод Хартри—Фока (ОХФ, RHF) не дает ответа даже минимально похожего на правильный. Более того, попытки уточнения этого решения без выхода за рамки однодетерминантного приближения (Хартри—Фока) приводят к крайне причудливым и без дополнительных пояснений непонятным если не сказать патологическим результатам. Эта особенность существенно нелинейных задач проявляет себя на графике, помещенном на с. 228, и настоящее Дополнение, в общем следующее работам [1,2], служит тому, чтобы прояснить представленную там картину.

Рассмотрим упоминавшуюся в основном тексте на с. 227 задачу об электронном строении молекулы  $\text{H}_2$  в базисе  $1s$ -функций в зависимости от межъядерного расстояния. Дополнительное упрощение достигается если предположить, что интеграл перекрывания между базисными  $1s$ -функциями равен нулю (или же, считать, что эти функции суть ортогонализированные по Лёвдину). В этом случае выражение для точной (в данной постановке) нормированной на единицу волновой функции основного состояния упрощается:

$$\begin{aligned} \Phi_{\text{CI}} = & \frac{u}{\sqrt{2}} \underbrace{\left\{ \hat{A}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{A}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] \right\}}_{\text{ионные (H}^-\text{H}^+ \text{ и H}^+\text{H}^-)} \\ & + \frac{w}{\sqrt{2}} \underbrace{\left\{ \hat{A}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{A}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right\}}_{\text{ковалентные}} \end{aligned} \quad (\text{Д.1})$$

$$u^2 + w^2 = 1 + c^2.$$

В обозначениях основного текста книги (см. формулу (8.20))

$$\begin{aligned} u &= \frac{1}{\sqrt{2}}(1 + c)/\sqrt{1 + c^2} \\ w &= \frac{1}{\sqrt{2}}(1 - c)/\sqrt{1 + c^2} \end{aligned} \quad (\text{Д.2})$$

Для того чтобы оценить коэффициенты разложения  $u$  и  $w$ , применим следующие полуэмпирические соображения. Прежде всего заметим, что в разложении точной волновой функции участвуют три синглетных базисных конфигурации двух электронов: две ионные и одна ковалентная. Примем за начало отсчета энергию синглетного состояния, в котором на каждом из атомов водорода находится по одному электрону, т. е. энергию ковалентной синглетной конфигурации. Это означает, что диагональный матричный элемент оператора Гамильтона нашей модельной системы по волновой функции

$$\frac{1}{\sqrt{2}} \left( \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right)$$

равен нулю. Для ионных функций диагональный матричный элемент оператора Гамильтона можно считать равным энергии взаимодействия двух электронов, занимающих одну и ту же  $1s$ -орбиталь на одном из атомов водорода. Обозначим эту энергию  $\gamma = [\chi_A\chi_A | \chi_A\chi_A]$ . Матричный элемент оператора Гамильтона между двумя ионными конфигурациями  $H^-H^+$  и  $H^+H^-$  по правилам Слэтера отличен от нуля, так как эти конфигурации отличаются всего лишь двумя спин-орбиталями. Однако соответствующий двухэлектронный интеграл  $[\chi_A\chi_A | \chi_B\chi_B]$  пропорционален квадрату дифференциального перекрытия орбиталей  $\chi_A$  и  $\chi_B$ , а потому в первом приближении может быть отброшен. Напротив, матричный элемент оператора Гамильтона между ковалентной и любой из ионных конфигураций отличен от нуля и (в силу симметрии задачи) одинаков, поэтому может также рассматриваться как еще один независимый параметр, который мы обозначим через  $-\sqrt{2}\beta$ . В этих обозначениях оператор Гамильтона в базисе трех синглетных конфигураций, которые можно построить в системе двух электронов на двух орбиталях, примет вид

$$\begin{pmatrix} 0 & -\sqrt{2}\beta & -\sqrt{2}\beta \\ -\sqrt{2}\beta & \gamma & 0 \\ -\sqrt{2}\beta & 0 & \gamma \end{pmatrix} \quad (\text{Д.3})$$

Эта конструкция носит название гамильтониана (Андерсона—)Хаббарда и обычно вводится с помощью операторов вторичного квантования, которые не используются в этой книге. Она широко применяется в физике твердого тела и по этой причине пример решения задачи о молекуле  $H_2$  можно найти в книге [3].

Переход к новым базисным состояниям в подпространстве ионных конфигураций, а именно к симметричной и антисимметричной относительно перестановки  $1s$ -орбиталей атомов водорода комбинации ионных конфигураций позволяет переписать матрицу оператора Гамильтона в этом новом базисе как:

$$\begin{pmatrix} 0 & -2\beta & 0 \\ -2\beta & \gamma & 0 \\ 0 & 0 & \gamma \end{pmatrix} \quad (\text{Д.4})$$

Решая соответствующее (сводящееся к квадратному) вековое уравнение, можно найти его низшие собственные значения:

$$E_S = \frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} + 4\beta^2} = \frac{\gamma}{2} \left( 1 - \sqrt{1 + \frac{16\beta^2}{\gamma^2}} \right) \approx -\frac{4\beta^2}{\gamma}$$

$$E'_S = \gamma; \quad E_T = 0,$$

где легко видеть, что энергия антисимметричной комбинации синглетных ионных конфигураций равна  $\gamma$ , а кроме того можно убедиться в том, что энергия триплетного состояния равна нулю.

Подставляя полученное выше собственное значение в матричное уравнение на собственные векторы, получаем также явные значения коэффициентов разложения основного состояния по конфигурациям:

$$\begin{pmatrix} w \\ u \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1 + \frac{\gamma}{\sqrt{\gamma^2 + 4\beta^2}}} \\ \sqrt{1 - \frac{\gamma}{\sqrt{\gamma^2 + 4\beta^2}}} \end{pmatrix}$$

В рамках этой конструкции легко промоделировать процесс диссоциации молекулы  $\text{H}_2$ . Энергия взаимодействия двух электронов на одном центре  $\gamma$  в первом приближении не зависит от межъядерного расстояния, в то время как параметр  $\beta$ , описывающий в том числе и перескок электрона между центрами, экспоненциально быстро обращается в нуль при бесконечном разведении ядер. Поэтому предел  $r \rightarrow \infty$  эквивалентен пределу  $\beta \rightarrow 0$ . Легко убедиться, что при бесконечном удалении атомов водорода (точная) волновая функция сводится к синглету, образованному из двух однократно заполненных орбиталей ( $u = 0; w = 1$ ), т. е. к ковалентной функции.

Теперь рассмотрим как с точки зрения точного решения выглядят различные хартри-фоковские решения этой задачи. В приближении ОХФ ответ полностью определяется на основе симметричных соображений: общая пространственная часть двух заполненных спин-орбиталей (отвечающих проекциям спина «вверх» и «вниз») имеет вид  $\varphi = \frac{1}{\sqrt{2}}(\chi_A + \chi_B)$ . При ее подстановке в однодетерминантную волновую функцию обнаруживается, что коэффициенты конфигураций  $u = w = \frac{1}{\sqrt{2}}$  отвечают точному решению при  $\gamma = 0$ , т. е. при отсутствии какого бы то ни было взаимодействия между электронами. Они остаются неизменными при любых значениях параметров модели в том числе и при бесконечном межъядерном расстоянии. Понятно, что это серьезный дефект по сравнению с точной волновой функцией, поскольку вклад в волновую функцию ионной конфигурации, имеющей положительную энергию, не обращается в нуль в пределе  $r \rightarrow \infty$ , как это должно быть у точной функции, и диссоциационный предел энергии оказывается неправильным. Заметим, что в выражении для энергии ОХФ состояния, но не его волновой функции (см. ниже), взаимодействие  $\gamma$  входит. Эта энергия, как и положено,

является оценкой сверху для точной энергии, причем она относительно тем хуже, чем меньше величина безразмерного параметра задачи  $\beta/\gamma$ . Это моделирует неправильный диссоциационный предел приближения ОХФ.

Для приближения неограниченного метода Хартри–Фока (НХФ) однодетерминантная волновая функция получается заполнением спин-орбиталей с *разными* пространственными частями для каждой из проекций спина:

$$\begin{aligned}\varphi_{\uparrow} &= x_{\uparrow}\chi_A + y_{\uparrow}\chi_B \\ \varphi_{\downarrow} &= x_{\downarrow}\chi_A + y_{\downarrow}\chi_B\end{aligned}\quad (\text{Д.5})$$

При этом нормировка заполненных орбиталей: как той, которая занята электроном с проекцией спина вверх, так и той, которая занята электроном с проекцией спина вниз, обеспечена условиями

$$1 = x_{\uparrow}^2 + y_{\uparrow}^2 = x_{\downarrow}^2 + y_{\downarrow}^2.$$

Коэффициенты  $x_{\sigma}$  и  $y_{\sigma}$  должны быть найдены диагонализацией операторов Фока для каждой из проекций спина:

$$F_{\sigma} \begin{pmatrix} x_{\sigma} \\ y_{\sigma} \end{pmatrix} = \begin{pmatrix} \gamma x_{-\sigma}^2 & -\beta \\ -\beta & \gamma y_{-\sigma}^2 \end{pmatrix} \begin{pmatrix} x_{\sigma} \\ y_{\sigma} \end{pmatrix} = \varepsilon_{\sigma} \begin{pmatrix} x_{\sigma} \\ y_{\sigma} \end{pmatrix}.$$

Подчиним диагональные матричные элементы плотности, равные квадратам соответствующих коэффициентов и входящие в свою очередь в определение оператора Фока, дополнительному, но физически вполне естественному условию: полная заселенность (электронами с проекциями спинов вверх и вниз) каждой из  $1s$ -орбиталей должна быть равна единице, с тем чтобы не возникла поляризация молекулы:

$$1 = x_{\uparrow}^2 + x_{\downarrow}^2 = y_{\uparrow}^2 + y_{\downarrow}^2.$$

Этим условиям можно удовлетворить, если положить:

$$\begin{aligned}x_{\uparrow}^2 &= \frac{1}{2} + \delta, & x_{\downarrow}^2 &= \frac{1}{2} - \delta \\ y_{\uparrow}^2 &= \frac{1}{2} - \delta, & y_{\downarrow}^2 &= \frac{1}{2} + \delta\end{aligned}$$

Подстановка этих формул в операторы Фока приводит к выражениям

$$\begin{aligned}F_{\uparrow} &= \begin{pmatrix} \gamma \left(\frac{1}{2} - \delta\right) & -\beta \\ -\beta & \gamma \left(\frac{1}{2} + \delta\right) \end{pmatrix} \\ F_{\downarrow} &= \begin{pmatrix} \gamma \left(\frac{1}{2} + \delta\right) & -\beta \\ -\beta & \gamma \left(\frac{1}{2} - \delta\right) \end{pmatrix},\end{aligned}\quad (\text{Д.6})$$

из которых становится очевидно, что собственные значения обоих операторов  $F_{\uparrow}$  и  $F_{\downarrow}$  совпадают (не зависят от проекции спина):

$$\varepsilon_{\uparrow\downarrow}^{\pm} = \varepsilon^{\pm} = \frac{\gamma}{2} \pm \sqrt{(\gamma\delta)^2 + \beta^2}, \quad (\text{Д.7})$$

так что и для проекции спина вверх и для проекции спина вниз должны быть заполнены спин-орбитали с одной и той же орбитальной энергией:  $\varepsilon^-$ . Заметим также, что задачи на собственные векторы для операторов Фока с проекциями спина вверх и вниз отличаются перестановкой строк и столбцов. Это позволяет записать

$$\begin{aligned} x_{\uparrow} &= x, & y_{\uparrow} &= y \\ x_{\downarrow} &= y, & y_{\downarrow} &= x, \end{aligned} \quad (\text{Д.8})$$

после чего, решая любую из двух задач на собственные векторы матриц  $F_{\sigma}$  с собственным значением  $\varepsilon^-$ , получаем выражения для коэффициентов  $x$  и  $y$ :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1 + \frac{\gamma\delta}{\sqrt{(\gamma\delta)^2 + \beta^2}}} \\ \sqrt{1 - \frac{\gamma\delta}{\sqrt{(\gamma\delta)^2 + \beta^2}}} \end{pmatrix}$$

Величина параметра  $\delta$  должна определяться из условия самосогласования, которое удобно получить, выражая  $\delta$  через квадраты получившихся при диагонализации коэффициентов МО ЛКАО:

$$2\delta = \frac{\gamma\delta}{\sqrt{(\gamma\delta)^2 + \beta^2}} \Rightarrow 4((\gamma\delta)^2 + \beta^2) = \gamma^2 \Rightarrow \delta = \pm \frac{1}{2} \sqrt{1 - \frac{4\beta^2}{\gamma^2}} \quad (\text{Д.9})$$

Из этих условий легко видеть, что всегда существует тривиальное решение (уравнения самосогласования):  $\delta = 0$ . В том случае, если энергия взаимодействия  $\gamma$  превышает критическую величину, парами (если  $\delta$  — решение, то  $-\delta$  — тоже решение) появляются нетривиальные решения, отвечающие различному (но произвольному) выбору того, на каком из концов молекулы будет концентрироваться электрон с проекцией спина вверх:

$$\begin{cases} \gamma \leq 2\beta; \delta \equiv 0 \\ \gamma > 2\beta; \delta = \pm \frac{1}{2} \sqrt{1 - \frac{4\beta^2}{\gamma^2}} \end{cases} \quad (\text{Д.10})$$

Это означает, что при значениях энергии взаимодействия  $\gamma$  меньших критического единственное решение задачи НХФ совпадает с решением задачи ОХФ. При большей энергии взаимодействия появляется дополнительное нетривиальное решение приближения НХФ. Смысл происходящего довольно прозрачен: в нашей упрощенной модели имеет место взаимодействие (отталкивание) только между электронами, находящимися на одной АО. Оно приводит к увеличению полной энергии. При этом электрон с данной проекцией спина взаимодействует только с электроном с противоположной проекцией спина. Поэтому, если электроны с разными проекциями спина соберутся в основном на разных концах молекулы  $\text{H}_2$ , это приведет к снижению энергии отталкивания. Однако одноэлектронная часть энергии (отрицательная) при этом уменьшается по абсолютной величине. В тот момент, когда выигрыш за счет разбегания электронов

с разными проекциями спина становится больше, чем проигрыш из-за потери порядка связи между атомами Н, возникает нетривиальное решение НХФ. С учетом условий симметрии (Д.8) волновая функция НХФ примет вид

$$\Phi_{\text{УНФ}} = xy\hat{A}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + xy\hat{A}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] + x^2\hat{A}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + y^2\hat{A}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)]$$

Подставляя выражения коэффициентов  $x$  и  $y$  в выражение для двухэлектронной волновой функции, получаем

$$\Phi_{\text{УНФ}} = \sqrt{\frac{1}{4} - \delta^2} \left\{ \hat{A}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{A}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] \right\} + \frac{1}{2} \left\{ \hat{A}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{A}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right\} + \delta \left\{ \hat{A}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] - \hat{A}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right\}$$

откуда становится очевидным основной дефект приближения НХФ: его волновая функция не отвечает никакому определенному значению полного спина. Первые два слагаемых представляют собой линейную комбинацию синглетных конфигураций двух электронов на двух орбиталях, уже появлявшуюся ранее. Третье слагаемое пропорционально компоненте триплетного состояния с нулевой проекцией полного спина. Это — та цена, которую в методе НХФ приходится платить за большую гибкость, т. е. за то, что волновая функция, оставаясь однодетерминантной, стала зависеть от параметров задачи.

Выписывая энергию системы в приближении НХФ, получаем

$$\varepsilon_{\text{УНФ}} = -4\beta\sqrt{\frac{1}{4} - \delta^2} + 2\gamma\left(\frac{1}{4} - \delta^2\right)$$

При значениях энергии взаимодействия выше критического это выражение точно равно  $-2\beta^2/\gamma$ , и в диссоциационном пределе оказывается равным полусумме точных энергий триплетного и синглетного состояний, что неудивительно, так как волновая функция НХФ в диссоциационном пределе есть суперпозиция синглетной и триплетной ковалентных конфигураций с равными весами.

Исследуем поведение энергии в окрестности критической точки, в которой появляется нетривиальное решение. Очевидно, что параметр  $\delta$  в критической точке непрерывен, но недифференцируем, так как имеет в ней вертикальную касательную. Формальная производная  $\delta'$ ,

$$\delta' = -2 \frac{\beta\beta'}{\gamma^2 \sqrt{1 - \frac{4\beta^2}{\gamma^2}}},$$

обращается в бесконечность. В свою очередь, в выражение для энергии  $\varepsilon_{\text{УНФ}}$  параметр  $\delta$  входит только квадратично, а потому производная энергии

$$\varepsilon'_{\text{УНФ}} = -4\beta'\sqrt{\frac{1}{4} - \delta^2} + 4\beta \frac{\delta\delta'}{\sqrt{\frac{1}{4} - \delta^2}} - 4\gamma\delta\delta'$$

содержит непрерывную комбинацию

$$\delta\delta' = -\frac{\beta\beta'}{\gamma^2},$$

поэтому энергия дифференцируема в критической точке. Ее производная равна  $-2\beta'$ . Кроме того, подставляя явное выражение для  $\delta$ , мы убеждаемся, что слагаемые, содержащие  $\delta'$ , взаимно уничтожаются. Это можно понять, так как в используемом нами модельном гамильтониане только параметр  $\beta$  зависит от расстояния, и производная гамильтониана в том же базисе, где

гамильтониан имеет вид (Д.4), принимает вид:  $\begin{pmatrix} 0 & -2\beta' & 0 \\ -2\beta' & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ , а в соот-

ветствии с теоремой Гельмана—Фейнмана производная энергии есть среднее от производной гамильтониана. Также можно убедиться, что первые производные энергий ОХФ и НХФ в критической точке совпадают. Высшие производные  $\varepsilon_{\text{УНФ}}$ , как можно убедиться, уже не будут существовать в критической точке.

Как указано в основном тексте, возможны два способа избавления от нежелательного триплетного вклада в волновую функцию метода НХФ. Исторически первый заключается в вычислении параметра  $\delta$  по формулам (Д.9), последующем отбрасывании триплетного вклада и перенормировании получившейся функции на единицу:

$$\begin{aligned} \Phi_{\text{ЕНФ}} &= \Phi_{\text{УНФ-СП}} = \\ &= \sqrt{\frac{1-4\delta^2}{2-4\delta^2}} \left\{ \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_A(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_B(2)\beta(2)] \right\} + \\ &+ \sqrt{\frac{1}{2-4\delta^2}} \left\{ \hat{\mathcal{A}}[\chi_A(1)\alpha(1)\chi_B(2)\beta(2)] + \hat{\mathcal{A}}[\chi_B(1)\alpha(1)\chi_A(2)\beta(2)] \right\}. \end{aligned}$$

Это отвечает приближению НХФ-СП («метод НХФ с последующей спиновой проекцией»). Получившаяся волновая функция имеет вид некоего разложения КВ, аналогичного (Д.1). Подстановка равновесного значения  $\delta$  дает следующие выражения для коэффициентов:

$$\begin{aligned} u &= \frac{2\beta}{\sqrt{\gamma^2 + 4\beta^2}} \\ w &= \frac{\gamma}{\sqrt{\gamma^2 + 4\beta^2}}, \end{aligned}$$

что приводит к выражению для энергии:

$$\varepsilon_{\text{УНФ-СП}} = -\frac{4\beta^2}{\gamma} \frac{1}{\left(1 + \frac{4\beta^2}{\gamma^2}\right)}$$

В асимптотической области ( $\beta \rightarrow 0$ ,  $r \rightarrow \infty$ ) имеем

$$\varepsilon_{\text{УНФ-СП}} \approx -\frac{4\beta^2}{\gamma} \left(1 - \frac{4\beta^2}{\gamma^2}\right),$$

т. е. кривая, отвечающая методу НХФ-СП, идет несколько выше, чем точная, причем относительная ошибка квадратична по малому параметру  $\frac{\beta}{\gamma}$ . Что же касается поведения энергии метода НХФ-СП в критической точке, то ее производная имеет вид

$$\varepsilon'_{\text{УНФ-СП}} = \frac{8\beta\gamma^3}{(4\beta^2 + \gamma^2)^2} \beta'$$

что в критической точке дает

$$\varepsilon'_{\text{УНФ-СП}} = -\beta'$$

и не совпадает с производными энергии в критической точке, которые получаются в методах ОХФ и НХФ. Поэтому кривая энергии метода НХФ-СП имеет конечный угол с остальными кривыми в критической точке, что и видно на рисунке (см. с. 228).

Альтернативой этому является расширенный метод Хартри—Фока (РХФ) в терминологии основного текста книги. В его рамках также производится проектирование волновой функции НХФ на подпространство синглетных состояний. Формально результат не отличается от метода НХФ-СП, однако в этом случае  $\delta$  рассматривают как независимую переменную. Это приводит к выражениям для амплитуд КВ:

$$u = \sqrt{\frac{1 - 4\delta^2}{2 - 4\delta^2}}; \quad w = \sqrt{\frac{1}{2 - 4\delta^2}},$$

что может рассматриваться как возможный способ параметризации коэффициентов полного разложения КВ в рассматриваемом конечном базисе. В этом смысле метод РХФ эквивалентен ему. Это объясняет реалистический вид соответствующего энергетического профиля на рис. (см. с. 228).

## Литература

1. Mayer I. Acta Phys. Hung. **54**, 249 (1983).
2. Löwdin P.-O., Mayer I. Adv. Quant. Chem. **24**, 79 (1992).
3. Ашкрофт Н., Мермин Х. *Физика твердого тела*, т. 2. — М.: Мир, 1983.





## П1. Отделение движения центра масс в классической механике

Система уравнений Ньютона для  $N$  точечных масс имеет вид

$$m_i \ddot{\vec{r}}_i = \vec{f}_i \quad i = 1, 2, \dots, N \quad (\text{П1.1})$$

где  $m_i$  — масса  $i$ -й частицы;  $\ddot{\vec{r}}_i$  — ускорение  $i$ -й частицы;  $\vec{r}_i$  — ее радиус-вектор и « $\cdot$ » обозначает дифференцирование по времени;  $\vec{f}_i$  — сила, действующая на  $i$ -ю частицу. Суммируя уравнения

$$\sum_i m_i \ddot{\vec{r}}_i = \sum_i \vec{f}_i = \vec{F}, \quad (\text{П1.2})$$

получаем  $\vec{F}$  — векторную сумму *внешних* сил, действующих на частицы. (Согласно третьему закону Ньютона, векторная сумма внутренних сил равна нулю.)

Введем полную массу системы  $M$ :

$$M = \sum_i m_i. \quad (\text{П1.3})$$

Умножив и разделив выражение (П1.2) на  $M$ , получим

$$M \sum_i \frac{m_i}{M} \ddot{\vec{r}}_i = \vec{F}, \quad \text{т. е.} \quad M \ddot{\vec{R}} = \vec{F}, \quad (\text{П1.4})$$

где

$$\vec{R} = \sum_i \frac{m_i}{M} \vec{r}_i, \quad (\text{П1.5})$$

— *радиус-вектор центра масс*, или масс-взвешенное среднее радиус-векторов отдельных частиц.

Если система замкнутая,  $\vec{F} = 0$  и

$$M \ddot{\vec{R}} = 0 \quad \longrightarrow \quad \dot{\vec{R}} = \text{const}, \quad (\text{П1.6})$$

т. е. центр масс замкнутой системы покоится или движется с постоянной скоростью в постоянном направлении.

Для изучения внутреннего движения системы введем относительные координаты  $\vec{r}_i'$ , определенные относительно центра масс:

$$\vec{r}_i' = \vec{r}_i - \vec{R}, \quad (\text{П1.7})$$

тогда

$$\ddot{\vec{r}}_i' = \ddot{\vec{r}}_i - \ddot{\vec{R}}. \quad (\text{П1.8})$$

Для замкнутой системы  $\ddot{\vec{R}} = 0$ , поэтому

$$\ddot{\vec{r}}_i' = \ddot{\vec{r}}_i = \frac{1}{m_i} \vec{f}_i, \quad (\text{П1.9})$$

т. е., как хорошо известно, в инерциальной системе отсчета, связанной с центром масс, верны те же самые уравнения Ньютона, что и в лабораторной системе. (Разница только в начальных условиях: в системе центра масс полный импульс системы должен быть равен нулю, так что можно выбрать только такие начальные скорости, которые удовлетворяют этому условию.)

Покажем, что кинетическая энергия становится суммой кинетических энергий внутреннего движения и движения центра масс. Обозначим  $\vec{v}_i = \dot{\vec{r}}_i$ ;  $\vec{v}_i' = \dot{\vec{r}}_i'$ ;  $\vec{V}_R = \dot{\vec{R}}$ ; тогда из определения  $\vec{r}_i' = \vec{r}_i - \vec{R}$  следует, что  $\vec{r}_i = \vec{r}_i' + \vec{R}$ ;  $\vec{v}_i = \dot{\vec{r}}_i = \dot{\vec{r}}_i' + \dot{\vec{R}} = \vec{v}_i' + \vec{V}_R$ . Поэтому мы можем записать

$$\frac{1}{2} \sum_i m_i \vec{v}_i^2 = \frac{1}{2} \sum_i m_i \vec{v}_i'^2 + \underbrace{\frac{1}{2} \sum_i m_i \vec{V}_R^2}_M + \sum_i m_i \vec{v}_i' \vec{V}_R. \quad (\text{П1.10})$$

Первые два слагаемых в правой части являются кинетической энергией внутреннего (относительного) движения и движения центра масс (имеющего массу  $M$ ), тогда как последнее слагаемое исчезает. В самом деле

$$\sum_i m_i \vec{v}_i' \vec{V}_R \equiv M \vec{V}_R \sum_i \frac{m_i}{M} \vec{v}_i'. \quad (\text{П1.11})$$

Множитель  $\sum_i \frac{m_i}{M} \vec{v}_i'$  равен скорости центра масс в системе отсчета, связанной с ним самим, а она, очевидно, равна нулю:

$$\sum_i \frac{m_i}{M} \vec{v}_i' = \sum_i \frac{m_i}{M} (\dot{\vec{r}}_i - \dot{\vec{R}}) = \underbrace{\sum_i \frac{m_i}{M} \dot{\vec{r}}_i}_{\dot{\vec{R}}} - \underbrace{\sum_i \frac{m_i}{M}}_1 \dot{\vec{R}} = \dot{\vec{R}} - \dot{\vec{R}} = 0. \quad (\text{П1.12})$$

Отделение движения центра масс возможно также и в квантовой механике, но достигается намного более сложным образом (см. гл. 1, разд. 1.2).

## П2. Переход от задачи двух тел к двум задачам одного тела в классической механике

Уравнения Ньютона для задачи двух тел имеют вид

$$\begin{aligned} m_1 \ddot{\vec{r}}_1 &= \vec{f}_1 \\ m_2 \ddot{\vec{r}}_2 &= \vec{f}_2 \end{aligned} \quad (\text{П2.1})$$

В замкнутой системе соблюдается третий закон Ньютона

$$\vec{f}_2 = -\vec{f}_1 \quad (\text{П2.2})$$

Масса системы равна

$$M = m_1 + m_2 \quad (\text{П2.3})$$

Движение центра масс:

Умножим оба уравнения движения на  $\frac{1}{M} = \frac{1}{m_1 + m_2}$  и просуммируем:

$$\ddot{\vec{R}} = \frac{m_1 \ddot{\vec{r}}_1 + m_2 \ddot{\vec{r}}_2}{m_1 + m_2} = 0 \quad \longrightarrow \quad \dot{\vec{R}} = \frac{m_1 \dot{\vec{r}}_1 + m_2 \dot{\vec{r}}_2}{m_1 + m_2} = \text{const} \quad (\text{П2.4})$$

(как и в приложении П1, центр масс ведет себя как одна частица).

Относительное движение:

Разделим уравнения движения на  $m_1$  и  $m_2$  соответственно:

$$\begin{aligned} \ddot{\vec{r}}_1 &= \frac{1}{m_1} \vec{f}_1 \\ \ddot{\vec{r}}_2 &= \frac{1}{m_2} \vec{f}_2 = -\frac{1}{m_2} \vec{f}_1 \end{aligned} \quad (\text{для замкнутой системы } \vec{f}_2 = -\vec{f}_1) \quad (\text{П2.5})$$

Вычитаем:

$$\ddot{\vec{r}}_1 - \ddot{\vec{r}}_2 = \frac{1}{m_1} \vec{f}_1 + \frac{1}{m_2} \vec{f}_1 = \underbrace{\left( \frac{1}{m_1} + \frac{1}{m_2} \right)}_{\frac{1}{\mu}} \vec{f}_1 \quad (\text{П2.6})$$

Введем определения:  $\vec{r} = \vec{r}_1 - \vec{r}_2$  — координата *относительного движения*;  $\mu$  — «*приведенная масса*»:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} = \frac{m_1 + m_2}{m_1 m_2}, \quad (\text{П2.7})$$

т. е.

$$\mu = \frac{m_1 m_2}{m_1 + m_2} = \frac{m_2}{1 + \frac{m_2}{m_1}}. \quad (\text{П2.8})$$

Важные частные случаи: а) если  $m_1 = m_2$ , то  $\mu = \frac{1}{2} m_1$ ; б) при  $m_1 \rightarrow \infty$   $\lim_{m_1 \rightarrow \infty} \mu = m_2$ .

Используя определения  $\vec{r}$  и  $\mu$ , перепишем (П2.6) как

$$\mu \ddot{\vec{r}} = \vec{f}_1, \quad (\text{П2.9})$$

т. е. получаем, что уравнение относительного движения формально является одностичной задачей.

Полная кинетическая энергия также равна сумме кинетических энергий центра масс и относительного движения, соответствующего приведенной массе  $\mu$ . В самом деле, согласно уравнениям (П1.10)–(П1.12) приложения П1, имеем

$$T = \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} = \frac{M V_R^2}{2} + \frac{m_1 v_1'^2}{2} + \frac{m_2 v_2'^2}{2}, \quad (\text{П2.10})$$

где  $v_1'$ ,  $v_2'$  — скорости частиц в системе центра масс.

Далее, полный импульс в системе центра масс исчезает, откуда следует, что

$$\vec{v}_2' = -\frac{m_1}{m_2} \vec{v}_1' \quad (\text{П2.11})$$

и тогда

$$T = \frac{M V_R^2}{2} + \frac{1}{2} \left( m_1 + \frac{m_1^2}{m_2} \right) v_1'^2 = \frac{M V_R^2}{2} + \frac{1}{2} m_1 \left( 1 + \frac{m_1}{m_2} \right) v_1'^2. \quad (\text{П2.12})$$

В то же время, согласно (П2.11), относительная скорость  $\vec{v} = \dot{\vec{r}} = \vec{v}_1 - \vec{v}_2 = \vec{v}'_1 - \vec{v}'_2$  равна

$$\vec{v} = \vec{v}'_1 - \left(-\frac{m_1}{m_2}\vec{v}'_1\right) = \left(1 + \frac{m_1}{m_2}\right)\vec{v}'_1. \quad (\text{П2.13})$$

В свою очередь это означает, что

$$\vec{v}'_1 = \left(1 + \frac{m_1}{m_2}\right)^{-1} \vec{v}. \quad (\text{П2.14})$$

Подставляя последнее выражение в (П2.12), получим

$$T = \frac{MV_R^2}{2} + \frac{1}{2} \frac{m_1}{1 + \frac{m_1}{m_2}} v^2 = \frac{MV_R^2}{2} + \frac{\mu v^2}{2}. \quad \text{Ч. т. д.} \quad (\text{П2.15})$$

### П3. Аналогия между дифференциалами и вариациями

$dy$

$\delta J$

(Первый) дифференциал функции  $y = f(x)$ .

Если  $\Delta x$  ( $= dx$ ) есть произвольное (малое) изменение независимой переменной  $x$ , тогда

$$y(x + \Delta x) - y(x) = dy + \mathcal{O}(\Delta x^2).$$

Можно написать

$$dy = \left. \frac{\partial}{\partial \alpha} f(x + \alpha \Delta x) \right|_{\alpha=0}$$

так как правая часть (как производная сложной функции  $\alpha$ )

$$\begin{aligned} &= f'(x + \alpha \Delta x) \Big|_{\alpha=0} \times \Delta x \\ &= f'(x) \Delta x = dy \end{aligned}$$

Вариация (первого порядка) функционала  $J = J(\Psi)$ .

Если  $\delta \Psi$  является произвольной (малой) вариацией (изменением в классе допустимых функций)  $\Psi$ , например,  $\delta \Psi = \eta \Phi$ ;  $\eta \rightarrow 0$  и  $\Phi$  есть ограниченная функция, удовлетворяющая граничным условиям задачи, тогда

$$J(\Psi + \delta \Psi) - J(\Psi) = \delta J + \mathcal{O}(\delta \Psi^2).$$

Можно принять как определение

$$\delta J = \left. \frac{\partial}{\partial \alpha} [J(\Psi + \alpha \delta \Psi)] \right|_{\alpha=0}$$

Эта аналогия указывает на то, что правила дифференцирования произведений, дробей и т. д. применимы также и для вычисления вариаций.

### П4. Теорема Эйлера об однородных функциях

Функция  $f = f(x_1, x_2, \dots, x_n)$  является однородной функцией степени  $k$  по переменным  $x_1, x_2, \dots, x_n$ , если

$$f = f(\alpha x_1, \alpha x_2, \dots, \alpha x_n) = \alpha^k f(x_1, x_2, \dots, x_n). \quad (\text{П4.1})$$

Для однородных функций выполняется *теорема Эйлера*:

$$\sum_{i=1}^n x_i \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} = k f(x_1, x_2, \dots, x_n) \quad (\text{П4.2})$$

*Доказательство.*

Определим функцию  $u = u(t)$  переменной  $t$  как

$$u(t) = f(tx_1^0, tx_2^0, \dots, tx_n^0) \quad (\text{П4.3})$$

где точка  $M(x_1^0, x_2^0, \dots, x_n^0)$  — произвольная внутренняя точка области определения функции  $f$ ; это означает, что все переменные  $x_i$  в  $f$  рассматриваются как функции переменной  $t$ :  $x_i = tx_i^0$ .

Дифференцируем сложную функцию:

$$\left. \frac{du}{dt} \right|_{t=1} = \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \right|_{t=1}, \quad (\text{П4.4})$$

однако

$$\left. \frac{\partial f}{\partial x_i} \right|_{t=1} = \frac{\partial f}{\partial x_i^0}; \quad \left. \frac{dx_i}{dt} \right|_{t=1} = \frac{d}{dt}(tx_i^0) \Big|_{t=1} = x_i^0. \quad (\text{П4.5})$$

Поэтому

$$\left. \frac{du}{dt} \right|_{t=1} = \sum_{i=1}^n x_i^0 \frac{\partial f}{\partial x_i^0}. \quad (\text{П4.6})$$

С другой стороны, так как  $f$  — однородная функция:

$$u(t) = f(tx_1^0, tx_2^0, \dots, tx_n^0) = t^k f(x_1^0, x_2^0, \dots, x_n^0). \quad (\text{П4.7})$$

Продифференцируем это выражение:

$$\left. \frac{du}{dt} \right|_{t=1} = k t^{k-1} f(x_1^0, x_2^0, \dots, x_n^0) \Big|_{t=1} = k f(x_1^0, x_2^0, \dots, x_n^0). \quad (\text{П4.8})$$

Сравним с предыдущей формулой:

$$\sum_{i=1}^n x_i^0 \frac{\partial f}{\partial x_i^0} = k f(x_1^0, x_2^0, \dots, x_n^0). \quad (\text{П4.9})$$

Так как  $M(x_1^0, x_2^0, \dots, x_n^0)$  является произвольной внутренней точкой области определения  $f$ , верхний индекс «0» можно опустить. Ч. т. д.

## П5. Теорема вириала в классической механике

### П5.1. Среднее полной производной физической величины по времени при финитном движении

Если система совершает *финитное движение* (т. е. все координаты и импульсы остаются ограниченными), то среднее по времени (обозначаемое горизонтальной чертой) любой физической величины, которую можно представить в виде *полной производной* по времени от какой-то другой величины, равно нулю, т. е. если  $f(t) = \frac{d}{dt}F$  и  $F(t)$  ограничено, то  $\bar{f} = 0$ :

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{dF}{dt} = \lim_{T \rightarrow \infty} \frac{F(T) - F(0)}{T} = 0. \quad (\text{П5.1})$$

## П5.2. Теорема вириала

Рассмотрим среднее по времени величины  $\frac{d}{dt} \sum_i \vec{r}_i \vec{p}_i$ , где суммирование ведется по всем частицам системы,  $\vec{r}_i$  есть радиус-вектор  $i$ -й частицы, и  $\vec{p}_i$  есть ее импульс. Если все  $\vec{r}_i$ ,  $\vec{p}_i$  остаются ограниченными, то

$$\overline{\frac{d}{dt} \sum_i \vec{r}_i \vec{p}_i} = 0. \quad (\text{П5.2})$$

С другой стороны,

$$\frac{d}{dt} \sum_i \vec{r}_i \vec{p}_i = \sum_i \vec{v}_i \vec{p}_i + \sum_i \vec{r}_i \dot{\vec{p}}_i. \quad (\text{П5.3})$$

Так как  $\vec{v}_i \vec{p}_i = m_i v_i^2$ , первая сумма равна  $2T$ , где  $T$  есть кинетическая энергия системы. Поэтому имеем

$$2\overline{T} = -\overline{\sum_i \vec{r}_i \dot{\vec{p}}_i}. \quad (\text{П5.4})$$

Последнюю сумму можно преобразовать с помощью уравнения Ньютона  $m_i \ddot{\vec{r}}_i = \vec{F}_i$ , записанного в виде  $\dot{\vec{p}}_i = \vec{F}_i$ :

$$2\overline{T} = -\overline{\sum_i \vec{r}_i \vec{F}_i}. \quad (\text{П5.5})$$

Величина  $\sum_i \vec{r}_i \vec{F}_i$  в правой части называется «вириалом» системы. Каждое слагаемое этой суммы зависит от начала координат; однако, если система замкнутая (все силы внутренние), полная сумма инвариантна. В самом деле, если перенести начало координат  $O$  в точку  $O'$ , имеющую радиус-вектор  $\vec{r}_0$ , то радиус-вектор в новой (штрихованной) системе будет  $\vec{r}'_i = \vec{r}_i - \vec{r}_0$ . Вириал в новой системе будет

$$\sum_i \vec{r}'_i \vec{F}_i = \sum_i (\vec{r}_i - \vec{r}_0) \vec{F}_i = \sum_i \vec{r}_i \vec{F}_i - \vec{r}_0 \sum_i \vec{F}_i = \sum_i \vec{r}_i \vec{F}_i, \quad (\text{П5.6})$$

так как  $\sum_i \vec{F}_i = 0$  для суммы внутренних сил. Это доказывает инвариантность вириала  $\sum_i \vec{r}_i \vec{F}_i$  по отношению к смещениям системы как целого.

Для консервативной системы  $\vec{F}_i = -\nabla_i U \equiv -\left(\vec{i} \frac{\partial U}{\partial x} + \vec{j} \frac{\partial U}{\partial y} + \vec{k} \frac{\partial U}{\partial z}\right) \equiv -\frac{\partial U}{\partial \vec{r}_i}$ , поэтому из (П5.5) имеем

$$2\overline{T} = \overline{\sum_i \vec{r}_i \frac{\partial U}{\partial \vec{r}_i}}. \quad (\text{П5.7})$$

Если потенциал  $U$  является однородной функцией степени  $k$  от радиус-векторов  $\vec{r}_i$ , то правая часть, в силу теоремы Эйлера о производных однородной функции, равна  $k\overline{U}$ , поэтому

$$2\overline{T} = k\overline{U}. \quad (\text{П5.8})$$

### Частные случаи:

Гармонический осциллятор:  $k = 2$ ,  $\overline{T} = \overline{U}$ ,  $E = 2\overline{U}$

Движение частиц под действием кулоновских сил:  $k = -1$ ,  $\overline{T} = -\frac{1}{2}\overline{U}$ ,  $E = \frac{1}{2}\overline{U} < 0$  (связанное движение с отрицательной полной энергией)

## П6. Электронное уравнение Шрёдингера в атомных единицах



В квантовой химии мы имеем дело преимущественно с электронным уравнением Шрёдингера. Для упрощения обычно опускают индексы « $e$ » и не указывают явно аргументы функций:

$$\hat{H}\Psi = E\Psi \quad (\text{П6.1})$$

Удобно использовать «атомную систему единиц», в которой, по определению,  $\hbar = 1$ ,  $e_0 = 1$ , и  $m_e = 1$ . Тогда единица длины есть боровский радиус («радиус» атома Н), обычно называемый «атомной единицей длины» (или просто «а. е.»)  $\cong 0.52917726 \text{ \AA}$ ; скорость света  $c \approx 137$  а. е., и атомная единица энергии («Хартри»; а.е.;  $E_h$ )  $\cong 27.2113957 \text{ эВ} \cong 627.509541 \text{ ккал/моль} \cong 2625.49992 \text{ кДж/моль} \cong 219474.625 \text{ см}^{-1} \cong 315773.213 \text{ К}$  (Кельвин — используется, как единица энергии, равная тепловой энергии  $kT$ , соответствующей температуре 1 К). В таких единицах энергия атома Н при фиксированном ядре равна *точно*  $-\frac{1}{2}$  а. е.

Если используются атомные единицы, то электронный гамильтониан молекулы принимает особенно простой вид:

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^{N_e} \Delta_i - \sum_{\alpha=1}^{N_N} \sum_{i=1}^{N_e} \frac{Z_{\alpha}}{r_{\alpha i}} + \sum_{i < j} \frac{1}{r_{ij}} + \sum_{\alpha < \beta} \frac{Z_{\alpha} Z_{\beta}}{R_{\alpha \beta}}. \quad (\text{П6.2})$$

Последнее слагаемое описывает взаимное отталкивание ядер. Оно не зависит от электронной волновой функции и представляет собой аддитивный вклад в энергию. Поэтому часто его не включают в электронный гамильтониан, и лишь впоследствии добавляют к энергии системы электронов (Обычно называют «электронной энергией» энергию, полученную без учета межъядерного отталкивания, и говорят о «полной энергии», если оно было учтено.)



## П7. «Бра-кет» формализм

### П7.1. «Бра»- и «кет»-векторы Дирака

Будем рассматривать волновые функции  $\varphi, \psi$  и т. д. как «кет-векторы» (векторы состояния)  $|\varphi\rangle, |\psi\rangle$  и т. д., и введем для каждого из них соответствующий сопряженный «бра-вектор»  $\langle\varphi|, \langle\psi|$  и т. д. Если за бра-вектором  $\langle\varphi|$  следует кет-вектор  $|\psi\rangle$ , то будем считать, что это автоматически образует скалярное произведение (интеграл перекрывания)  $\langle\varphi|\psi\rangle$ . (Названия «бра» и «кет» происходят от английского слова «bracket» (скобка) и были введены Дираком.) Эта система обозначений исключительно гибка и удобна, так как позволяет использовать очень компактные обозначения для проекционных операторов и т. д., и облегчает практическое обращение с ними.

Например, проекция некоторой волновой функции  $\psi(x)$  на нормированную волновую функцию  $\varphi(x)$  по определению задана интегральным оператором

$$\hat{P}_{\varphi}\psi(x) = \varphi(x) \int \varphi^*(x') \psi(x') dx' \quad (\text{П7.1})$$

$$= \int P_{\varphi}(x, x') \psi(x') dx' \quad (\text{П7.2})$$

с ядром  $P_\varphi(x, x') = \varphi(x)\varphi^*(x')$ , где  $x, x'$  обозначают все переменные, от которых зависят  $\varphi$  и  $\psi$ , и интегрирование определяется соответственно. Однако в общепринятых обозначениях скалярного произведения

$$\int \varphi^*(x')\psi(x')dx' = \langle \varphi | \psi \rangle, \quad (\text{П7.3})$$

поэтому, вводя кет-символ  $|\varphi\rangle$  для волновой функции  $\varphi$ , можно написать

$$\hat{P}_\varphi|\psi\rangle = |\varphi\rangle\langle\varphi|\psi\rangle. \quad (\text{П7.4})$$

Это означает, что для проекционного оператора имеем простую запись:

$$\hat{P}_\varphi = |\varphi\rangle\langle\varphi|. \quad (\text{П7.5})$$

Таким образом, получается очень гибкий формализм, в котором нужно лишь соблюдать некоторые простые формальные правила.

1. Операторы действуют на кет-векторы слева, на бра-векторы справа, например

$$\begin{aligned} |\psi_1\rangle &= \hat{A}|\varphi_1\rangle \\ \langle\psi_2| &= \langle\varphi_2|\hat{B} \end{aligned} \quad (\text{П7.6})$$

( $\hat{A}$  и  $\hat{B}$  являются какими-то линейными операторами.)

2. «Между двумя угловыми скобками» имеем скаляр, а «между двумя вертикальными линиями» — оператор:

$$\begin{aligned} \langle u | \dots | v \rangle &= \text{скаляр} \\ | u \rangle \dots \langle v | &= \text{оператор}. \end{aligned} \quad (\text{П7.7})$$

В первом выражении точки обозначают операторы или скаляры, а во втором случае часть, обозначаемая точками, может, разумеется, содержать выражения с операторами, но в целом она должна являться скаляром.

Аналогично, выражения, «открытые слева» являются кет-векторами, а «открытые справа» — бра-векторами:

$$\begin{aligned} \dots | u \rangle &= \text{кет} \\ \langle v | \dots &= \text{бра}. \end{aligned} \quad (\text{П7.8})$$

3. Можно менять местами бра и кет формы и получать эквивалентные выражения, производя сопряжение:

$$\begin{aligned} \text{векторы } |\varphi\rangle &\leftrightarrow \langle\varphi| \quad (\text{кет} \leftrightarrow \text{бра}) \\ \text{скаляры (числа) } a &\leftrightarrow a^* \quad (\text{комплексное сопряжение}) \\ \text{операторы } A &\leftrightarrow A^\dagger \quad (\text{эрмитово сопряжение}). \end{aligned} \quad (\text{П7.9})$$

(Как частный случай скаляра, скалярное произведение преобразуется, как  $\langle u | v \rangle \leftrightarrow \langle v | u \rangle = \langle u | v \rangle^*$ .)

Например, уравнение на собственные значения  $\hat{A}|\varphi\rangle = \lambda|\varphi\rangle$  полностью эквивалентно уравнению  $\langle\varphi|\hat{A}^\dagger = \lambda^*\langle\varphi|$ ; если оператор  $\hat{A}$  эрмитов, то последнее уравнение сводится к  $\langle\varphi|\hat{A} = \lambda\langle\varphi|$ .

Важно заметить, что, в соответствии с правилом  $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$ , имеем

$$(|u\rangle\langle v|)^\dagger = |v\rangle\langle u| \quad (\text{П7.10})$$

и вообще, при составлении сопряженных форм более сложных выражений, необходимо идти «задом наперед». Например, предполагая, что  $\hat{A}$  эрмитов, а  $\hat{B}$  — нет, имеем следующее преобразование:

$$(\hat{B}|u\rangle\langle\varphi|\hat{A}|\psi\rangle\langle a|b\rangle\langle v|)^\dagger = |v\rangle\langle b|a\rangle\langle\psi|\hat{A}|\varphi\rangle\langle u|\hat{B}^\dagger. \quad (\text{П7.11})$$



Можно оценить гибкость этого формализма, заметив, что то же самое выражение (П7.11) можно также записать в виде  $Q|u\rangle\langle v|B^\dagger$ , где  $Q = \langle b|a\rangle\langle\psi|\hat{A}|\varphi\rangle$  является скаляром, или в виде  $P\hat{X}\hat{Y}\hat{B}^\dagger$ , где  $P = \langle\psi|\hat{A}|\varphi\rangle$  — скаляр, а  $\hat{X} = |v\rangle\langle b|$  и  $\hat{Y} = |a\rangle\langle u|$  — операторы, и т. д.

## П7.2. Аналогия с матричным формализмом

Правила использования бра и кет подобны правилам матричного исчисления, которым мы пользуемся в случае, когда разные функции и операторы выражены в полном ортонормированном базисе.

Переписать различные выражения из одного формализма в другой очень просто. Кет-вектор  $|\varphi\rangle$  соответствует вектор-столбцу  $\mathbf{c}$ , образованному из коэффициентов разложения  $|\varphi\rangle$  по данному ортонормированному базису  $\{|\chi_i\rangle\}$ :  $|\varphi\rangle = \sum_i c_i |\chi_i\rangle$ .

Бра-вектор  $\langle\varphi|$  соответствует вектор-строке  $\mathbf{c}^\dagger$ , сопряженной к  $\mathbf{c}$ . Скалярное произведение  $\langle\varphi|\psi\rangle$  соответствует скалярному произведению  $\mathbf{c}^\dagger \mathbf{d}$  (предполагается, что  $\mathbf{c}$  и  $\mathbf{d}$  содержат коэффициенты разложения  $|\varphi\rangle$  и  $|\psi\rangle$ , соответственно); операторы соотносятся с матрицами, представляющими их в данном базисе, например, если  $|\varphi\rangle \sim \mathbf{c}$ ,  $|\psi\rangle \sim \mathbf{d}$  (здесь соответствие обозначим как « $\sim$ »), то

$$|\varphi\rangle\langle\psi| \sim \mathbf{c}\mathbf{d}^\dagger = \begin{pmatrix} \vdots \\ c \\ \vdots \end{pmatrix} (\dots d^\dagger \dots) \quad (\text{П7.12})$$

есть матрица (диадное произведение вектора-строки и вектора-столбца). В то же время

$$\langle\psi|\varphi\rangle = \mathbf{d}^\dagger \mathbf{c} = (\dots d^\dagger \dots) \begin{pmatrix} \vdots \\ c \\ \vdots \end{pmatrix}, \quad (\text{П7.13})$$

т. е.  $\langle\psi|\varphi\rangle$  и  $\mathbf{d}^\dagger \mathbf{c}$  являются скалярами, которые равны друг другу, если  $\langle\psi| \sim \mathbf{d}^\dagger$  и  $|\varphi\rangle \sim \mathbf{c}$ .

Можно получить матричный эквивалент любого выражения, записанного в формализме бра-кет, вводя явно разложение по выбранному базису и собирая соответствующие коэффициенты, например

$$\begin{aligned} |\varphi\rangle &= \sum_i c_i |\chi_i\rangle = \sum_i c_i |\chi_i\rangle \sim \mathbf{c} \\ \langle\psi| &= \langle\sum_i d_i \chi_i| = \sum_i d_i^* \langle\chi_i| \sim \mathbf{d}^\dagger \end{aligned} \quad (\text{П7.14})$$

$$\langle\psi|\varphi\rangle = \langle\sum_i d_i \chi_i| \sum_j c_j \chi_j\rangle = \sum_{i,j} d_i^* \underbrace{\langle\chi_i|\chi_j\rangle}_{\delta_{ij}} c_j = \sum_i d_i^* c_i = \mathbf{d}^\dagger \mathbf{c}$$

$$|\varphi\rangle\langle\psi| = |\sum_i c_i \chi_i\rangle \langle\sum_j d_j \chi_j| = \sum_{i,j} c_i d_j^* |\chi_i\rangle\langle\chi_j| \sim \mathbf{c}\mathbf{d}^\dagger.$$

Как замечено раньше, проекционный оператор на состояние  $|\varphi\rangle$  есть  $P_\varphi = |\varphi\rangle\langle\varphi|$ ; он связан с векторами коэффициентов разложения следующим образом:

$$\hat{P}_\varphi = |\varphi\rangle\langle\varphi| = |\sum_i c_i \chi_i\rangle \langle\sum_j c_j \chi_j| = \sum_{i,j} c_i c_j^* |\chi_i\rangle\langle\chi_j| \sim \mathbf{c}\mathbf{c}^\dagger = \mathbf{P}_\varphi. \quad (\text{П7.15})$$

В самом деле, проекция произвольной функции, описываемой кет-вектором  $|\psi\rangle = \sum_j d_j |\chi_j\rangle$ , на одномерное подпространство  $|\varphi\rangle \sim \mathbf{c}$ , согласно (П7.5), равна

$$\begin{aligned}\hat{P}_\varphi |\psi\rangle &= |\varphi\rangle \langle \varphi | \psi \rangle = \left| \sum_i c_i \chi_i \right\rangle \left\langle \sum_j c_j \chi_j \right| \sum_k d_k \chi_k \rangle \\ &= \sum_{i,j,k} c_i c_j^* d_k |\chi_i\rangle \langle \chi_j | \chi_k \rangle = \sum_{i,j,k} c_i c_j^* d_k \delta_{jk} |\chi_i\rangle \\ &= \sum_{i,j} c_i c_j^* d_j |\chi_i\rangle = \sum_i (\mathbf{c} \mathbf{c}^\dagger \mathbf{d})_i |\chi_i\rangle,\end{aligned}\quad (\text{П7.16})$$

в согласии с известным фактом, что диадное произведение  $\mathbf{c} \mathbf{c}^\dagger$  равно матрице проекции на подпространство, натянутое на вектор  $\mathbf{c} \sim |\varphi\rangle$ . Также и матричные элементы  $\langle \chi_k | \hat{P}_\varphi | \chi_l \rangle$  оператора  $\hat{P}_\varphi$  в ортонормированном базисе  $\{|\chi_k\rangle\}$  равны матричным элементам этого диадного произведения:

$$\begin{aligned}(\mathbf{P}_\varphi)_{kl} &= \langle \chi_k | \hat{P}_\varphi | \chi_l \rangle = \langle \chi_k | \varphi \rangle \langle \varphi | \chi_l \rangle \\ &= \langle \chi_k | \sum_i c_i \chi_i \rangle \left\langle \sum_j c_j \chi_j \right| \chi_l \rangle = \sum_{i,j} c_i c_j^* \delta_{ki} \delta_{jl} = c_k c_l^* = (\mathbf{c} \mathbf{c}^\dagger)_{kl}.\end{aligned}\quad (\text{П7.17})$$

### П7.3. Неортогональный базис

Можно обобщить правила соответствия между бра-кет формализмом и матричным исчислением на случай неортонормированного базиса  $\{|\chi_i\rangle\}$ , но тогда выражения становятся более сложными, потому что появляется нетривиальная матрица перекрывания  $\mathbf{S}$  с элементами  $S_{ij} = \langle \chi_i | \chi_j \rangle$ . Например:

$$\begin{aligned}\langle \varphi | \psi \rangle &= \left\langle \sum_i c_i \chi_i \right| \sum_j d_j \chi_j \rangle = \sum_{i,j} c_i^* \langle \chi_i | \chi_j \rangle d_j \\ &= \sum_{i,j} c_i^* S_{ij} d_j = \mathbf{c}^\dagger \mathbf{S} \mathbf{d}.\end{aligned}\quad (\text{П7.18})$$

В случае неортогональных базисов надо различать матрицу линейного оператора  $\hat{A}$  с элементами

$$A_{ij} = \langle \chi_i | \hat{A} | \chi_j \rangle \quad (\text{П7.19})$$

и матрицу, описывающую линейное преобразование, соответствующее этому оператору.

Пусть  $|\varphi\rangle \sim \mathbf{c}$  и  $|\psi\rangle \sim \mathbf{d}$  и предположим, что  $|\varphi\rangle = \hat{A}|\psi\rangle$ . Тогда имеем

$$|\varphi\rangle = \left| \sum_i c_i \chi_i \right\rangle = \hat{A} \left| \sum_j d_j \chi_j \right\rangle. \quad (\text{П7.20})$$

Умножая выражение (П7.20) на  $\langle \chi_l |$  ( $l = 1, 2, \dots$ ), получим

$$\langle \chi_l | \sum_i c_i \chi_i \rangle = \langle \chi_l | \hat{A} \left| \sum_j d_j \chi_j \right\rangle, \quad (\text{П7.21})$$

т. е.

$$\sum_i S_{li} c_i = \sum_j A_{lj} d_j \quad (\text{П7.22})$$

или в матричных обозначениях

$$\mathbf{S}\mathbf{c} = \mathbf{A}\mathbf{d} . \quad (\text{П7.23})$$

Умножая выражение (П7.23) слева на матрицу  $\mathbf{S}^{-1}$ , обратную к матрице перекрывания, получим

$$\mathbf{c} = \mathbf{S}^{-1}\mathbf{A}\mathbf{d} . \quad (\text{П7.24})$$

Этот результат показывает, что в случае неортогонального базиса ( $\mathbf{S} \neq \mathbf{1}$ ), матрица линейного преобразования, описывающая действие линейного оператора  $\hat{A}$ , равна  $\mathbf{S}^{-1}\mathbf{A}$ , где элементы матрицы  $\mathbf{A}$  определяются по формуле (П7.19). Для проектора  $\hat{P}_\varphi$  этот результат означает, что матрица оператора  $\hat{P}_\varphi$ , как легко видеть, равна  $\mathbf{S}\mathbf{c}\mathbf{c}^\dagger\mathbf{S}$ , тогда как матрица проекционного оператора равна  $\mathbf{c}\mathbf{c}^\dagger\mathbf{S}$ . (Именно эта последняя матрица обладает характерными свойствами идемпотентности и т. д., обсуждаемыми ниже для проекционных операторов.)

## П7.4. Пример использования бра-кет формализма: гипервириальная теорема

Рассмотрим матричный элемент  $\langle \Psi_i | \hat{H}\hat{Q} - \hat{Q}\hat{H} | \Psi_j \rangle$  коммутатора гамильтониана  $\hat{H}$  с произвольным оператором  $\hat{Q}$ , вычисленный для точных решений  $\Psi_i$  и  $\Psi_j$  не зависящего от времени уравнения Шрёдингера. Уравнение Шрёдингера для точного состояния (кет-вектора)  $|\Psi_j\rangle$  имеет вид

$$\hat{H}|\Psi_j\rangle = E_j|\Psi_j\rangle, \quad (\text{П7.25})$$

а сопряженное уравнение Шрёдингера для точного бра-вектора  $\langle \Psi_i |$  имеет вид

$$\langle \Psi_i | \hat{H} = E_i \langle \Psi_i | . \quad (\text{П7.26})$$

Получаем с помощью этих соотношений (стрелки указывают направления, в которых действует гамильтониан)

$$\begin{aligned} \langle \Psi_i | \hat{H}\hat{Q} - \hat{Q}\hat{H} | \Psi_j \rangle &= \langle \Psi_i | \hat{H} \hat{Q} | \Psi_j \rangle - \langle \Psi_i | \hat{Q} \hat{H} | \Psi_j \rangle \\ &= E_i \langle \Psi_i | \hat{Q} | \Psi_j \rangle - E_j \langle \Psi_i | \hat{Q} | \Psi_j \rangle = (E_i - E_j) Q_{ij} . \end{aligned} \quad (\text{П7.27})$$

Это так называемая («недиагональная», если  $i \neq j$ ) «гипервириальная теорема»; она используется в теории интенсивности спектральных линий.

В частном случае  $i = j$  ( $\Psi_i = \Psi_j = \Psi$ ) имеем «диагональную» гипервириальную теорему, согласно которой для произвольного оператора  $\hat{Q}$  и *точной* волновой функции  $\Psi$  для коммутатора этого оператора с гамильтонианом выполняется равенство

$$\langle \Psi | \hat{H}\hat{Q} - \hat{Q}\hat{H} | \Psi \rangle = 0 . \quad (\text{П7.28})$$

Вывод теоремы вириала в гл. 2, разд. 3.2 можно было бы непосредственно проделать на основе этого результата.

## П7.5. Проекционные операторы

Как обсуждалось выше, оператор проекции на одномерное подпространство, натянутое на вектор  $|\varphi\rangle$ , равен  $\hat{P}_\varphi = |\varphi\rangle\langle\varphi|$ . Аналогично, если имеется  $n$ -мерное

подпространство ортонормированных векторов  $\{|\varphi_i\rangle\}$  ( $i = 1, 2, \dots, n$ ), то соответствующий проекционный оператор на это подпространство равен сумме

$$\hat{P} = \sum_{i=1}^n |\varphi_i\rangle\langle\varphi_i|. \quad (\text{П7.29})$$

Легко проверить два основных свойства проекционных операторов.

1. Проекционные операторы эрмитовы — это следует из определения (П7.29) и правил обращения с бра- и кет-векторами, обсуждавшихся выше.

2. Проекционные операторы *идемпотентны*,  $\hat{P}^2 = \hat{P}$ . В самом деле:

$$\begin{aligned} \hat{P}^2 &= \sum_{i=1}^n |\varphi_i\rangle\langle\varphi_i| \sum_{j=1}^n |\varphi_j\rangle\langle\varphi_j| \\ &= \sum_{i,j=1}^n |\varphi_i\rangle \underbrace{\langle\varphi_i|\varphi_j\rangle}_{\delta_{ij}} \langle\varphi_j| = \sum_{i=1}^n |\varphi_i\rangle\langle\varphi_i| = \hat{P}. \end{aligned} \quad (\text{П7.30})$$

Легко видеть, что любой вектор, который можно представить в виде линейной комбинации векторов  $|\varphi_i\rangle$ , растягивающих рассматриваемое подпространство, остается неизменным при действии на него проекционного оператора  $\hat{P}$ , в то время как любой вектор, ортогональный к этому подпространству, полностью уничтожается. Другими словами, любой вектор, полностью лежащий в подпространстве  $\{|\varphi_i\rangle\}$ , является собственным вектором оператора  $\hat{P}$  с собственным значением 1, а любой вектор, ортогональный к этому подпространству, является собственным вектором с собственным значением 0.

Проекционный оператор  $\hat{P}$  остается инвариантным, если заменить векторы  $\{|\varphi_i\rangle\}$  любым другим набором ортонормированных векторов  $\{|\psi_i\rangle\}$ , которые полностью лежат в том же самом подпространстве. Так как оба набора векторов  $\{|\varphi_i\rangle\}$  и  $\{|\psi_i\rangle\}$  ортонормированы, они должны быть связаны некоторым *унитарным* преобразованием  $\mathbf{U}$ :

$$|\psi_j\rangle = \sum_{k=1}^n U_{kj} |\varphi_k\rangle. \quad (\text{П7.31})$$

Тогда проекционный оператор  $\hat{P}'$ , построенный из  $n$  функций  $|\psi_j\rangle$ , имеет вид

$$\begin{aligned} \hat{P}' &= \sum_{j=1}^n |\psi_j\rangle\langle\psi_j| = \sum_{j=1}^n \left| \sum_{k=1}^n U_{kj} \varphi_k \right\rangle \left\langle \sum_{l=1}^n U_{lj} \varphi_l \right| \\ &= \sum_{j,k,l=1}^n U_{kj} U_{lj}^* |\varphi_k\rangle\langle\varphi_l| = \sum_{k,l=1}^n \underbrace{\left( \sum_{j=1}^n U_{kj} (U^\dagger)_{jl} \right)}_{\delta_{kl}} |\varphi_k\rangle\langle\varphi_l| \\ &= \sum_{k=1}^n |\varphi_k\rangle\langle\varphi_k| = \hat{P}. \end{aligned} \quad (\text{П7.32})$$

Часто необходимо рассматривать подпространство, определенное некоторым набором *неортогональных* векторов  $\{|\varphi_i\rangle\}$ . Тогда можно построить проекционный оператор на это подпространство, определяя набор ортонормированных векторов в том же подпространстве. Предположим, что матрица перекрытия векторов  $\{|\varphi_i\rangle\}$ , растягивающих подпространство, равна  $\sigma$ , т. е.

$$\langle\varphi_i|\varphi_j\rangle = \sigma_{ij}. \quad (\text{П7.33})$$

Проще всего построить ортонормированный набор векторов  $\{|\psi_k\rangle\}$ , растягивающих то же подпространство, что и  $\{|\varphi_i\rangle\}$ , с помощью симметричной ортогонализации по Лёвдину (гл. 3, разд. 2):

$$|\psi_k\rangle = \sum_{i=1}^n (\sigma^{-1/2})_{ik} |\varphi_i\rangle. \quad (\text{П7.34})$$

Тогда проекционный оператор на подпространство имеет вид:

$$\begin{aligned} \hat{P} &= \sum_{k=1}^n |\psi_k\rangle \langle \psi_k| = \sum_{k=1}^n \left| \sum_{i=1}^n (\sigma^{-1/2})_{ik} |\varphi_i\rangle \right\rangle \left\langle \sum_{j=1}^n (\sigma^{-1/2})_{jk} \varphi_j \right| \\ &= \sum_{i,j,k=1}^n |\varphi_i\rangle (\sigma^{-1/2})_{ik} (\sigma^{-1/2})_{kj} \langle \varphi_j| = \sum_{i,j=1}^n |\varphi_i\rangle (\sigma^{-1})_{ij} \langle \varphi_j|, \end{aligned} \quad (\text{П7.35})$$

где использована эрмитовость матрицы  $\sigma^{-1/2}$ .

## П7.6. Разложение единицы

Если векторы  $\{|\varphi_i\rangle\}$  образуют полный ортонормированный набор функций (базис) в данном (гильбертовом) пространстве конечной или бесконечной размерности, то каждую функцию  $|\psi\rangle$  можно представить («почти всюду») в виде

$$|\psi\rangle = \sum_i c_i |\varphi_i\rangle \quad (\text{П7.36})$$

со значениями коэффициентов (Фурье-амплитудами)

$$c_i = \langle \varphi_i | \psi \rangle \quad (\text{П7.37})$$

Подставляя коэффициенты (П7.37) в разложение (П7.36), имеем

$$\begin{aligned} |\psi\rangle &= \sum_i c_i |\varphi_i\rangle = \sum_i \langle \varphi_i | \psi \rangle |\varphi_i\rangle = \sum_i |\varphi_i\rangle \langle \varphi_i | \psi \rangle \\ &= \left( \sum_i |\varphi_i\rangle \langle \varphi_i| \right) |\psi\rangle. \end{aligned} \quad (\text{П7.38})$$

Это означает, что выражение в скобках представляет собой единичный оператор  $\hat{I}$ :

$$\hat{I} = \sum_i |\varphi_i\rangle \langle \varphi_i|, \quad (\text{П7.39})$$

записанный в форме, называемой «разложением единицы» по функциям полного базиса. (Так как условия (П7.38)-(П7.39) выражают полноту набора  $\{|\varphi_i\rangle\}$ , равенство (П7.38) называется также «условием полноты».)

Так как  $\hat{P} = |\varphi_i\rangle \langle \varphi_i|$  является оператором проекции на базисный вектор  $|\varphi_i\rangle$ , разложение единицы (П7.39) выражает тот факт, что каждый вектор можно представить в виде суммы его проекций на все базисные векторы. Другими словами, проекция каждого вектора на полное пространство равна ему самому, что находится в согласии с общим свойством, что любой вектор, полностью лежащий в рассматриваемом (под)пространстве, является собственным вектором с собственным значением 1 оператора проекции на это (под)пространство.

## П7.7. Спектральное разложение эрмитовых операторов

Рассмотрим эрмитов оператор  $\hat{A}$ . Как известно, его собственные векторы  $|\varphi_i\rangle$  образуют полный ортонормированный набор; используя их, можно, следовательно, построить разложение единицы:

$$\hat{I} = \sum_i |\varphi_i\rangle\langle\varphi_i|. \quad (\text{П7.40})$$

Теперь рассмотрим оператор  $\hat{A}$ , действующий на произвольную функцию  $|\psi\rangle$ , и вставим  $\hat{I}$  после  $\hat{A}$ :

$$\begin{aligned} \hat{A}|\psi\rangle &= \hat{A}\hat{I}|\psi\rangle = \hat{A}\left(\sum_i |\varphi_i\rangle\langle\varphi_i|\right)|\psi\rangle \\ &= \left(\sum_i a_i |\varphi_i\rangle\langle\varphi_i|\right)|\psi\rangle, \end{aligned} \quad (\text{П7.41})$$

где мы учли, что  $|\varphi_i\rangle$  является собственным вектором  $\hat{A}$ , т. е.  $\hat{A}|\varphi_i\rangle = a_i|\varphi_i\rangle$  с собственным значением  $a_i$ . Выражение (П7.41) означает, что любой эрмитов оператор можно представить в виде «спектрального разложения»

$$\hat{A} = \sum_i a_i |\varphi_i\rangle\langle\varphi_i|. \quad (\text{П7.42})$$

с помощью его собственных векторов и собственных значений.

## П7.8. Неэрмитовы операторы — биортогональные наборы функций

В большинстве случаев можно обобщить соотношение (П7.42) на неэрмитовы операторы, пользуясь биортогональностью их правых и левых собственных векторов. Исключениями являются лишь «патологические» случаи, в которых число собственных векторов рассматриваемого оператора меньше, чем размерность линейного пространства, в котором он определен. (В таких случаях матрицу оператора нельзя диагонализировать преобразованием подобия. Мы не будем рассматривать такие сингулярные случаи, представляющие скорее математический, чем физический интерес. Более того, такие случаи невозможно даже опознать в практической численной работе, так как уже бесконечно малое изменение матричных элементов устраняет это специальное свойство матрицы.)

Пусть  $\hat{A}$  является неэрмитовым оператором с правым и левым собственными векторами  $\{|\varphi_i\rangle\}$  и  $\{\langle\psi_j|\}$  соответственно

$$\begin{aligned} \hat{A}|\varphi_i\rangle &= a_i|\varphi_i\rangle \\ \langle\psi_j|\hat{A} &= b_j\langle\psi_j|. \end{aligned} \quad (\text{П7.43})$$

Теперь мы поступим так, как обычно делают при доказательстве ортогональности собственных векторов эрмитовых операторов и матриц. Умножим второе из уравнений (П7.43) справа на  $|\varphi_i\rangle$ :

$$\langle\psi_j|\hat{A}|\varphi_i\rangle = b_j\langle\psi_j|\varphi_i\rangle \quad (\text{П7.44})$$

и заметим, что  $\hat{A}|\varphi_i\rangle = a_i|\varphi_i\rangle$ , согласно первому из уравнений (П7.43). Получаем

$$a_i\langle\psi_j|\varphi_i\rangle = b_j\langle\psi_j|\varphi_i\rangle. \quad (\text{П7.45})$$

Уравнение (П7.45) показывает, что либо

а)  $a_i \neq b_j$  и тогда  $\langle \psi_j | \varphi_i \rangle = 0$ , или

б)  $a_i = b_j$  и тогда в общем случае,  $\langle \psi_j | \varphi_i \rangle \neq 0$ . В этом случае мы можем выбрать нормировку  $\langle \psi_j |$  и  $|\varphi_i \rangle$  так, чтобы обеспечить  $\langle \psi_j | \varphi_i \rangle = 1$ .

Из этих результатов следует, что левые и правые собственные значения  $\hat{A}$  попарно равны, а левые и правые собственные векторы образуют взаимно *биортогональные* наборы. (В случае вырожденных значений  $a_i = b_j$ , их можно выбрать таким же образом.) Обозначим знаком  $\sim$  над символом  $\varphi$  вектор, биортогональный к вектору  $|\varphi \rangle$ :

$$|\tilde{\varphi}_j \rangle = |\psi_j \rangle. \quad (\text{П7.46})$$

Он обладает свойством

$$\langle \tilde{\varphi}_j | \varphi_i \rangle = \delta_{ij}, \quad (\text{П7.47})$$

а также

$$\langle \tilde{\varphi}_j | \hat{A} = a_j \langle \tilde{\varphi}_j |. \quad (\text{П7.48})$$

Разложение единицы можно построить и с помощью биортогональных наборов функций:

$$\hat{I} = \sum_i |\varphi_i \rangle \langle \tilde{\varphi}_i|. \quad (\text{П7.49})$$

В самом деле, рассмотрим действие суммы в правой части выражения (П7.49) на произвольную функцию  $|\psi \rangle = \sum_j c_j |\varphi_j \rangle$ :

$$\sum_i |\varphi_i \rangle \langle \tilde{\varphi}_i | \sum_j c_j |\varphi_j \rangle = \sum_{i,j} c_j |\varphi_i \rangle \underbrace{\langle \tilde{\varphi}_i | \varphi_j \rangle}_{\delta_{ij}} = \sum_j c_j |\varphi_j \rangle, \quad (\text{П7.50})$$

что подтверждает разложение (П7.49). (Заметим, что форма  $\hat{I} = \sum_i |\tilde{\varphi}_i \rangle \langle \varphi_i |$  также верна.)

Теперь можно записать для произвольной  $|\psi \rangle$ :

$$\begin{aligned} \hat{A}|\psi \rangle &= \hat{A}\hat{I}|\psi \rangle = \hat{A} \sum_i |\varphi_i \rangle \langle \tilde{\varphi}_i | \psi \rangle \\ &= \sum_i \hat{A}|\varphi_i \rangle \langle \tilde{\varphi}_i | \psi \rangle = \left( \sum_i a_i |\varphi_i \rangle \langle \tilde{\varphi}_i | \right) |\psi \rangle, \end{aligned} \quad (\text{П7.51})$$

т. е. спектральное разложение неэрмитова оператора по его правым и левым (биортогональным) собственным векторам  $\langle \tilde{\varphi}_i |$  и  $|\varphi_i \rangle$  имеет вид

$$\hat{A} = \sum_i a_i |\varphi_i \rangle \langle \tilde{\varphi}_i|. \quad (\text{П7.52})$$

Для эрмитова оператора левые и правые собственные векторы сопряжены друг к другу (т. е.  $\langle \tilde{\varphi}_i | = \langle \varphi_i |$ ), и выражение (П7.52) сводится к (П7.42).

## П7.9. След проекционного оператора

След проекционного оператора  $\hat{P}$  равен размерности  $n$  подпространства, на которое он проектирует. Рассмотрим

$$\hat{P} = \sum_{i=1}^n |\varphi_i \rangle \langle \varphi_i|, \quad (\text{П7.53})$$

который определен в некотором  $m$ -мерном линейном пространстве ( $m$  может быть и бесконечным), заданном ортонормированным базисом  $\{|\psi_j\rangle\}$ , и поэтому векторы  $|\varphi_i\rangle$  можно разложить по этому базису:

$$|\varphi_i\rangle = \sum_{j=1}^m c_j^i |\psi_j\rangle. \quad (\text{П7.54})$$

След  $\hat{P}$  тогда имеет вид

$$\text{Tr}\{\hat{P}\} = \sum_{k=1}^m \langle \psi_k | \hat{P} | \psi_k \rangle = \sum_{k=1}^m \langle \psi_k | \sum_{i=1}^n |\varphi_i\rangle \langle \varphi_i | \psi_k \rangle. \quad (\text{П7.55})$$

Подставляя это выражение в разложение (П7.54), получим

$$\begin{aligned} \text{Tr}\{\hat{P}\} &= \sum_{k=1}^m \langle \psi_k | \sum_{i=1}^n \sum_{j=1}^m c_j^i \psi_j \rangle \langle \sum_{l=1}^m c_l^i \psi_l | \psi_k \rangle \\ &= \sum_{i=1}^n \sum_{j,k,l=1}^m c_j^i c_l^{i*} \delta_{kj} \delta_{lk} = \sum_{i=1}^n \sum_{k=1}^m c_k^i c_k^{i*} = \sum_{i=1}^n 1 = n, \end{aligned} \quad (\text{П7.56})$$

где мы воспользовались ортонормированностью базиса  $\{|\psi_k\rangle\}$  и тем фактом, что каждая из функций  $|\varphi_i\rangle$  нормирована на единицу.

Как легко заметить, такой же результат верен для следа матрицы  $\mathbf{P}$ , представляющей оператор  $\hat{P}$  в ортонормированном базисе  $\{|\psi_k\rangle\}$ .

## П8. Сводка формул теории возмущений Рэля—Шрёдингера (невыврожденный случай)

$\hat{H} = \hat{H}^0 + \hat{V}$  — возмущенный гамильтониан

$\hat{H}\Psi_i = E_i\Psi_i$  — возмущенное уравнение Шрёдингера

$\hat{H}\Psi_l^0 = E_l^0\Psi_l^0$  — невозмущенное уравнение Шрёдингера

Энергия  $i$ -го состояния

$$E_i = E_i^0 + \sum_{n=1}^{\infty} \varepsilon_i^{(n)},$$



где (до четвертого порядка включительно)

$$\begin{aligned}
 \varepsilon_i^{(1)} &= V_{ii} \\
 \varepsilon_i^{(2)} &= - \sum_{\substack{j \\ (j \neq i)}} \frac{|V_{ij}|^2}{E_j^0 - E_i^0} \\
 \varepsilon_i^{(3)} &= \sum_{\substack{j,k \\ (j,k \neq i)}} \frac{V_{ik} V_{kj} V_{ji}}{(E_k^0 - E_i^0)(E_j^0 - E_i^0)} - \sum_{\substack{k \\ (k \neq i)}} \frac{V_{ii} |V_{ik}|^2}{(E_k^0 - E_i^0)^2} \\
 \varepsilon_i^{(4)} &= - \sum_{\substack{j,k,l \\ (j,k,l \neq i)}} \frac{V_{ik} V_{kj} V_{jl} V_{li}}{(E_k^0 - E_i^0)(E_j^0 - E_i^0)(E_l^0 - E_i^0)} \\
 &\quad + \sum_{\substack{j,k \\ (j,k \neq i)}} \frac{|V_{ij}|^2 |V_{ik}|^2}{(E_j^0 - E_i^0)(E_k^0 - E_i^0)^2} - \sum_{\substack{k \\ (k \neq i)}} \frac{V_{ii}^2 |V_{ki}|^2}{(E_k^0 - E_i^0)^3} \\
 &\quad + \sum_{\substack{j,k \\ (j,k \neq i)}} \frac{V_{ii} V_{ik} V_{kj} V_{ji}}{(E_k^0 - E_i^0)(E_j^0 - E_i^0)} \left( \frac{1}{E_k^0 - E_i^0} + \frac{1}{E_j^0 - E_i^0} \right)
 \end{aligned}$$

(см. также формулы в гл. 4, разд. 1.4.2).

*Волновая функция  $i$ -го состояния*

$$\Psi_i = \Psi_i^0 + \sum_{n=1}^{\infty} \Phi_i^{(n)}; \quad \Phi_i^{(n)} = \sum_{\substack{k \\ (k \neq i)}} p_{ik}^{(n)} \Psi_k^0,$$

где (до третьего порядка включительно)

$$\begin{aligned}
 p_{ik}^{(1)} &= \frac{-V_{ki}}{E_k^0 - E_i^0} \\
 p_{ik}^{(2)} &= \sum_{\substack{j \\ (j \neq i)}} \frac{V_{kj} V_{ji}}{(E_k^0 - E_i^0)(E_j^0 - E_i^0)} - \frac{V_{ii} V_{ki}}{(E_k^0 - E_i^0)^2} \\
 p_{ik}^{(3)} &= - \sum_{\substack{j,l \\ (j,l \neq i)}} \frac{V_{kj} V_{jl} V_{li}}{(E_k^0 - E_i^0)(E_j^0 - E_i^0)(E_l^0 - E_i^0)} \\
 &\quad + \sum_{\substack{j \\ (j \neq i)}} \frac{|V_{ij}|^2 V_{ki}}{(E_j^0 - E_i^0)(E_k^0 - E_i^0)^2} - \frac{V_{ii}^2 V_{ki}}{(E_k^0 - E_i^0)^3} \\
 &\quad + \sum_{\substack{j \\ (j \neq i)}} \frac{V_{ii} V_{kj} V_{ji}}{(E_k^0 - E_i^0)(E_j^0 - E_i^0)} \left( \frac{1}{E_k^0 - E_i^0} + \frac{1}{E_j^0 - E_i^0} \right).
 \end{aligned}$$

## П9. Прямое произведение матриц

Пусть  $\mathbf{A}$  и  $\mathbf{B}$  являются матрицами размерностей  $m_A \times n_A$  и  $m_B \times n_B$ , с элементами  $A_{ij}$  и  $B_{ij}$ , соответственно. Тогда матрицу  $\mathbf{C}$  размерности  $m_A m_B \times n_A n_B$  с элементами

$$C_{IJ} = C_{(ij)(kl)} = A_{ik} B_{jl} \quad (\text{П9.1})$$

называют прямым произведением матриц **A** и **B** и обозначают как

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} . \quad (\text{П9.2})$$

В определении (П9.1) индекс  $I$  нумерует пары индексов  $(i, j)$  в последовательности  $(1,1), (1,2), \dots, (1, m_B), (2,1), (2,2), \dots, (m_A, m_B)$ , а  $J$  нумерует пары  $(k, l)$  в последовательности  $(1,1), (1,2), \dots, (1, n_B), (2,1), (2,2), \dots, (n_A, n_B)$ . В частном случае векторов (матриц с одним столбцом) прямое произведение

$$\mathbf{c} = \mathbf{a} \otimes \mathbf{b} \quad (\text{П9.3})$$

векторов **a** и **b** размерностей  $m_A$  и  $m_B$ , соответственно, есть вектор размерности  $m_{A \otimes B}$  с элементами

$$c_I = c_{(ij)} = a_i b_j , \quad (\text{П9.4})$$

где определение  $I$  совпадает с данным выше.

Можно легко проверить соотношения:

$$\begin{aligned} (\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) &= \mathbf{AC} \otimes \mathbf{BD} \\ (\mathbf{A} \otimes \mathbf{B})(\mathbf{c} \otimes \mathbf{d}) &= \mathbf{Ac} \otimes \mathbf{Bd} . \end{aligned} \quad (\text{П9.5})$$

(Эти формулы, разумеется, имеют смысл, только если число столбцов матрицы **A** равно числу строк матрицы **C** или вектора **c**, и такие же условия должны выполняться для матриц **B** и **D**, или для матрицы **B** и вектора **d**.)

Прямое произведение трех (или больше) матриц можно определить ассоциативно

$$\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) \text{ и т. д.} \quad (\text{П9.6})$$

При этом каждый матричный элемент нумеруется тройкой (четверкой и т. д.) индексов, соответствующих отдельным матрицам **A**, **B**, **C**.

Рассмотрим две линейные вариационные задачи, определенные в линейных пространствах функций  $\{\varphi_i^A\}$  и  $\{\varphi_j^B\}$  размерностей  $m_A$  и  $m_B$ , соответственно, и их вековые уравнения:

$$\begin{aligned} \mathbf{H}^A \mathbf{c}^A &= E^A \mathbf{c}^A \\ \mathbf{H}^B \mathbf{c}^B &= E^B \mathbf{c}^B . \end{aligned} \quad (\text{П9.7})$$

Можно ввести пространство прямого произведения размерности  $m_A m_B$  с базисным набором функций  $\varphi_I^{AB} = \varphi_{(ij)}^{AB} = \varphi_i^A \varphi_j^B$ . Если уравнения (П9.7) выполняются для двух наборов  $\{\varphi_i^A\}$ ,  $\{\varphi_j^B\}$ , взятых по отдельности, то в прямом произведении пространств также выполняется уравнение на собственные значения:

$$\mathbf{H}^{AB} \mathbf{c}^{AB} = (E^A + E^B) \mathbf{c}^{AB} , \quad (\text{П9.8})$$

где

$$\begin{aligned} \mathbf{H}^{AB} &= \mathbf{H}^A \otimes \mathbf{1}^B + \mathbf{1}^A \otimes \mathbf{H}^B ; \\ \mathbf{c}^{AB} &= \mathbf{c}^A \otimes \mathbf{c}^B , \end{aligned} \quad (\text{П9.9})$$

а  $\mathbf{1}^A$ ,  $\mathbf{1}^B$  являются единичными матрицами размерностей  $m_A \times m_A$  и  $m_B \times m_B$  соответственно. В самом деле

$$\begin{aligned} \mathbf{H}^{AB} \mathbf{c}^{AB} &= (\mathbf{H}^A \otimes \mathbf{1}^B)(\mathbf{c}^A \otimes \mathbf{c}^B) + (\mathbf{1}^A \otimes \mathbf{H}^B)(\mathbf{c}^A \otimes \mathbf{c}^B) \\ &= \mathbf{H}^A \mathbf{c}^A \otimes \mathbf{1}^B \mathbf{c}^B + \mathbf{1}^A \mathbf{c}^A \otimes \mathbf{H}^B \mathbf{c}^B = (E^A + E^B) \mathbf{c}^A \otimes \mathbf{c}^B . \end{aligned} \quad (\text{П9.10})$$

В случае неортогональных базисных функций  $\{\varphi_i^A\}$  и  $\{\varphi_j^B\}$  в пространствах  $A$  и  $B$  выполняются обобщенные уравнения на собственные значения

$$\begin{aligned}\mathbf{H}^A \mathbf{c}^A &= E^A \mathbf{S}^A \mathbf{c}^A; \\ \mathbf{H}^B \mathbf{c}^B &= E^B \mathbf{S}^B \mathbf{c}^B.\end{aligned}\quad (\text{П9.11})$$

Как легко видеть, матрица перекрывания (метрики) в прямом произведении пространств имеет вид

$$\mathbf{S}^{AB} = \mathbf{S}^A \otimes \mathbf{S}^B, \quad (\text{П9.12})$$

и обобщенное уравнение на собственные значения становится таким:

$$\mathbf{H}^{AB} \mathbf{c}^{AB} = (E^A + E^B) \mathbf{S}^{AB} \mathbf{c}^{AB}, \quad (\text{П9.13})$$

где

$$\mathbf{H}^{AB} = \mathbf{H}^A \otimes \mathbf{S}^B + \mathbf{S}^A \otimes \mathbf{H}^B. \quad (\text{П9.14})$$

#### Замечание

Можно отметить, что два квантово-механические взаимодействия строго аддитивны, если они соответствуют случаям, описываемым матрицами гамильтониана (П9.9) или (П9.14). Это выполняется для объединенных систем, состоящих из невзаимодействующих подсистем. В других случаях обычно наблюдаются некоторые отклонения от строгой аддитивности<sup>1</sup>. Иногда эффекты неаддитивности можно описать, как возмущение.

## П10. Перестановки

Перестановка  $\hat{P}$  есть операция, изменяющая порядок расположения некоторых объектов; если их исходная последовательность до перестановки задавалась номерами  $1, 2, \dots, N$ , то после перестановки порядок номеров изменится на  $P_1, P_2, \dots, P_N$ . (Набор  $\{P_1, P_2, \dots, P_N\}$  тот же самый, что и  $\{1, 2, \dots, N\}$ , однако элементы расположены в другом порядке.) Каждая перестановка  $\hat{P}$  определяется последовательностью  $P_1, P_2, \dots, P_N$ . Это означает, что, применяя  $\hat{P}$ , необходимо ставить на первое место элемент, который был на  $P_1$ -м месте, на второе место — тот, который был на  $P_2$ -м, и т. д.; вообще,  $k$ -я позиция будет занята элементом, который исходно был  $P_k$ -м. (Часто данное расположение  $P_1, P_2, \dots, P_N$  элементов  $1, 2, \dots, N$  также называется перестановкой.)

Хорошо известно, что число различных перестановок  $N$  объектов (включая «тождественную перестановку»  $\hat{I}$ , не изменяющую порядок расположения объектов) есть  $N! = 1 \cdot 2 \cdot 3 \dots (N-1) \cdot N$ . Если мы рассмотрим две перестановки  $\hat{P}$  и  $\hat{Q}$ , где  $\hat{P}$  характеризуется последовательностью  $\hat{P}_1, \hat{P}_2, \dots, \hat{P}_N$ , а  $\hat{Q}$  — последовательностью  $\hat{Q}_1, \hat{Q}_2, \dots, \hat{Q}_N$  — тогда «произведение» перестановок  $\hat{P}$  и  $\hat{Q}$  есть третья перестановка  $\hat{R} = \hat{P}\hat{Q}$ , которая получается при действии сначала перестановки  $\hat{Q}$ , а затем

<sup>1</sup> Речь идет об аддитивности взаимодействий, описываемых самими матрицами гамильтонианов  $\mathbf{H}^A$  и  $\mathbf{H}^B$ , а не о взаимодействии между какими-то системами  $A$  и  $B$ . Однако возможен и такой подход, в котором  $\mathbf{H}^A$  и  $\mathbf{H}^B$  описывают *внутренние* взаимодействия двух систем — тогда они аддитивны, если системы невзаимодействующие; наличие (слабых) взаимодействий между ними можно попытаться учесть по теории возмущений.

перестановки  $\hat{P}$ . (В общем случае  $\hat{P}\hat{Q} \neq \hat{Q}\hat{P}$ .) Действуя перестановкой  $\hat{Q}$ , мы получим последовательность  $Q_1, Q_2, \dots, Q_N$ ; когда перестановка  $\hat{P}$  действует на эту последовательность, то необходимо поставить в первую позицию  $P_1$ -й элемент этой последовательности, равный  $Q_{(P_1)}$ . Вообще элемент  $Q_{(P_k)}$  перемещается в  $k$ -ю позицию, т. е. перестановка  $\hat{R} = \hat{P}\hat{Q}$  дает последовательность  $R_1, R_2, \dots, R_N = Q_{(P_1)}, Q_{(P_2)}, \dots, Q_{(P_N)}$ .

Дальнейшие важные свойства перестановок, которыми мы будем пользоваться:

а) Каждая перестановка  $\hat{P}$  имеет единственную обратную к ней перестановку  $\hat{P}^{-1}$ , такую, что их произведение есть тождественная перестановка  $\hat{I}$ . Перестановка  $\hat{Q} = \hat{P}^{-1}$  характеризуется расположением  $Q_1, Q_2, \dots, Q_N$ , где  $Q_k$  определяется, как позиция, в которой элемент  $k$  находится в последовательности  $P_1, P_2, \dots, P_N$ . Обратные «слева» и «справа» совпадают, что легко видеть, умножая равенство  $\hat{Q}\hat{P} = \hat{I}$  слева на  $\hat{P}$ :  $\hat{P}\hat{Q}\hat{P} = \hat{P}$ , откуда  $\hat{P}\hat{Q} = \hat{I}$ .

б) Если применить все  $N!$  перестановок к любой исходной последовательности  $Q_1, Q_2, \dots, Q_N$ , то получится  $N!$  разных перестановок; это означает, что множество полученных последовательностей не зависит от исходной — от нее зависит только их порядок. Таким образом, множество всех перестановок  $\hat{R} = \hat{P}\hat{Q}$  для данной (фиксированной)  $\hat{Q}$  то же, что и множество всех  $\hat{P}$ .

в) Каждой перестановке сопоставляется ее *четность*. Четность перестановки  $\hat{P}$  равна  $+1$ , если последовательность  $P_1, P_2, \dots, P_N$  можно получить четным числом транспозиций (перестановок двух элементов), исходя из последовательности  $1, 2, \dots, N$ , и она равна  $-1$ , если число транспозиций нечетно. Четность  $\hat{P}$  можно обозначить, как  $(-1)^p$ , где  $p$  есть число транспозиций  $\hat{P}$ . Четность произведения  $\hat{R} = \hat{P}\hat{Q}$  есть произведение четностей  $\hat{P}$  и  $\hat{Q}$ :  $(-1)^r = (-1)^{p+q}$ .

## П11. Алгоритм ортогонализации

Существуют случаи, когда дан некоторый вектор в  $N+1$ -мерном пространстве и необходимо построить ортонормированный базис, в котором данный вектор будет первым вектором базиса. Разумеется, классическая ортогонализация Грама—Шмидта подходит для этой цели. Однако она неэкономична и имеет тот недостаток, что разные векторы описываются несимметричным образом. Вместо метода Грама—Шмидта можно использовать следующий простой алгоритм.

Задача состоит в том, чтобы построить полный набор ортонормированных векторов, ортогональных к заданному вектору

$$\mathbf{q} = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix} \quad (\text{П11.1})$$

Предположим, что вектор  $\mathbf{q}$  нормирован ( $\sum_{i=0}^N |q_i|^2 = 1$ ), и определим:

$$\begin{aligned} \cos\varphi &= q_0 \\ \sin\varphi &= \left( \sum_{i=1}^N |q_i|^2 \right)^{1/2} \\ x_i &= \frac{q_i}{\sin\varphi} \quad (i = 1, 2, \dots, N) \end{aligned} \quad (\text{П11.2})$$

Тогда коэффициенты нового базиса можно задать с помощью унитарной матрицы

$$\begin{pmatrix} \cos\varphi & -x_1\sin\varphi & -x_2\sin\varphi & \dots & -x_i\sin\varphi & \dots \\ x_1\sin\varphi & 1+x_1^2(\cos\varphi-1) & x_2x_1(\cos\varphi-1) & \dots & x_ix_1(\cos\varphi-1) & \dots \\ x_2\sin\varphi & x_1x_2(\cos\varphi-1) & 1+x_2^2(\cos\varphi-1) & \dots & x_ix_2(\cos\varphi-1) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_i\sin\varphi & x_1x_i(\cos\varphi-1) & x_2x_i(\cos\varphi-1) & \dots & 1+x_i^2(\cos\varphi-1) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_N\sin\varphi & x_1x_N(\cos\varphi-1) & x_2x_N(\cos\varphi-1) & \dots & x_ix_N(\cos\varphi-1) & \dots \end{pmatrix} \quad (\text{П11.3})$$

Легко проверить, что первый столбец этой матрицы идентичен вектору  $\mathbf{q}$ , а также то, что все ее столбцы ортонормированы.



Ansatz, 35, 310

P-матрица, 242

– свойство идемпотентности  $PSP = P$ , 244

Адиабатическое приближение, 28

Адиабатические реакции, 28

Алгебраическое дополнение, 169–170, 179–180

Алгоритм, основанный на теореме

*Бриллюэна*, 202

Анализ заселенностей, 271

Аналогия между дифференциалами и вариациями, 361

Антисимметризация разделенных подсистем, 162

Антисимметризованное произведение строго ортогональных геминалей, 310

Асимптотическое поведение волновой функции на больших расстояниях, 338

Атомоподобная система, 20

Атомные единицы, 365

Базисные наборы, 241

Бесспиновая матрица плотности, 272

Биортогональные функции, 373

*Борна—Оппенгеймера* приближение, 23, 27

*Борна—Оппенгеймера* разделение переменных, 23–29

Бра-кет формализм, 366

– аналогия с матричным формализмом, 367

Бра векторы, 366

*Бриллюэна* теорема, 198, 266

– алгоритм, основанный на ней, 202

– использование, 203, 223, 238

– и вариационный принцип, 198

– для детерминанта со стационарной энергией, 200

– для детерминанта с абсолютным минимумом энергии, 199

*Бриллюэна—Вигнера* теория возмущений, 118

*Вайберга* индекс, 280

– для систем с открытой оболочкой, 280

Вариационная энергия

– точность, 34

Вариационно-пертурбационный метод, 110

Вариационный принцип, 32

– как эквивалент уравнения Шрёдингера, 34

– для детерминантной волновой функции, 198

– для возбужденных состояний, 37

– для основного состояния, 32

– полезная форма, 34

Вариации

– аналогия с дифференциалами, 361

Вековое уравнение, 65

Вектор-столбец орбитальных коэффициентов, 238

Вертикальные потенциалы ионизации, 218

ВЗМО, 220

*Вигнера*  $2n + 1$ -теорема, 102, 104, 111

– точность среднего значения, 102

– расчет поправок высших порядков к энергии, 109

Вириал

– инвариантность, 364

– внешних сил, 52

Вириальное отношение, 57

Вириала теорема

– и химическая связь, 58

– связь с масштабированием, 51

– для систем с кулоновским взаимодействием, 49

– в классической механике, 47–49

– в приближении Борна—Оппенгеймера, 52, 57

Виртуальные орбитали, 218, 303

Внутренние координаты, 17

Внутрипарная корреляция, 303

Внутри-наружу корреляция, 303

Возбуждения энергия, 222

Возбужденные состояния, 37, 222

– синглет и триплет, 234

Волновой функции силы, 45, 260

Временная зависимость физической величины, 47

Выражения для поправок теории возмущений к энергии

– сводка формул, 376

– использование  $2n + 1$ -теоремы Вигнера, 104

Вырожденная теория возмущений

Рэлея—Шрёдингера, 115

*Гельмана—Фейнмана* силы, 260

Гельмана—Фейнмана теорема, 40, 56, 60  
– условия ее действия, 44  
– дифференциальная, 41  
– интегральная, 46  
– проинтегрированная форма, 46  
Геминиали, 310  
Гипервириальная теорема, 370  
Гомолитическая диссоциация, 227  
Градиент энергии  
– вывод с использованием теоремы  
Бриллюэна, 260  
– вывод с использованием множителей  
Лагранжа, 264



Дважды возбужденные детерминанты  
– их важность, 301  
Дважды занятые орбитали, 223  
Двух тел проблема  
– в классической механике, 357  
– в квантовой механике, 11  
Двухэлектронные операторы  
– матричные элементы для  
неортогональных орбиталей, 178  
– матричные элементы для ортогональных  
орбиталей, 180  
Диагонализация, 66  
Диборана молекула, 291  
Динамическая корреляция, 299  
Дирака формализм бра-кет, 366  
Диссоциационная катастрофа в методе  
ОХФ, 227  
Дублетное состояние, 257

Занятые орбитали, 217–218  
Занятых орбиталей подпространство, 217  
«Зацепляющие» слагаемые в приближении  
Борна—Оппенгеймера, 27

Инвариантность оператора Фока, 206  
Индекс валентности, 282, 294  
– связь с индексами порядков связей, 283  
Индекс порядков связи, 281, 294  
– для систем с открытыми оболочками, 281  
– инвариантность, 281  
– связь с обменной плотностью, 288  
Индекс свободной валентности, 284  
– связь с матрицей спиновой плотности, 284  
Ионные слагаемые, 305  
Истинное решение НХФ, 233

Карадакова расширенная теорема  
парности, 188  
КВ разложение многоэлектронной  
волновой функции, 296  
КВ метод, 306  
Кет-векторы, 366

Кинетическая энергия, 49, 363  
– электронная, 24  
– ядерная, 24  
Ковалентные слагаемые, 305  
Коммутационные свойства оператора  
антисимметризации, 160  
Коммутатор, 47  
Конфигурационного взаимодействия  
разложение, 296  
Корреляционная нормировка, 93, 300  
Корреляционная энергия  
– определение, 299  
Коулсона—Фишер волновая функция, 311  
Коэффициенты связи в методе ОХФО, 252  
– их значения, 257  
Кратность связи, 278  
Критическая точка в методе НХФ, 230  
Кулоновский оператор, 204  
– эрмитовость, 207  
Купманса теорема, 216, 266  
– приближенные ионизационные  
потенциалы, 217  
Кутриевича метод приведения к  
эрмитовому виду, 37–38, 246, 254, 267

Лагранжа множители, 35, 38, 68, 210, 253, 264, 266  
– задачи, связанные с ними, 212, 266  
Лапласа теорема, 170  
– обобщенная, 171  
Лёвдина базис, 73  
Лёвдина дилемма симметрии, 234  
Лёвдина каноническая ортогонализация, 84–86  
– использование в уравнениях на  
собственные значения, 86  
Лёвдина метод разбиения, 139  
Лёвдина общая формула матричных  
элементов с неортогональными  
орбиталями, 168  
Лёвдина оператор спиновой проекции, 232  
Лёвдина симметричная ортогонализация, 71  
– двумерный пример, 79  
– свойство стационарности, 75  
– использование в уравнениях на  
собственные значения, 72  
Лёвдина теорема парности, 184  
ЛКАО энергия, выражение, 250  
ЛКАО метод, 237

Маллишена полная атомная заселенность, 275  
– инвариантность, 276  
Маллишена полная орбитальная  
заселенность, 274  
Маллишена чистая атомная заселенность, 275

- инвариантность, 278
- Малликена* чистая орбитальная заселенность, 274
- Малликена* заселенность перекрытия, 275
- инвариантность, 278
- Малликена* анализ заселенностей, 274, 298
- недостаток, 293
- Масштабирование, 52–53
- связь с теоремой вириала, 53
- Матричная задача на собственные значения, 65
- Матричное уравнение на собственные значения, 65
- обобщенное, 65
- «стандартное», 65
- Матричные элементы между детерминантными волновыми функциями, 168, 174, 178
- Матричные элементы для неортогональных орбиталей, 174, 178
- Матрица проекции на подпространство занятых орбиталей, 243
- Межпарная корреляция, 302
- Межэлектронные расстояния, 297, 299, 329
- Меллера—Плессета* разбиение, 311
- Меллера—Плессета* теория возмущений, 311
- Метод
  - $\Delta$ SCF, 218
  - CASSCF, 310
  - CCA, 320
  - CISD, 308
  - – утрата размерной согласованности, 308
  - EHF, 233
  - GVB, 311
  - MBPT, 311
  - RMP, 312
  - UMP, 315
  - АПСГ, 310
  - КВ в базисе однократных возбуждений, 222
  - НХФ, 230
  - ОХФ, 223
  - ОХФО, 251
  - РОРС, 225
  - РХФ, 233
  - ХФ, 197
  - ХФР, 237
- Миноры детерминанта матрицы перекрытия, 170–172, 180
- МК ССП метод, 309
- Многочастичная теория возмущений, 311
- Многоэлектронные спиновые состояния, 149
- их число, 149, 151



- Невырожденная теория возмущений Рэлея—Шрёдингера, 90
- алгебраическое разложение, 93
- компактные матричные обозначения, 97
- использование приведенной резольвенты, 98
- Независимый электрон
  - модель, 197
- Независимые гамильтонианы
  - сумма, 13
- Независимые переменные, 13
- Неограниченный метод *Хартри—Фока*, 230, 267
  - выражение для энергии, 211
  - уравнения, 204
  - – для последовательной оптимизации орбиталей, 245
  - – каноническая форма, 208
- Несбета* теорема, 301
- Нормировки интеграл для детерминантов *Слэтера*
  - ортогональные орбитали, 172
- НСМО, 220
- Обмен и химическая связь, 288
- Обменная энергия и индекс Вайберга, 290
- Обменный оператор, 204
  - эрмитовость, 207
- Обменная парная плотность, 287
- Обобщенная теорема Бриллюэна, 200, 266
- Обобщенная проблема собственных значений, 65
  - использование канонической ортогонализации по Лёвдину, 86
  - использование симметричной ортогонализации по Лёвдину, 72
- Обобщенный метод валентных связей, 311
- Ограниченный метод *Хартри—Фока*, 223
  - выражение для энергии, 211
  - симметрично-адаптированная волновая функция, 226
  - уравнения, 224
- Симметрия и метод ОХФ, 225
- Ограниченный метод *Хартри—Фока* для открытых оболочек, 251
  - выражение для энергии, 252
  - уравнения, 253
- Одноэлектронные операторы
  - матричные элементы для неортогональных орбиталей, 174
  - матричные элементы для ортогональных орбиталей, 176
- Однородные функции
  - *Эйлера* теорема, 362
- Оператор антисимметризации, 155
  - приложение к спин-орбитальным индексам, 159
  - коммутационные свойства, 160



– проекционный характер, 156  
 Оператор проекции на подпространство занятых орбиталей, 242  
 Орбитальная энергия, 217  
 Орбитальные коэффициенты, 238  
 Орбитальные энергии  
 – связь с полной энергией, 219  
 Ортогонализации алгоритм, 381  
 Открытая оболочка в синглетном состоянии, 257  
 Относительное движение, 13  
 Относительные координаты, 17

**Парная корреляционная энергия, 302**  
 Парности теорема, 184  
 – расширенная, 188  
*Паули* матрицы, 147  
 Перекрывание детерминантных волновых функций  
 – случай ортогональных орбиталей, 172  
 – общий случай неортогональных орбиталей, 168  
 Перекрывание между детерминантами Слэтера, 168  
 Перекрывания заселенность, 275  
 – инвариантность, 278  
 Перекрывания плотность, 274  
 Перестановки  
 – свойства, 380  
 Подобия преобразование, 66  
 Подпространство занятых орбиталей,  
 Полная атомная заселенность, 275  
 – инвариантность, 276  
 Полная орбитальная заселенность, 274  
 Полного КВ метод, 306  
 Порядок связи, 279  
 Порядки связи в трехцентровых системах, 291  
 Потенциал ионизации, 217  
 – вертикальный, 218  
 Потенциальная энергия, 24, 49  
 Потенциальной энергии  
 (гипер)поверхность, 23  
 Право-левая корреляция, 303  
 Преобразование  $S^{-1/2}$ , 72  
 Приведенная масса, 12  
 Приведенная резольвента, 98  
 – инвариантность, 102  
 Принцип Aufbau, 220  
 Проблема собственных значений матрицы, 65  
 – использование канонической ортогонализации по Лёвдину, 86  
 – использование преобразования Лёвдина, 72  
 Проекторы в методе ЛКАО, 245  
 Проекционные операторы, 370  
 Промежуточная нормировка, 93, 300



Прямого КВ метод, 309  
 Прямое произведение матриц, 378

**Радиальная корреляция, 303**  
 Разложение единицы, 372  
 Размерная расширяемость, 122  
 Размерная согласованность, 122  
 Размерная согласованность теории возмущений Рэлея—Шрёдингера  
 – формальное рассмотрение, 124  
 – разложение теории возмущений, 126  
 – доказательство, 126  
 Размерной согласованности проблема, 121  
 Разных орбиталей для разных спинов метод, 225  
 Расширенная теорема парности, 188  
 Расширенный метод Хартри—Фока, 233  
 Разложение  $1/Z$ , 332  
*Ридберга* постоянная, 15  
*Ритца* метод, 63  
 Рэлея коэффициент, 32  
 Рэлея—Шрёдингера теория возмущений  
 – алгебраическое разложение, 93  
 – компактные матричные обозначения, 97  
 – вырожденная, 115  
 – невырожденная, 90  
 – размерная согласованность, 124, 126  
 – использование приведенной резольвенты, 97

Самоотталкивания слагаемое, 206, 220–221, 236, 245, 262, 289, 316  
 Самосогласованное поле, 197  
 Свободное движение центра масс, 13  
 Свободные атомы, 20  
 Связанных кластеров подход, 317, 320  
 – размерная согласованность, 326  
 Сдвига уровня метод, 220  
 Седловая точка, 225  
 Силы, действующие на ядра, 42  
 Сильная ортогональность, 310  
 Симметрии дилемма, 234  
 Симметрическая ортогонализация, 71  
 Симметрично-неограниченная функция ХФ,  
 Симметрично-ограниченная функция ХФ,  
 Синглетная нестабильность, 226  
 Синглетное возбужденное состояние, 235  
 Синглетные и триплетные возбуждения из основного состояния ОХФ, 235  
 След проекционного оператора, 374  
*Слэтера* детерминант  
 – определение, 152  
 – выражение через оператор антисимметризации, 156  
 – свойства инвариантности, 165, 173  
 – матричные элементы, 172, 176, 180  
 – двухэлектронные примеры, 158



Слетера правила

- одноэлектронные операторы, 176
- перекрытия, 172
- двухэлектронные операторы, 181, 182

Смешивание занятых орбиталей, 165, 169

Состояние отсчета, 300

Спектральное разрешение, 373

Спин-орбитали, 146

Спин-спроецированный расширенный

метод *Хартри—Фока*, 233, 267

Спиновая плотность, 273

Спиновое загрязнение, 232

Спиновой плотности матрица, 273

Спиновой проекции оператор, 232

Спиновые операторы, 147

Спиновые функции, 147

– ортогональность, 148



Спроецированный гамильтониан, 38, 307

ССП, 197

ССП алгоритм

– обязательно сходящийся, 203

ССП расчеты

– сходимости, 220

ССП процедура, 210

Статическая корреляция, 299

Суммирование по спиновым переменным, 148

Сумма независимых гамильтонианов, 14

Сходимость расчетов СПП, 220

Теорема о произведении гауссовых функций, 339

Теорема об орбиталях специальной структуры, 191

– теорема существования, 191

– определение *a posteriori*, 193

Теория возмущений, 311

Теория *Хартри—Фока* в конечном базисе, 237

Точечная разрывность, 226

Точка бифуркации в методе НХФ, 230

Точность вариационной энергии, 34

Трехцентровая двухэлектронная связь, 291

Трехцентровая четырехэлектронная связь, 291

Триплетная нестабильность, 225

Триплетное возбужденное состояние, 235

Триплетное состояние, 257

Угловая корреляция, 303

Унитарное преобразование занятых орбиталей, 167

Упорядоченная конфигурация, 297

Условие излома

– электрона, 337

– ядра, 335

Факторизация детерминанта матрицы перекрытия, 170

*Фока* матрица, 247

*Фока* оператор, 206

– эрмитовость, 208

– инвариантность, 206

Фокиан, 206

Формула энергии

– RMP2, 312

– UMP2, 315

*Хартри—Фока* энергии выражение, 211

– в конечном базисе, 250

*Хартри—Фока* уравнения, 203

– вывод с использованием множителей Лагранжа, 210

– вывод с использованием специальных вариаций, 244

– вывод из теоремы Бриллюэна,

*Хартри—Фока* метод, 197

*Хартри—Фока—Рутана* уравнения, 237, 240

– вывод, 238–240

– выражение для энергии, 208–209

– формальное введение, 241

*Хиллерааса* функционал, 113

*Холецкого* разложение, 73

Центр масс

– в классической механике, 357

– в квантовой механике, 11, 16

Чистая атомная заселенность, 275

– инвариантность, 278

Чистая орбитальная заселенность, 274

Чисто синглетный характер волновой функции ОХФ, 231

*Эйлера* теорема об однородных функциях, 362

*Эккарта* неравенство, 36

Экспоненциальный Ansatz в теории связанных кластеров, 325

Электронная корреляция, 296, 298, 304, 311

– и теорема Несбета, 298

– методы учета, 306

– двухэлектронный пример, 304

Электронная плотность, 44, 271

– оператор, 271

Электронное уравнение Шрёдингера, 25

Электростатическая сила, действующая на ядро, 44

Ядерное уравнение Шрёдингера, 25

*Якоби* координаты, 16

# Оглавление

Предисловие редактора перевода .....	5
Предисловие автора к русскому изданию .....	7
Предисловие .....	9
Глава 1. Гамильтониан Борна—Оппенгеймера .....	11
1. Отделение движения центра масс в квантовой механике .....	11
1.1. Переход от задачи двух тел к двум задачам одного тела .	11
1.2. Центр масс в квантовой механике .....	16
1.3. Свободные атомы и атомоподобные системы.....	20
2. Приближение Борна—Оппенгеймера .....	23
2.1. Вводные замечания .....	23
2.2. Разделение электронных и ядерных переменных по Борну—Оппенгеймеру .....	24
2.3. Почему разделение переменных по Борну—Оппенгеймеру не является точным .....	26
2.4. Приближенное разделение уравнений .....	27
2.5. Замечания по поводу разделения переменных по Борну—Оппенгеймеру .....	29
Библиографические заметки .....	30
Литература .....	31
Глава 2. Общие теоремы и принципы .....	32
1. Вариационный принцип .....	32
1.1. Среднее значение .....	32
1.2. Вариационный принцип для основного состояния .....	32
1.3. Вариационный принцип как эквивалент уравнения Шрёдингера. Полезная формулировка вариационного принципа .....	34
1.4. Неравенство Эккарта .....	36
1.5. Возбужденные состояния .....	37
2. Теорема Гельмана—Фейнмана .....	40
2.1. Дифференциальная теорема Гельмана—Фейнмана .....	40
2.2. Интегральная теорема Гельмана—Фейнмана .....	45
3. Теорема вириала в квантовой механике .....	47
3.1. Зависимость физической величины от времени .....	47
3.2. Теорема вириала .....	48
3.3. Масштабирование — связь с вариационным принципом ...	50
3.4. Теорема вириала в приближении Борна—Оппенгеймера ..	52

3.5. Теорема вириала и химическая связь .....	58
Библиографические заметки .....	59
Литература .....	61
<b>Глава 3. Метод Ритца и ортогонализация по Лёвдину .....</b>	<b>63</b>
1. Линейный вариационный метод (метод Ритца) .....	63
2. Симметричная ортогонализация по Лёвдину .....	71
2.1. Матрица $\mathbf{S}^{-1/2}$ .....	71
2.2. Преобразование $\mathbf{S}^{-1/2}$ .....	72
2.3. Лёвдиновский базис .....	73
2.4. Свойство экстремальности симметричной ортогонализации по Лёвдину .....	74
2.5. Ортогонализация по Лёвдину. Двумерный пример .....	78
3. Линейная независимость базиса и каноническая ортогонализация по Лёвдину .....	82
3.1. Собственные значения матрицы интегралов перекрывания. Мера линейной независимости базиса .....	82
3.2. Каноническая ортогонализация по Лёвдину .....	84
Библиографические заметки .....	87
Литература .....	88
<b>Глава 4. Метод возмущений .....</b>	<b>90</b>
1. Невырожденная теория возмущений Рэлея—Шрёдингера .....	90
1.1. Формулировка задачи .....	90
1.2. «Алгебраическое» разложение .....	93
1.3. Использование приведенной резольвенты в теории возмущений Рэлея—Шрёдингера .....	97
1.4. $2n + 1$ -Теорема Вигнера .....	102
1.4.1. Точность среднего значения энергии, вычисленного с волновой функцией $n$ -го порядка теории возмущений .....	102
1.4.2. Вычисление точных поправок к энергии вплоть до порядка $2n + 1$ с использованием первых $n$ поправок к волновой функции .....	104
1.4.3. Выражения для поправок к энергии $E^{(2n)}$ и $E^{(2n+1)}$ .....	109
2. Вариационный метод и теория возмущений. Функционал Хиллерааса .....	110
3. Вырожденная теория возмущений Рэлея—Шрёдингера .....	115
4. Теория возмущений Бриллюэна—Вигнера .....	118
4.1. Проблема размерной согласованности .....	121
5. Размерная согласованность теории возмущений Рэлея—Шрёдингера .....	123
5.1. Формальное рассмотрение, основанное на свойствах степенного ряда .....	124
5.2. Размерная согласованность разложений в ряд теории возмущений .....	126

6. Метод разбиения по Лёвдину .....	139
Библиографические заметки .....	143
Литература .....	144
 Глава 5. <b>Детерминантные волновые функции</b> .....	146
1. Спин-орбитали .....	146
2. Многоэлектронные спиновые состояния .....	149
3. Детерминанты Слэтера .....	152
3.1. Двухэлектронные примеры .....	153
4. Оператор антисимметризации .....	155
4.1. Оператор антисимметризации как проекционный опера- тор .....	156
4.2. Коммутационные свойства оператора антисимметризации .....	160
4.3. Антисимметризация координат электронов в простран- ственно удаленных подсистемах .....	162
5. Инвариантность детерминантной волновой функции по отно- шению к «смешиванию» занятых орбиталей .....	165
6. Матричные элементы между детерминантными волновыми функциями: общие формулы Лёвдина для неортогональных ор- биталей .....	168
6.1. Перекрывание .....	168
6.1.1. Общий случай .....	168
6.1.2. Факторизация .....	169
6.1.3. Частный случай ортонормированных орбиталей (правила Слэтера) .....	172
6.2. Одноэлектронные операторы .....	174
6.2.1. Общий случай .....	174
6.2.2. Частный случай ортонормированных орбиталей (правила Слэтера) .....	176
6.3. Двухэлектронные операторы .....	178
6.3.1. Общий случай .....	178
6.3.2. Частный случай ортонормированных орбиталей (правила Слэтера) .....	180
7. Теорема парности Лёвдина и ее обобщение .....	183
7.1. Теорема парности Лёвдина .....	184
7.2. Расширенная теорема парности Карадакова .....	188
8. Теорема о существовании орбиталей специальной структуры .....	191
8.1. Теорема существования .....	192
8.2. <i>A posteriori</i> определение .....	193
Библиографические заметки .....	195
Литература .....	196

Глава 6. Метод Хартри—Фока .....	197
1. Вариационный принцип для однодетерминантных волновых функций: теорема Бриллюэна .....	198
1.1. Теорема Бриллюэна для детерминанта, дающего абсолютный минимум энергии .....	199
1.2. Теорема Бриллюэна для детерминанта, имеющего стационарную энергию .....	200
1.3. Алгоритм для решения задачи Хартри—Фока, основанный на теореме Бриллюэна .....	202
2. Уравнения Хартри—Фока .....	203
2.1. Неограниченные по спину уравнения Хартри—Фока .....	203
2.2. Альтернативный вывод с использованием множителей Лагранжа .....	210
2.3. Альтернативный вывод с использованием специальных вариаций .....	214
3. Теорема Купманса .....	216
3.1. Орбитальные энергии и полная энергия .....	218
4. Метод ОХФ .....	222
4.1. Схемы ОХФ и НХФ .....	222
4.2. Симметрия и метод ОХФ .....	225
4.3. «Диссоциационная катастрофа» и различные варианты метода Хартри—Фока .....	227
4.4. Синглетные и триплетные возбуждения .....	234
5. Теория Хартри—Фока в конечном базисе .....	237
5.1. Уравнения Хартри—Фока—Рутана .....	237
5.2. Матрица $P$ .....	242
5.3. Пример использования проекционных операторов в методе ЛКАО. Уравнения НХФ для последовательной оптимизации орбиталей .....	244
5.4. Матрица Фока и энергия .....	247
6. Ограниченный метод Хартри—Фока для открытых оболочек ..	250
6.1. Уравнения ОХФО .....	252
6.2. Коэффициенты связи для некоторых систем с открытой оболочкой .....	256
7. Градиент энергии .....	258
Библиографические заметки .....	265
Литература .....	267
Глава 7. Анализ заселенностей .....	269
1. Анализ заселенностей по Малликену .....	270
1.1. Электронная плотность .....	270
1.2. Анализ заселенностей .....	273
2. Порядки связей и валентности .....	278
2.1. Индекс порядка связи .....	278
2.2. Индексы валентности .....	281
2.3. Обменная плотность и индекс порядка связи .....	284

2.4. Порядки связей в трехцентровых системах .....	290
Библиографические заметки .....	292
Литература .....	294
<b>Глава 8. Электронная корреляция .....</b>	<b>295</b>
1. Разложение КВ .....	295
2. Корреляционная энергия и теорема Несбета .....	297
3. Методы учета электронной корреляции .....	305
3.1. Метод КВ .....	305
3.2. Методы МК ССП .....	308
3.3. Методы многочастичной теории возмущений или теории возмущений Меллера—Плессета .....	310
3.4. Метод связанных кластеров .....	316
Библиографические заметки .....	327
Литература .....	329
<b>Глава 9. Разное .....</b>	<b>331</b>
1. Разложение по степеням $1/Z$ .....	331
2. Граничные условия для волновой функции в точках сингулярности потенциала .....	333
2.1. Сингулярность потенциала взаимодействия электрона с ядром .....	333
2.2. Сингулярность потенциала электрон-электронного взаимодействия .....	336
3. Асимптотическое поведение волновой функции на больших расстояниях .....	337
4. Базисные функции: теорема о произведении гауссовых функций .....	338
5. Проблема преобразования интегралов .....	341
Библиографические заметки .....	342
Литература .....	343
<b>Дополнение редактора перевода .....</b>	<b>345</b>
<b>Приложения .....</b>	<b>353</b>
П1. Отделение движения центра масс в классической механике ...	353
П2. Переход от задачи двух тел к двум задачам одного тела в классической механике .....	354
П3. Аналогия между дифференциалами и вариациями .....	356
П4. Теорема Эйлера об однородных функциях .....	356
П5. Теорема вириала в классической механике .....	357
П5.1. Среднее полной производной физической величины по времени при финитном движении .....	357
П5.2. Теорема вириала .....	358
П6. Электронное уравнение Шрёдингера в атомных единицах ....	359
П7. «Бра-кет» формализм .....	359
П7.1. «Бра»- и «кет»-векторы Дирака .....	359
П7.2. Аналогия с матричным формализмом .....	361

П7.3. Неортогональный базис .....	362
П7.4. Пример использования бра-кет формализма: гипервири- альная теорема .....	363
П7.5. Проекционные операторы .....	363
П7.6. Разложение единицы .....	365
П7.7. Спектральное разложение эрмитовых операторов .....	366
П7.8. Неэрмитовы операторы — биортогональные наборы функций .....	366
П7.9. След проекционного оператора .....	367
П8. Сводка формул теории возмущений Рэлея—Шрёдингера (невыво- рожденный случай) .....	368
П9. Прямое произведение матриц .....	369
П10. Перестановки .....	371
П11. Алгоритм ортогонализации .....	372
<b>Предметный указатель .....</b>	<b>374</b>







*Минимальные системные требования определяются соответствующими требованиями программ Adobe Reader версии не ниже 11-й либо Adobe Digital Editions версии не ниже 4.5 для платформ Windows, Mac OS, Android и iOS; экран 10"*

*Учебное электронное издание*

**Майер Иштван**

## **ИЗБРАННЫЕ ГЛАВЫ КВАНТОВОЙ ХИМИИ: ДОКАЗАТЕЛЬСТВА ТЕОРЕМ И ВЫВОД ФОРМУЛ**

Ведущий редактор канд. хим. наук *Т. И. Почкаева*

Художник *Н. В. Зотова*

Оригинал-макет подготовлен *М. Б. Дарховским* в пакете L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>

Подписано к использованию 04.12.20.

Формат 145×225 мм

Издательство «Лаборатория знаний»

125167, Москва, проезд Аэропорта, д. 3

Телефон: (499) 157-5272

e-mail: [info@pilotLZ.ru](mailto:info@pilotLZ.ru), <http://www.pilotLZ.ru>

Это учебник высокого уровня, где приведены явные выводы основных квантово-химических утверждений с использованием методов квантовой механики.

Читатель должен быть готов следовать за автором, осознавая, что нет смысла заучивать формулировки теорем без изучения их доказательств.

Имея начальные представления по квантовой химии, с помощью этого учебника читатель достигнет более глубокого понимания фундаментальных концепций.

Для студентов, аспирантов, преподавателей и научных работников.

