

Министерство образования и науки Российской Федерации
Южно-Уральский государственный университет
Кафедра компьютерного моделирования и нанотехнологий

539.2(07)
В756

А.Г. Воронцов, А.Н. Соболев, С.А. Созыкин

КОМПЬЮТЕРНЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ МАТЕРИАЛОВ

Учебное пособие

Челябинск
Издательский центр ЮУрГУ
2018

УДК 539.21:004(075.8)
В756

*Одобрено
учебно-методической комиссией
института естественных и точных наук*

*Рецензенты:
Е.А. Беленков, А.А. Юрьев*

Воронцов А.Г.

В756 Компьютерные методы моделирования материалов: учебное пособие / А.Г. Воронцов, А.Н. Соболев, С.А. Созыкин. - Челябинск: Издательский центр ЮУрГУ, 2018. – 121 с.

Данное пособие является введением в теорию и практику компьютерного моделирования материалов. Рассмотрены методы моделирования структуры и свойств материалов «из первых принципов», т.е. на основе квантово-механических расчетов. Первая часть пособия является теоретическим введением в систему приближений и конкретных вычислительных алгоритмов, которые и составляют основу первопринципных методов. Во второй части изложены основы программного пакета VASP, являющегося на сегодняшний день самым известным и признанным для выполнения первопринципных вычислений. В третьей части рассматриваются сценарии использования VASP для решения разных задач на основе конкретных примеров. При выборе примеров основное внимание уделено задачам, демонстрирующим весь спектр возможностей пакета.

Пособие предназначено для аспирантов направлений «Физика и астрономия» и «Физическая химия», специализирующихся в области вычислительного материаловедения. Данное пособие будет полезно всем специалистам в области физики, химии и материаловедения, желающим познакомиться с теорией первопринципного компьютерного моделирования материалов и пакетом VASP.

УДК 539.21:004(075.8)

© Издательский центр ЮУрГУ, 2018

Оглавление

Введение	4
1. Теоретические основы методов компьютерного моделирования материалов	
1.1. Общие принципы моделирования материалов	5
1.2. Общие принципы нахождения электронной структуры	11
1.3. Потенциалы и базисы	27
1.4. Симметрия системы и базис плоских волн	38
2. Начало работы с пакетом VASP	
2.1. Общая характеристика пакета VASP	49
2.2. Установка VASP	50
2.3. Входные файлы	53
2.4. Выходные файлы	75
2.5. Наиболее важные параметры, источники ошибок	78
3. Сценарии использования пакета VASP	
3.1. Самосогласованный расчет	85
3.2. Моделирование STM изображения	86
3.3. Зонная структура	87
3.4. Электронная плотность	88
3.5. Оптимизация структуры	90
3.6. Фононные спектры	91
3.7. Учёт магнетизма	91
3.8. Молекулярная динамика	92
3.9. Переходное состояние	93
3.10. Спин-орбитальное взаимодействие	95
4. Вспомогательное программное обеспечение	
4.1. p4v	98
4.2. Phonopy	103
Некоторые параметры файла INCAR	111
Библиографический список	119

Введение

Хорошо известно, что первые попытки квантово-механического моделирования ковалентной связи позволяли получить лишь качественное объяснение ее природы, но не смогли получить правильное значение ее энергии. Однако в наше время химики с высокой точностью могут определять энергию ковалентных связей, их равновесную длину, силовые константы, упругость и поляризуемость. Физика конденсированных сред разработала ряд методов, позволяющих с высокой точностью определять свойства материалов, такие как решеточные параметры или упругие константы. Такие методы, единственной требуемой информацией для которых является задание типов (атомных номеров) атомов, входящих в состав рассматриваемого вещества, принято называть *ab initio* (первопринципными) методами.

Последние десятилетия были ознаменованы прорывом в создании теоретических основ и развитии алгоритмов и программ, позволившим говорить о создании первопринципного квантово-механического моделирования материалов на атомном уровне. В результате возникли вычислительные технологии, позволяющие перебросить мост через существовавший ранее разрыв между фундаментальными науками и инженерным материаловедением. Компьютерное моделирование электронной структуры – одно из важнейших направлений физики конденсированных сред, в котором прогресс теории и вычислительных методов наиболее динамичен. В настоящее время методология компьютерного моделирования конденсированных систем на атомном уровне достигла стадии, где уже возможно ее использование в области прикладного материаловедения для разработки и проектирования высококачественных материалов.

Данное учебное пособие разработано в соответствии ФГОС ВО по направлению подготовки 03.06.01 «Физика и астрономия» (уровень аспирантуры). Цель представленного пособия – дать читателю основные представления о современной теории и методах компьютерного моделирования, а также подготовить аспиранта к решению ряда профессиональных задач в области компьютерного моделирования материалов.

1. Теоретические основы методов компьютерного моделирования материалов

1.1. Общие принципы моделирования материалов

Конечной задачей теории твердого тела является установление свойств материалов на основе последовательного квантовомеханического описания многочастичной системы электронов. Для предсказания атомной и электронной структуры конденсированных тел требуется рассчитать полную квантово-механическую энергию системы, и затем минимизировать ее по электронным и ядерным координатам. Расчеты полной энергии системы могут быть проведены только при использовании ряда существенных упрощений и приближений. Рассмотрим подробнее приближения, лежащие в основе методов первопринципного моделирования, и методы эффективного численного решения получаемых задач.

1.1.1. Приближение Борна–Оппенгеймера

Квантовомеханическое состояние любой системы описывается зависящим от времени уравнением Шредингера:

$$i\hbar \frac{\partial \Phi(\mathbf{r}, \mathbf{R})}{\partial t} = \hat{H} \Phi(\mathbf{r}, \mathbf{R}), \quad (1.1)$$

где $\Phi(\mathbf{r}, \mathbf{R})$ — волновая функция, зависящая от положения ионов \mathbf{R} и электронов \mathbf{r} ; \hat{H} — оператор Гамильтона:

$$\begin{aligned} \hat{H} = & - \sum_{I=1}^N \frac{\hbar^2}{2M_I} \frac{\partial^2}{\partial \mathbf{R}_I^2} - \sum_{i=1}^n \frac{\hbar^2}{2m} \frac{\partial^2}{\partial \mathbf{r}_i^2} - \sum_{I=1}^N \sum_{i=1}^n \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_i|} + \\ & + \frac{1}{2} \sum_{I=1}^N \sum_{\substack{J=1 \\ J \neq I}}^N \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} + \frac{1}{2} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} = \quad (1.2) \\ = & T_I + T_e + V_{Ie}(\mathbf{r}, \mathbf{R}) + V_{II}(\mathbf{R}) + V_{ee}(\mathbf{r}). \end{aligned}$$

Здесь M_I и Z_I — масса и заряд ионов, m и e — масса и заряд электрона. Первое и второе слагаемое — это кинетическая энергия ионов и электронов, третье, четвертое и пятое слагаемые — это потенциальная энергия

ион–электронного, ион–ионного и электрон–электронного кулоновского взаимодействия.

В общем виде уравнение (1.1) не решается, т.к. содержит очень большое число неизвестных, однако ситуацию можно упростить, разделив ионные и электронные переменные. Масса ионов и электронов различается примерно на 3 порядка, поэтому электроны быстро приходят в положение равновесия, при этом ионы практически неподвижны (адиабатическое приближение). Волновая функция при этом преобразуется в произведение электронной $\Psi(\mathbf{r}, \mathbf{R})$ и ионной $\xi(\mathbf{R})$ волновых функций:

$$\Phi(\mathbf{r}, \mathbf{R}) = \Psi(\mathbf{r}, \mathbf{R}) \cdot \xi(\mathbf{R}). \quad (1.3)$$

Электронная волновая функция зависит от положения ионов, как от параметров. Воздействие электронов на ионы, в свою очередь, определяется потенциальной энергией системы электронов. Решение общей задачи (1.1) сводится к решению задач отдельно для ионных и электронных степеней свободы:

$$\begin{cases} [T_I + V_{II}(\mathbf{R}) + E_e(\mathbf{R})] \cdot \xi(\mathbf{R}) = E_I \cdot \xi(\mathbf{R}) \\ [T_e + V_{Ie}(\mathbf{r}, \mathbf{R}) + V_{ee}(\mathbf{r})] \cdot \Psi(\mathbf{r}, \mathbf{R}) = E_e(\mathbf{R}) \cdot \Psi(\mathbf{r}, \mathbf{R}). \end{cases} \quad (1.4)$$

Собственные состояния электронов $E_e(\mathbf{R})$ параметрически зависят от положения ионов. Считая, что электроны занимают основное состояние $E_0(\mathbf{R})$, и пренебрегая квантовыми эффектами при движении ионов, получаем задачу в приближении Борна–Оппенгеймера:

$$\begin{cases} -\frac{\partial}{\partial \mathbf{R}_I} [V_{II}(\mathbf{R}) + E_0(\mathbf{R})] = M_I \ddot{\mathbf{R}}_I \\ \hat{H}_{BO} \Psi(\mathbf{r}, \mathbf{R}) = E_0(\mathbf{R}) \cdot \Psi(\mathbf{r}, \mathbf{R}), \end{cases} \quad (1.5)$$

где $\hat{H}_{BO} = T_e + V_{Ie}(\mathbf{r}, \mathbf{R}) + V_{ee}(\mathbf{r})$ — электронный гамильтониан Борна–Оппенгеймера. Первое уравнение системы (1.5) — классическое уравнение движения иона в поле кулоновских сил от других ионов и эффективном поле электронного газа, находящегося в основном состоянии.

Силы, действующие на ионы со стороны электронов, можно определить по теореме Гельмана–Фейнмана:

$$\begin{aligned} \frac{\partial E_0(\mathbf{R})}{\partial \mathbf{R}_I} &= \frac{\partial}{\partial \mathbf{R}_I} \langle \Psi | \hat{H}_{BO} | \Psi \rangle = \langle \frac{\partial \Psi}{\partial \mathbf{R}_I} | \hat{H}_{BO} | \Psi \rangle + \\ &\langle \Psi | \frac{\partial \hat{H}_{BO}}{\partial \mathbf{R}_I} | \Psi \rangle + \langle \Psi | \hat{H}_{BO} | \frac{\partial \Psi}{\partial \mathbf{R}_I} \rangle = \langle \Psi | \frac{\partial \hat{H}_{BO}}{\partial \mathbf{R}_I} | \Psi \rangle. \end{aligned} \quad (1.6)$$

Первое и третье слагаемое в сумме равны нулю для основного электронного состояния в силу вариационного принципа. Силы, действующие на

ионы, определяются матричными элементами градиента электронного гамильтониана в основном состоянии. Однако основное состояние должно быть определено очень точно, т.к. при ошибке в полной энергии во 2-м порядке, силы будут иметь ошибки в 1-м порядке.

Также можно найти тензор напряжений σ_{ij} при деформации t_{ij} векторов кристаллической решетки \mathbf{a} . Записываем компоненты векторов решетки после деформации: $a'_i = \sum_j (\delta_{ij} + t_{ij}) a_j$, и находим тензор напряжений:

$$\sigma_{ij} = -\frac{\partial E_0(\mathbf{a}')}{\partial t_{ij}} = -\langle \Psi | \frac{\partial \hat{H}_{BO}(\mathbf{a}')}{\partial t_{ij}} | \Psi \rangle .$$

1.1.2. Моделирование движения ионов

Оставим на время вопрос нахождения энергии основного состояния электронной подсистемы $E_0(\mathbf{R})$ и обратимся к проблеме решения первого из уравнений (1.5).

Уравнение движения ионов представляет собой классическую задачу ньютоновской механики для движения системы тел под действием сил:

$$\mathbf{F}_I = -\frac{\partial}{\partial \mathbf{R}_I} [V_{II}(\mathbf{R}) + E_0(\mathbf{R})] .$$

Это выражение в англоязычной литературе называется «force field» — силовое поле.

При заданном силовом поле интегрирование указанной системы проводят методом классической молекулярной динамики [1], обычно с помощью алгоритма Верле:

$$\mathbf{R}_I^{(n+1)} = 2\mathbf{R}_I^{(n)} - \mathbf{R}_I^{(n-1)} + \frac{(\Delta t)^2}{M_I} \mathbf{F}_I .$$

Основным параметром здесь является шаг интегрирования по времени Δt . Алгоритм Верле при интегрировании уравнений движения привносит ошибку порядка $(\Delta t)^4$.

Реже для моделирования движения атомов используют метод Монте-Карло [2]. В этом случае смещения атомов носят случайный характер, но вероятности смещений задаются так, чтобы удерживать систему вблизи определенного уровня полной энергии. В результате получается набор состояний системы, пригодный для изучения ее свойств методами статистической физики.

Проблему задания силового поля решают разными способами. Перечислим их в порядке убывания строгости.

1. Первопринципная (*ab initio MD*) молекулярная динамика: находится основное состояние электронной подсистемы и силы вычисляются по теореме Гелмана–Фейнмана (1.6).
2. Метод Кара–Паринелло (*Car–Parinello MD*): для электронов записываются классические уравнения движения и реализуется одновременная динамика ионов и электронов.
3. Метод подгонки сил (*force matching*): после вычисления электронной структуры силовое поле параметризуется численно или аналитически, в дальнейшем электронные вычисления не проводятся.
4. Эмпирические потенциалы взаимодействия: силы вычисляются по некоторым эмпирическим формулам.

Первопринципная молекулярная динамика

Непосредственное решение системы (1.5) проводят по следующему алгоритму [3]. Сначала берется произвольное расположение ионов и для него из второго уравнения находится электронная структура (энергия основного состояния электронной подсистемы $E_0(\mathbf{R})$). Координаты ионов \mathbf{R} входят в это уравнение в качестве постоянных параметров. Затем определяется изменение этой энергии при небольшом смещении ионов. Это изменение определяет силы, действующие на ионы (1.6). Ионы слегка перемещаются под действием сил, и снова получается задача для вычисления электронной структуры. Затем это процедура повторяется снова. Основная проблема для реализации этой процедуры заключена в определении электронной структуры — решении уравнения Шредингера для системы электронов. Именно недостаточные быстродействие и точность этой части расчетов долгое время являлись препятствием для распространения «первопринципных» методов расчета.

Достичь желаемой цели долгое время не представлялось возможным по причине огромного объема требуемых компьютерных ресурсов. В настоящее время стали доступны для моделирования системы, состоящие из нескольких сотен атомов, что достаточно для исследования ограниченного, но весьма важного круга задач. Такой прогресс был обеспечен развитием весьма эффективных численных алгоритмов для реализации указанного метода, которые позволили почти на порядок увеличить доступный размер системы. Один из них, развитый Каром и Паринелло (1985) [4], изменил путь поиска решения квантовомеханической задачи. Вместо самосогласованного решения системы двух дифференциальных уравнений в частных производных методом разложения по полному базису с дальнейшей диагонализацией они используют минимизацию единственной функции методом «симуляции отжига» (*simulated annealing*).

Молекулярная динамика Кара–Паринелло

Современная *ab initio* молекулярная динамика фактически появилась в работе Кара и Паринелло [4]. Их работа внесла две новые важные идеи. Первая идея состоит в том, что решение уравнения Шредингера может быть проведено в различных математических формах. Вторая идея состоит в том, что нет необходимости проводить порознь расчеты электронной и атомной подсистем. Их самосогласованность может быть достигнута одновременной минимизацией функционала свободной энергии по электронным и ионным степеням свободы.

В методе Кара–Паринелло энергия электронной подсистемы $E_0(\mathbf{R}, \mathbf{C})$ рассматривается как функция от положения ионов \mathbf{R} и набора коэффициентов $\mathbf{C} \equiv \{C_i\}$ разложения волновых функций ψ_i по некоторому базису. Каждый коэффициент C_i можно рассматривать как некоторую координату классической «частицы». Чтобы минимизировать функционал энергии при решении уравнения Шредингера, этим частицам приписывается некоторая кинетическая энергия, и система постепенно охлаждается, пока набор координат не достигнет значений \mathbf{C}_0 , которые минимизируют функционал энергии. Таким образом, проблема нахождения собственных состояний гамильтониана сводится к решению некоторого количества классических уравнений движения. Следует отметить, однако, что функционал энергии имеет при этом физический (квантово-механический) смысл только после достижения минимального значения.

Лагранжиан электронной подсистемы был определен следующим образом:

$$L = \sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle - E_0(\mathbf{R}, \mathbf{C}), \quad (1.7)$$

где μ — фиктивная масса, связанная с электронными волновыми функциями. Слагаемое с кинетической энергией в лагранжиане описывает фиктивную динамику электронных степеней свободы. Функционал энергии играет роль потенциальной энергии в обычной формулировке механики Лагранжа.

Электронные волновые функции ограничиваются условием ортонормировки:

$$\int \psi_i^*(r) \psi_j(r) dr = \delta_{ij}. \quad (1.8)$$

Это ограничение вносится в лагранжеву динамику при помощи метода неопределенных множителей Лагранжа. В результате лагранжиан принимает вид:

$$L = \sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle - E_0(\mathbf{R}, \mathbf{C}) + \sum_{i,j} \Lambda_{i,j} \left\{ \int \psi_i^*(r) \psi_j(r) dr - \delta_{ij} \right\}. \quad (1.9)$$

Неопределенные множители Лагранжа $\Lambda_{i,j}$ можно рассматривать как дополнительные силы, которые, действуя на волновые функции, обеспечивают сохранение их ортонормировки. Уравнения движения электронных состояний находятся из уравнений в методе Лагранжа,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\psi}_i} \right) = \frac{\partial L}{\partial \psi_i}, \quad (1.10)$$

откуда находим

$$M_I \ddot{\mathbf{R}}_I = - \frac{\partial E_0(\mathbf{R}, \mathbf{C})}{\partial \mathbf{R}_I}, \quad \mu \ddot{\psi}_i = - \hat{H} \psi_i + \sum_j \Lambda_{i,j} \psi_j, \quad (1.11)$$

где \hat{H} — гамильтониан электронной подсистемы. Сила $(-\hat{H} \psi_i)$ представляет собой градиент энергетического потенциала в точке фазового пространства, которая соответствует волновой функции ψ_i . Уравнения движения представляют собой классические уравнения Ньютона. Массы ионов совпадают с их физическими массами, а фиктивная масса электронов μ выбирается так, чтобы обеспечить максимально быструю сходимость алгоритма. Таким образом, идея Кара–Паринелло заключается в том, что не обязательно находить основное состояние электронной подсистемы для запуска движения ионов, достаточно некоторого приближения.

Метод Кара–Паринелло позволяет проводить расчеты для систем размера порядка сотни атомов. Выяснилось, что метод Кара–Паринелло хорошо работает для неметаллических систем (полупроводники и диэлектрики), у которых уровень Ферми лежит в области запрещенной зоны. Для металлических же систем, где в области уровня Ферми существует множество состояний с близкими собственными значениями, легко возникает перекачка энергии из электронных к ионным степеням свободы и наоборот. Это вызывает макроскопические осцилляции электронной плотности при очень небольших изменениях полной энергии. Такие скачки сильно затрудняют сходимость алгоритма. Единственная возможность решить эту проблему — уменьшение временного шага, что резко снижает производительность метода. Заметим, что величина минимального вектора обратной решетки определяется размером ячейки, поэтому с ростом

числа частиц в системе данная трудность становится непреодолимой и требуется переход к более эффективным алгоритмам.

Для систем из большого числа частиц более эффективной оказывается методика, при которой динамика атомных и электронных степеней свободы описывается различными алгоритмами. Отказ от элегантной единой методики Кара и Паринелло позволяет получить большой выигрыш во времени из-за отсутствия необходимости использования малого временного шага для ионной подсистемы. Для каждой атомной конфигурации, полученной методом молекулярной динамики для ионной подсистемы, полная энергия электронной подсистемы находится прямой минимизацией функционала энергии методом сопряженных градиентов [5]. Это устраняет указанные проблемы в молекулярно-динамическом подходе Кара–Паринелло. При этом остается возможность ускорить сходимость полного расчета, если находить не точный минимум электронной энергии, а ограничиваться некоторым приближением. В результате получается первопринципная молекулярная динамика с контролируемым уровнем точности вычисления силового поля.

Контрольные вопросы

1. В чем заключается приближение Борна – Оппенгеймера? Чем оно отличается от адиабатического приближения? Приведите пример, когда оно не справедливо.
2. Как моделируется движение ионов, перечислите разные способы.
3. В чем достоинство первопринципной молекулярной динамики? В чем ее недостатки? Приведите пример, когда целесообразно ее использование.
4. В чем достоинство классической молекулярной динамики? В чем ее недостатки? Приведите пример, когда целесообразно ее использование.
5. В чем достоинство молекулярной динамики Кара – Паринелло? В чем ее недостатки? Приведите пример, когда целесообразно ее использование.

1.2. Общие принципы нахождения электронной структуры

Вычисление электронной структуры при известном расположении атомов сводится к решению второго уравнения системы (1.5) с гамильтонианом Борна–Оппенгеймера. Существует современный метод квантово-

го Монте-Карло [6], при помощи которого можно непосредственно подобрать n волновых функций электронов, являющихся решениями этого уравнения. Однако для реализации подобных вычислений требуются невероятные объемы компьютерной памяти и временные затраты. На сегодняшний день совместное получение нескольких волновых функций уже считается успехом. Конечно, данные методы позволяют глубже взглянуть в особенности взаимодействия электронов, но для расчета электронной структуры реальных систем они пока мало пригодны. Поэтому пока основным приближенным путем решения многочастичной задачи о поведении плотного электронного газа в решетке является ее сведение к одночастичной задаче о движении электрона в эффективном поле ионов и остальных электронов, которое пытаются выбрать так, чтобы по возможности учесть эффекты взаимодействия электронов между собой.

1.2.1. Одночастичное приближение (метод Хартри–Фока)

Первые работы по замене многочастичной задачи на одночастичную принадлежат Хартри (1928). Он предложил разбить задачу для многих электронов на одноэлектронные задачи и скомпоновать многоэлектронную волновую функцию как произведение одноэлектронных:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) \dots \psi_n(\mathbf{r}_n). \quad (1.12)$$

Такое приближение получило название приближение Хартри или приближение невзаимодействующих электронов. При таком подходе мы получаем газ электронов, не оказывающих друг на друга никакого влияния (кроме кулоновского поля). Волновые функции отдельных электронов полностью независимы, что приводит к разделению уравнения Шредингера на n независимых уравнений

$$\left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial \mathbf{r}_i^2} - \sum_{I=1}^N \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_i|} + V_H(\mathbf{r}_i) \right) \psi_i(\mathbf{r}_i) = \varepsilon_i \psi_i(\mathbf{r}_i). \quad (1.13)$$

где потенциал Хартри $V_H(\mathbf{r}_i)$ есть потенциальная энергия электрона i в усредненном электростатическом поле, созданном другими электронами:

$$V_H(\mathbf{r}_i) = e^2 \sum_{j \neq i} \int \frac{|\psi_j(\mathbf{r}')|^2}{|\mathbf{r}_i - \mathbf{r}'|} d\mathbf{r}', \quad (1.14)$$

а энергия всей системы складывается из сумм энергий каждого электрона за вычетом слагаемого, компенсирующего двойной учет электрон-электронной кулоновской энергии в сумме одноэлектронных собственных

энергий:

$$E_{tot} = \sum_i \varepsilon_i - \frac{e^2}{2} \int \int \psi_i^*(\mathbf{r}) \psi_i(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j^*(\mathbf{r}') \psi_j(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (1.15)$$

Дальнейшее уточнение относительно вида волновой функции сделал В.А. Фок. Конструируя общую волновую функцию из одночастичных, он не ограничился простым произведением, а сразу использовал принцип Паули и построил антисимметричную волновую функцию — детерминант Слэтера:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \dots & \psi_1(\mathbf{r}_n) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \dots & \psi_2(\mathbf{r}_n) \\ \dots & \dots & \dots & \dots \\ \psi_n(\mathbf{r}_1) & \psi_n(\mathbf{r}_2) & \dots & \psi_n(\mathbf{r}_n) \end{vmatrix}. \quad (1.16)$$

Это приводит к появлению дополнительного вклада в энергию (1.15):

$$E_X = \frac{e^2}{2} \sum_{j \neq i} \int \int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (1.17)$$

называемого обменной энергией.

Отметим, что в процессе решения уравнения (1.13) в методе Хартри–Фока необходимо добиться совпадения исходных функций, входящих в потенциал (1.14), и результатов, полученных после решения уравнения. Эта процедура называется самосогласованием, а найденное решение — самосогласованным. Самосогласованное решение обычно строится методом итераций и требует многократного повторения решения одноэлектронного уравнения Шредингера с известным потенциалом. В случае систем с большим числом электронов самосогласование — трудоемкий, но единственный реально возможный способ получения электронной структуры. Приближение Хартри–Фока позволяет учесть эффекты взаимодействия электронов между собой. Эти эффекты — двоякого рода. Во-первых, это сильное электростатическое взаимодействие с другими валентными электронами, описываемое эффективным потенциалом Хартри $V_H(\mathbf{r}_i)$. Во-вторых, квантовомеханический принцип Паули, который заставляет электроны с параллельными спинами держаться порознь, и тем самым понижает кулоновское взаимодействие на величину обменной энергии. Тем не менее метод Хартри–Фока, широко используемый в квантовой химии, для случая твердого тела работает очень плохо (он приводит к бесконечной скорости Ферми и очень широкой зоне sp-металлов). Это связано с тем, что приближение Хартри–Фока не учитывает динамические корреляции даже между электронами с антипараллельными спинами, которые дают вклад в так называемую корреляционную энергию.

1.2.2. Теория функционала плотности

Наиболее полно схема самосогласования и построения одночастичного потенциала была разработана Коном и Шэмом [7]. Большим успехом теории стало открытие, что сложные обменные и корреляционные явления могут быть достаточно точно описаны путем добавления локального обменно-корреляционного потенциала $V_{XC}(\mathbf{r})$ к потенциалу Хартри $V_H(\mathbf{r})$. Получаемая при этом система самосогласованных уравнений Кона–Шэма, сводящая многоэлектронную задачу к одночастичной, получила название теории функционала плотности (ТФП). Данная теория стала общепринятой при рассмотрении электронного газа в твердом теле, заменив собой множество существовавших ранее разнородных подходов и значительно улучшив согласие между теорией и экспериментом. Теория функционала плотности позволяет обеспечивать высокую точность описания полной энергии как кластеров, так и твердых тел, и за ее развитие Вальтер Кон был удостоен Нобелевской премии в 1998 году.

Теоремы Хохенберга и Кона

Основным параметром, определяющим свойства системы в теории функционала плотности, является электронная плотность:

$$\rho(\mathbf{r}) \equiv |\Psi(\mathbf{r})|^2. \quad (1.18)$$

Использование одной скалярной функции плотности вместо N -частичной волновой функции упрощает процедуру нахождения самосогласованного решения уравнения Шредингера. Значительное сокращение затрат на вычисления без существенной потери точности является основной причиной того, что теория функционала плотности повсеместно используется в области традиционного материаловедения.

Возможность использования плотности в качестве единственного параметра основана на следующих теоремах Хохенберга и Кона [8].

Теорема 1:

В невырожденном основном состоянии плотность взаимодействующих электронов $\rho(\mathbf{r})$ взаимно однозначно соответствует внешнему (явно не зависящему от плотности) потенциалу $V(\mathbf{r})$.

Теорема 2:

Существует функционал $E[\rho]$, зависящий только от плотности $\rho(\mathbf{r})$, такой, что его минимум $E[\rho_0]$ в точности равен энергии основного состояния системы.

Доказательство теоремы 1

Изложим доказательство, следуя оригинальной работе Хохенберга и Кона [8]. Рассмотрим систему из N электронов, движущихся во внешнем поле $v(\mathbf{r})$, которая обладает невырожденным основным состоянием $|\Psi\rangle$. В атомных единицах гамильтониан системы имеет вид:

$$H = T + U + V, \quad (1.19)$$

где

$$T = \frac{1}{2} \int \nabla \Psi^*(\mathbf{r}) \nabla \Psi(\mathbf{r}) d\mathbf{r}, \quad (1.20)$$

$$U = \frac{1}{2} \int \frac{\Psi^*(\mathbf{r}) \Psi^*(\mathbf{r}') \Psi(\mathbf{r}') \Psi(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (1.21)$$

$$V = \int v(\mathbf{r}) \Psi^*(\mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r}, \quad (1.22)$$

Ясно, что $|\Psi\rangle$ есть однозначный функционал от $v(\mathbf{r})$, следовательно, наблюдаемая электронная плотность

$$\rho(\mathbf{r}) = |\Psi(\mathbf{r})|^2 = |\Psi'(\mathbf{r})|^2 \quad (1.23)$$

также является однозначным функционалом от $v(\mathbf{r})$.

Докажем, что справедливо и обратное: $|\Psi\rangle$ и $v(\mathbf{r})$ однозначно определяются распределением плотности $\rho(\mathbf{r})$. Доказательство проводится от противного. Предположим, что потенциал $v'(\mathbf{r})$, отличный от $v(\mathbf{r})$, приводит к основному состоянию $|\Psi'\rangle$ с той же плотностью $\rho(\mathbf{r})$. Очевидно, что $|\Psi\rangle \neq |\Psi'\rangle$, т.к. эти состояния представляют собой собственные состояния разных гамильтонианов. Обозначим гамильтонианы и энергии основного состояния, соответствующие функциям $|\Psi\rangle$ и $|\Psi'\rangle$, через H , H' , E , E' . Тогда в силу требования минимальности энергии E основного состояния гамильтониана H :

$$E = \langle \Psi | H | \Psi \rangle < \langle \Psi' | H | \Psi' \rangle = \langle \Psi' | H' + V - V' | \Psi' \rangle. \quad (1.24)$$

Отсюда, при условии равенства двух плотностей следует, что

$$E < E' + \int [v(\mathbf{r}) - v'(\mathbf{r})] \Psi^*(\mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r} = E' + \int [v(\mathbf{r}) - v'(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r}. \quad (1.25)$$

Из аналогичного рассмотрения энергии основного состояния гамильтониана H' :

$$E' < E + \int [v'(\mathbf{r}) - v(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r}. \quad (1.26)$$

При сложении (1.25) и (1.26) приходим к противоречию: $E + E' < E' + E$. Таким образом в основном состоянии двум разным потенциалам не может

соответствовать одна и та же плотность электронов (кроме случая, когда $v'(\mathbf{r}) - v(\mathbf{r}) = const$). Отсюда следует, что поскольку плотность однозначно определяет потенциал, то она однозначно определяет и основное состояние, а также все остальные электронные свойства системы.

Доказательство теоремы 2 (вариационный принцип для энергии)

Поскольку $|\Psi\rangle$ есть функционал от $\rho(\mathbf{r})$, то функционалом от электронной плотности будет также и величина

$$F[\rho(\mathbf{r})] \equiv \langle \Psi | T + U | \Psi \rangle, \quad (1.27)$$

которая представляет собой сумму кинетической и кулоновской энергий. Этот функционал играет ключевую роль в ТФП.

Рассмотрим систему взаимодействующих электронов в некотором внешнем поле $v(\mathbf{r})$. В такой системе определим энергию как следующий функционал от плотности

$$E[\rho(\mathbf{r})] \equiv \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho(\mathbf{r})]. \quad (1.28)$$

Покажем теперь, что при истинном распределении плотности $\rho(\mathbf{r})$ значение $E_v[\rho(\mathbf{r})]$ меньше, чем при любом другом распределении $\rho'(\mathbf{r})$ с тем же полным числом частиц.

Используем обычный функционал энергии Рэлея–Ритца

$$E[\Psi'] = \langle \Psi' | V | \Psi' \rangle + \langle \Psi' | T + U | \Psi' \rangle, \quad (1.29)$$

который для истинного состояния $|\Psi\rangle$ имеет меньшую величину, чем для любого другого $|\Psi'\rangle$ при том же числе частиц. Пусть $|\Psi'\rangle$ есть состояние, отвечающее некоторой другой плотности $\rho'(\mathbf{r})$. Тогда в соответствии с принципом Рэлея–Ритца и уравнениями (1.27) и (1.28) имеем:

$$E[\Psi'] = \int v(\mathbf{r})\rho'(\mathbf{r})d\mathbf{r} + F[\rho'(\mathbf{r})] > E[\Psi] = \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho(\mathbf{r})] \quad (1.30)$$

Это и доказывает требуемый принцип минимума для $E[\rho'(\mathbf{r})]$: $E[\rho'(\mathbf{r})] > E[\rho(\mathbf{r})]$. Из выражений (1.27) и (1.28) ясно также, что минимальное значение энергии $E[\rho(\mathbf{r})]$ есть также точная энергия основного состояния для внешнего поля $v(\mathbf{r})$ и числа частиц N .

На практике полезность нового вариационного принципа определяется тем, существуют ли такие приближения для $F[\rho'(\mathbf{r})]$, которые бы было достаточно просто применять и которые бы обладали требуемой

точностью. Для точности в 10% такие приближения существуют. Если же требуются более аккуратные вычисления, то трудности связанные с выводом и использованием высокоточных выражений для функционала энергии растут крайне быстро.

Самосогласованные уравнения Кона–Шэма

Теоремы Хохенберга–Кона заложили фундамент ТФП, однако ее успешное применение оказалось возможным только после преобразования многоэлектронной задачи в одночастичную. Получаемая при этом система самосогласованных уравнений Кона–Шэма является основой численного алгоритма решения многочастичного уравнения Шредингера.

Удобно разделить функционал $F[\rho(\mathbf{r})]$ на части:

$$F[\rho] = T_0[\rho] + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{XC}[\rho] \equiv T_0[\rho] + E_H[\rho] + E_{XC}[\rho], \quad (1.31)$$

где $T_0[\rho]$ — кинетическая энергия невзаимодействующего электронного газа плотности ρ , находящегося в основном состоянии, второе слагаемое $E_H[\rho]$ — классическая кулоновская энергия (энергия Хартри) и остаток — $E_{XC}[\rho]$, который называется обменно-корреляционной энергией. Теперь мы можем записать функционал энергии (1.28) в виде:

$$E[\rho(\mathbf{r})] = T_0[\rho(\mathbf{r})] + \int v(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + E_H[\rho(\mathbf{r})] + E_{XC}[\rho(\mathbf{r})]. \quad (1.32)$$

Минимизация $E[\rho(\mathbf{r})]$ по отношению к $\rho(\mathbf{r})$ приводит к уравнению:

$$\int \delta\rho(\mathbf{r}) \left\{ \frac{\delta T_0[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} + v(\mathbf{r}) + v_H(\mathbf{r}) + \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \right\} d\mathbf{r} = 0, \quad (1.33)$$

где потенциал Хартри $v_H(\mathbf{r})$ — это потенциал кулоновского взаимодействия электрона остальных зарядов: $v_H(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$. Поскольку число электронов фиксировано, то $\int \delta\rho(\mathbf{r}) d\mathbf{r} = 0$. Отсюда следует, что в (1.33) выражение в фигурных скобках равно некоторой постоянной величине, т.е.:

$$\frac{\delta T_0[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} + v(\mathbf{r}) + v_H(\mathbf{r}) + \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} = const. \quad (1.34)$$

Если применить вариационный принцип к системе из N невзаимодействующих электронов с эффективным потенциалом

$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + v_H(\mathbf{r}) + \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}, \quad (1.35)$$

то приходим к уравнению совпадающему с (1.34). Таким образом приходим к важному заключению — вариационные свойства исследуемого функционала энергии и функционала энергии системы невзаимодействующих электронов, находящихся в эффективном поле (1.35) совпадают.

Таким образом, можно найти распределение истинной плотности электронного газа через задачу о невзаимодействующих электронах с эффективным потенциалом $v_{eff}(\mathbf{r})$. Для этого достаточно решить уравнение Шредингера, аналогичное одночастичным уравнениям Хартри–Фока (1.13):

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial \mathbf{r}^2} + v_{eff}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}), \quad (1.36)$$

и для N низших по энергии собственных состояний определить электронную плотность системы:

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2. \quad (1.37)$$

Получившаяся система (1.35), (1.36) и (1.37) получила название «самосогласованных уравнений Кона–Шэма».

Для вычисления полной энергии системы заметим, что

$$\sum_i \varepsilon_i = \sum_i \langle \psi_i | -\frac{1}{2} \frac{\partial^2}{\partial \mathbf{r}^2} + v_{eff}(\mathbf{r}) | \psi_i \rangle = T_0[\rho] + \int v_{eff}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}. \quad (1.38)$$

Тогда согласно (1.32) и (1.35) имеем

$$E_{tot} = \sum_i \varepsilon_i - E_H[\rho(\mathbf{r})] + E_{XC}[\rho(\mathbf{r})] - \int \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \rho(\mathbf{r}) d\mathbf{r}. \quad (1.39)$$

Следует подчеркнуть, что собственные значения уравнений Кона–Шэма могут отличаться от истинных, физических энергий электронов в исследуемой системе, поскольку являются решениями эффективного, а не истинного гамильтониана. При этом энергия и электронная плотность основного состояния, согласно теоремам Хохенберга–Кона, конечно, совпадают с истинными.

Приближения для обменно-корреляционной энергии

Практическое применение теории функционала плотности полностью определяется наличием способа построения приемлемого выражения для обменно-корреляционной энергии $E_{XC}[\rho(\mathbf{r})]$.

- Приближение локальной плотности.

Кон и Шэм сумели предложить такой способ, который получил название приближения локальной плотности (LDA) [9]. В этом приближении обменно-корреляционная энергия может быть записана следующим образом:

$$E_{XC}[\rho(\mathbf{r})] = \int \varepsilon_{XC}[\rho(\mathbf{r})]\rho(\mathbf{r}) d\mathbf{r}, \quad (1.40)$$

где $\varepsilon_{XC}[\rho(\mathbf{r})]$ — обменно-корреляционная энергия на одну частицу в однородном электронном газе, имеющем плотность $\rho(\mathbf{r})$. Тем самым задача об обмене и корреляции в неоднородном электронном газе сводится к задаче расчета плотности обменно-корреляционной энергии однородного электронного газа, которая решается уже гораздо легче. Приближение локальной плотности является точным в пределе плавно меняющейся плотности электронов и особенно хорошо работает в случае металлов. Практические расчеты показали, что энергетические уровни атомов и твердых тел определяются с большой точностью, так, в настоящее время когезионные свойства всех металлов получены теоретически с точностью до 10%.

- Обобщенное градиентное приближение (GGA)

Дальнейшее развитие метода построения обменно-корреляционной энергии связано с обобщенным градиентным приближением (GGA) (Generalized Gradient Approximation) [9]. В этом приближении обменно-корреляционный функционал зависит не только от плотности, но и от ее первой пространственной производной. Для расчета твердых тел наиболее широко используется обобщенно-градиентное приближение, предложенное группой Педью (Perdew) [10, 11, 12]. Такое приближение удачно исправляет многие дефекты LDA. Так, например, с помощью GGA было верно определено основное состояние для ферромагнетика Fe и антиферромагнетиков Cr и Mn. Также были правильно предсказаны магнитообъемные и магнитоструктурные эффекты. Однако существуют случаи, где использование GGA ведет к отсутствию связывания между атомами. Например, в случае димеров атомов благородных газов и молекулярного кристалла N₂ использование GGA приводит к абсурдному результату: отсутствию связи в данных структурах.

1.2.3. Вычислительная схема метода ТФП

Самосогласование уравнений Кона–Шэма представляет собой трудоемкую и непростую задачу. Опишем в общих чертах схему самосогласованного расчета, представленную на рис. 1.1. Сначала тем или иным

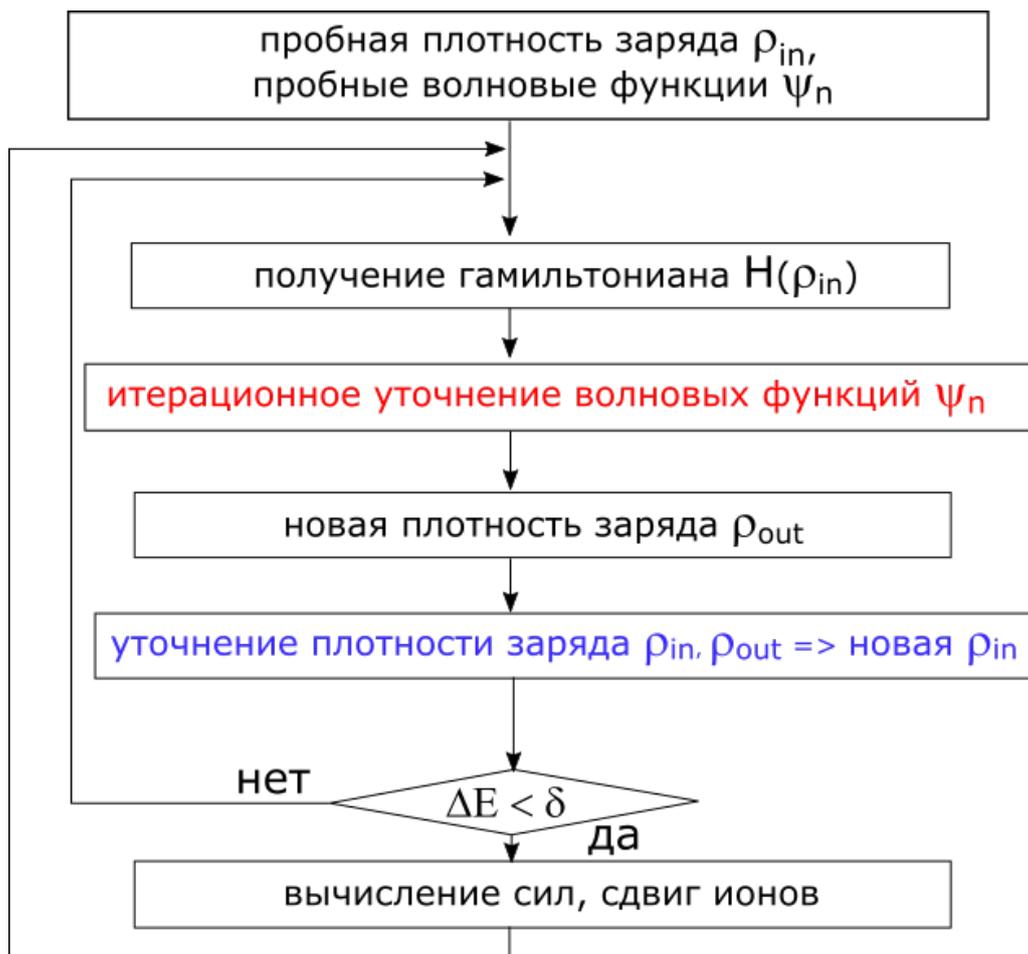


Рис. 1.1. Блок-схема самосогласованного метода расчета электронной структуры на основе теории функционала плотности.

способом строится начальное приближение для электронной плотности и волновой функции. Обменно-корреляционный потенциал определяется в рамках одного из приближений, описанных в 1.2.2. С помощью решения уравнений Кона–Шема (1.36) находятся новые волновые функции, по которым снова определяется электронная плотность. Электронная плотность «подмешивается» к старой и получается ее значение для следующей итерации. Критерий самосогласованности заключается в требовании, чтобы электронная плотность, получающаяся в результате m -ой итерации, равнялась, в пределах заданной точности, электронной плотности на ее входе. Условие согласования обычно заменяется более простым условием постоянства полной энергии системы при итерациях.

Наиболее трудоемкой частью расчета в рамках ТФП, безусловно, является решение уравнений (1.36) и получение новой электронной плотности. Поэтому эффективность схемы полностью определяется эффективностью выбранного метода решения одноэлектронной задачи. Распространены два подхода к поиску самосогласованного решения:

- В традиционном подходе основную роль играет электронная плотность. Именно она передается от итерации к итерации и модифицируется на каждом шаге. Для поиска следующего значения решается задача нахождение собственных значений гамильтониана с известным потенциалом:

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial \mathbf{r}^2} + v_{eff}(\mathbf{r}, \rho(\mathbf{r}')) \right) \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}).$$

Полученные функции дают новую плотность и, если она не совпадает с исходной, вычисления продолжаются.

- Можно непосредственно искать минимум функционала энергии, градиент которого дается выражением

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial \mathbf{r}^2} + v_{eff}(\mathbf{r}, \psi_i(\mathbf{r}')) - \varepsilon_i \right) \psi_i(\mathbf{r}).$$

При этом основными параметрами являются волновые функции или коэффициенты их разложения по базису. Они передаются между итерациями и для их изменения используются методы оптимизации. Эта точка зрения на задачу поиска основного электронного состояния появилась благодаря методу Кара–Паринелло, описанному в разделе 1.1.2. Они показали, что для быстрой работы с системами, содержащими много электронов, эффективней находить приближенное распределение электронов, которое, впрочем, может быть сколь угодно близко к истинному.

На сегодняшний день предложен ряд эффективных вычислительных алгоритмов как для поиска собственных значений, так и для поиска минимума энергии. Они во многом схожи. Рассмотрим их.

Решение задачи на собственные значения

Большинство методов расчета электронной структуры, использующих одночастичное приближение, имеют схожую структуру. В ее основе лежит вариационный принцип Рэля–Ритца. Будем искать решение уравнения Кона–Шэма (1.36) в виде линейной комбинации функций из некоторого базиса φ_m с неизвестными коэффициентами C_{mi} , которые подлежат определению:

$$|\psi_i\rangle = \sum_m C_{mi} |\varphi_m\rangle. \quad (1.41)$$

Если выполнены условия

$$\sum_m \langle \varphi_n | \hat{H} - \varepsilon_i \hat{S} | \varphi_m \rangle C_{mi} = 0, \quad (1.42)$$

то собственные значения ε_i и собственные функции (1.41) дают наилучшее приближение для истинных волновых функций в подпространстве, натянутом на набор векторов $\{\varphi_m\}$.

Если базис является полным, то после определения коэффициентов C_{mi} получается набор истинных волновых функций и значений энергии. На практике необходимо ограничиться конечным числом членов этого ряда. Это приводит к тому, что волновая функция не будет точно задаваться суммой (1.41), а полученные значения энергии будут несколько больше истинных.

Мерой отклонения найденных значений ψ_i от минимальных может служить вектор невязки $R(\psi_i)$.

$$|R(\psi_i)\rangle = (\hat{H} - \varepsilon_i \hat{S}) |\psi_i\rangle, \quad (1.43)$$

где приближенные собственные значения это

$$\varepsilon_i = \frac{\langle \psi_i | \hat{H} | \psi_i \rangle}{\langle \psi_i | \hat{S} | \psi_i \rangle}. \quad (1.44)$$

Вектор невязки равен нулю, если найдено точное значение минимума, и отличен от нуля в противном случае, т.е. его норма — это мера отклонения значений от истины. Вектор невязки используют для изменения волновой функции

$$\psi_i = \psi_i + \lambda R(\psi_i), \quad (1.45)$$

а алгоритмы отличаются лишь способом, которым это делается.

Эффективность итерационной процедуры повышается, если выровнять модули собственных значений матрицы $\hat{H} - \varepsilon_i \hat{S}$, что сделает число обусловленности матрицы меньше. Для этого можно умножить матрицу на предобуславливатель \hat{K} — любую положительно определенную матрицу. Условие обнуления невязки (1.43) для минимальных функций при этом не меняется. Для метода итераций идея состоит в том, чтобы найти матрицу, которая умножаясь на вектор невязки, дает точную ошибку в волновой функции. В идеале, после добавления такого произведения к текущей итерации (формула (1.45)), мы получим точный результат. Формально предобуславливатель — это функция Грина, которая может быть записана так:

$$\hat{K} = \frac{1}{\hat{H} - \varepsilon_i \hat{S}},$$

где ε_i — точное собственное значение энергии интересующей зоны. Конечно, получение этой матрицы невозможно без решения задачи на собственное значения, однако можно найти простое приближение для нее.

Приближенная матрица может быть не очень точной, но важно, чтобы она была положительно определена, т.е. получающиеся значения не обращались в ноль.

В случае базиса плоских волн с волновыми векторами G известно, что для больших G -векторов в гамильтониане преобладают диагональные элементы, т.е. кинетическая энергия квадратично растет с увеличением G :

$$H_{G,G'} \rightarrow \delta_{G,G'} \frac{\hbar^2}{2m} G^2.$$

Матрицу \hat{K} можно аппроксимировать диагональной матрицей, элементы которой сходятся к $\frac{2m}{\hbar^2 G^2}$ для больших G -векторов и постоянны для малых G -векторов. Форма такой функции была предложена Тетером [13] и модифицирована Крессе [14]:

$$K = - \sum_q \frac{2|q\rangle\langle q|}{3/2E_{kin}(R)} \frac{27 + 18x + 12x^2 + 8x^3}{27 + 18x + 12x^2 + 8x^3 + 16x^4},$$

где $x = \frac{\hbar^2}{2m} \frac{G^2}{3/2E_{kin}(R)}$, а $E_{kin}(R)$ — кинетическая энергия, вычисленная для вектора невязки.

Предобусловленный вектор невязки задается формулой

$$|R_{Ki}\rangle = \hat{K} |R_i\rangle = \hat{K} \left(\hat{H} - \varepsilon_i \hat{S} \right) |\psi_i\rangle. \quad (1.46)$$

Реализация алгоритмов

Самым прямым способом поиска волновых функций является метод простой итерации Девидсона. Для каждой зоны выбирается пробный вектор ψ_i и вычисляется предобусловленный вектор невязки R_{Ki} , что приводит к набору $b_i = \{\psi_i, R_{Ki}\}$ размером $2 \cdot N_{bands}$. В этом подпространстве вычисляются N_{bands} низших собственных функций через решение задачи на собственные значения:

$$\langle b_i | \hat{H} - \varepsilon_j \hat{S} | b_j \rangle = 0. \quad (1.47)$$

Эта процедура называется оптимизацией Релея–Ритца или вращением подпространства. Найденные волновые функции используются на следующем шаге.

Итерационная схема Девидсона оптимизирует все зоны одновременно. Это неудобно, если зон много. Последовательная оптимизация зон позволяет сохранить память и применить градиентные методы поиска.

Градиент можно получить, если преобусловленный вектор невязки (1.46) ортогонализировать к текущему набору волновых функций:

$$|g_i\rangle = \left(1 - \sum_j |\psi_j\rangle \langle \psi_j| \hat{S}\right) \hat{K} \left(\hat{H} - \varepsilon_i \hat{S}\right) |\psi_i\rangle, \quad (1.48)$$

а затем найти линейную комбинация нового направления $|g_i\rangle$ и текущей волновой функции $|\psi_i\rangle$, которая минимизирует математическое ожидание гамильтониана. Это требует решения задачи (1.47) размерности 2×2 с базисным набором $b_i = \{\psi_i, g_i\}$. Найденное значение используется для коррекции волновой функции (1.45).

Схему, предложенную выше, можно сделать более эффективной, если на каждом шаге итерации расширять базисный набор, упомянутый в предыдущем разделе. Т.е. на N -м шаге итерации задача на собственные значения (1.47) решается с использованием базисного набора $b_i = \{\psi_i, g_i^1, g_i^2, g_i^3, \dots\}$, а новый собственный вектор с минимальным значением энергии используется для вычисления нового вектора поиска g_i^N по формуле (1.48).

Вместо предыдущей схемы итераций можно оптимизировать математическое ожидание гамильтониана с помощью метода сопряженных градиентов. Первый шаг этого метода соответствует поиску градиента, как описано выше. На всех последующих этапах градиент g_i^N сопряжен с предыдущим направлением поиска. Как показал Крессе [14], метод сопряженных градиентов лишь немного уступает в эффективности блочному методу Девидсона, описанному далее.

Метод сопряженных градиентов является довольно стабильным и эффективным, но у него есть несколько недостатков. Во-первых, необходимо проводить процедуру ортогонализации на каждом шаге итерации, что существенно замедляет вычисления для больших систем. Во-вторых, этот метод находит не сами собственные вектора, а некоторые их линейные комбинации. Это не существенно для диэлектриков и полупроводников, т.к. дальнейшее суммирование происходит по всей зоне Бриллюэна (см. раздел 1.4.1). Но для металлов нам нужны истинные энергии и состояния, чтобы отделить заполненную часть зоны Бриллюэна от незаполненной. В этом случае обычно добавляют еще один шаг, который называется поворот подпространства. На нем окончательно найденные собственные функции ψ_i при помощи унитарного преобразования поворачивают так, чтобы матрица Гамильтона стала диагональной в подпространстве, натянутом на эти функции. Для этого необходимо найти матричное представление гамильтониана и, возможно, матрицы перекрытия, а также решить стандартную задачу на собственные значения (1.47) размерности N_{bands} . В

результате получается линейное замедлению работы с увеличением большего из размеров (супер)ячейки.

Оптимальная схема вычислений поиска собственных векторов это блочная схема Девидсона (block Davidson scheme). В ней все зоны разбиваются на блоки и для каждого блока запускается процесс оптимизации методом Рэлея–Ритца (1.47) с расширяющимся базовым набором, как описано в разделе 1.2.3. Затем блок функций заменяется на оптимизированный и процесс запускается для следующего блока. В конце алгоритма проводится оптимизация Рэлея–Ритца (поворотом подпространства) для всех полученных функций вместе. Алгоритм метода представлен ниже.

- Выбираем из N_{band} функций подмножество размерности n_1 : $\{\phi_i^1\}$.
 - Добавляем к подмножеству ортогонализированные вектора невязки (1.48) и получаем набор векторов размерности $2n_1$: $\{\phi_i^1, g_i^1\}$.
 - Применяем оптимизацию Рэлея–Ритца в пространстве, натянутом на эти векторы и определяем из них n_1 низших по энергии: $\{\phi_i^2\}$.
 - Добавляем к множеству векторов новые вектора невязки, вычисленные для полученных функций, и получаем набор размерности $3n_1$: $\{\phi_i^1, g_i^1, g_i^2\}$.
 - применяем оптимизацию Рэлея–Ритца в пространстве размерности $3n_1$, натянутом на эти векторы и опять определяем из них n_1 низших по энергии: $\{\phi_i^3\}$.
 - Продолжаем процесс, добавлением четвертого и т.д. набора векторов, если потребуется. Если итерация завершена, волновые функции $\{\phi_i\}$ в полном наборе заменяются на найденные оптимизированные функции $\{\phi_i^m\}$.
 - Выбираем следующий набор векторов $\{\phi_j^1\}$ и повторяем для него шаги оптимизации, описанные выше.
- После оптимизации всех блоков выполняется оптимизация Рэлея–Ритца во всем искомом подпространстве размерности N_{band} .

Блочный метод Девидсона приблизительно в 1,5–2 раза медленнее, чем метод DIIS, описанный ниже, но он всегда стабилен.

Рассмотрим еще один эффективный метод поиска нижних собственных значений. Это прямое обращение (части матрицы) в итерационном подпространстве (DIIS). Для этого реализуется двух шаговая процедура.

- Начинаем с произвольного набора векторов (волновых функций) размерности N_{bands} .
- Применяем оптимизацию Рэлея–Ритца для подпространства натянутого на эти вектора (поворот подпространства).
- По очереди оптимизируем эти вектора независимо друг от друга 2 или 3 раза.

В простейшем случае можно рассматривать оптимизацию Рэлея–Ритца, как поиск унитарного преобразования для набора векторов в подпространстве: $\psi_i = \sum_{i'} U_{ii'} \psi_{i'}$, которое приводит подматрицу матрицы гамильтониана к диагональному виду: $\langle \psi_i | \hat{H} | \psi_j \rangle = \delta_{ij} \varepsilon_i$, при условии, что исходный набор векторов ортогональный. Вычисление элементов подматрицы требует порядка N_{band}^2 операций, что довольно медленно.

Улучшить производительность можно, если вместо оптимизации математического ожидания гамильтониана, как это делалось в предыдущих методах (1.2.3–1.2.3), минимизировать норму вектора невязки методом прямого обращения в итерационном пространстве (RMM–DIIS). При этом каждый вектор оптимизируется индивидуально, что легко реализуется в параллельных программах, т.к. каждый узел работает с собственным набором функций. При этом имеется линейное масштабирование N_{bands} .

Это приводит к итерационной схеме аналогичной схеме из раздела 1.2.3, но с другой задачей на собственные значения. Норма вектора невязки дается выражением

$$\langle R_i | R_i \rangle = \langle \phi_n | (H - \varepsilon)^+ (H - \varepsilon) | \phi_n \rangle,$$

и обладает квадратичным минимумом на каждой собственной функции ϕ_n без дополнительных ограничений. В этом заключается главное его преимущество перед традиционными методами оптимизации. Этот алгоритм можно использовать без сведений о других волновых функциях и, следовательно, без явной ортогонализации предобусловленного вектора невязки.

Можно рассматривать вектор невязки как функционал электронной плотности: $R[\rho_{in}] = \rho_{out}[\rho_{in}] - \rho_{in}$. DIIS алгоритм используется для оптимизации нормы вектора невязки. Если использовать линейное представление для вектора невязки $R[\rho_{in}]$ около минимального значения ρ_0 (теория линейного отклика), то получится $R[\rho] = -J(\rho - \rho_0)$ с диэлектрической функцией $J = 1 - \xi U$, что дает $R[\rho_{in}] = \rho_{out}[\rho_{in}] - \rho_{in} = J(\rho_{in} - \rho_0)$.

Спектр значений J определяет сходимость: чем шире спектр, тем медленнее сходимость. Для полупроводников и изоляторов ширина спектра собственных значений постоянна и не зависит от размера. Для металлов, напротив, спектр расходится и его ширина пропорциональна квадрату максимального размера (супер)ячейки: для коротких волн $J \approx 1$, для длинных волн $J \approx 1/q^2 \propto L^2$.

Обычно для старта в большинстве систем используют модельную диэлектрическую функцию вида $J^{-1} \propto G_q^1 = \max(\frac{q^2 AMIX}{q^2 + BMIX}, AMIN)$. Это начальное значение для диэлектрической матрицы улучшается на каждом шаге (DIIS).

В этом случае, после определения всех зон, необходима ортогонализация Грама–Шмидта для получения нового ортогонального пробного базисного набора. Без явной ортогонализации к текущему набору пробных волновых функций все остальные алгоритмы имеют тенденцию сходиться к нижней зоне, независимо от того, из какой зоны они начинаются.

У метода DIIS есть недостаток: для больших систем можно потерять некоторые собственные значения и не существует простого способа узнать, что это произошло. Иногда потерю собственных значений можно вообще не заметить, а в «лучшем» случае можно заметить, что нет сходимости, или сходимость может замедлиться после некоторого числа шагов. В этом случае нужно использовать блочную схему Девидсона.

Контрольные вопросы

1. Перечислите способы сведения многоэлектронной задачи к одноэлектронной.
2. Почему метод теории функционала плотности наиболее востребован сегодня?
3. Сформулируйте две теоремы Хохенберга – Кона. Какое преимущество они дают в сравнении с другими методами?
4. Запишите согласованные уравнения Кона – Шэма.
5. Перечислите приближения, используемые для записи обменно-корреляционного функционала энергии.
6. Перечислите основные этапы решения уравнений Кона–Шэма. Какие вычислительные алгоритмы используются на каждом этапе?

1.3. Потенциалы и базисы

Для быстрого решения уравнения Кона–Шэма необходимо, чтобы разложение (1.41) содержало как можно меньше слагаемых при обеспечении удовлетворительного качества описания истинной волновой функции. Второе требование — возможность быстрого определения слагаемых, входящих в эффективный потенциал (1.35) или в функционал энергии (1.32). Для этого важно, чтобы функции базиса были близки по форме к истинным волновым функциям, т.е. эффективный потенциал должен подходить к выбранному базису и наоборот. Оптимальный выбор базиса является непростой задачей, поскольку внутри атомного остова волновая функция и потенциал близки к своим значениям для свободного атома (т.е. обладают сферической симметрией и сильной зависимостью от радиального расстояния r , тогда как в межатомной области потенциал и волновая функция меняются очень плавно. Отсюда возникают трудности

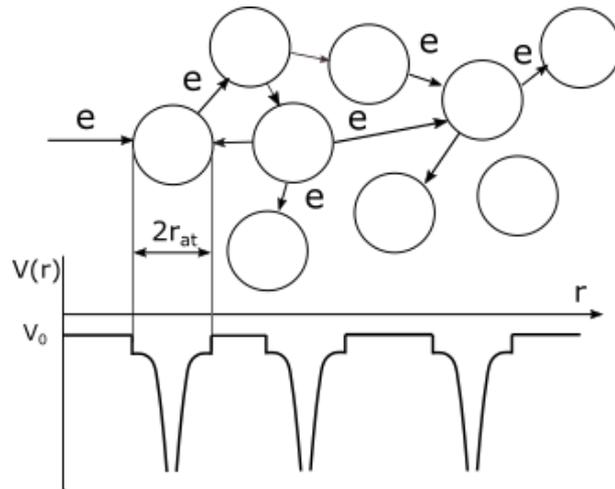


Рис. 1.2. Схематическое изображение поиска электронной структуры, как задачи рассеяния электронов на атомных сферах, и примерный вид потенциала (потенциал в межсферной области постоянный).

единого аналитического описания волновых функций и потенциала во всем пространстве.

Выходом из ситуации является отдельное рассмотрение сферических областей определенного радиуса вокруг ионов (атомных сфер) и оставшейся части (межатомного пространства). В атомных сферах потенциал принимается сферически симметричным и близким к потенциалу свободного иона. За пределами сфер потенциалы ионов перекрываются, но образованный таким образом потенциал межсферной области является плавным и достаточно однородным. Таким образом, приходим к картине потенциала представленной на рис. 1.2, где в области постоянного или слабо меняющегося потенциала расположены сферически симметричные потенциальные ямы ионов. Движение электронов в этом потенциале можно описать с помощью плоских волн, которые распространяются в промежуточной области и многократно рассеиваются на сферических ямах ионных потенциалов. Такой способ построения потенциала включает как представление о свободных электронах, распространяющихся в межсферном пространстве, так и влияние сильных ионных потенциалов, т.е. является универсальным и применим для любых материалов, что и обусловило его широкое применение для расчетов зонной структуры твердых тел.

Конкретные реализации подобного разбиения основаны на общих представлениях о свойствах тех или иных материалов. Можно выделить три основные группы методов:

- Методы псевдопотенциала: потенциал внутри сфер заменяется более плавным, экранированным электронами близкими к ядру, что позволя-

ет использовать единый набор плавных базисных функций — плоских волн во всем пространстве. Метод псевдопотенциала будет подробно рассмотрен чуть ниже.

- Методы сильной связи: межсферная область исключается из рассмотрения, а сами атомные сферы задаются перекрывающимися. Полный потенциал представляется непосредственно в виде суммы атомных потенциалов. При этом базисными функциями являются атомные орбитали, локализованные на различных центрах, которые вычисляются через произведение функции углового момента и радиальной функции. Такой выбор базиса позволяет построить полную волновую функцию, похожую на совокупность волновых функций отдельных атомов. Это обеспечивает быструю сходимость к искомой волновой функции в случае локализованных электронов. Хорошо локализованные орбитали позволяют, в принципе, получить линейное масштабирование при увеличении размера системы. Однако базис зависит от положений ионов, поэтому при вычислении необходимо корректировать вычисление сил Гельмана–Фейнмана. Трудно контролировать полноту базиса и сходимость результата.
- Методы присоединенных волн: и потенциал и волновые функции содержат 2 части с разным аналитическим описанием, которые сшиваются на границе каждой из сфер. Для ячеечного потенциала уравнение Шредингера может быть точно решено и в области сферического ионного потенциала и в межсферной области с плавным потенциалом. Поэтому можно сразу получить базисные функции – очень похожие на точные решения уравнения Шредингера. Использование небольшого числа этих функций в вариационной процедуре уже позволяет определять электронную структуру с хорошей точностью. Существует две разновидности данного метода в которых сферические волны внутри атомных сфер сшиваются в межатомной области с плоскими волнами (метод присоединенных плоских волн, ППВ или APW) или со сферическими волнами, центрированными на других узлах (метод маффин-тин орбиталей, МТО).

При очевидных достоинствах у такого подхода есть недостаток: при построении базиса нужно задать уровень энергии (для решения уравнения Шредингера), поэтому полученный базис зависит от энергии. Это делает задачу нахождения волновой функции нелинейной. Для возвращения задачи к привычному виду базисные функции раскладываются в ряд по энергии до первого порядка. Получаются так называемые линейные методы LAPW и LMTO.

Современные версии метода LAPW определяют электронную плотность во всем пространстве, в том числе и в сложной межатомной

области, самосогласованным способом. Этот метод получил название LAPW с полным потенциалом (FP-LAPW). Он на сегодняшний день является самым точным из методов расчета, однако он же является и самым трудоемким. Базис APW можно назвать промежуточным между плоскими волнами и атомными орбиталями, что определяет его универсальность. К его недостаткам можно отнести трудность построения и сложность проведения компьютерных расчетов из-за большого массива данных и невысокой эффективности вычислительных алгоритмов.

1.3.1. Метод псевдопотенциала

Базис из гладких функций на первый взгляд слабо подходит для твердых тел, состоящих из ионов и электронов, которые взаимодействуют посредством сильного кулоновского потенциала. Тем не менее согласно теории ферми-жидкости электронные возбуждения вблизи энергии Ферми можно рассматривать как слабо взаимодействующие квазичастицы с перенормированной массой. Таким образом, кроме своей фундаментальной роли в теории металлов, теория ферми-жидкости обеспечивает базу для введения эффективных плавных потенциалов (псевдопотенциалов, PP). Мы можем ожидать, что потенциал Хартри и обменно-корреляционный потенциал, представляющие собой действие на выделенный валентный электрон со стороны остальных валентных электронов, значительно слабее, чем «голый» кулоновский потенциал, и слаб настолько, что во многих случаях валентные электроны можно считать почти свободными. Однако остается еще сильное взаимодействие внутренних (остовных) электронов с ядрами. В большинстве случаев остовные электроны сильно связаны и слабо откликаются на движение внешних валентных электронов, поэтому они могут рассматриваться как фиксированные, не зависящие от остовных состояний и от взаимного расположения ионов. Именно в этом и состоит приближение псевдопотенциала. Изменением состояний остовных электронов нельзя пренебречь в некоторых случаях, но его можно учесть дополнительно, например в методе проективно присоединенных волн (PAW), который будет рассмотрен отдельно.

Сильный потенциал внутри остова V (помимо кулоновского потенциала он включает потенциал Хартри от электронов остова и обменно-корреляционный потенциал от валентных электронов) заменяется на слабый псевдопотенциал \tilde{V} , как показано на рисунке 1.3. Волновая функция основного состояния $\tilde{\varphi}$, полученная для псевдопотенциала, повторяет поведение полноэлектронной волновой функции за пределами сферы

радиуса r_c . Для большинства элементов результирующая псевдоволновая функция $\tilde{\varphi}$ является весьма плавной и может быть с хорошей точностью представлена несколькими плоскими волнами с низшими значениями волнового вектора. По этой причине плоские волны являются простым и эффективным базисом для псевдоволновой функции. Конечно, за это приходится заплатить определенную цену. Эта цена – необходимость генерировать и использовать псевдопотенциал, вместо реального кристаллического потенциала. Действительно, большинство трудностей переносится с этапа получения электронной структуры на этап генерации псевдопотенциала.

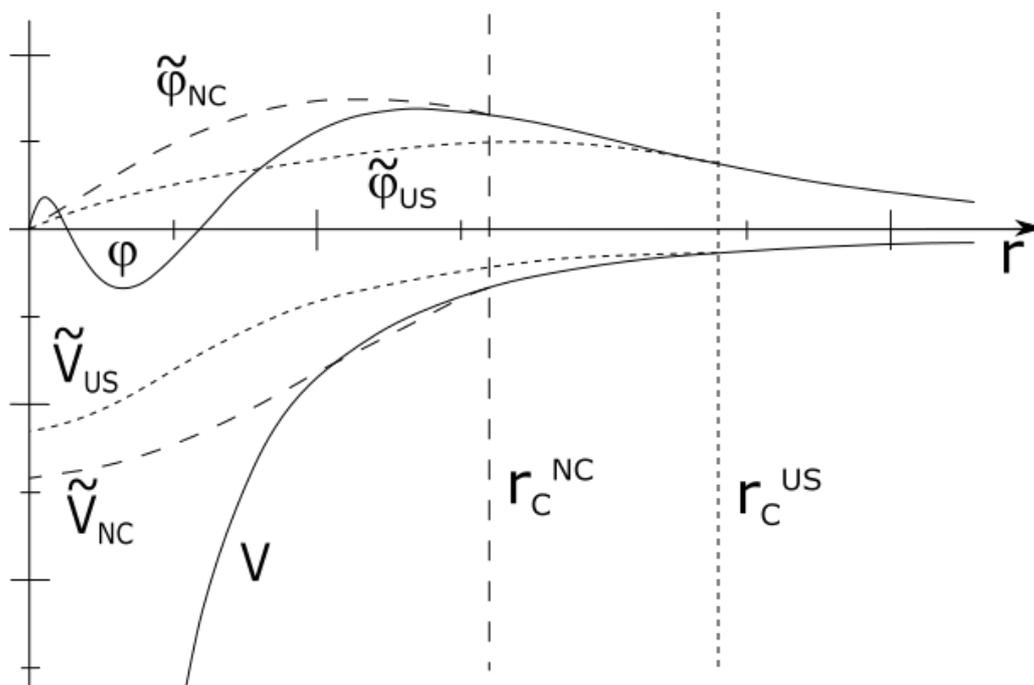


Рис. 1.3. Схема замены потенциала V и полнэлектронной волновой функции φ на псевдопотенциал \tilde{V} и псевдоволновую функцию $\tilde{\varphi}$. Показаны различия в построении нормосохраняющего (NC) и ультрамягкого (US) псевдопотенциалов (обсуждение дальше в тексте).

Построение Филипса–Клейнмана.

Подход псевдопотенциала возник из метода ортогонализированных плоских волн (OPW) Херринга, в котором валентные волновые функции раскладывались по базису плоских волн ϕ_i , дополнительно ортогонализированных к нижележащим основным состояниям φ_c :

$$\phi'_i = \phi_i - \sum_{\alpha,c} \langle \varphi_c | \phi_i \rangle \varphi_c, \quad (1.49)$$

где сумма берется по всем остовным состояниям c и всем атомам α . Использование такого базиса можно рассматривать, как решение обычной задачи с некоторым модифицированным гамильтонианом.

ПП в рамках этого подхода может быть построен следующим образом. Пусть \hat{H} – исходный гамильтониан с остовными φ_c и валентными φ_v волновыми функциями. Рассмотрим теперь псевдосостояния

$$\tilde{\varphi} = \varphi_v + \sum_{\alpha,c} a_{vc} \varphi_c, \quad (1.50)$$

где $a_{vc} = \langle \varphi_c | \tilde{\varphi} \rangle$. Действуя на такое состояние гамильтонианом \hat{H} , мы получаем

$$\hat{H}\tilde{\varphi} = \varepsilon_v \varphi_v + \sum_{\alpha,c} a_{vc} \varepsilon_c \varphi_c = \varepsilon_v \tilde{\varphi} + \sum_{\alpha,c} a_{vc} (\varepsilon_c - \varepsilon_v) \varphi_c, \quad (1.51)$$

где ε_c и ε_v собственные значения энергии остовных и валентных электронов соответственно. Тогда, пользуясь определением величины a_{vc} , находим

$$\left[\hat{H} + \sum_{\alpha,c} a_{vc} (\varepsilon_c - \varepsilon_v) |\varphi_c\rangle \langle \varphi_c| \right] \tilde{\varphi} = \tilde{\varepsilon} \tilde{\varphi}. \quad (1.52)$$

Таким образом, псевдосостояния удовлетворяют уравнению Шредингера с дополнительным вкладом \tilde{V} в гамильтониан,

$$\tilde{V} = \sum_{\alpha,c} (\varepsilon_c - \varepsilon_v) |\varphi_c\rangle \langle \varphi_c|, \quad (1.53)$$

где \tilde{V} отличается от обычных вкладов в потенциальную энергию тем, что это слагаемое зависит от энергии через величины ε_c . Добавляя данный вклад к исходному потенциалу V , содержащемуся в гамильтониане \hat{H} , получаем псевдопотенциал Филлипса–Клейнмана:

$$V_{PK} = V + \tilde{V}. \quad (1.54)$$

За пределами области, занимаемой остовом, V_{PK} становится равным V , поскольку здесь остовные волновые функции обращаются в нуль. Таким образом, вокруг атома существует некий радиус, за пределами которого вклад атома в \tilde{V} пренебрежимо мал. Более того, конструкция псевдопотенциала линейна в том смысле, что вклады от каждого отдельного атома независимы друг от друга и аддитивны. Наиболее важным при этом является то, что псевдопотенциал значительно слабее, чем исходный потенциал, что приводит к быстрой сходимости ряда плоских волн для псевдволновой функции.

Современные псевдопотенциалы

Сложность и эффективность псевдопотенциала существенно выросли со времени работы Филиппса и Клейнмана. Это развитие мотивировалось достижением следующих целей:

1. Прежде всего псевдопотенциал должен быть мягким (soft), что означает его плавное («плоское») поведение в пространстве, позволяющее использовать разложение по небольшому количеству плоских волн;
2. Псевдопотенциал должен быть трансферабельным, т.е. псевдопотенциал сгенерированный для одной конфигурации атомов, должен правильно воспроизводить и другие конфигурации;
3. Псевдоволновая плотность (зарядовая плотность, построенная по псевдоволновой функции) должна воспроизводить валентную плотность заряда настолько точно, насколько это возможно.

Концепция нормсохраняющих (norm conserving, NC) трансферабельных псевдопотенциалов представляет важное продвижение к этим целям. На сегодняшний день для построения псевдопотенциала выполняют следующие операции [15], которые удобно проследить на рисунке 1.3:

1. Проводится полноэлектронный расчет для атома или иона на некотором уровне энергии ε .
2. Определяется радиус отсечки r_c захватывающий узлы волновой функции верхнего состояния, относящегося к остову.
3. Строится непрерывная и непрерывно дифференцируемая псевдоволновая функция $\tilde{\varphi}_l$, которая совпадает с полноэлектронной волновой функцией φ_l для $r > r_c$, но не имеет узлов в области $r < r_c$. Конечно, волновые функции и собственные значения для различных значений орбитального момента l электрона различаются, и это приводит к тому, что нужно строить разные потенциалы для разных значений l .
4. Если вспомнить, что волновая функция это совокупность интерферирующих электронов, испытывающих рассеяние на множестве атомных центров, то псевдопотенциал можно рассматривать как замену рассеивающего центра для движущегося почти свободного электрона на эквивалентный, во всяком случае обладающий теми же рассеивающими свойствами вне сферы радиуса r_c . При рассеянии сдвиг фазы у обычной φ_l и псевдоволновой $\tilde{\varphi}_l$ функций будет совпадать (по модулю 2π), если будут совпадать их логарифмические производные при $r = r_c$:

$$\left. \frac{\partial \ln(\tilde{\varphi}_l(r, \varepsilon))}{\partial r} \right|_{r=r_c} = \left. \frac{\partial \ln(\varphi_l(r, \varepsilon))}{\partial r} \right|_{r=r_c} \quad (1.55)$$

5. При построении должно выполняться условие сохранения нормы волновой функции внутри сферы радиуса r_c :

$$\int_0^{r_c} |\tilde{\varphi}_l(r)|^2 \cdot r^2 dr = \int_0^{r_c} |\varphi_l(r)|^2 \cdot r^2 dr, \quad (1.56)$$

где волновые функции записаны в системе отсчета связанной с центром атома. Сохранение нормы означает сохранение заряда внутри выбранной сферы, а значит и сохранение электростатического потенциала вне сферы, согласно теореме Гаусса.

В качестве удобной меры трансферабельности может служить условие 1.55. Это условие выполняется точно при ε , равном собственному значению внутри атома. Тогда мы можем охарактеризовать трансферабельность величиной области энергии, внутри которой 1.55 выполняется достаточно хорошо. Хамман [15] показал, что условие сохранения нормы 1.56 приводит к совпадению производных по энергии, т.к.

$$-\frac{\partial}{\partial E} \frac{\partial}{\partial r} \ln \varphi(r_c, E) = \frac{1}{r_c^2 |\varphi(r_c, E)|^2} \int_0^{r_c} |\varphi(r, E)|^2 \cdot r^2 dr. \quad (1.57)$$

Это означает, что совпадение логарифмических производных будет сохраняться и в некоторой окрестности энергии уровня. Или, другими словами, разность логарифмических производных псевдоволновой и полноэлектронной волновых функций есть величина второго порядка малости по отклонению энергии от значения референсного уровня и это позволяет обосновать трансферабельность нормсохраняющих псевдопотенциалов. Методы конструирования «мягких», полулокальных, нормсохраняющих псевдопотенциалов были развиты Хаманном [15] и затем усовершенствованы в работе Башеле, Хаманна и Шлютера [16], где они были протабулированы для всех элементов периодической системы. Табуляция псевдопотенциалов явилось очень важным шагом в развитии теории электронной структуры материалов, поскольку открыло дорогу *ab initio* расчетам, основанным на методе псевдопотенциала.

Несмотря на очевидные преимущества нормсохраняющих псевдопотенциалов (NC-PP) они не являются достаточно гладкими для *p*- и *d*-элементов. Поэтому для их использования приходится применять много базисных функций. Решить эту проблему можно применяя "ультрамягкие псевдопотенциалы (US-PP) [17]. Ультрамягкие псевдопотенциалы не сохраняют норму (см. рисунок 1.3), поэтому можно получать более "плоские" псевдоволновые функции, что очень удобно для *2p* и *3d* состояний. Это позволяет уменьшить размер базиса, т.е. увеличить производительность [18]. Расплачиваться за это приходится введением корректирующей

процедуры для добавления исключенного заряда внутрь каждой сферы. Построение ультрамягких потенциалов, как оказалось, является частным случаем более общего подхода — метода проективно присоединенных волн, который обсуждается ниже.

1.3.2. Метод проективно присоединенных волн (PAW)

Идея метода PAW предложена Башле [16]. В методе PAW удается сочетать точность полноэлектронных методов, таких как FP-LAPW с эффективностью псевдопотенциальных методов. Кратко его опишем.

Разделим все электроны системы на валентные электроны и электроны остова. Будем искать волновые функции в виде:

$$|\psi_i\rangle = |\tilde{\psi}_i\rangle + \sum_{l,m,\varepsilon}^{sites} (|\varphi_{lm\varepsilon}\rangle - |\tilde{\varphi}_{lm\varepsilon}\rangle) \langle \tilde{p}_{lm\varepsilon} | \tilde{\psi}_i \rangle, \quad (1.58)$$

где $|\psi_i\rangle$ и $|\tilde{\psi}_i\rangle$ — волновая и псевдоволновая функции, определенные во всем объеме, $|\varphi_{lm\varepsilon}\rangle$ и $|\tilde{\varphi}_{lm\varepsilon}\rangle$ — функции, центрированные на положениях ионов и определенные внутри не перекрывающихся сфер с радиусами r_c (атомные функции), $\langle \tilde{p}_{lm\varepsilon} |$ — проекционная функция, l, m — квантовые числа для момента импульса и его проекции у атомных функций, ε — некоторый уровень энергии.

Приближенное поведение произвольной псевдоволновой функции внутри конкретной сферы можно найти домножив ее на проекционный оператор и получив коэффициенты разложения $c_{lm\varepsilon}$:

$$c_{lm\varepsilon} = \langle \tilde{p}_{lm\varepsilon} | \tilde{\psi}_i \rangle. \quad (1.59)$$

Проекционный оператор проектирует произвольную функцию на состояния $|\tilde{\varphi}_{lm\varepsilon}\rangle$, т. е. должно выполняться условие $\langle \tilde{p}_{lm\varepsilon} | \tilde{\varphi}_{l'm'\varepsilon'} \rangle = \delta_{ll'} \delta_{mm'} \delta_{\varepsilon\varepsilon'}$.

В результате проектирования получим разложение псевдоволновой функции по набору атомных функций, которое справедливо в сфере радиуса r_c :

$$|\tilde{\psi}_i\rangle^{r < r_c} = (\approx) \sum_{l,m,\varepsilon} |\tilde{\varphi}_{lm\varepsilon}\rangle c_{lm\varepsilon} \quad (1.60)$$

Точность представления (1.60) зависит от полноты набора атомных функций. Подставляя (1.59) и (1.60) в (1.58) получаем, что истинная волновая

функция внутри сфер также раскладывается по набору атомных функций (других), но с такими же коэффициентами:

$$|\psi_i\rangle^{r < r_c} \approx \sum_{l,m,\varepsilon} |\varphi_{lm\varepsilon}\rangle c_{lm\varepsilon}. \quad (1.61)$$

В результате получаем следующее представление волновой функции (скобки опущены):

$$\psi_i = \tilde{\psi}_i - \sum_{l,m,\varepsilon}^{\text{sites}} \tilde{\varphi}_{lm\varepsilon} c_{lm\varepsilon} + \sum_{l,m,\varepsilon}^{\text{sites}} \varphi_{lm\varepsilon} c_{lm\varepsilon}. \quad (1.62)$$

Можно рассматривать эту формулу следующим образом: чтобы получить истинную волновую функцию, определенную во всем пространстве, мы берем гладкую псевдоволновую функцию и практически зануляем ее внутри атомных сфер при помощи второго слагаемого. Затем в атомные сферы добавляем сильно меняющиеся части, чтобы получить полную волновую функцию. Схематично эти операции показаны на рисунке 1.4.

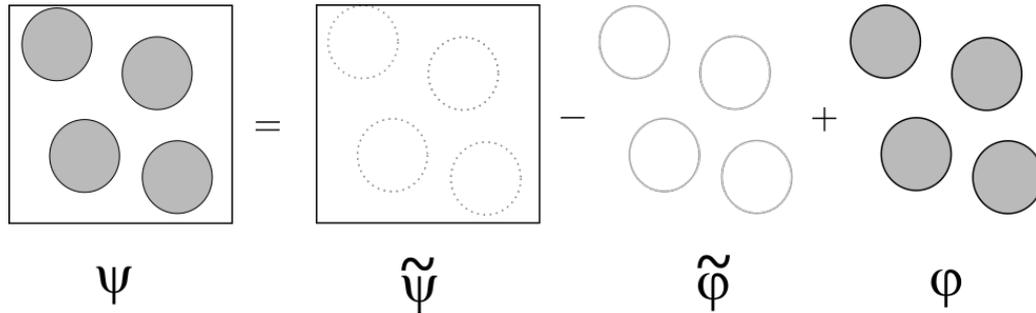


Рис. 1.4. Области определения функций, входящих в волновую функцию (1.62).

Схема построения функций, показанная на рисунке 1.4, применяется для всех величин, входящих в энергию и потенциалы: ρ , T , E_H , E_{XC} , V_H , V_{XC} ... Например для кинетической энергии $T = \sum_i f_i \langle \psi_i | -\Delta | \psi_i \rangle$ после подстановки волновой функции (1.62) получаем:

$$T = \tilde{T} - \tilde{T}^* + T^* = \sum_i f_i \langle \tilde{\psi}_i | -\Delta | \tilde{\psi}_i \rangle - \sum_{a,b}^{\text{site}} \rho_{ab} \langle \tilde{\varphi}_a | -\Delta | \tilde{\varphi}_b \rangle + \sum_{a,b}^{\text{site}} \rho_{ab} \langle \varphi_a | -\Delta | \varphi_b \rangle, \quad (1.63)$$

где f_i — числа заполнения, $\rho_{ab} = \sum_i f_i c_a c_b$ — матрица плотности внутри атомных сфер, а a и b наборы коэффициентов l, m, ε .

Так как волновые и псевдоволновые функции имеют разную норму внутри сфер, то необходимо скорректировать электростатический потенциал, т.е. добавить корректирующий заряд \hat{n} в энергию Хартри:

$$E_H = \tilde{E}_H - \tilde{E}_H^* + E_H^* = E_H[\tilde{n} + \hat{n}] - \sum_{site} E_H[\tilde{n}^* + \hat{n}^*] + \sum_{site} E_H[n^* + \hat{n}^*], \quad (1.64)$$

где \tilde{n}^* — псевдо-заряд в сфере, а \hat{n}^* — компенсационный заряд в сфере.

Каждое из слагаемых функционала энергии раскладывается на три части, поэтому полный функционал так же состоит из трех частей [19]: $E = \tilde{E} - \tilde{E}^* + E^*$. \tilde{E} вычисляется на обычной сетке, для вычисления функционала Кона–Шема используется базис плоских волн с добавочным компенсационным зарядом, чтобы исправить норму волновой функции. \tilde{E}^* и E^* вычисляются на сферических сетках, расположенных вокруг каждого атома. При этом вычисления на стандартных и сферических сетках полностью независимы, что обеспечивает высокую эффективность алгоритма.

Ультратрагкие потенциалы (US–PP), предложенные Вандербильдом [17] и обсуждаемые выше, не имели на тот момент достаточно строгого обоснования. Оказалось [20], что выражения для US–PP можно строго получить из функционала PAW после линеаризации сферических слагаемых \tilde{E}^* и E^* около исходной атомной конфигурации. Это показывает тесную связь между этими двумя подходами, но также позволяет прогнозировать проблемы методе US–PP. Можно заключить, что чем больше окружение отдаляется от точки отсчета, тем менее аккуратным будет результат US–PP. Самой проблемной ситуацией является наличие магнетизма, однако и в этом случае оба метода дают практически совпадающие результаты.

Контрольные вопросы

1. Какого вида потенциал используется для конденсированных систем?
2. Перечислите наиболее распространенные базисные наборы. Каковы последствия выбора того или иного базиса?
3. Сформулируйте критерии хорошего базисного набора.
4. В чем заключается метод псевдопотенциала?
5. В чем разница между классическими, ультратрагкими и нормосохраняющими псевдопотенциалами?
6. В чем заключается метод проективно присоединенных волн? Выделите 3 его достоинства.

1.4. Симметрия системы и базис плоских волн

Абсолютное большинство методов расчета структуры конденсированных сред используют базис плоских волн. Это происходит по нескольким причинам:

- Историческая причина в том, что многие материалы дают зонную структуру, которая хорошо описывается приближением почти свободных электронов (металлические s- и p-элементы). Электроны при этом описываются плоскими волнами. Изначально метод псевдопотенциала разрабатывался для плоских волн.
- Практическая причина в том, что некоторые слагаемые в полной энергии и гамильтониане просто записываются для базиса плоских волн.
- Вычислительная причина — действие гамильтониана на базисные функции легко вычисляется при помощи быстрого преобразования Фурье (FFT).

Для периодических систем, какими являются кристаллы, волновой вектор является хорошим квантовым числом, т.е. плоские волны являются практически независимыми. Все электронные состояния такой системы $\Psi_{\mathbf{k}}$ удобно перенумеровать набором индексов \mathbf{k} . Одноэлектронные состояния $\psi_{i,\mathbf{k}}$ дополнительно имеют второй индекс — номер зоны.

Теорема Блоха утверждает, что в периодической структуре электронная волновая функция может быть записана в виде произведения периодической части $f_i(\mathbf{r})$ на плосковолновую часть:

$$\psi_{i,\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\mathbf{R}} f_i(\mathbf{r}), \quad (1.65)$$

где \mathbf{R} любой вектор трансляции, оставляющий гамильтониан системы неизменным. Периодическая часть волновой функции может быть разложена в дискретный ряд по плоским волнам:

$$f_i(\mathbf{r}) = \sum_G c_{i,G} \cdot e^{i\mathbf{G}\mathbf{r}}. \quad (1.66)$$

При этом волновые вектора \mathbf{G} , равным векторам обратной решетки рассматриваемого кристалла, т.е. выполняется условие $\mathbf{G} \cdot \mathbf{R} = 2\pi n$. В силу этого, каждая электронная волна может быть записана в виде суммы плоских волн вида:

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = \sum_G c_{i,\mathbf{k}+\mathbf{G}} \cdot e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}. \quad (1.67)$$

Вектора \mathbf{k} для удобства приводят к первой зоне Бриллюэна — элементарной ячейке в обратном пространстве, условно говоря ограничиваются

случаем $\mathbf{k} < \mathbf{G}$. Теорема Блоха, таким образом, заменяет проблему расчета бесконечного числа волновых функций расчетом конечного числа электронных волновых функций для каждой k -точки.

Внутри зоны Бриллюэна электронные состояния могут существовать только для набора k -точек, определяемого граничными условиями, наложенными на весь кристалл. При этом плотность разрешенных k -точек пропорциональна объему кристалла. Например бесконечное число электронов в твердом теле без границ дает бесконечное число k -точек равномерно распределенных по зоне Бриллюэна, причем каждой k -точке соответствует конечное число электронов. На практике можно ограничиться конечным числом точек, обеспечивающим приемлемую точность.

1.4.1. Интегрирование по зоне Бриллюэна

Для многих целей (вычисления плотности состояний, матричных элементов, функций отклика и т.д.) необходимо вычислять средние значения функций в зоне Бриллюэна:

$$I(\varepsilon) = \frac{1}{\Omega_{BZ}} \int_{BZ} F(\varepsilon_{i,\mathbf{k}}) d\mathbf{k}. \quad (1.68)$$

Удобнее всего, также как в любых методах численного интегрирования, заменить интеграл суммой вида $I(\varepsilon) = \sum_n \omega_{\mathbf{k}_n} F(\varepsilon_{i,\mathbf{k}_n})$, где \mathbf{k}_n — дискретный набор k -точек, а $\omega_{\mathbf{k}_n}$ — весовые коэффициенты.

Для выбора точек и коэффициентов придерживаются следующего алгоритма, предложенного Монкхорстом и Паком [21]:

- генерируем равномерную сетку в k -пространстве: $k_{prs} = u_p b_1 + u_r b_2 + u_s b_3$, где b_i — вектора обратной решетки, а $u_r = \frac{2r - q_r - 1}{2q_r}$ — набор коэффициентов и r меняется от 1 до q_r ;
- сдвигаем начало координат, если необходимо;
- применяем все операции симметрии решетки Браве ко всем k -точкам;
- по наложению k -точек выделяем неприводимые k -точки (неприводимая часть зоны Бриллюэна, IBZ);
- по количеству совпадений точек при операциях симметрии вычисляем веса $\omega_{\mathbf{k}_n}$ k -точек.

Начало координат желательно поместить в точку (центр зоны Бриллюэна), в противном случае симметрия системы может быть нарушена. Пример построения k -точек и вычисления их весов приведен на рисунке 1.5.

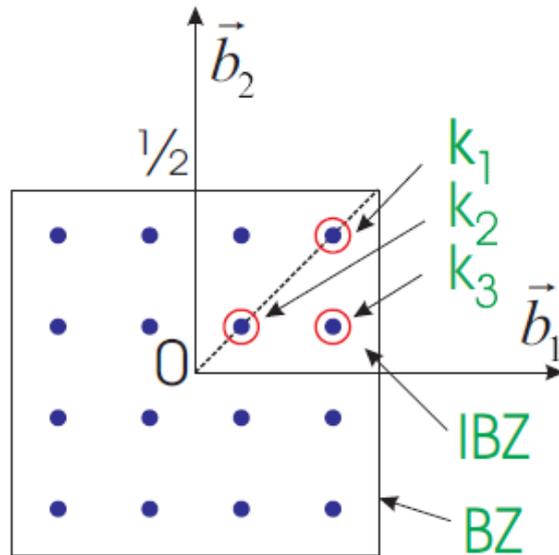


Рис. 1.5. Схематическая иллюстрация построения k -точек и выделения неприводимой части зоны Бриллюэна. Веса точек равны: $k_1 - 0,25$, $k_2 - 0,25$, $k_3 - 0,5$

Случай частично заполненной зоны

Для металлов зона Бриллюэна заполнена частично, до некоторого уровня энергии, поэтому возникает проблема при вычислении интеграла (1.68). Если учитывать только электроны с энергией меньше определенного уровня μ (уровень Ферми), то в зоне Бриллюэна возникает поверхность сложной формы (поверхность Ферми), вне которой состояния не заполнены, т.е. не должны учитываться в интеграле. Для математической записи пользуются ступенчатой функцией Дирака $\Theta(\varepsilon_{i,k} - \mu)$:

$$I(\varepsilon) = \sum_n \omega_{k_n} F(\varepsilon_{i,k_n}) \Theta(\varepsilon_{i,k} - \mu).$$

При использовании конечного числа k -точек точно описать положение поверхности Ферми не удастся, поэтому используется прием, при котором граница размывается, т.е. вместо ступенчатой функции Дирака используется гладкая функция $f(\varepsilon_{i,k})$. В этом случае достаточно небольшого числа k -точек, но заселенность каждой k -точки будет разной (частичная заселенность). Оказывается, что при такой замене полная энергия E (см. выражение (1.39)) больше не является вариационной (или минимальной). Поэтому необходимо заменить полную энергию некоторой обобщенной свободной энергией

$$F = E - \sum_{ik} \omega_k \sigma S(f_{ik}), \quad (1.69)$$

для которой можно применять вариационный принцип. В противном случае силы, действующие на ионы, будут вычисляться с ошибками. Для реализации частичной заселенности было предложено несколько методов, обсуждающихся далее.

Метод конечной температуры

В этом подходе функция f является функцией Ферми–Дирака:

$$f\left(\frac{\varepsilon - \mu}{\sigma}\right) = \frac{1}{\exp\left(\frac{\varepsilon - \mu}{\sigma}\right) + 1}. \quad (1.70)$$

При размытии границы функцией Ферми–Дирака F можно интерпретировать как свободную энергию электронов, S — как энтропию системы невзаимодействующих электронов, а σ — как некоторую конечную температуру:

$$F = E - \sum_n \sigma S(f_n), \quad (1.71)$$

$$S(f) = -[f \cdot \ln(f) + (1 - f) \cdot \ln(1 - f)], \quad (1.72)$$

$$\sigma = k_B T. \quad (1.73)$$

Метод вычисления для конечных температур был предложен Мерминым [22], а согласованность предложенного подхода можно увидеть, если заметить, что

$$\frac{\partial S}{\partial f_n} = \ln \frac{1 - f}{f}, \quad \frac{\partial E}{\partial f_n} = \varepsilon_n.$$

Применим вариационный принцип:

$$\begin{aligned} 0 &= \frac{\partial}{\partial f_n} \left[F - \mu \left(\sum_n f_n - N \right) \right] = \frac{\partial E}{\partial f_n} - \sigma \frac{\partial S}{\partial f_n} - \mu = \\ &= \varepsilon_n - \sigma \ln \frac{1 - f_n}{f_n} - \mu, \end{aligned} \quad (1.74)$$

т.е. функция f_n имеет вид (1.70). Таким образом свободная энергия F является вариационным функционалом, определяющим основное состояние электронов и силы, действующие на ионы.

Функционал F зависит от волновых функций ψ , частичных заселенностей f и положений ионов R . Электронное основное состояние (для фиксированного положения ионов) определяется вариационным свойством свободной энергии:

$$0 = \delta F(\psi, f, R) = \frac{\partial F}{\partial \psi} \delta \psi + \frac{\partial F}{\partial f} \delta f.$$

Для произвольных вариаций ψ и f эта величина равна нулю только тогда, когда

$$\frac{\partial F}{\partial \psi} = 0, \quad \frac{\partial F}{\partial f} = 0,$$

что приводит к системе уравнений, которая определяет основное электронное состояние. Силы, действующие на ионы в приближении локальной электронной плотности равны производной от обобщенной свободной энергии относительно ионных положений, т.е.

$$\text{СИЛА} = \frac{dF(\psi, f, R)}{dR} = \frac{\partial F}{\partial \psi} \frac{\partial \psi}{\partial R} + \frac{\partial F}{\partial f} \frac{\partial f}{\partial R} + \frac{\partial F}{\partial R} = \frac{\partial F}{\partial R}$$

Последнее равенство справедливо, т.к. в основном состоянии первые два слагаемых равны нулю. Таким образом можно зафиксировать ψ и f на значениях, полученных для основного состояния, и вычислить частную производную свободной энергии только по положениям ионов, что является относительно простой задачей.

Метод размытия функциями Гаусса

Можно уширять энергетические уровни функциями Гаусса, тогда f — это интеграл от функции Гаусса:

$$f\left(\frac{\varepsilon - \mu}{\sigma}\right) = \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{\varepsilon - \mu}{\sigma}\right) \right]. \quad (1.75)$$

У функции ошибок $\operatorname{erf}()$ нет аналитического выражения для обратной функции, поэтому энтропию и свободную энергию нельзя выразить через f .

$$S\left(\frac{\varepsilon - \mu}{\sigma}\right) = \frac{1}{2\sqrt{\pi}} \exp\left[-\left(\frac{\varepsilon - \mu}{\sigma}\right)^2\right]. \quad (1.76)$$

В случае использования функций Гаусса, физический смысл σ в функционале свободной энергии (1.69) остается неясным. Функционал свободной энергии $F(\sigma)$ отличен от точного $E(0)$, поэтому силы, вычисленные

через вариацию $F(\sigma)$ могут отличаться от истинных сил для основного электронного состояния.

Эту проблему можно обойти, т.к. для функций Гаусса можно получить точную экстраполяцию для $\sigma \rightarrow 0$ из результатов при конечных σ . Как это сделать видно из следующей цепочки выражений:

$$\begin{aligned}
 F(\sigma) &\approx E(0) + \gamma\sigma^2 \\
 F(\sigma) &= E(\sigma) - \sigma S(\sigma) \\
 S(\sigma) &= -\frac{\partial F(\sigma)}{\partial \sigma} \approx -2\gamma\sigma \\
 E(\sigma) - \sigma(-2\gamma\sigma) &\approx E(0) + \gamma\sigma^2 \rightarrow E(\sigma) \approx E(0) - \gamma\sigma^2 \\
 E(0) &\approx \frac{1}{2}(F(\sigma) + E(\sigma)). \quad (1.77)
 \end{aligned}$$

Таким способом мы можем получить «физическую» величину $E(0)$ из расчета при конечной σ , а метод размытия функцией Гаусса служит математическим инструментом для получения более быстрой сходимости по отношению к числу k -точек.

Метод, описанный в выше, имеет два недостатка:

- Силы являются производными от свободной энергии электронов F . Поэтому силы не могут быть использованы для получения равновесного основного состояния, которое соответствует минимуму энергии $E(\sigma \rightarrow 0)$. Тем не менее ошибка в силах обычно мала и приемлема. В принципе можно оценить производные $E(\sigma \rightarrow 0)$ по ионным координатам, но это непросто и требует дополнительного компьютерного времени.
- Выбор параметра σ требует большой осторожности. Если σ слишком велико, то энергия $E(\sigma \rightarrow 0)$ будет сходиться к неправильному значению даже для бесконечного числа k -точек, если σ слишком мала сходимость будет ухудшаться. Единственный способ получить хороший параметр σ — это выполнить несколько расчетов с различными сетками в обратном пространстве и разными σ .

Метод Метфесселя–Пакстона

Проблему выбора σ можно решить, выбирая функцию $f(\varepsilon_{nk})$ в несколько другом виде. Разложим ступенчатую функцию по полному ортонормированному набору функций и будем рассматривать ее последовательные приближения (метод Метфесселя и Пакстона [23]). Гауссова функция — это только первое приближение ($N = 0$) к ступенчатой

функции, но легко получить дальнейшие последовательные приближения ($N = 1, 2, \dots$):

$$f_0(x) = \frac{1}{2}(1 - \operatorname{erf}(x)),$$

$$f_N(x) = f_0(x) + \sum_{m=1}^N A_m H_{2m-1}(x) e^{-x^2},$$

$$S_N(x) = \frac{1}{2} A_N H_{2N}(x) e^{-x^2}, \quad (1.78)$$

где $A_n = \frac{(-1)^n}{n! 4^n \sqrt{\pi}}$, а H_N — полином Эрмита степени N . Вид функций $S_N(x)$ и производных можно увидеть на рисунке 1.6.

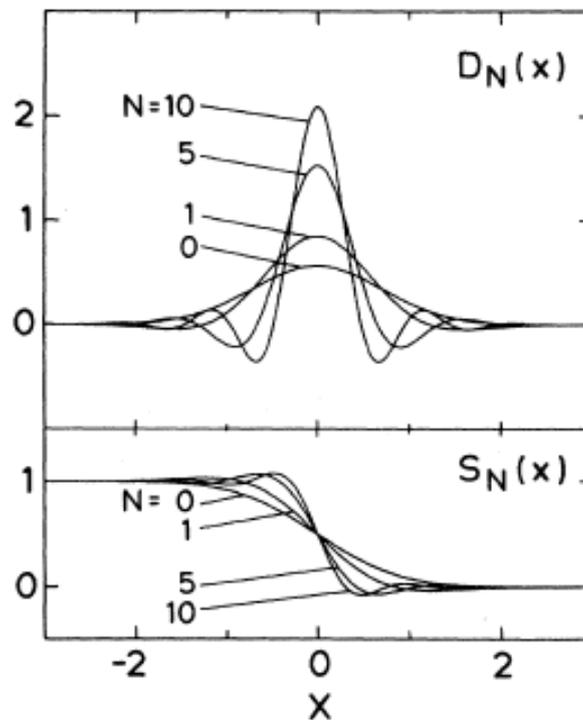


Рис. 1.6. Вид аппроксимирующих функций $S_N(x)$ и их производных с обратным знаком $D_N(x)$.

Преимуществами такого подхода являются: отклонение $F(\sigma)$ от $E(0)$ имеет порядок $2 + N$ по σ , экстраполяция $\sigma \rightarrow 0$ обычно не требуется, но она возможна:

$$E(0) \approx \frac{1}{N+2} ((N+1)F(\sigma) + E(\sigma))$$

Функции Метфесселя и Пакстона порядка N дают исчезающе малые ошибки в случае, если подынтегральная функция является полиномом степени $2N$ вблизи энергии Ферми. На практике достаточно положить N равное 1 или 2. Для высоких порядков N размытие σ можно

существенно увеличить без потери точности. «Энтропийное слагаемое» $S = \sigma \sum_n S_N(f_n)$ определяет отклонение $F(\sigma)$ от $E(\sigma)$, поэтому если S меньше нескольких мэВ, то $F(\sigma) \approx E(0)$, т.е. силы находятся точно. В отличие от гауссовского метода энтропийный член будет очень мал для разумных значений σ , поэтому σ можно увеличивать, пока он не станет слишком большим.

Линейный метод тетраэдров

В рамках линейного метода тетраэдров зона Бриллюэна разбивается на тетраэдры, а подынтегральная функция $X(\mathbf{k})$ линейно интерполируется внутри них: $X(\mathbf{k}) = \sum_j c_j(\mathbf{k})X(\mathbf{k}_j)$, где j — номера k -точек. После этого находятся эффективные веса k -точек: $\omega_j = \frac{1}{\Omega_{BZ}} \int_{BZ} F(\varepsilon_{\mathbf{k}})c_j(\mathbf{k})d\mathbf{k}$ и находится результат: $\sum_j \omega_j X(\mathbf{k})$. К недостаткам линейного метода относится то, что тетраэдры могут нарушать симметрию решетки Браве, необходимо по меньшей мере 4 k -точки и то, что линейная интерполяция дает завышенный или заниженный результат.

В модификации этого метода, предложенной Блечлем [19], вводится коррекция, которая устраняет квадратичную ошибку, присущую линейному методу тетраэдров:

$$\delta\omega_k = \sum_T \frac{1}{40} D_t(\mu) \sum_{j=1}^4 (\varepsilon_j - \varepsilon_{\mathbf{k}}),$$

где j — угловые k -точки тетраэдра, $D_t(\mu)$ — плотность электронных состояний на уровне Ферми. Линейный метод тетраэдров с коррекцией Блечля считается более-менее апробированным и он требует минимальных настроек. Главный его недостаток заключается в том, что если включены корректирующие вклады, то веса ω изменяются при изменении положений ионов. Поэтому для поиска сил приходится находить вариацию весов относительно положения ионов. В противном случае в силах возникает ошибка порядка нескольких процентов. Если требуется точно рассчитать силы, то вместо метода тетраэдров рекомендуется использовать метод конечной температуры.

1.4.2. Вычисление функционала энергии с использованием прямого и обратного пространства

Плоские волны являются собственными функциями оператора кинетической энергии T . Следовательно, мы легко можем получить решение для «пустой» решетки. Зонные энергии в атомных единицах есть просто

$(\mathbf{k} + \mathbf{G})^2$, а волновые функции — плоские волны $e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}$, модуль которых определяется условием нормировки. Теперь предположим, что к гамильтониану добавляется малое периодическое возмущение $H = T + \Delta V(\mathbf{r})$. Тогда волновые функции уже не будут чистыми плоскими волнами, а станут их суперпозицией. Тем не менее, из-за того, что энергетические знаменатели, которые возникают при разложении в ряд теории возмущения быстро увеличиваются, суперпозиция будет включать лишь несколько низших по энергии состояний. Таким образом, нет нужды включать большое число плоских волн, если $\Delta V(\mathbf{r})$ — слабое возмущение. Обычно достаточно рассмотреть состояния у которых $\frac{1}{2}(\mathbf{k} + \mathbf{G})^2 < E_{cutoff}$, где E_{cutoff} — уровень обрезания энергии. Поэтому плоские волны идеально подходят для модели почти свободных электронов и метода PAW.

Для большинства вычислений по схеме, приведенной на рисунке 1.1, используют следующий порядок действий, который также пояснен на рисунке 1.7:

- Находим обменно-корреляционный потенциал в прямом пространстве и переводим его в обратное пространство. $V_{XC}[\rho_r] \rightarrow V_{XC}[\rho_G]$.
- В обратном пространстве решаем уравнение Пуассона для периодического псевдопотенциала и находим потенциал Хартри $V_H[\rho_G] = \frac{4\pi e^2}{|\mathbf{G}|^2} \rho_G$.
- Получаем полный потенциал и переводим его в прямое пространство $V_G = V_{H,G} + V_{XC,G} + V_{ext,G} \rightarrow V_r + V_{loc,r}$.
- Находим вектор невязки R_G и переводим его в обратное пространство.
- Добавляем невязку R_G к волновой функции и находим следующую итерацию для электронной плотности.

При переходах между пространствами могут возникнуть так называемые ошибки наложения. Ошибки наложения возникают при использовании FFT сетки не достаточно большого размера. Это связано с тем, что каждая операция умножения в обратном пространстве или свертки в прямом пространстве увеличивает область значений G . В общем случае свертка двух функций f_1 и f_2 с максимальными Фурье компонентами G_1 и G_2 соответственно дает функцию с максимальным Фурье компонентом $G_1 + G_2$. То же самое происходит при перемножении функций, заданных в обратном пространстве. При недостаточном размере FFT сетки, в силу граничных условий, волны соседних ячеек с большими G будут влиять на расчетную ячейку.

Можно показать, что подобные ошибки не возникают, если FFT сетка для зарядовой плотности содержит все компоненты G до $2G_{cut}$, где G_{cut} — максимальная длина волны в базисном наборе ψ . Это можно увидеть на рисунке 1.7, если проследить, при каких операциях происходит увеличе-

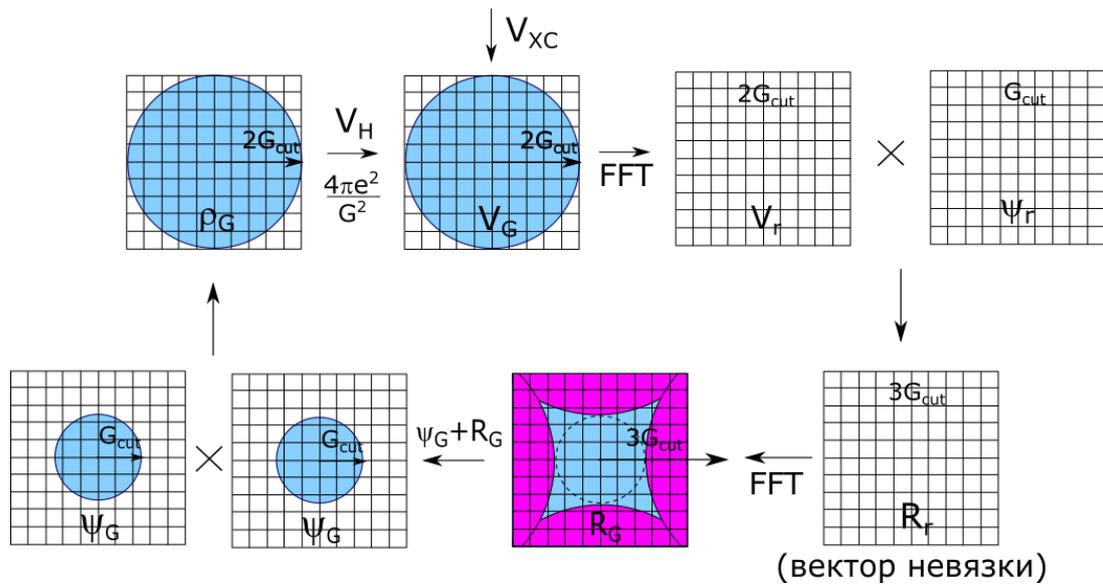


Рис. 1.7. Общая схема переходов между прямым и обратным пространством

ние границ G . Если волновая функция имеет границу G_{cut} , то электронная плотность, как произведение двух функций, имеет границу $2G_{cut}$. Еще одно увеличение границы происходит при действии потенциала (граница $2G_{cut}$) на волновую функцию (граница G_{cut}), которое вычисляется в прямом пространстве. Итого невязка R , переведенная в обратное пространство, имеет границу $3G_{cut}$. Для того, чтобы область векторов G_{cut} , необходимых для вычисления волновой функции не была перекрыта соседними ячейками при вычислении невязки, нужно выбрать минимальный размер сетки равный $2G_{cut}$, как показано на рисунке 1.7.

1.4.3. Метод суперячейки. Непериодические структуры

Теорема Блоха не может быть применена ни к кристаллам, содержащим дефекты, ни к жидкостям ни к молекулярным системам. Следовательно, в этих случаях базис плоских волн становится континуумом и не зависимо от того, какой бы малой мы ни выбрали энергию обрезания, базисный ряд становится бесконечным. По этой причине вычисления в базисе плоских волн могут проводиться для таких систем только при использовании периодических суперячеек.

Суперячейка содержит фрагмент структуры, окруженный областью пустого пространства, а в случае дефекта — идеального кристалла. Примеры суперячеек для структур разной размерности показаны на рисунке 1.8. На суперячейку накладываются периодические граничные условия,

так что суперячейка после многократного повторения заполняет все пространство, как показано на рисунке. Область пустого пространства (или идеального кристалла), должна быть настолько большой, чтобы исключить взаимодействие структур в соседних ячейках после трансляции. Это позволит рассматривать энергию суперячейки, как энергию одного фрагмента структуры. Отсутствие взаимодействия структур в соседних ячейках можно проверить путем увеличения объема суперячейки и наблюдая за сходимостью удельной энергии системы.

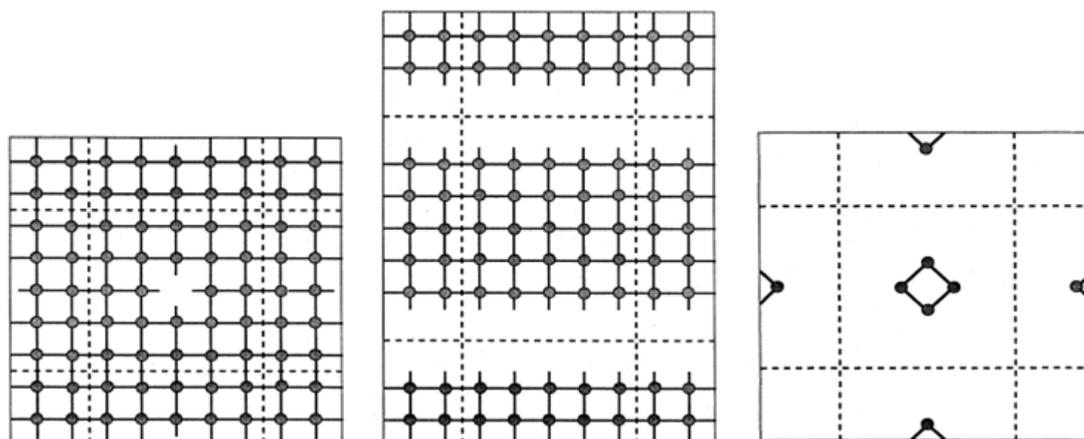


Рис. 1.8. Схематическая иллюстрация геометрии суперячейки для точечного дефекта (вакансии), поверхности в кристалле и молекулы. Суперячейка представляется областью, ограниченной пунктирными линиями

Контрольные вопросы

1. Сформулируйте теорему Блоха.
2. В чем преимущество базиса плоских волн? В каком случае это преимущество наиболее выгодно?
3. Что такое зона Бриллюэна? Для чего нужно вычислять интегралы по зоне Бриллюэна?
4. У каких веществ зона Бриллюэна заполнена не полностью? Какие вычислительные трудности при этом возникают?
5. Перечислите методы интегрирования по зоне Бриллюэна в случае частично заполненной зоны. В чем преимущества и недостатки каждого из них?
6. Что такое ошибки наложения при работе в прямом и обратном пространстве?
7. Можно ли использовать базис плоских волн для неперiodических структур? Как проводится моделирование поверхности?

2. Начало работы с пакетом VASP

2.1. Общая характеристика пакета VASP

В настоящее время существуют десятки программных пакетов, реализующих методы теории функционала электронной плотности, причем каждый из них имеет свои возможности и ограничения. Очевидно, что в точности удовлетворять всем требованиям, выдвигаемым решаемой задачей, может только программа, написанная самостоятельно. Однако это тяжёлый (и неблагодарный) труд; поэтому для того, чтобы определить, какой программный пакет использовать, необходимо хорошо представлять себе, чего мы хотим достичь, проводя компьютерное моделирование. После этого можно выбрать инструментарий, балансируя при этом на тонкой грани между точностью и эффективностью расчётов.

В этом пособии мы опишем использование мощного программного пакета VASP (Vienna Ab initio Simulation Package), созданного Г. Крессе и Ю. Хафнером [24] и развитого затем Г. Крессе и Дж. Фуртмюллером [14] в Техническом университете Вены. В последнее десятилетие VASP за счет своей точности и быстродействия стал де-факто стандартным пакетом для проведения расчетов методами теории функционала плотности; возможность моделирования систем из сотен атомов на десятке вычислительных узлов привела к тому, что пакет используют более 1000 научных групп по всему миру, и у основополагающей работы [14] уже более 34000 цитирований в базе данных Scopus.

Пакет VASP относится к классу пакетов, реализующих теорию функционала электронной плотности (см. главу 1), т.е. он предназначен для вычисления различных свойств материалов в основном состоянии (полной энергии, зонной структуры, плотности электронных состояний, фононных спектров и т.д.) путем решения уравнений Кона–Шэма. Эти уравнения решаются самосогласованно в базисе плоских волн; в программе реализован широкий выбор алгоритмов решения уравнений и параметров, влияющих на скорость получения результата. Пакет VASP работает для любых систем — атомов, молекул, объемных тел, поверхностей, атомных кластеров и т.д. Он эффективно распараллеливается на вплоть до 32 узлов, при этом он может обсчитывать до 4000 валентных электронов.

VASP — очень сложная программа. В ней за много лет было добавлено, исправлено, обновлено и изменено множество алгоритмов, схем,

подходов и методологий. VASP может выдать различные ответы на одну и ту же задачу в зависимости от пути её решения и исходных данных, использующихся для запуска программного пакета. При этом при запуске VASP с верными параметрами моделирования магически достигается требуемый баланс между точностью и эффективностью моделирования, однако для этого новичку обычно требуется экспертная помощь учёного с большим опытом.

2.2. Установка VASP

Программный пакет VASP, как и большинство аналогичных программных продуктов, распространяется в виде исходных кодов. Для скачивания пакета необходимо приобрести лицензию на его использование и публикацию результатов, полученных с его участием. Эта лицензия приобретается у авторов пакета; после приобретения в электронном письме покупателю присылается логин и пароль к закрытой части портала <https://www.vasp.at/>, в которой хранятся ссылки на исходный код пакета. В данном пособии описывается работа с пакетом VASP версии 5.4.4, которая является последней на момент написания пособия.

Перед установкой пакета убедитесь, что вы имеете нужные файлы:

- `vasp.5.4.4.tar.gz` — исходные файлы пакета¹;
- `potraw_LDA.54.tar.gz` — файлы псевдопотенциалов (см. раздел 1.3.1), созданные в рамках приближения локальной плотности (см. раздел 1.2.2);
- `potraw_PBE.54.tar.gz` — файлы псевдопотенциалов (см. раздел 1.3.1), созданные в рамках обобщенного градиентного приближения (см. раздел 1.2.2).

Программный код основной части пакета VASP написан на языке Fortran. Кроме этого, некоторые вставки написаны на C, поэтому перед компиляцией пакета убедитесь, что в системе установлены компиляторы для обоих языков. Успешная работа пакета также зависит от наличия в системе нескольких хорошо оптимизированных библиотек для численного моделирования:

- BLAS/LAPACK — библиотека линейной алгебры;
- SCALAPACK — параллельная библиотека линейной алгебры;
- FFT — быстрое преобразование Фурье;
- MPI — параллельное выполнение кода.

¹Начиная с версии 5.4.1, процедуры из сторонних библиотек, необходимые для запуска VASP, которые находились в файле `vasp.5.lib.tar.gz`, распространяются в основном архиве в директории `src/lib`.

Существуют различные версии этих библиотек. Среди них можно выделить набор открытых библиотек, распространяемых под лицензией GPL, и программное обеспечение, разрабатываемое компанией Intel (компиляторы C и Fortran, математические библиотеки Math Kernel Library (MKL)). И те, и другие библиотеки можно использовать для компиляции и установки VASP (ПО от Intel обычно устанавливается на суперкомпьютеры, в то время как GPL библиотеки входят во все современные версии дистрибутивов Linux). Разберем установку программного пакета более подробно.

После распаковки архива директория с файлами VASP имеет несколько поддиректорий (`arch`, `bin`, `build` и `src`) и файлы `README` и `makefile`. Для компиляции программы необходимо создать в этой директории высокоуровневый `makefile.include`, в котором будут описаны все параметры, используемые для компиляции. Директория `arch` содержит несколько таких файлов с наборами параметров для разных сред компиляции. Значения параметров, заданные в этих файлах, ведут к получению оптимизированного выполняемого файла программы.

2.2.1. Библиотеки и компиляторы GPL

Прежде, чем начать компиляцию пакета, убедитесь, что у вас на компьютере установлены нужные пакеты (названия пакетов приведены для ОС Ubuntu; в других ОС они могут быть другими):

- Компиляторы и программы для сборки: `make`, `g++`, `gfortran`, `build-essential`
- Библиотеки линейной алгебры и численных методов: `libblas-dev`, `liblapack-dev`, `libfftw3-dev`, `libfftw3-3`
- Библиотеки параллельных вычислений: `libopenmpi-dev`, `libscalapack-mpi-dev`, `libscalapack-openmpi`

После этого скопируем файл `makefile.include` в основную директорию VASP. С помощью редактора откроем файл и отредактируем его. В нем надо заменить пути к библиотеке FFTW (переменная `FFTW`) и установленным библиотекам линейной алгебры (переменная `LIBDIR`). После замены пути на верный можно выйти из редактора и запустить команду `make`.

В случае возникновения ошибок их следует прочитать и понять. Ошибки обычно связаны с невозможностью найти заголовочные файлы или библиотеки (если неверно прописаны пути к ним в `makefile.include`). Если ошибок нет, то при окончании процесса компиляции в поддиректории `bin` возникнут три файла:

- `vasp_std` — стандартный исполняемый файл VASP;

- `vasp_gam` — исполняемый файл, оптимизированный для расчетов с одной Γ -точкой (см. раздел 1.4.1)
- `vasp_nc1` — исполняемый файл VASP, предназначенный для расчета спин-орбитального взаимодействия и неколлинеарной спиновой плотности.

2.2.2. Библиотеки и компиляторы Intel

Рекомендованный авторами способ компиляции VASP — с использованием компиляторов и математических библиотек производства компании Intel. Обычно данные компиляторы уже предустановлены на суперкомпьютерах; если же этого недостаточно, то можно установить компиляторы и библиотеки на собственный компьютер. Лицензии на эти продукты для студентов можно получить бесплатно сроком на 1 год. Для этого необходимо выбрать нужный продукт на странице <https://software.intel.com/en-us/qualify-for-free-software/student> и заполнить анкету. После заполнения анкеты на электронную почту придет письмо с ссылками на скачивание выбранного продукта и его серийным номером, который надо будет ввести при установке.

VASP 5.4.4 использует директивы MPI, появившиеся в стандарте MPI-3.0, вследствие чего для его компиляции нужна версия библиотеки Intel MPI не ниже 5.0-beta1, в которой появилась поддержка этого стандарта. При установке VASP на суперкомпьютер убедитесь, что на нем установлены и работают компиляторы и библиотеки Intel нужных версий. Если это так, то для компиляции VASP, как и в случае GPL компиляторов, просто скопируем необходимый файл `makefile.include` в основную директорию VASP:

```
cp arch/makefile.include.linux_intel ./makefile.include
```

и запустим `make`. Ошибки компиляции могут быть связаны с отсутствием значений у переменных среды, задаваемых компиляторами. Эти переменные задаются в скрипте `compilervars.sh`, распространяемом вместе с компиляторами. Его надо выполнить в текущем окружении следующим образом:

```
# переход в директорию, содержащую компилятор Intel
cd /opt/intel/compilers_and_libraries/linux/bin
source ./compilervars.sh linux64
```

`linux64` здесь — строка, описывающая архитектуру компьютера. После этого команда `make` должна выполняться без ошибок.

2.2.3. Расчеты на графических ускорителях с CUDA

Начиная с версии 5.4, программный пакет VASP поддерживает перенос части вычислений на графические ускорители NVidia с использованием технологии CUDA.

2.3. Входные файлы

Для запуска любого расчета необходимо подготовить по крайней мере четыре входных файла. Они должны иметь определенные названия: INCAR, POSCAR, POTCAR и KPOINTS. Дадим их краткое описание.

- POSCAR содержит информацию о структуре и динамике изучаемой системы;
- KPOINTS определяет разбиение обратной зоны Бриллюэна на k-точки;
- POTCAR содержит информацию, позволяющую восстановить псевдопотенциал для каждого сорта атомов, входящих в изучаемую систему;
- INCAR — основной файл расчета, содержит все его параметры.

2.3.1. POSCAR (расположение атомов)

Файл POSCAR содержит описание векторов трансляции расчетной ячейки и позиции атомов. Дополнительно могут быть заданы начальные скорости и предшествующие координаты в случае использования схемы «предиктор–корректор» для интегрирования уравнений движения атомов в молекулярной динамике. Примеры:

Cubic BN	Cubic BN
3.57	3.57
0.0 0.5 0.5	0.0 0.5 0.5
0.5 0.0 0.5	0.5 0.0 0.5
0.5 0.5 0.0	0.5 0.5 0.0
1 1	1 1
Selective dynamics	Direct
Cartesian	0.00 0.00 0.00
0.00 0.00 0.00 T T F	0.25 0.25 0.25
0.25 0.25 0.25 F F F	
Cartesian	
0.01 0.01 0.01	
0.00 0.00 0.00	

Первая строка воспринимается пакетом VASP как комментарий (в ней можно задать «название» моделируемой системы). Положительное действительное число во второй строке указывает параметр для масштабирования базисных векторов и координат атомов. Если указанное значение

меньше нуля, оно интерпретируется как объем расчетной ячейки. Следующие три строки содержат базисные вектора расчетной ячейки. Шестая строка определяет число атомов каждого сорта (по одному числу для каждого из типов атомов). Порядок типов атомов должен соответствовать указанному в файлах POSCAR (см. стр. 57) и INCAR (см. стр. 57). Седьмая строка позволяет переключиться в режим выборочного движения атомов «Selective dynamics» (в строке учитывается только первый символ, который должен быть «S» или «s»). Этот режим позволяет для каждого атома указать дополнительные параметры, указывающие, необходимо ли изменять относительные координаты конкретного атома в процессе оптимизации геометрии системы. Эта возможность полезна, если необходимо отрелаксировать позиции определенных атомов около дефекта или слои вблизи поверхности. Параметр «Selective dynamics» является необязательным.

Седьмая строка (или восьмая, если выбран режим «Selective dynamics») указывает, в декартовых или в дробных координатах заданы позиции атомов. Лишь первый символ строки интерпретируется пакетом VASP. Для задания декартовой системы координат достаточно указать один из четырех символов: «C», «c», «K» или «k». Следующие строки содержат координаты каждого атома. В режиме дробных координат позиции атомов задаются выражением $\mathbf{R} = x_1 \cdot \mathbf{a}_1 + x_2 \cdot \mathbf{a}_2 + x_3 \cdot \mathbf{a}_3$, где \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 — базисные вектора, x_1 , x_2 , x_3 — значения указанные в файле. В режиме декартовых координат позиции атома масштабируются с использованием параметра s из второй строки файла POSCAR:

$$\mathbf{R} = s \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Порядок этих строк должен соответствовать порядку в перечислении количеств атомов каждого сорта в шестой строке. Для контроля корректности входных данных можно проверить файл OUTCAR, содержащий компоненты вектора \mathbf{R} . Если активирован режим «Selective dynamics», то после тройки координат каждого атома указываются логические флаги, определяющие, допускается ли изменение соответствующей координаты (в левом примере третья координата первого атома и все координаты второго атома фиксированы). Если строка «Selective dynamics» отсутствует, подобные флаги игнорируются (принимаются равными .T.). Следует учесть, что флаги относятся к позициям ионов в дробных координатах независимо от того, какая система координат выбрана. В приведенном выше примере первый атом может перемещаться в направлении первого и второго базисного вектора.

На этом файл POTCAR может быть завершен, если нет необходимости задавать начальные скорости. В режиме молекулярной динамики скорости задаются случайным образом в соответствии с распределением Максвелла–Больцмана при заданной начальной температуре TEBEG (см. стр. 116). Явное указание скоростей встречается редко.

2.3.2. KPOINTS (точки в зоне Бриллюэна)

Файл KPOINTS должен содержать координаты и веса всех k -точек для интегрирования по зоне Бриллюэна (см. раздел 1.4.1), либо размерность сетки в обратном пространстве для автоматической генерации списка k -точек. В VASP 5.4 файл KPOINTS может отсутствовать, а описание сетки k -точек может быть перенесено в файл INCAR.

Наиболее простой способ задания списка k -точек состоит в разбиении k -пространства автоматически по алгоритму Монкхорста и Пака (см. раздел 1.4.1). Для задания такого разбиения можно использовать файл KPOINTS следующего формата:

```

1 Automatic mesh
2 0
3 Auto
4 10

```

Как и в случае файла POSCAR, первая строка файла KPOINTS представляет собой комментарий. Вторая строка представляет собой требуемое количество k -точек. Ноль в этой строке включает процедуру автоматической генерации сетки. В третьей строке важен лишь первый символ; в случае «A» (или «a») включается полностью автоматическая процедура, которая генерирует центрированную на Γ -точке сетку из равноудаленных между собой k -точек с координатами

$$\mathbf{k} = \mathbf{b}_1 \frac{n_1}{N_1} + \mathbf{b}_2 \frac{n_2}{N_2} + \mathbf{b}_3 \frac{n_3}{N_3}. \quad (2.1)$$

Здесь \mathbf{b}_i — векторы обратной решётки, N_i — число k -точек вдоль каждого обратного вектора решётки, которое может быть рассчитано по формуле $N_i = \max(1, l \cdot |\mathbf{b}_i| + 0.5)$. l — число из четвертой строки файла KPOINTS.

Для большего контроля над процедурой генерации можно в явном виде задать числа N_i . Тогда файл будет выглядеть следующим образом:

```

1 Automatic mesh
2 0
3 Gamma
4 4 4 4
5 0. 0. 0.

```

В этом случае третья строка начинается с «G». Это запускает процедуру генерации сетки с центром на Γ -точке и известными N_i (заданными на четвертой строке). Если первый символ третьей строки — «M», используется оригинальная процедура Монкхорста–Пака. В ней при четных N_i центр сетки сдвинут относительно Γ -точки на $\mathbf{b}_i/2N_i$:

$$\mathbf{k} = \mathbf{b}_1 \frac{n_1 + 0.5}{N_1} + \mathbf{b}_2 \frac{n_2 + 0.5}{N_2} + \mathbf{b}_3 \frac{n_3 + 0.5}{N_3}. \quad (2.2)$$

Пятая строка такого файла является необязательной, и в случае ее наличия содержит сдвиг сетки относительно центра, выраженный в единицах \mathbf{b}_i/N_i . Таким образом, выбор «M» без сдвига эквивалентен выбору «G» со сдвигом в 0.5 по каждому вектору обратной решетки и наоборот.

В случае, если третья строка не начинается с «G» или «M», сетка k -точек задается вручную: явным образом прописываются координаты всех точек сетки. Этот способ задания сетки должен использоваться только в том случае, если пользователь программы знает, что он делает, и автоматически созданная сетка его не устраивает.

Зонная структура Для расчета зонной структуры, т.е. зависимости энергии от положения k -точек вдоль определенного направления, требуется файл KPOINTS несколько другого формата (пример использования находится в разделе 3.3). Пример файла KPOINTS приведён ниже.

```

1 k-points along high symmetry lines
2 30
3 Line-mode
4 Cartesian
5 0 0 0 ! \Gamma
6 0 0 1 ! X
7
8 0 0 1 ! X
9 0.5 0 1 ! W
10
11 0.5 0 1 ! W
12 0 0 1 ! \Gamma

```

Здесь третья строка начинается с «L» (line-mode, режим определения отрезков, соединяющих специальные точки зоны Бриллюэна). Специальные точки могут быть заданы в декартовых или в обратных координатах. Система координат определяется первым символом четвертой строки («C» или «K» для декартовых, «R» — для обратных). Строки после четвертой идут парами. Каждая пара строк — координаты концов отрезка в зоне Бриллюэна. VASP генерирует равное количество k -точек для каждого из

заданных отрезков так, чтобы общее количество k -точек было равно числу из второй строки файла. Так, в примере выше для каждого из заданных отрезков будет сгенерировано по 10 k -точек.

2.3.3. POTCAR (задание псевдопотенциалов)

Файл POTCAR содержит псевдопотенциалы (см. раздел 1.3.1) для каждого типа атомов, участвующего в расчетах. Файлы псевдопотенциалов распространяются вместе с пакетом в виде архивов `potraw_*` (см. раздел 2.2). В случае, если количество типов больше одного, следует соединить файлы POTCAR для всех типов участвующих в расчете, в один *в том порядке, в котором типы атомов заданы в файле POSCAR*. Новые пользователи VASP часто допускают ошибку в порядке следования файлов. Для соединения файлов друг с другом в Linux используется команда `cat`:

```
cat ~/pot/Al/POTCAR ~/pot/C/POTCAR ~/pot/H/POTCAR > POTCAR
```

Данная команда объединит три файла из разных поддиректорий папки `~/pot` в один, находящийся в текущей директории. Файлы POTCAR также содержат энергию обрезания псевдопотенциалов (ENMAX и ENMIN). В случае, если энергия обрезания не задана в параметре ENCUT файла INCAR (см. раздел 2.3.4), ее значение будет взято как максимум из значений параметра ENMAX всех используемых псевдопотенциалов.

2.3.4. INCAR (параметры расчета)

Файл INCAR — основной входной файл VASP, который определяет весь ход расчета. В файле INCAR может быть определено много параметров, большинство из которых имеет достаточно разумные значения по умолчанию. Изменять их следует только тогда, когда пользователь программы понимает, что он делает. В дальнейшем мы опишем параметры, наиболее часто встречающиеся в файле. Кроме этого, часть параметров будет описана в Приложении. Примеры файлов для решения некоторых типов задач приведены в главе 3.

Параметры в файле INCAR описываются следующим образом:

```
ENCUT = 500
```

Здесь ENCUT — *тэг* параметра, а 500 — его *значение*. Тэг отделяется от значения знаком “=”. Перед и после знака “=” может быть любое количество пробелов. Следует избегать использования в файле INCAR символов табуляции: они неверно считываются программой.

На одной строке может быть определено несколько тэгов; в этом случае они отделяются друг от друга точкой с запятой. Если строка заканчивается обратным слэшем \, следующая строка будет считаться продолжением предыдущей. Комментарии в файле INCAR обычно отделяются «решёткой» (#), однако в большинстве случаев комментарии можно написать либо на отдельной строке без знака “=”, либо сразу после значения параметров (тогда следует избегать использования в комментарии точки с запятой).

SYSTEM (комментарии)

Параметр SYSTEM включает в себя пользовательский комментарий об изучаемой системе. Он не используется нигде в коде VASP, а только считывается из файла INCAR и записывается в файл OUTCAR.

ENCUT (предел энергии)

Параметр ENCUT определяет энергию обрезания для базиса плоских волн в эВ (см. раздел 1.4.2):

$$|\mathbf{G} + \mathbf{k}| < G_{cut}; E_{cut} = \frac{\hbar^2}{2m} G_{cut}^2. \quad (2.3)$$

Количество используемых плоских волн различно для каждой k -точки; это позволяет проводить расчеты уравнения состояния изучаемой системы. При изменении объема ячейки количество плоских волн изменяется плавным образом. Значение параметра по умолчанию берется согласно параметру PREC; несмотря на это, авторы пакета рекомендуют явным образом указывать используемую энергию обрезания в файле INCAR и не изменять её в сериях расчетов, чтобы результаты нескольких расчетов были сравнимы друг с другом.

PREC (точность)

Параметр PREC определяет точность расчетов. Допустимые значения параметра:

- Low
- Medium
- High (использование не рекомендуется)
- Normal (значение по умолчанию)
- Single
- Accurate

Изменение значения параметра PREC меняет, как показано в таблице 2.1, значения по умолчанию для ряда параметров (ENCUT — энергии обрезания; NGX, NGY, NGZ — числа точек Фурье-сетки для функций, заданных во всем объеме; NGXF, NGYF, NGZF — числа точек Фурье-сетки для функций, центрированных на атомах; ROPT — параметра оптимизации проекционной функции в прямом пространстве). Такого же эффекта можно добиться изменением этих параметров в файле INCAR.

Таблица 2.1

Параметры расчетов, изменяемые с помощью тэга PREC ($x = X, Y, Z$)

PREC	ENCUT	NG x	NG x F	ROPT
Normal	max(ENMAX)	$1.5G_{cut}$	2NG x	$-5 \cdot 10^{-4}$
Single	max(ENMAX)	$1.5G_{cut}$	NG x	$-5 \cdot 10^{-4}$
Accurate	max(ENMAX)	$2G_{cut}$	2NG x	$-2.5 \cdot 10^{-4}$
Low	max(ENMIN)	$1.5G_{cut}$	$3G_{aug}$	-0.01
Medium	max(ENMAX)	$1.5G_{cut}$	$4G_{aug}$	-0.002
High	$1.3 \cdot \max(\text{ENMAX})$	$2G_{cut}$	$16/3 \cdot G_{aug}$	$-4 \cdot 10^{-4}$

В таблице 2.1 max(ENMAX/ENMIN) соответствует максимальному значению параметра ENMAX либо ENMIN в файле POTCAR, значение по умолчанию для ENAUG задается максимумом EAUG в файле POTCAR. Максимальные волновые вектора для функций, заданных во всем объеме и центрированных на атомах (G_{cut} и G_{aug}), определяются следующим образом (см. раздел 1.4.2):

$$\frac{\hbar}{2m_e} |G_{cut}|^2 = \text{ENCUT}; \quad \frac{\hbar}{2m_e} |G_{aug}|^2 = \text{ENAUG}.$$

Таким образом, параметр PREC влияет на выбор энергии обрезания базиса плоских волн ENCUT и параметры сетки точек для быстрого преобразования Фурье, если они не заданы в файле INCAR явным образом.

Для PREC = Low параметр ENCUT устанавливается равным максимуму значений ENMIN из файла POTCAR. При PREC = Medium или Accurate параметр ENCUT устанавливается равным максимальному значению ENMAX из файла POTCAR. Для PREC = high параметр ENCUT устанавливается равным 130% от максимального значения ENMAX из файла POTCAR. В общем случае увеличение энергии обрезания свыше max(ENMAX) необходимо только для оценки величин, связанных с тензором напряжений (например, упругих свойств). Использование PREC =

high в настоящее время не рекомендуется (и оставлено только для обратной совместимости). Для вычисления точного тензора напряжений энергию обрезания необходимо увеличить вручную, и, если также нужно получить точное значение сил на атомах, то вдобавок к увеличению энергии обрезания можно использовать `PREC = Accurate`. Точность сил может быть дополнительно улучшена путем указания `ADDGRID = .TRUE.`

Параметры Фурье-сетки выбираются таким образом, чтобы при выборе `PREC = High` и `PREC = Accurate` ошибки наложения (см. раздел 1.4.2) не возникали бы. При этих параметрах используется в два раза более детализированная сетка, чем для представления псевдоволновых функций. `PREC = Accurate` увеличивает требования к памяти, но его необходимо использовать, если требуются определение сил с большой точностью (вычисление фононов или вторых производных от энергии). При `PREC = Low, Medium, Normal` сетка FFT уменьшается до 3/4 от необходимой. Обычно `Medium` и `Normal` достаточны для достижения погрешности менее 1 мэВ/атом.

Кроме этого, параметр `PREC` изменяет тэги `NGXF, NGYF, NGZF`, определяющие число узлов Фурье-сетки для работы с проекционными (центрированными на атомах) зарядами в атомных сферах \hat{n} , и тэг `ROPT`, контролирующей оптимизацию, т.е. вырезание высокочастотных компонентов, проекционной функции в прямом пространстве. Более подробную информацию вы можете найти в руководстве пользователя VASP.

ISTART (начальные волновые функции)

Параметр `ISTART` определяет характер инициализации *волновых функций* при запуске расчета VASP. Возможные значения:

- 0 — начать расчет сначала, инициализировать волновые функции случайными числами.
- 1 — продолжить расчет с постоянным значением энергии обрезания.
- 2 — продолжить расчет с постоянным набором базисных функций.
- 3 — полный перезапуск расчета, включая механизм предсказания зарядовой плотности на следующем шаге.

Значение по умолчанию: 1, если существует файл `WAVECAR`, и 0 в обратном случае.

ICHARG (начальная электронная плотность)

Параметр `ICHARG` определяет характер инициализации *зарядовой плотности* при запуске расчета VASP. Возможные значения:

- 0 — расчёт зарядовой плотности из первоначальных волновых функций. Если `ISTART = 0` вследствие отсутствия валидного файла `WAVECAR`, то `ICHARG` изменит значение на 2.
- 1 — чтение зарядовой плотности из файла `CHGCAR` и экстраполяция её на новые атомные позиции, используя линейную комбинацию зарядовых плотностей на атомах.
- 2 — использование линейной комбинации зарядовых плотностей на атомах.
- 4 — чтение потенциалов из файлов `POT` (практически не используется; см. руководство пользователя `VASP`).

Значение по умолчанию: 2 в случае `ISTART = 0`, 0 в остальных случаях.

ISMEAR, SIGMA (интегрирование по зоне Бриллюэна)

Параметр `ISMEAR` определяет, как будут вычислены коэффициенты частичного заполнения орбиталей для интегрирования по зоне Бриллюэна. Методики определения этих коэффициентов описаны в разделе 1.4.1. Для метода конечной температуры параметр `SIGMA` определяет ширину размытия в эВ. Параметр `ISMEAR` может принимать следующие значения (описание методов смотри в разделе 1.4.1):

- -1 — размытие функциями Ферми
- 0 — размытие функциями Гаусса
- 1..N — метод Метфесселя–Пакстона порядка N. Имейте в виду, что для этого метода часть коэффициентов заполнения могут быть отрицательными.
- -2 — коэффициенты заполнения считываются из файла `WAVECAR` (или `INCAR`), и остаются постоянными при работе программы. Если коэффициенты заполнения выбраны постоянными, то в файл `INCAR` нужно добавить тег `FERWE = f1 f2 f3 ... f(NBANDS)`, а для спинполяризованных вычислений дополнительно тег `FERDO = f1 f2 f3 ... f(NBANDS)`, куда записать коэффициенты заполнения для всех зон и k -точек. При записи параметров быстрее всего меняется индекс номера зоны. Коэффициенты заполнения обязательно должны быть между 0 и 1. Имейте в виду, что коэффициенты заполнения также записываются в файл `OUTCAR`, но в этом случае они умножаются на 2, т.е. они лежат между 0 и 2.
- -3 — производится циклический перебор параметров размытия в последовательности, указанной в файле `INCAR`. В этом случае в файл `INCAR` нужно добавить тег `SMEARINGS = ismear1 sigma1 ismear2`

σ_2 ..., в котором приведены различные параметры размытия. Также нужно установить таги $IBRION = -1$, $NSW = ismeari$, где $ismeari$ — число указанных значений «ismear». Первый цикл в этом случае выполняется с использованием метода тетраэдров с коррекцией Блечла.

- -4 — метод тетраэдров без коррекции Блечла (используйте сетку, центрированную в Γ -точке)
- -5 — метод тетраэдров с коррекцией Блечла (используйте сетку, центрированную в Γ -точке)

Для нахождения полной энергии объемных материалов рекомендуется выбирать метод тетраэдров с коррекцией Блечла ($ISMEAR=-5$). Этот метод также дает хорошие значения для электронной плотности состояния (DOS). Единственный недостаток — это то, что не является вариационным по отношению к частичным заселенностям, поэтому силы и тензор напряжения, вычисленные с его помощью могут содержать ошибку до 5–10% для металлов. Для вычисления частот фононов через силы рекомендуется использовать метод Метфесселя–Пакстона ($ISMEAR > 0$).

Метод Метфесселя–Пакстона дает аккуратное описание полной энергии, но нужно тщательно подойти к выбору ширины размытия ($SIGMA$) (см. 1.4.1). Слишком большое размытие может привести к неправильной полной энергии, малое размытие требует сетки с большим числом k -точек. $SIGMA$ должно быть как можно больше при условии, что в файле OUTCAR разница между свободной энергией и полной энергией (т.е. энтропийное слагаемое $T \cdot S$) сохраняется незначительной (≈ 1 мэВ/атом). В большинстве случаев $N = 1$ и $N = 2$ приводят к очень похожим результатам. Метод Метфесселя–Пакстона также стоит выбрать для больших суперячеек (см. 1.4.3), т.к. метод тетраэдров оказывается неприменим, если число k -точек меньше 3. Имейте в виду, что следует избегать использовать $ISMEAR > 0$ для полупроводников и диэлектриков, т.к. это часто приводит к неправильным результатам (коэффициенты заполнения некоторых состояний могут быть больше или меньше 1). Для изоляторов используйте $ISMEAR = 0$ или $ISMEAR = -5$.

Метод размытия функциями Гаусса также дает приемлемые результаты в большинстве случаев. В этом методе необходимо экстраполировать результаты для конечных $SIGMA$ к точке $SIGMA = 0$. В этом случае в файле OUTCAR появляется дополнительная строка «energy($SIGMA \rightarrow 0$)» в которой приведен результат экстраполяции. Большое значение $SIGMA$ приводит к ошибкам схожим с ошибками в схеме Метфесселя–Пакстона, но, в отличие от схемы Метфесселя–Пакстона, здесь нельзя определить величину ошибки путем систематического уменьшения

SIGMA. Поэтому метод Метфесселя–Пакстона более удобен, чем размытие функциями Гаусса.

Для взвешенного выбора метода размытия нужно подробно ознакомиться с разделами 1.4.1. Кратко рекомендации таковы:

- Для полупроводников и диэлектриков используйте метод тетраэдров ($ISMEAR = -5$), если ячейка очень большая (или если вы используете только одну или две k -точки) используйте $ISMEAR = 0$ в комбинации с малым значением $SIGMA = 0.05$.
- Для релаксации металлов всегда используйте $ISMEAR = 1$ или $ISMEAR = 2$ и соответствующее значение $SIGMA$ (энтропийное слагаемое должно быть меньше 1 мэВ/атом). Имейте в виду, что нужно избегать использования $ISMEAR > 0$ для полупроводников и изоляторов, т.к. это может вызвать проблемы. Для металлов разумное значение $SIGMA = 0.2$ (установлено по умолчанию).
- Для вычисления плотности состояний (DOS) и очень точных расчетов полной энергии (не релаксации металлов) используйте метод тетраэдров ($ISMEAR = -5$).

IBRION, NFREE (движение ионов)

Параметр IBRION определяет способ обновления положения ионов в расчете. Допустимые значения параметра IBRION:

- -1 — позиции ионов не изменяются. Данный вариант нужно использовать для обновления электронной структуры или получения электронных свойств (плотности электронных состояний или зонной структуры). Обратите внимание, что NSW шагов моделирования все равно выполняются, поэтому есть смысл задавать $NSW = 0$ или $NSW = 1$.
- 0 — стандартная первопринципная молекулярная динамика. Для интегрирования уравнений Ньютона используется алгоритм Верле (см. раздел 1.1.2) или метод предиктор–корректор четвертого порядка, если VASP был скомпилирован с опцией `stepprecor.o`. Дополнительно нужно задать параметры POTIM — шаг по времени в фемтосекундах и SMASS — выбор статистического ансамбля.
- 1 — Релаксация структуры при помощи квазиньютоновского алгоритма RMM–DIIS [25]. Для определения направлений сдвига атомов используются силы и тензор напряжений (полная энергия не рассматривается). Алгоритм очень быстрый и эффективный вблизи локальных минимумов, но оказывается непригодным, если стартовые позиции были выбраны неудачно. Поскольку алгоритм основан на приближении

матрицы Гессе (матрица вторых производных от энергии), то ему требуются очень точно рассчитанные силы, в противном случае алгоритм не сходится.

- 2 — Релаксация структуры методом сопряженных градиентов. Данный выбор является оптимальным для сложных случаев оптимизации атомной структуры.
- 3 — Метод молекулярной динамики с затуханием (вязким трением) обычно эффективен при изучении систем, изначально далеких от равновесной конфигурации.
- 5 — Для вычисления матрицы вторых производных от энергии (Гессиан) используется метод конечных разностей. Для каждого иона рассматриваются смещения вдоль каждой из осей декартовой системы координат. Гессиан вычисляется по силам, действующим на ионы. Данный способ имеет высокую точность, но сопряжен с большими вычислительными затратами.
- 6 — То же самое, что $IBRION = 5$, но при смещениях рассматриваются только неэквивалентные (с точки зрения симметрии) из них, а оставшаяся часть Гессиана восстанавливается из соображений симметрии.
- 7 — Для вычисления матрицы вторых производных от энергии используется теория возмущений в рамках теории функционала электронной плотности. Симметрия системы не учитывается. Получаемые данные аналогичны случаю $IBRION = 5$.
- 8 — То же самое, что $IBRION = 7$, но при вычислениях учитывается симметрия системы и рассматриваются только неэквивалентные смещения.
- 44 — Поиск переходного состояния по усовершенствованному методу димера. В данном случае осуществляется поиск седловых точек поверхности энергии, путем моделирования движения гантели.

По умолчанию $IBRION = -1$ для числа шагов моделирования $NSW = 0$ или 1, и $IBRION = 0$ во всех остальных случаях.

Количество шагов перемещения ионов задается при помощи параметра NSW . Для методов релаксации $IBRION = 1, 2, 3$ можно задать предельную разницу энергий $EDIFFG$, которая будет вызывать остановку программы. Флаг $ISIF$ определяет, какие из возможных степеней свободы используются при релаксации (изменяются ли координаты ионов, размер и форма расчетной ячейки, и т.д.). Для молекулярной динамики $IBRION = 0$ изменение формы расчетной ячейки возможно только при использовании модуль Томаса Буко (при компиляции установлен флаг $-Dtbdyn$).

Для всех алгоритмов оптимизации $IBRION = 1, 2, 3$ параметр масштабирования сил $POTIM$ должен быть определен в файле $INCAR$. $POTIM$ не имеет физического смысла, а служит множителем масштабиро-

вания. Для большинства систем оптимальным значением ROTIM является 0.5. Так как квази-ньютоновский (IBRION = 1) и релаксационный (IBRION = 3) алгоритмы чувствительны к выбору этого параметра, используйте IBRION = 2, если вы не уверены каково должно быть оптимальное значение этого параметра.

Ниже приведены особенности использования некоторых режимов движения ионов и ключевые параметры, необходимые для работы.

IBRION = 1 Алгоритм использует модификацию метода Ньютона, т.е. находится аппроксимация поверхности энергии полиномом второго порядка происходит смещение в точку минимума. Метод быстрый и эффективный вблизи локальных минимумов, но оказывается непригодным, если стартовые позиции были выбраны неудачно. Вторая проблема — матрица вторых производных должна быть вычислена очень точно, в противном случае алгоритм не сходится. Для его использования рекомендуется установить минимальное число шагов электронной оптимизации между ионными шагами NELMIN от 4 (для простых объемных материалов) до 8 (в случае сложной поверхности). Это обеспечит надежное определение сил перед каждым смещением ионов.

Реализованный алгоритм называется RMM-DIIS [25]. Его применение для оптимизации электронной структуры обсуждалось в разделе 1.2.3. В данном случае вычисляется приближение для обратной матрицы Гессе с учетом информации из предыдущих итераций. Количество векторов, хранящихся в истории, не должно превышать числа степеней свободы системы равное $3*(NIONS-1)$. Оно может быть существенно уменьшено за счет учета симметрии и ограничений в оптимизации координат атомов. Кроме того, информация с предшествующих шагов моделирования (которая может привести к линейной зависимости векторов) должна удаляться из истории итераций по мере необходимости. Существует два алгоритма удаления сведений о предшествующих итерациях.

1) Если в файле INCAR задан параметр NFREE, хранится информация не более чем о NFREE предшествующих шагах моделирования (при этом ранг матрицы Гессеана не превосходит NFREE).

2) Если в файле INCAR не задан параметр NFREE, то история итераций удаляется если одно из собственных значения обратной матрицы Гессеана становится больше 8.

Для сложных случаев NFREE стоит выбрать достаточно большим (например, от 10 до 20), однако для систем пониженной размерности необходимо тщательно подбирать этот параметр (или точно рассчитывать число степеней свободы). Значения NFREE больше 20 редко улучшает схо-

димось. При слишком больших значениях NFREE алгоритм RMM-DIIS может расходиться.

Параметр ROTIM в данном случае имеет смысл весового коэффициента, определяющего абсолютную величину сдвига ионов. Выбор адекватного значения ROTIM важен и может значительно ускорить вычисления. Способ оценки оптимального значения ROTIM можно посмотреть в случае IBRION = 2.

IBRION = 2 При IBRION = 2 для поиска основного состояния используется метод сопряженных градиентов. Его действия подчиняются следующему алгоритму:

1. На первом шаге позиции ионов (и форма ячейки) изменяются вдоль направления наискорейшего уменьшения энергии, определенного с помощью сил и тензора напряжений. Величина шага масштабируется с учетом параметра ROTIM. Затем энергия и силы вычисляются еще раз.
2. Приближенный минимум полной энергии вычисляется по кубической (или квадратичной) интерполяции, учитывающей изменение полной энергии и сил. Выполняется шаг.
3. Вновь рассчитывается энергия, силы и проверяется имеют ли силы значительную составляющую в направлении параллельном предыдущему смещению. Если это так, то минимум энергии уточняется дополнительным шагом с использованием варианта алгоритма Брента.

Пункты 2 и 3 повторяются до тех пор, пока не будет найдено основное состояние.

Рассмотрим данные, выдаваемые программой. Следующие строки выводятся после каждого шага коррекции (обычно после нечетного шага):

```
1 trial: gam= .00000 g(F)= .152E+01 g(S)= .000E+00 ort = .000E+00  
2 (trialstep = .82)
```

Величина «gam» — параметр сопряжения к предыдущему шагу, «g(F)» и «g(S)» — нормы силы и тензора напряжений, соответственно. Величина «ort» является индикатором того, является ли текущее направление поиска ортогональным к предыдущему (для оптимального шага эта величина должна быть намного меньше $(g(F) + g(S))$). Величина «trialstep» представляет собой текущую величину среднего размера шага, приводившего к минимуму на предыдущем ионном шаге. Оптимальный ROTIM для других методов релаксации можно определить, умножив текущее значение ROTIM на величину «trialstep».

После выполнения шага моделирования в выходной поток выводится следующая информация:

```
1 trial-energy change: -1.153185 1.order -1.133 -1.527 -.739
2 step: 1.7275(harm= 2.0557) dis= .12277
3 next Energy= -1341.57 (dE= -.142E+01)
```

Рассмотрим данные по строкам.

1. Значение «trial-energy change» показывает изменение энергии за шаг. Первое значение после «1.order» — ожидаемое изменение энергии, оцененное на основании сведений о силах $((F(start) + F(trial))/2 \times \Delta R_I)$. Второе и третье значения соответствуют $F(start) \times \Delta R_I$ и $F(trial) \times \Delta R_I$.
2. Первое значение во второй строке равно величине шага в выбранном направлении, приводящего к минимуму. Это значение вычисляется из интерполяционной формулы третьего порядка по значениям сил и энергий на текущем и предыдущем шагах. Величина «harm» — оптимальный шаг, оцененный по интерполяционной формуле второго порядка (или гармонической). Для гармонической интерполяции используется только информация о силах (и их производных). Вблизи минимума первое и второе числа должны быть близки. Величина «dis» — максимальное смещение ионов в дробных координатах.
3. Величина «next Energy» является предположением относительно значения энергии, которое будет получено на следующем шаге. Величина «dE» — оценка изменения энергии.

В файл OUTCAR в конце каждого шага моделирования записывается несколько иные данные:

```
1 trial-energy change: -1.148928 1.order -1.126 -1.518 -.735
2 (g-g1).g = .152E+01 g.g = .152E+01 g1.g1 = .000E+00
3 g(Force) = .152E+01 g(Stress)= .000E+00 ortho = .000E+00
4 gamma = .00000
5 opt step = 1.72745 (harmonic = 2.05575) max dist = .12277085
6 next E = -1341.577507 (d E = 1.42496)
```

Рассмотрим данные по строкам.

1. Первая строка аналогична строке, выведенной в стандартный поток вывода (см. выше).
2. Вспомогательные данные.
3. Величина «g(Force)» соответствует «g(F)», выводимой в стандартный поток; «g(Stress)» соответствует «g(S)»; «ortho» соответствует «ort».
4. «gamma» соответствует «gam».
5. Строка аналогична строке, выведенной в стандартный поток вывода (см. выше).

6. Строка аналогична строке, выведенной в стандартный поток вывода (см. выше).

IBRION = 3 Молекулярная динамика с трением. Работа данного алгоритма регулируется двумя параметрами: POTIM — масштаб сил α и SMASS — коэффициент трения μ .

Для обновления ионных степеней свободы используется уравнение движения второго порядка:

$$\ddot{\mathbf{R}} = -2\alpha\mathbf{F} - \mu\dot{\mathbf{R}}.$$

Для интегрирования этого уравнения используется алгоритм Верле:

$$\mathbf{V}_{N+1/2} = ((1 - \mu/2) \mathbf{V}_{N-1/2} - 2\alpha\mathbf{F}_N) / (1 + \mu/2)$$

$$\mathbf{R}_{N+1} = \mathbf{R}_N + \mathbf{V}_{N+1/2}.$$

Очевидно, что $\mu = 2$ соответствует максимальному трению, а алгоритм превращается в метод наискорейшего спуска. При $\mu = 0$ получается обычная молекулярная динамика. Разумным выбором для μ обычно является значение 0.4.

Для выбора оптимальных параметров коэффициента трения и временного шага рекомендуется сначала зафиксировать μ (например, 1) и найти оптимальный POTIM (максимальный из всех, при которых наблюдается сходимость энергии). Затем следует зафиксировать POTIM и уменьшать SMASS. При правильном выборе параметров POTIM и SMASS данный метод сходится приблизительно в два раза быстрее метода сопряженных градиентов. Обратите внимание, что параметр POTIM имеет смысл шага по времени. Поэтому слишком большой POTIM приводит к расходимости, слишком маленький замедляет скорость работы алгоритма.

Если параметр SMASS не задан или задан отрицательным, то используется метод «заморозки скоростей»: скорость иона обнуляется, если она направлена в сторону противоположную силе.

IBRION = 5 и IBRION = 6 Значения **IBRION = 5** и **IBRION = 6** появились, соответственно, в 4.5 и 5.1 версиях VASP. Оба значения позволяют рассчитать Гессиан (матрицу вторых производных энергии по координате) и колебательный спектр системы. По умолчанию находятся частоты, относящиеся к центральной (Γ -точке) обратной пространства. Частоты записываются в выходной файл после фразы: «Eigenvectors and eigenvalues of the dynamical matrix». Борновские эффективные заряды, пьезоэлектрические константы и ионный вклад в тензор диэлектрической проницаемости также могут быть вычислены, если дополнительно

включены параметры теории линейного отклика `LEPSILON = .TRUE.` или конечного внешнего поля `LCALCEPS = .TRUE.`.

Для расчета Гессиана используются конечные разности, т.е. для каждого иона рассматриваются смещения вдоль каждой из осей декартовой системы координат и определяется изменение силы. Отличие между `IBRION = 5` и `IBRION = 6` заключается в том, что в случае `IBRION = 5` не учитывается симметрия системы и рассматриваются все возможные смещения атомов вдоль осей. В случае `IBRION = 6` симметрия учитывается и рассматриваются только неэквивалентные смещения атомов. Необходимо сразу сделать замечание: для сходимости тензора напряжений параметр обрезания по энергии для плоских волн `ENCUT` должен быть очень большим. Обычно значение по умолчанию нужно увеличить на 30%, но рекомендуется проверить это, исследовав сходимость с некоторым шагом (например 15%).

На расчет матрицы Гессе влияют два параметра: `NFREE` — количество смещений вдоль каждого направления от текущего узла в методе конечных разностей; `POTIM` — расстояние между узлами (размер шага). Если во входном файле указан очень большой размер шага, он приравнивается 0.015 \AA (начиная с версии VASP 5.1). При четных `NFREE` используются центральные разности, т.е. количество узлов слева и справа от центрального x_I одинаково. Для `NFREE = 2` и x координаты иона это:

$x_I - POTIM, x_I, x_I + POTIM$.

Аналогично для координат y и z . Для `NFREE = 4` выполняется четыре смещения:

$x_I - 2 \cdot POTIM, x_I - POTIM, x_I, x_I + POTIM, x_I + 2 \cdot POTIM$.

Для `NFREE = 1` рассматривается только одно смещение, дающее большую погрешность, поэтому разработчики VASP не рекомендуют использовать это значение.

Режим «Selective dynamics» (см. описание в разделе 2.3.1) настоящее время поддерживается только для параметра `IBRION = 5`. В этом режиме рассчитываются только компоненты Гессиана, соответствующие координатам атомов, для которых установлен параметр `.TRUE.` в файле POSCAR. Режим «Selective dynamics» всегда относится к декартовым компонентам матрицы Гессе. Например, для следующего POSCAR файла, второй атом смещается только в направлении оси x и рассчитывается только x -компонента матрицы Гессе для второго атома.

```
1 Cubic BN
2 3.57
3 0.0 0.5 0.5
4 0.5 0.0 0.5
5 0.5 0.5 0.0
```

```

6 | 1 1
7 | selective
8 | Direct
9 | 0.00 0.00 0.00 F F F
10| 0.25 0.25 0.25 T F F

```

Режим `IBRION = 6` при условии, что разрешено вычисление тензора упругости (`ISIF > 2`), позволяет рассчитать упругие константы. Тензор упругости определяется на основании связи между напряжением и деформацией по шести различным деформациям решетки. Тензор упругости определяется для фиксированного положения ионов и для ионов, которые могут релаксировать при изменении решетки. Модули упругости для фиксированного положения ионов выводятся после строки: «SYMMETRIZED ELASTIC MODULI (kBar)». Ионный вклад в тензор упругости определяется через умножение обратной матрицы к матрице Гессе на тензор внутренних деформаций. Соответствующие вклады выводятся после строки: «ELASTIC MODULI CONTR FROM IONIC RELAXATION (kBar)». Окончательные значения модулей упругости, состоящие из вкладов от неподвижных ионов и от релаксировавших ионов, приводятся после строки «TOTAL ELASTIC MODULI (kBar)».

В некоторых предыдущих версиях NSW (число шагов ионов) нужно было устанавливать 1 в файле INCAR, т.к. при `NSW = 0` параметр `IBRION` автоматически устанавливался равным `-1`, независимо от значения в файле INCAR. В VASP 4.6 для параметров `IBRION = 5-6` не было автоматического изменения набора k -точек после смещение атомов, поэтому вычисления по методу конечных разностей могли приводить к неточным результатам.

IBRION = 44 При `IBRION = 44` VASP переключается в режим поиска переходного состояния (седловой точки на поверхности энергии) по усовершенствованному методу димера (УМД, IDM), предложенному Хейде-ном и др. [26].

Алгоритм метода УМД состоит в циклическом повторении следующих шагов:

- методом конечных разностей вычисляется кривизна поверхности энергии вдоль направления оси димера. Начальное направление димера должно быть задано пользователем (см. ниже).
- ось димера поворачивается в направлении параллельном направлению с максимальной по модулю отрицательной кривизной.

- делается шаг оптимизации, при этом ищется максимум потенциальной энергии вдоль оси димер (нестабильного направления) одновременно с минимумом энергии по всем другим направлениям.

Важная информация о ходе оптимизации записывается в файл OUTCAR после фразы «DIMER METHOD». В частности, полезно проверить кривизну вдоль направления димера, которая должна быть отрицательна (длинная последовательность положительных чисел обычно указывает на то, что алгоритм не сходится).

Кроме установки параметра `IBRION = 44` пользователь должен указать направление поиска переходного состояния (нестабильной моды). Направление нестабильной колебательной моды может быть получено путем предварительного анализа колебаний при `IBRION = 5` и определения x-, y- и z- компоненты мнимой колебательной моды параллельной направлению реакции (после деления на квадратный корень из массы). Чтобы эта информация выводилась на печать, необходимо установить `NWRITE = 3`. Тогда после строки «Eigenvectors after division by SQRT(mass)» будут находиться нужные данные.

Направление поиска в виде $3N$ -мерного вектора задается в файле POSCAR после строк с координатами атомов и отделяется пустой строкой. Отметим, что это направление автоматически нормируется, т.е. модуль вектора не имеет значения. Ниже дан пример входного POSCAR файла для данного метода.

```

1 ammonia flipping
2 1.
3 6. 0. 0.
4 0. 7. 0.
5 0. 0. 8.
6 H N
7 3 1
8 cart
9 -0.872954 0.000000 -0.504000 ! координаты первого атома
10 0.000000 0.000000 1.008000
11 0.872954 0.000000 -0.504000
12 0.000000 0.000000 0.000000 ! координаты $N$-го атома
13 ! пробное направление:
14 0.000001 0.522103 -0.000009 ! компоненты для атома 1
15 -0.000006 0.530068 0.000000
16 -0.000005 0.522067 -0.000007
17 0.000001 -0.111442 0.000001 ! компоненты для атома N

```

Как и в случае других методов оптимизации структуры в VASP, сходимость контролируется параметром EDIFFG. Опытные пользователи могут

повысить эффективность метода димера, изменяя значения следующих параметров.

- Параметр `FINDIFF` определяет тип разностной формулы численного дифференцирования для расчета кривизны: 1 (forward) или 2 (центральные). Значение по умолчанию: `FINDIFF = 1`.
- Действительное значение `DIMER_DIST` задает размер шага для численного дифференцирования в Å. Значение по умолчанию: `DIMER_DIST = 0.01`.
- Действительное значение `MINROT` определяет минимальный угол (в радианах), на который может быть повернут димер. Значение по умолчанию: `MINROT = 0.01`
- Действительное значение `STEP_SIZE` задает размер шага оптимизации в Å. Значение по умолчанию: `STEP_SIZE = 0.01`
- Действительное значение `STEP_MAX` определяет верхнюю границу для шага оптимизации в Å. Значение по умолчанию: `STEP_MAX = 0.1`

Важное замечание: текущая реализация алгоритма не позволяет выполнять оптимизацию решетки (`ISIF > 2`) и может быть использована только для релаксации положений атомов.

NELM, NELMIN и NELMDL (количество шагов самосогласования для электронов)

Целочисленный параметр `NELM` (значение по умолчанию: 60) задает максимально возможное количество шагов самосогласования электронной структуры на каждом шаге моделирования движения ионов. Как правило этот параметр менять нет необходимости: если критерии сходимости не были достигнуты за 40 шагов, то скорее всего они не будут достигнуты никогда. В этом случае следует обратить внимание на параметры `IALGO`, (`ALGO`), `LDIAG` и параметры смешивания.

Целочисленный параметр `NELMIN` (значение по умолчанию: 2) определяет минимальное количество шагов самосогласования на каждом шаге моделирования. Обычно изменять этот параметр не нужно. В некоторых случаях (например, при расчетах методом молекулярной динамики или при оптимизации геометрии) можно увеличить этот параметр до 4–8.

Целочисленный параметр `NELMDL` задает количество шагов вычисления электронной структуры вначале шага моделирования, которые выполняются без самосогласования. Значения по умолчанию:

- -5 при `ISTART = 0`, `INIWAV = 1` и `IALGO=8`;
- -12 при `ISTART=0`, `INIWAV=1` и `IALGO=48` (VASP 4.4);
- 0 — во всех остальных случаях

При случайном способе инициализации орбиталей они мало похожи на реальные. Это же можно сказать и о полученной с их использованием плотности заряда. По этой причине имеет смысл на протяжении нескольких первых шагов не изменять Гамильтониан, рассчитывая его по суперпозиции атомных зарядовых плотностей. Применение такого подхода оказывается полезным во всех случаях, когда сходимость в цикле самосогласования медленная (например, при исследовании поверхностей, кластеров, атомных цепочек). Параметр NELMDL может быть как положительным (шаги без самосогласования выполняются на каждом шаге моделирования — в общем случае неподходящий способ), так и отрицательным (шаги без самосогласования выполняются только на первом шаге моделирования).

EDIFF (точность согласования электронов)

Действительный параметр EDIFF (значение по умолчанию: 10^{-4}) определяет критерий выхода из цикла самосогласования электронной подсистемы. Цикл самосогласования завершается когда изменения и полная энергия системы, и собственных значений энергии в зонной структуре от шага к шагу становится меньше EDIFF. Для EDIFF = 0 всегда будут выполняться NELM шагов самосогласования. Обратим внимание на то, что в большинстве случаев время работы экспоненциально зависит от EDIFF. Так, если вы хотите получить значение энергии с четырьмя достоверными (с точки зрения расчетной схемы, а не в сравнении с экспериментом) цифрами после запятой, установите EDIFF = 10^{-4} . Использование меньших значений обычно не имеет смысла.

NSW (число шагов моделирования)

Целочисленный параметр NSW (значение по умолчанию: 0) задает максимальное количество шагов моделирования для движения ионов. Обратим внимание на то, что на каждом шаге моделирования может выполняться до NELM шагов самосогласования, если ранее не будет достигнут критерий EDIFF. Силы Гельмана-Феймана и механические напряжения вычисляются на каждом шаге моделирования.

POTIM (масштаб сдвига ионов)

Действительная переменная POTIM определяет шаг по времени в фемтосекундах и должна быть задана явно при IBRION = 0. При IBRION = 1, 2 или 3 параметр POTIM является множителем, влияющим на величины сил и имеет значение по умолчанию 0.5.

ISIF (опции при движении ионов)

Параметр ISIF (допустимые значения: 0, 1, 2, 3, 4, 5, 6) определяет дополнительные опции, использующиеся при перемещении ионов. Таблица ниже поясняет значения параметра ISIF. Значение по умолчанию: 0 при запуске молекулярной динамики (`IBRION = 0`), 2 для всех остальных случаев.

Параметр ISIF определяет какие из возможных степеней свободы используются при релаксации (изменяются ли координаты ионов, размер и форма расчетной ячейки, и т.д.), а также рассчитываются ли силы и тензор напряжений. Расчет тензора напряжений относительно затратная по времени процедура и поэтому по умолчанию в методе молекулярной динамики не применяется. Силы тем не менее рассчитываются всегда.

На момент написания учебного пособия изменение объема и формы ячейки поддерживается только при релаксации структуры (`IBRION = 1, 2, 3`), но не при моделировании методом молекулярной динамики (`IBRION = 0`).

ISIF	расчет сил	расчет тензора напряжений	оптимизация ионов	изменение формы ячейки	изменение объема ячейки
0	да	нет	да	нет	нет
1	да	только след *	да	нет	нет
2	да	да	да	нет	нет
3	да	да	да	да	да
4	да	да	да	да	нет
5	да	да	нет	да	нет
6	да	да	нет	да	да
7	да	да	нет	нет	да

«Только след*» означает, что только значение полного давления, указанное в строке «`external pressure = ... kBar`», является корректным. Отдельные компоненты тензора напряжений в этом режиме являются недостоверными. Поэтому значение флага `ISIFG = 1` нужно использовать с осторожностью.

Замечание: прежде чем проводить релаксацию с изменением объема или формы ячейки вы должны разобраться с напряжением Пулей (фиктивные напряжения, возникающие из-за неполноты электронного базиса) и его коррекцией. Обычно изменение объема должно сопровождаться небольшим увеличением энергии обрезания (примерно в 1.3 раза от обычного значения, или переключением в режим `PREC = high` в VASP 4.4).

2.4. Выходные файлы

VASP в результате расчетов создает большое количество выходных файлов. Кратко перечислим и опишем их, а затем остановимся на каждом более подробно.

Выходные файлы VASP:

- OUTCAR — файл с подробным выводом результатов расчета VASP;
- OSZICAR — файл с информацией о ходе расчета, его сходимости и скорости;
- CONTCAR — файл с информацией о положениях атомов в формате POSCAR;
- CHGCAR — файл с информацией о параметрах решетки, координатах атомов и зарядовой плотности;
- WAVECAR — двоичный файл с информацией о волновых функциях;
- и др.

2.4.1. OUTCAR

В файле OUTCAR находятся промежуточные и окончательные результаты каждого запуска пакета VASP, включающие в себя:

- используемые входные параметры;
- информация об электронных состояниях системы;
- тензоры напряжений;
- силы, действующие на атомы;
- локальные заряды и магнитные моменты на атомах;
- диэлектрические свойства и т.д.

Что именно и в какой момент записывается в файл OUTCAR, можно регулировать с помощью параметра NWRITE, изменяемого в файле INCAR.

2.4.2. OSZICAR

Информация о скорости сходимости и о текущем шаге моделирования записывается в стандартный выходной поток и в файл OSZICAR.

Файл OSZICAR содержит строки вида:

```
reading files
WARNING: wrap around errors must be expected
entering main loop
      N      E          dE          d eps      ncg      rms      rms(c)
CG : 1  -.13238703E+04  -.132E+04  -.934E+02  56  .28E+02
CG : 2  -.13391360E+04  -.152E+02  -.982E+01  82  .54E+01
CG : 3  -.13397892E+04  -.653E+00  -.553E+00  72  .13E+01  .14E+00
CG : 4  -.13400939E+04  -.304E+00  -.287E+00  84  .48E+00  .39E-01
```

```

CG : 5 -.13401306E+04 -.366E-01 -.322E-01 69 .35E+00 .17E-01
CG : 6 -.13401489E+04 -.183E-01 -.169E-01 75 .74E-01 .66E-02
CG : 7 -.13401516E+04 -.267E-02 -.250E-02 68 .47E-01 .37E-02
CG : 8 -.13401522E+04 -.567E-03 -.489E-03 53 .15E-01 .90E-03
1 F= -.13401522E+04 E0= -.13397340E+04 d E = -.13402E+04
trial: gam= .00000 g(F)= .153E+01 g(S)= .000E+00 ort = .000E+00
charge predicted from atoms
charge from overlapping atoms
      N      E              dE          d eps      ncg      rms      rms(c)
CG : 1 -.13400357E+04 -.134E+04 -.926E+01 56 .97E+01

```

Здесь N — номер электронного шага, E — текущее значение свободной энергии системы, dE — изменение свободной энергии между предпоследним и последним шагом, $d \text{ eps}$ — то же, но для энергии зонной структуры. Столбец ncg показывает сколько раз рассчитывался гамильтониан, rms — оценка ошибки в пробных волновых функциях, $rms(c)$ — разность между входной и результирующей плотностью заряда.

Следующая строка содержит информацию о полной энергии после достижения критерия сходимости. Первое значение равно полной свободной энергии F , E_0 — энергия при $\sigma \rightarrow 0$ (1.4.1), а $d E$ — изменение полной энергии на текущем шаге по сравнению с предыдущим шагом.

В случае расчета методом молекулярной динамики ($IBRION = 0$) эта строка имеет другой вид:

```

1 T= 1873.0 E= -.133821E+04 F= -.134015E+04 E0= -.133973E+04
EK= .19368E+01 SP= .00E+00 SK= .00E+00

```

T — текущая температура моделируемой системы, E — полная свободная энергия (включая кинетическую энергию движения ионов и энергию термостата Нозе). Смысл величин F и E_0 не изменился. EK — кинетическая энергия, SP — потенциальная энергия термостата Нозе, SK — соответствующая кинетическая энергия.

2.4.3. CONTCAR

Файл CONTCAR записывается после каждого ионного шага и в конце расчета. Этот файл имеет такой же формат как файл POSCAR и может быть использован для продолжения расчета.

В случае моделирования методом молекулярной динамики ($IBRION = 0$) файл CONTCAR содержит текущие координаты, скорости и предиктор-корректор координаты, необходимые в качестве входных параметров для следующих шагов моделирования.

2.4.4. CHGCAR

Файл CHGCAR содержит вектора трансляции расчетной ячейки, координаты атомов, полную плотность электрического заряда, умноженную на объем $\rho(r) \cdot V_{cell}$ на FFT сетке, одноцентровые PAW заселенности. Файл CHGCAR используется при перезапуске пакета VASP (для визуализации используется файл CHG). В пакете VASP плотность выводится следующей командой (Fortran):

```
WRITE(IU,FORM) (((C(nx,ny,nz),nx=1,NGXC),ny=1,NGYZ),nz=1,NGZC)
```

Наиболее быстро изменяющийся индекс соответствует оси «x», наиболее медленно меняющийся — оси «z». В случае проведения спинполяризованного расчета в файл CHGCAR записывается два набора данных. Первый содержит полную плотность заряда (соответствующую сумме вкладов от спина up и спина down), второй содержит плотность магнитного момента (разность вкладов от спина up и спина down). Для расчета с неколлинеарным спином файл CHGCAR содержит полную плотность заряда и плотность магнитного момента в проекциях на оси x, y и z.

При использовании метода молекулярной динамики ($IBRION = 0$) плотность заряда в файле CHGCAR соответствует стартовой плотности для следующего шага моделирования, т.е. совместим с данными в файле CONTCAR, но не соответствует координатам в файле OUTCAR. Это позволяет файлы CHGCAR и CONTCAR использовать совместно для продолжения расчета методом молекулярной динамики.

2.4.5. WAVECAR

Бинарный файл WAVECAR содержит ряд характеристик моделируемой системы, среди которых:

NBAND — количество зон

ENCUT1 — энергия обрезания

AX — базисные вектора, определяющие суперъячейку

CELEN — собственные значения

SPTRWFP — волновые функции.

При моделировании методом молекулярной динамики ($IBRION = 0$) файлы WAVECAR, CHGCAR и CONTCAR могут совместно использоваться для продолжения расчета. Для статических расчетов и релаксации ($IBRION = -1, 1, 2$) записанные в файл волновые функции являются решением уравнения Шредингера для последнего шага моделирования.

2.5. Наиболее важные параметры, источники ошибок

Все входные параметры описаны в руководстве пользователя VASP, однако корректно задать значения всех параметров непросто. В этом разделе сконцентрируемся на наиболее важных.

2.5.1. Количество зон NBANDS

Значение параметра NBANDS необходимо выбрать таким образом, чтобы в расчет было включено значительное количество пустых зон (по меньшей мере одна). VASP выдаст предупреждение, если это не так.

NBANDS также важен с технической точки зрения: в схемах итеративной диагонализации матриц собственные векторы с энергией, близкой к максимальной, сходятся намного медленнее, чем соответствующие наименьшей энергии. Не включение в расчет незаполненных зон может привести к значительной потере производительности. Рекомендованное разработчиками VASP значение для большинства систем равно $NELECT/2 + NIONS/2$. В некоторых случаях для больших систем можно уменьшить количество дополнительных зон до $NIONS/4$ без потери производительности. Существуют и исключения: переходные металлы требуют гораздо большего количества пустых зон (до $2 \cdot NIONS$).

Чтобы проверить этот параметр, выполните несколько расчетов для фиксированного потенциала ($ICHARG = 12$) с увеличением числа зон (например, начиная с $NELECT/2 + NIONS/2$). Точность 10^{-6} должна быть получена за 10–15 итераций. Имейте в виду, что схема RMM-DIIS ($IALGO = 48$) более чувствительна к числу зон, чем используемый по умолчанию алгоритм сопряженных градиентов ($IALGO = 8$).

2.5.2. Высокоточные количественные и качественные расчеты

Обсудим различие между «высокоточными количественными» (параметр PREC должен быть равен high) и просто «количественными» расчетами (параметр PREC может быть равен medium или low). «Высокоточные количественные» расчеты необходимы, если рассчитываются малые разности энергий (менее 10 мэВ) двух конкурирующих фаз, которые не могут быть описаны в одинаковой суперячейке.

Термин «такая же суперячейка» обозначает ячейку, содержащую такое же количество атомов и обладающую практически такой же геометрией

(т.е. базисные векторы решетки должны быть практически одинаковыми). Для расчета энергетических различий между двумя конкурирующими объемными фазами во многих случаях невозможно найти суперячейку, отвечающую этому критерию. Для вычисления небольших разностей энергии требуется проверить сходимость по всем параметрам (k -точки, FFT-сетки, а иногда и энергию обрезания). В большинстве случаев эти три параметра независимы, поэтому сходимость по каждому из параметров может быть проверена независимо.

Поверхности — довольно сложные объекты. Расчет энергии поверхности относится к «высокоточным количественным» вычислениям. Для ее определения необходимо вычесть из энергии пластины (slab) энергию объемной фазы. Обе энергии должны быть рассчитаны с высокой точностью. Если пластина содержит 20 атомов, ошибка 5 мэВ на каждый атом в объемном образце приведет к ошибке 100 мэВ на каждый поверхностный атом. Но если необходимо рассчитать энергию адсорбции молекул, то точные результаты (с ошибкой в несколько мэВ) могут быть получены при `PREC = Medium`, если исходная энергия сорбента (плоскости) и исходная энергия сорбата (молекулы) вычисляются в той же суперячейке, что и после адсорбции для сорбата и сорбента вместе.

Расчеты, выполненные в рамках метода *ab initio* молекулярной динамики не попадают в категорию высокоточных количественных, потому что форма ячейки и количество атомов остаются постоянными во время вычисления. Большинство расчетов *ab initio* молекулярной динамики могут быть выполнены с `PREC = Low`. Об исключении из этого правила будет сказано позже при обсуждении влияния сетки k -точек.

2.5.3. Обзор типов «технических» ошибок

Технические ошибки можно разделить на четыре категории.

- Ошибки, обусловленные разбиением обратного пространства на k -точки. Эти ошибки не переносимы, т.е. при использовании сетки $9 \times 9 \times 9$ k -точек для решеток ГЦК, ОКЦ, ГПУ и т.д. ошибки оказываются совершенно разными.
- Ошибки, обусловленные параметром `ENCUT (cut-off)`. Эти ошибки в большой степени переносимы, т.е. значение по умолчанию (считывается из файла `POTCAR`) в большинстве случаев является подходящим значением и можно ожидать, что разброс энергий не будет превышать несколько мэВ. Исключением является тензор напряжений, который, как известно, сходится медленно с расширением базисного набора плоских волн.

- Ошибки обращения при Фурье-преобразовании. Эти ошибки обусловлены недостаточной Фурье-сеткой и они ведут себя не так хорошо, как ошибки, связанные с выбором энергии обрезания ENCUT. Если вы используете значение параметра ENCUT по умолчанию (считанное из файла POTCAR), то ошибки преобразований обычно очень малы (несколько мэВ на атом), даже если Фурье-сетка недостаточно хороша. Причина в том, что энергии отсечки по умолчанию в VASP довольно велики.
- Проектирование в прямом пространстве всегда вводит дополнительные (небольшие) ошибки. Эти ошибки также довольно хорошо себя ведут, т.е. если все время использовать одни и те же операторы проектирования, ошибки являются почти постоянными. Во всяком случае, следует избегать оценки разностей энергий между вычислениями с `LREAL = .FALSE.` и `LREAL = On / .TRUE` (см. раздел описание флага LREAL в руководстве VASP). Имейте в виду, что для `LREAL = On` (рекомендуемая настройка) операторы реального пространства оптимизируются VASP в соответствии с ENCUT и PREC и ROPT, то есть при изменении ENCUT или PREC получают разные операторы.

Общий рецепт следующий. Чтобы свести к минимуму ошибки, следует использовать те же настройки для ENCUT, ENAUG, PREC, LREAL и ROPT во всех вычислениях (которые вы планируете сравнивать между собой), и эти флаги должны быть явно указаны в файле INCAR. Кроме того, также желательно использовать одну и ту же суперячейку для всех вычислений, когда это возможно.

2.5.4. Выбор энергии обрезания и разбиения FFT-сетки

В общем случае энергия отсечки ENCUT должна выбираться в соответствии с псевдопотенциалом. Все файлы POTCAR содержат значение энергии отсечки по умолчанию. Используйте это отсечение энергии, но выполняйте тестовые расчеты с различными энергиями отсечки, чтобы проверить корректность результатов. Значение по умолчанию, указанное в файле POTCAR, обычно приводит к ошибке в энергии когезии, которая меньше 10 мэВ. В настоящее время рекомендуется всегда указывать ENCUT вручную для избежания несовместимости между различными вычислениями.

Необходимо понимать разницу между абсолютной и относительной сходимостью. Абсолютная сходимость по отношению к энергии обрезания ENCUT — это скорость сходимости полной энергии, тогда как относительная сходимость — это скорость сходимости разностей энергий между различными фазами (например, энергия ГЦК-структуры минус энер-

гия ОЦК-структуры). Различия в энергии сходятся гораздо быстрее, чем полная энергия. Это особенно верно, если обе ситуации довольно схожи (например, ГПУ–ГЦК). В этом случае ошибка, связанная с конечностью энергии отсечки, «переносима» из одной ситуации в другую. Если две конфигурации сильно отличаются друг от друга (различное распределение s- p- и d- электронов, различная гибридизация), абсолютная сходимость становится все более критичной.

Существуют некоторые эмпирические правила, которым необходимо следовать при каждом расчете: для объемных материалов число плоских волн на атом должно быть от 50 до 100. Меньший базисный набор может привести к серьезным ошибкам. Большой базисный набор редко необходим, и это указывает на плохо оптимизированный псевдопотенциал. Если включить в расчетную ячейку большой вакуум, число плоских волн будет больше (т.е. если 50% суперячейки занимает вакуум → число плоских волн необходимо увеличить в 2 раза).

Выбор FFT-сетки более проблематичен, чем выбор ENCUT, поскольку связанная с ней ошибка плохо переносится от расчета к расчету. Для точного расчета FFT-сетка должна содержать все волновые векторы до $2G_{cut}$, если $E_{cut} = \frac{\hbar^2}{2m}G_{cut}^2$, E_{cut} является энергией обрезания. Увеличение FFT-сетки от этого значения не изменяет результаты, за исключением, возможно, очень небольшого изменения из-за измененного обменно-корреляционного потенциала. Тем не менее не всегда возможно и необходимо использовать такую большую FFT-сетку. В целом только «высокоточные количественные» вычисления требуют такие сетки. Для большинства вычислений, и в частности для псевдопотенциалов с отсечкой по умолчанию, достаточно установить NGX, NGY и NGZ на 3/4 требуемых значений (параметры PREC = Medium или PREC = Low в файле INCAR перед запуском утилиты makeparam или VASP 4.X). Значения, которые гарантируют отсутствие ошибок связанных с FFT-сеткой приведены в выходном OUTCAR файле:

```
WARNING: wrap around error must be expected
set NGX to 22
WARNING: wrap around error must be expected
set NGY to 22
WARNING: wrap around error must be expected
set NGZ to 22
```

Указать на использование достаточно детализированной FFT сетки могут строки

```
soft charge-density along one line
  0      1      2      3      4      5      6      7      8
x 32.0000 -.7711 1.9743 .0141 .3397 -.0569 -.0162 -.0006 .0000
```

```

y 32.0000 6.7863 .0205 .2353 .1237 -.1729 -.0269 -.0006 .0000
z 32.0000 -.7057 -.7680 -.0557 .1610 -.2262 -.0042 -.0069 .0000

```

записанные в файл OUTCAR (можно найти из по строке 'along'). Эти строки содержат зарядовую плотность в обратном пространстве на позициях

$$\mathbf{G} = 2\pi m_x \mathbf{g}^{(x)}, \mathbf{G} = 2\pi m_y \mathbf{g}^{(y)}, \mathbf{G} = 2\pi m_z \mathbf{g}^{(z)}.$$

Последнее число в каждой строке всегда равно нулю (оно явно задано VASP), предшествующие ему значения, деленные на общее число электронов, должно быть меньше $10^{-3} \div 10^{-4}$.

Еще одним критерием являются силы. Если значение «total drift» велико (больше величины необходимой для сходимости силы, например, 0.1 eV/\AA), необходимо увеличить FFT-сетку. В файле OUTCAR эта величина приведена под строкой TOTAL-FORCE:
total drift: -.00273 -.01048 .03856.

Для представления проекционной (центрированной на положениях атомов) плотности заряда используется вторая, более детализированная, FFT-сетка. Как правило, время, затрачиваемое на вычисления с использованием этой сетки, относительно невелико, поэтому в большинстве случаев можно полагаться на значения по умолчанию (утилиту makeparam). В редких случаях, таких как Cu, Fe_pv с чрезвычайно быстро меняющейся («жесткой») проекционной плотностью заряда, может потребоваться увеличить NGXf по сравнению с настройкой по умолчанию. Это можно сделать установив NGXf в файле param.inc или указав значение ENAUG в файле INCAR.

Мягкая (псевдопотенциальная) часть плотности заряда, и полная плотность заряда (сумма проекционной плотности и мягкой части) записываются в файл OUTCAR:

```

total charge-density along one line
  0      1      2      3      4      5      6      7      8
x 32.0000 -.7711 1.9743 .0141 .3397 -.0569 -.0162 -.0006 .0000
y 32.0000 6.7863 .0205 .2353 .1237 -.1729 -.0269 -.0006 .0000
z 32.0000 -.7057 -.7680 -.0557 .1610 -.2262 -.0042 -.0069 .0000

```

Для полного заряда должны выполняться те же критерии, которые выполняются для мягкой части зарядовой плотности. Недостаточная вторая сетка так же может стать причиной некорректных сил (увеличивая значение «total drift»). Имейте ввиду, что вторая сетка используется только

совместно с ультра мягкими псевдопотенциалами (US-PP). Для нормосохраняющих псевдопотенциалов (NC-PP) ни зарядовая плотность, ни локальный потенциал не пересчитываются с использованием более подробной сетки. В этом случае необходимо приравнять $NG(X,Y,Z)F$ к NGX,Y,Z или 1.

2.5.5. Явное определение ENCUT (ENAUG)

В большинстве случаев можно безопасно использовать значения по умолчанию для ENCUT и ENAUG, которые считываются из файла POTCAR. Но есть случаи, когда это может привести к небольшим, легко предотвращаемым неточностям. Например, если вас интересует разность энергий между объемными фазами с различными составами (например, Co-CoSi-Si). В этом случае ENCUT по умолчанию будет отличаться для расчетов чистого Co и чистого Si, но предпочтительно использовать одно и то же значение энергии обрезания для всех вычислений. В этом случае необходимо определить максимальные ENCUT и ENAUG из файлов POTCAR и использовать это значение для всех вычислений.

Другим примером является расчет энергии адсорбции молекул на поверхностях. Чтобы минимизировать непереносимые ошибки, нужно вычислить энергию изолированной молекулы, поверхности и комплекса сорбат-поверхность в одной и той же суперячейке, используя одно и то же значение энергии обрезания. Обычно это требует корректировки ENCUT и ENAUG в файле INCAR. Если вы также хотите использовать оптимизацию координатного пространства ($LREAL = On$), рекомендуется использовать $LREAL = On$ для всех трех вычислений (флаг ROPT также должен быть одинаковым для всех вычислений).

2.5.6. Выбор количества k -точек и метода сглаживания волновых функций

Информация в этом разделе основывается на теории из раздела 1.4.1.

Количество k -точек, необходимых для расчета, зависит от того, насколько расчет должен быть точен, и от того, является ли система металлической. Для расчета металлических систем требуется на порядок большее количество k -точек, чем для расчета полупроводников и диэлектриков. Кроме этого, число k -точек зависит еще и от используемого метода размытия границы поверхности Ферми: при использовании различных методов скорость сходимости в зависимости от количества k -точек будет различной. Ошибка при использовании одной и той же сетки k -точек оказывается совершенно разной для систем с различной симметрией. Таким

образом, для каждой системы в начале ее изучения требуется проводить расчет ее энергии в зависимости от количества используемых k -точек. Единственным исключением из этого правила являются соизмеримые суперячейки. Если для разных расчетов можно использовать одинаковую суперячейку, для них всегда можно использовать один и тот же набор k -точек.

Для диэлектриков для достижения точности около 10 мэВ обычно достаточно использовать 100 k -точек на атом в зоне Бриллюэна. Для металлов необходимо около 1000 k -точек для достижения той же точности. В сложных случаях (переходные металлы с быстро меняющейся DOS на уровне Ферми) может потребоваться увеличить число k -точек до 5000 на атом. Обычно это приводит к уменьшению ошибки до 1 мэВ на атом. Все приведенные числовые значения относились ко всей зоне Бриллюэна. Число неприводимых k -точек может быть существенно меньше. Для ГЦК, ОЦК и простой кубической решеток разбиение зоны Бриллюэна на $11 \times 11 \times 11$ k -точек (1331 k -точек) соответствует 56 неприводимым k -точкам

Для расчета DOS и очень точного определения энергии желательно воспользоваться рекомендациями, приведенными в разделе 2.3.4, где описывается параметр ISMEAR.

Вопросы и задания

1. Подготовьте файл POSCAR, содержащий ГЦК-структуру железа. Используйте параметр решетки, равный 3.515 Å.
2. Создайте файл INCAR для расчета энергии структуры.
3. Создайте файл KPOINTS, содержащий центрированную на Г сетку из $3 \times 3 \times 3$ k -точек.
4. Запустите расчет с подготовленными файлами; для этого скопируйте файл POTCAR из директории PBE/Fe. Рассмотрите изменение полной энергии системы.
5. Измените подготовленный в задании 1 файл так, чтобы он содержал информацию о поверхности железа вдоль плоскости (100). Определите, при какой сетке k -точек такой расчет сойдется.

3. Сценарии использования пакета VASP

3.1. Самосогласованный расчет

В первую очередь при проведении расчетов физико-химических свойств с помощью пакета VASP обычно проводится самосогласованный расчет для определения волновых функции и зарядовой плотности структуры в основном состоянии. Атомная структура для такого расчета может быть получена различными способами: оптимизацией в пакете VASP (см. стр. 90), методом молекулярной динамики (см. стр. 92) или по данным эксперимента. Ниже приведен пример наиболее простого варианта файла INCAR, который может использоваться для расчета ГЦК решетки кремния.

```
1 SYSTEM = fcc Si
2 ISTART = 0
3 ICHARG = 2
4 ISMEAR = 0
5 SIGMA = 0.1
```

Файл POSCAR для этого расчета будет выглядеть так:

```
1 fcc Si
2 3.9
3 0.5 0.5 0.0
4 0.0 0.5 0.5
5 0.5 0.0 0.5
6 1
7 cartesian
8 0 0 0
```

Структура файла POSCAR описана в разделе 2.3.1. Поскольку расчетная ячейка имеет малый размер, необходимо использовать детализированное разбиение обратного пространства. Построим центрированную на Γ сетку из $11 \times 11 \times 11$ k -точек:

```
1 k-points
2 0
3 Gamma
4 11 11 11
5 0 0 0
```

Наибольший интерес в таком расчете представляет значение полной энергии системы, так как с ее использованием можно определить много

важных количественных характеристик. Например, энергию образования такой структуры кремния.

3.2. Моделирование STM изображения

Пакет VASP предоставляет возможность смоделировать STM (сканирующая туннельная микроскопия) изображение исследуемого объекта. Это может оказаться полезным при сравнении результатов с экспериментом. Рассмотрим в качестве примера ГЦК решетку кремния. Файлы POSCAR и KPOINTS можно выбрать такими же как для расчета основного состояния в разделе 3.1.

По сравнению с самосогласованным расчетом отметим возможные изменения в файле INCAR. Нужно изменить значение переменной ISTART с 0 на 1. Это позволит считать сведения об орбиталях из файла WAVECAR и запустить расчет в режиме перезапуска с постоянной энергией обрезания («restart with constant energy cut-off»). При дополнительном расчете нет необходимости записывать в файл сведения об орбиталях, поэтому нужно задать LWAVE = .FALSE.. Параметр LPARD необходимо задать равным .TRUE. для расчета частичных (для определенных зон и k -точек) плотностей заряда. Интервал энергий, в котором эти зарядовые плотности рассчитываются задается после ключевого слова EINT. Если вместо двух действительных чисел задано одно, то второе принимается равным энергии Ферми. В этом примере будем учитывать область энергий в 1 эВ ниже энергии Ферми. Для того, чтобы указывать энергии относительно уровня Ферми параметр NBMOD следует принять равным -3. Учитывая указанные замечания, файл INCAR будет иметь следующий вид:

```
1 SYSTEM = fcc Si
2 ICHARG = 2
3 ISMEAR = 0
4 SIGMA = 0.1
5
6 ISTART = 1
7 LWAVE=.FALSE.
8 LPARD=.TRUE.
9 EINT=-1 0
10 NBMOD=-3
```

3.3. Зонная структура

Расчет плотности электронных состояний (DOS) может быть проведен двумя способами. Наиболее простым из них является проведение статического ($NSW = 0$, $IBRION = -1$) самосогласованного расчета с последующим использованием полученных файлов DOSCAR и vasprun.xml. Файл vasprun.xml можно визуализировать с помощью программы r4v (см. стр. 98).

Указанный простой способ применим не во всех случаях. Для расчета DOS с высокой точностью может потребоваться использование очень детализированной сетки в обратном пространстве: до $24 \times 24 \times 24$ k -точек для малых и до $6 \times 6 \times 6$ для больших расчетных ячеек. Схожая проблема возникает при расчете зонной структуры, где для получения достоверных результатов требуется рассчитать собственные значения вдоль определенного высокосимметричного направления в зоне Бриллюэна (не менее 10 k -точек).

Поскольку сходимость зарядовой плотности и эффективного потенциала достигается достаточно быстро при увеличении количества k -точек, часто оказывается полезно найти самосогласованную плотность заряда с использованием небольшого числа k -точек. На следующем этапе проводится несамосогласованный расчет с использованием рассчитанного ранее файла CHGCAR ($ICHARG = 11$). Такая возможность в настоящее время реализована только для теории функционала электронной плотности.

При расчете методом теории функционала электронной плотности с параметром $ICHARG = 11$ расчет для всех k -точек проводится независимо а потенциал остается фиксированным. Существует возможность провести расчет собственных значений отдельно для каждой k -точки, но на современных вычислительных платформах это редко используется.

Для гибридных функционалов и метода Хартри–Фока зонная структура может быть рассчитана с учетом дополнительных k -точек с нулевым весом из файла KPOINTS. Для этого необходимо провести расчет с использованием стандартного файла KPOINTS. Полученный файл IBZKPT, содержащий список использованных k -точек, переименовать в файл KPOINTS и добавить необходимые k -точки с нулевым весом. Поскольку VASP использует итеративный метод диагонализации матриц и так как добавленные k -точки не влияют на энергию, необходимо выполнить как минимум 5 итераций до получения достоверных значений энергий в k -точках с нулевым весом ($NELMIN = 5$).

Рассмотрим входные файлы для расчета зонной структуры на примере ГЦК решетки кремния. Файл POSCAR будет идентичным файлу, исполь-

зованному для самосогласованного расчета энергии ГЦК-Si (см. раздел 3.1). Файл INCAR может выглядеть так:

```

1 System = fcc Si
2 ICHARG = 11
3 ISMEAR = 0
4 SIGMA = 0.1
5 LORBIT = 11

```

Структура файла KPOINTS будет указывать на то, что необходимо провести расчет зонной структуры:

```

1 k-points for bandstructure L-G-X-U K-G
2 10
3 line
4 reciprocal
5 0.50000 0.50000 0.50000 1
6 0.00000 0.00000 0.00000 1
7
8 0.00000 0.00000 0.00000 1
9 0.00000 0.50000 0.50000 1
10
11 0.00000 0.50000 0.50000 1
12 0.25000 0.62500 0.62500 1
13
14 0.37500 0.7500 0.37500 1
15 0.00000 0.00000 0.00000 1

```

3.4. Электронная плотность

Плотность электронных состояний (Density Of States, DOS, ПЭС) \bar{n} определяется отношением

$$\bar{n}(\varepsilon_i) = \frac{N(\varepsilon_i) - N(\varepsilon_{i-1})}{\Delta\varepsilon},$$

где $\Delta\varepsilon$ — разность энергий двух соседних точек в файле DOSCAR, для которых вычисляется разность интегральных DOS, $N(\varepsilon_i)$ — интегральная DOS

$$N(\varepsilon_i) = \int_{-\infty}^{\varepsilon_i} n(\varepsilon)d\varepsilon.$$

При таком расчете ПЭС интеграл от полученной функции будет равен общему числу электронов в системе.

Плотность электронных состояний содержится в файле DOSCAR.

Диапазон энергий, для которых рассчитывается DOS, задается действительными параметрами EMIN и EMAX, выраженными в эВ. Число

энергий внутри указанного диапазона, для которых рассчитывается DOS задается целочисленным параметром NEDOS.

Файл DOSCAR содержит полную и интегральную плотность состояний, которые выражены в состояниях на ячейку. При расчете динамическим методом или оптимизации геометрии в файл DOSCAR записываются усредненная и интегральная DOS. После дополнительной информации, которая занимает первые несколько строк файла DOSCAR содержатся строки из трех чисел:

E DOS integrated-DOS

Для случая спинполяризованного расчета каждая строка содержит пять чисел:

E DOS(up) DOS(dn) integrated-DOS(up) integrated-DOS(dn)

«up» — значение для спина вверх, «dn» — значение для спина вниз. Если параметры RWIGS или LORBIT (радиус Вигнера-Зейтца, см. приложение 4.2.3) заданы в файле INCAR, дополнительно вычисляются проекции DOS на атомные s-, p-, d- орбитали, центрированные на положениях атомов и записываются в файл DOSCAR. Для каждого атома записывается NDOS строк в следующем формате

E s-DOS p-DOS d-DOS

или (с учетом спина)

E s-DOS(up) s-DOS(dn) p-DOS(up) p-DOS(dn) d-DOS(up) d-DOS(dn)

Для случая неколлинеарного расчета общая DOS имеет следующий формат:

E DOS(total) integrated-DOS(total)

«total» — полная плотность. Информация для компонент спина доступна только в виде проекций плотности электронных состояний на атомные орбитали. Эта информация представляется в следующем формате:

E s-DOS(total) s-DOS(mx) s-DOS(my) s-DOS(mz) p-DOS(total) ...

Каждая такая строка содержит информацию о полной плотности электронных состояний и намагниченности в проекциях на оси x — (mx), y — (my) и z — (mz).

Единицей измерения парциальной плотности состояний (на орбитальный момент или на атом) является состояний/атом/энергия.

Парциальная плотность состояний в расчете на атом не рассчитывается в следующих случаях:

- версия VASP 4.5, параметр NPAR6 = 1
- версия VASP 4.6, параметры NPAR6 = 1, LORBIT = 0-5.

В версии VASP 4.6 парциальная плотность состояний (на атом) может быть рассчитана для `LORBIT = 10-12` даже если `NPAR` не равен 1 (в отличие от предыдущей версии программы).

Обратите внимание, что при оптимизации геометрии файл `DOSCAR` обычно не представляет интереса. Для получения достоверной DOS конечной конфигурации необходимо переименовать файл `CONTCAR` в `POSCAR` и провести расчет в статическом режиме (`ISTART = 1; NSW = 0`).

3.5. Оптимизация структуры

На практике перед тем, как рассчитывать свойства модельных систем необходимо определить равновесную (отвечающую минимуму энергии) конфигурацию системы. Рассмотрим пример решения такой задачи для ГЦК поверхности никеля (100). Выберем параметр трансляции в направлении оси z настолько большим, чтобы атомы расчетной ячейки не взаимодействовали со своим периодическим продолжением. Расчетная ячейка содержит 5 атомов, координаты трех из них будем считать фиксированными. Файл `POSCAR` будет выглядеть так:

```
1 fcc (100) surface
2 3.53
3 .50000 .50000 .00000
4 -.50000 .50000 .00000
5 .00000 .00000 5.00000
6 5
7 Selective Dynamics
8 cartesian
9 .00000 .00000 .00000 F F F
10 .00000 .50000 .50000 F F F
11 .00000 .00000 1.00000 F F F
12 .00000 .50000 1.50000 T T T
13 .00000 .00000 2.00000 T T T
```

Рассмотрим отличия в файле `INCAR`, которые может потребоваться внести по сравнению с самосогласованным расчетом. Явно зададим энергию обрезания равной 270 эВ. Изменим глобальное условие окончания самосогласованного расчета (`EDIFF = 1E-6`). Изменим алгоритм оптимизации (`ALGO = Fast`). Изменим алгоритм учета частичных заселенностей орбиталей (`ISMEAR = 2`). Параметр `SIGMA` возьмем равным значению по умолчанию (удалим его из входного файла). Укажем параметры оптимизации. Зададим максимальное число шагов оптимизации атомной структуры (`NSW = 100`) и параметр (`POTIM = 0.8`) квазиньютоновского алгоритма (`IBRION = 1`). Файл `INCAR` будет выглядеть следующим образом:

```

1 ISTART = 0;
2 ICHARG = 2
3
4 general:
5 SYSTEM = clean Ni(100) surface
6 ENCUT = 270
7 ISMEAR = 2
8 ALGO = Fast
9 EDIFF = 1E-6
10
11 dynamic:
12 NSW = 100
13 POTIM = 0.8
14 IBRION = 1

```

Так как моделируется атомная плоскость и параметр решетки в координатном пространстве в направлении оси z большой, то в этом направлении достаточно рассмотреть всего одну k -точку. Файл KPOINTS может выглядеть следующим образом:

```

1 k-points
2 0
3 Monkhorst-Pack
4 9 9 1
5 0 0 0

```

3.6. Фононные спектры

К сожалению, процедуры для расчета фононных спектров не включены в VASP. Однако, существует множество сторонних программ, основанных на VASP и реализующих расчеты фононных спектров. В настоящем пособии мы остановимся на одной из таких программ, называемых PhonoPy (см. раздел 4.2).

3.7. Учёт магнетизма

Для учета магнитных явлений в INCAR используются параметры ISPIN и MAGMOM. Рассмотрим пример: расчет парциальной плотности электронных состояний для ферромагнитной ГЦК структуры никеля. Файл POSCAR предлагается использовать такой:

```

1 fcc:
2 -10.93
3 0.5 0.5 0.0

```

```

4 | 0.0 0.5 0.5
5 | 0.5 0.0 0.5
6 | 1
7 | Cartesian
8 | 0 0 0

```

Для проведения спинполяризованного расчета приравняем $ISPIN = 2$. Начальный магнитный момент каждого атома примем равным единице ($MAGMOM = 1.0$).

Можно использовать следующий файл INCAR:

```

1 | SYSTEM = Ni fcc bulk
2 | ISTART = 0
3 | ISPIN = 2
4 | MAGMOM = 1.0
5 | ISMEAR = -5
6 | VOSKOWN = 1
7 | LORBIT = 11

```

Для расчета будем использовать уже полюбившуюся нам сетку $11 \times 11 \times 11$ k -точек, KPOINTS файл для которой описан в разделе 3.1.

3.8. Молекулярная динамика

Рассмотрим молекулярно–динамическое моделирование молекулы H_2O . В файле POSCAR укажем декартовы координаты каждого из трех атомов молекулы. z -компоненты координат атомов будем считать фиксированными. Для моделирования поведения изолированной от окружения молекулы, параметры трансляции выберем достаточно большими, чтобы исключить взаимодействие атомов расчетной ячейки со своими периодическими изображениями. Для таких расчетных ячеек достаточно рассматривать одну Γ -точку в обратном пространстве. Файл POSCAR будет выглядеть следующим образом:

```

1 | H2O molecular dynamics
2 | 0.52918
3 | 12 0 0
4 | 0 12 0
5 | 0 0 12
6 | 1 2
7 | select
8 | cart
9 | 0.00      0.00      0.00 T T F
10 | 1.10     -1.43      0.00 T T F
11 | 1.10      1.43      0.00 T T F

```

Шестая строка файла POSCAR задает количество атомов каждого сорта в системе. Файл POTCAR должен содержать сначала псевдопотенциал для кислорода, а затем — для водорода.

Тип расчета (молекулярная динамика) задается значением переменной IBRION = 0. Значением переменной SMASS = -3 выберем микроканонический ансамбль. В данном примере рассмотрим моделирование динамики молекулы при постоянной температуре. Для этого зададим начальную температуру (TEBEG) равной конечной (TEEND). При проведении расчетов методом молекулярной динамики необходимо исключить учет симметрии (ISYM = 0). Шаг по времени выберем равным 1 фс. Это значение типично для расчетов неэмпирической молекулярной динамики. Файл INCAR с параметрами расчета может быть следующим:

```
1  PREC = Normal
2  ENMAX = 400
3  ISMEAR = 0
4  SIGMA = 0.1
5  ISYM = 0
6  IBRION = 0
7  NSW = 100
8  POTIM = 1.0
9  SMASS = -3
10 TEBEG = 2000
11 TEEND = 2000
```

3.9. Переходное состояние

Рассчитаем переходное состояние для инверсии азота в молекуле аммиака. Все рассмотренные здесь расчеты выполняются с использованием функционала PBE, зона Бриллюэна содержит только Γ -точку. Исходная структура для IDM-моделирования (усовершенствованный метод димера) должна быть близка к состоянию перехода.

Файл POSCAR для аммиака:

```
1  ammonia flipping
2  1.
3  6. 0. 0.
4  0. 7. 0.
5  0. 0. 8.
6  H N
7  3 1
8  Cartesian
9  -0.872954 0.000000 -0.504000
10 0.000000 0.000000 1.008000
```

```
11 0.872954 0.000000 -0.504000
12 0.000000 0.000000 0.000000
```

В качестве входного сигнала для метода димера необходимо направление неустойчивой моды (ось димера). Его можно получить, выполнив анализ колебаний. Файл INCAR должен содержать следующие строки:

```
1 NSW = 1
2 Prec = Normal
3 IBRION = 5
4 NFREE = 2
5 POTIM = 0.02
6 NWRITE = 3
```

После выполнения анализа колебаний, необходимо найти самую ярко выраженную мнимую моду в файле OUTCAR:

```
12 f/i= 23.224372 THz 145.9230 2PiTHz 774.6816 cm-1 96.0483 meV
X Y Z dx dy dz
5.127046 0.000000 7.496000 0.000001 0.522103 -0.000009
0.000000 0.000000 1.008000 -0.000006 0.530068 0.000000
0.872954 0.000000 7.496000 -0.000005 0.522067 -0.000007
0.000000 0.000000 0.000000 0.000001 -0.111442 0.000001
```

и использовать последние три столбца для описания оси димера в файле POSCAR:

```
1 ammonia flipping
2 1.
3 6. 0. 0.
4 0. 7. 0.
5 0. 0. 8.
6 H N
7 3 1
8 cart
9 -0.872954 0.000000 -0.504000 ! координаты атома 1
10 0.000000 0.000000 1.008000
11 0.872954 0.000000 -0.504000
12 0.000000 0.000000 0.000000 ! координаты атома N
13 ! определяем ось димера:
14 0.000001 0.522103 -0.000009 ! компоненты для атома 1
15 -0.000006 0.530068 0.000000
16 -0.000005 0.522067 -0.000007
17 0.000001 -0.111442 0.000001 ! компоненты для атома N
```

Для выполнения IDM расчета INCAR файл должен содержать следующие строки:

```
1 NSW = 100
2 Prec = Normal
```

```
3 IBRION = 44 ! использовать метод димера
4 EDIFFG = -0.03
```

После этой настройки алгоритм сходится всего за несколько шагов оптимизации. Необходимо провести дальнейший анализ колебаний, чтобы доказать, что полученная структура действительно является седловой точкой первого порядка (одна мнимая частота).

3.10. Спин-орбитальное взаимодействие

Как показывает практика, с использованием обменно-корреляционных функционалов LDA и GGA редко удается описать физические свойства систем с локализованными (сильно коррелированными) d- и f- электронами. В ряде случаев это можно исправить введя сильное внутриатомное взаимодействие, подобное потенциалу Хартри–Фока. Этот подход называют методом LDA+U. В пакете VASP он активируется значением переменной LDAU = .TRUE.. Кроме того, для некоторых атомных структур экспериментальные данные воспроизводятся существенно лучше, если провести расчет с неколлинеарным спином. Примером такого материала является оксид никеля. Структуру NiO можно задать следующим POSCAR файлом.

```
1 NiO
2 4.17
3 1.0 0.5 0.5
4 0.5 1.0 0.5
5 0.5 0.5 1.0
6 2 2
7 Cartesian
8 0.0 0.0 0.0
9 1.0 1.0 1.0
10 0.5 0.5 0.5
11 1.5 1.5 1.5
```

Файл KPOINTS может иметь следующий вид:

```
1 k-points
2 0
3 gamma
4 4 4 4
5 0 0 0
```

Файл INCAR может иметь следующий вид:

```
1 NiO LDA+U SOC
2 SYSTEM = "NiO"
3
```

```

4 | Electronic minimization
5 | ENCUT      = 450
6 | EDIFF      = 1E-5
7 | LORBIT     = 11
8 | LREAL      = .False.
9 | ISTART     = 0
10 | ISYM       = -1
11 | NELMIN     = 6
12 | LSORBIT    = .True.
13 | LWAVE      = .False.
14 | LCHARG     = .False.
15 |
16 | DOS
17 | ISMEAR     = -5
18 |
19 | Magnetism
20 | ISPIN      = 2
21 | MAGMOM     = 0 0 2 0 0 -2 6*0
22 | SAXIS      = 2 2 2
23 |
24 | Orbital Moment
25 | LORBMOM    = T
26 |
27 | Mixer
28 | AMIX       = 0.2
29 | BMIX       = 0.00001
30 | AMIX_MAG   = 0.8
31 | BMIX_MAG   = 0.00001
32 |
33 | LDA+U
34 | LDAU       = .TRUE.
35 | LDAUTYPE   = 2
36 | LDAUL      = 2 -1
37 | LDAUU      = 5.00 0.00
38 | LDAUJ      = 0.00 0.00
39 | LDAUPRINT  = 2
40 | LMAXMIX    = 4

```

Часть параметров файла INCAR вам уже знакомы, описание остальных приведено в руководстве пользователя пакета VASP. Остановимся лишь на некоторых из них. С помощью параметра MAGMOM задаются проекции магнитного момента каждого атома в файле POSCAR. Отсутствие магнитного момента на атомах кислорода кратко обозначено 6*0 (шесть компонент). Параметр SAXIS задает направление квантования для неколлинеарного спина (три компоненты). Варьирование этого параметра позволяет найти направление магнитного момента, минимизирующее полную энергию системы. Параметр LORBMOM = T активирует вывод орби-

тального момента в файл OUTCAR. На скорость (а иногда и наличие) сходимости самосогласованного расчета влияют параметры из раздела «Mixer». В большинстве случаев можно использовать значения по умолчанию, и только в случае отсутствия сходимости изменять приведенные в файле INCAR параметры.

Вопросы и задания

1. Определите равновесный параметр решетки ГЦК меди и оптимальное разбиение обратного пространства на k - точки для случаев определения энергии с точностью 10^{-1} эВ/атом и 10^{-2} эВ/атом. Совпадает ли параметр решетки с литературными данными?
2. Модуль объемного сжатия может быть представлен в следующем виде:

$$K = \left[V \frac{\partial^2 E}{\partial V^2} \right]_{V=V_0},$$

где V_0 — объем моделируемой системы, соответствующий минимуму энергии E . Проведите серию расчетов полной энергии ГЦК меди для десяти значений параметра решетки (с шагом 0,002 нм). Определите модуль объемного сжатия ГЦК меди. Результат сравните с литературными данными.

3. Найдите сведения о зонной структуре меди в литературных источниках. Расчитайте зонную структуру меди для тех же направлений в обратном пространстве, что и в найденных литературных источниках. Сравните результаты и сделайте вывод.
4. Расчитайте еще как минимум одну характеристику меди (упомянутую или не упомянутую в данном пособии).
5. Подготовьте отчет по проведенному исследованию. Отчет должен содержать обзор литературных источников о свойствах меди, полученные вами результаты, описание трудностей, с которыми вы столкнулись при проведении исследования.

4. Вспомогательное программное обеспечение

4.1. p4v

4.1.1. Установка

Программа p4vasp распространяется на бесплатной основе и доступна для скачивания на сайте <http://www.p4vasp.at>. Дистрибутив можно распаковать командой

```
tar -xvzf p4vasp-X.Y.Z.tgz
```

После этого необходимо установить все необходимые для работы p4vasp дополнительные пакеты (зависимости). Этот шаг зависит от используемой вами операционной системы. Например, если на вашем компьютере установлена Ubuntu, можно запустить скрипт `install/install-ubuntu-dependencies.sh`. В остальных случаях для получения подробной информации необходимо следовать указаниям в файле README.

В зависимости от того, устанавливаете ли вы p4vasp в одну из поддиректорий вашей домашней директории или для всех пользователей компьютера, необходимо запустить скрипты `install-local.sh` или `install.sh`, соответственно.

Запустить установленную программу можно командой

```
p4v
```

4.1.2. Графический интерфейс программы p4vasp

Перед началом описания деталей работы с программой p4vasp хочется отметить, что на момент написания этого учебного пособия программа p4vasp не в полной мере работоспособна. Об этом свидетельствует то, что после знакомства с руководством пользователя остается неясным назначение части элементов графического интерфейса. Опишем лишь реализованный функционал программы.

На рис. 4.1 приведено главное окно программы p4vasp. Оно состоит из горизонтального меню в верхней части, предоставляющего доступ ко всем возможностям программы. Часть функционала доступна посредством вертикального ряда кнопок в левой части окна. В поле System указывается выходной файл пакета Vasp, который обрабатывается в данный

момент. Если этот файл еще не выбран, то в поле System отображается текст “???(vasprun.xml)”. Наполнение остальной части экрана зависит от выполненных пользователем действий.

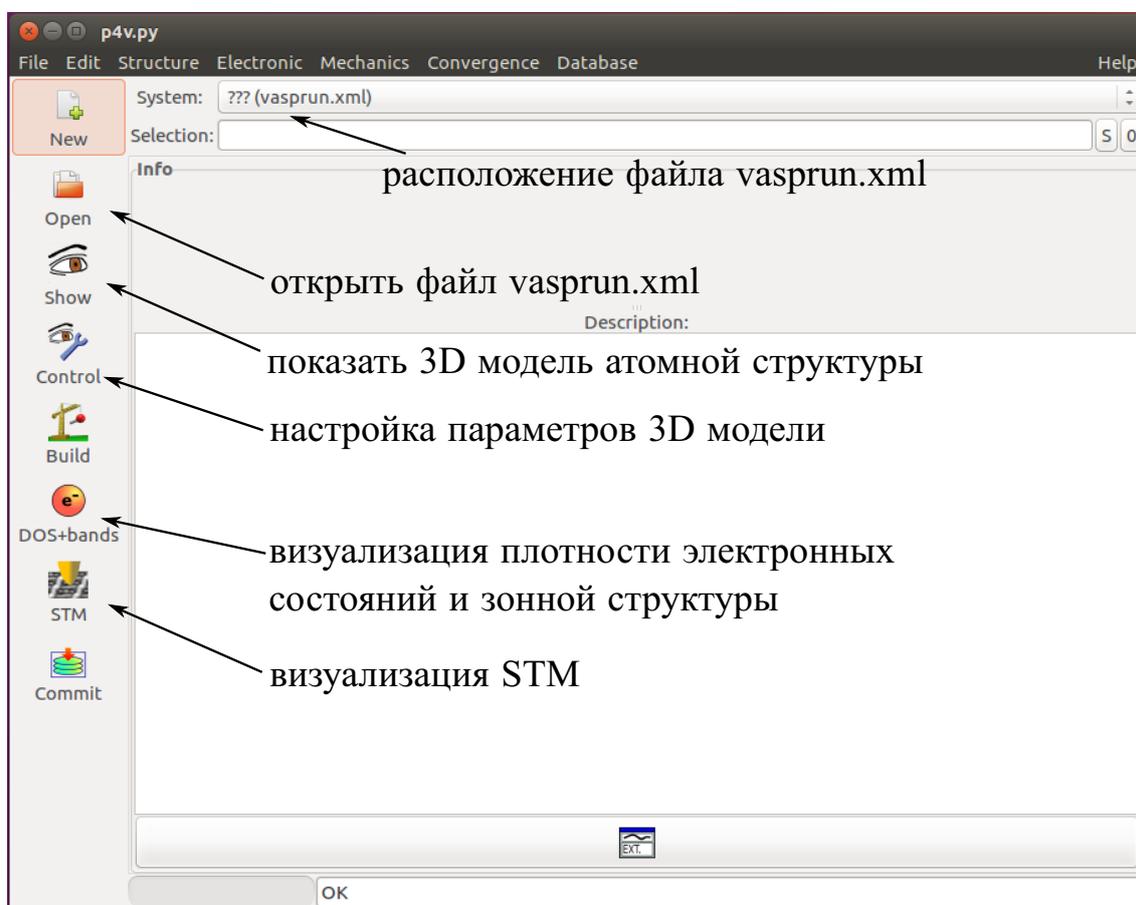


Рис. 4.1. Окно программы p4vasp

4.1.3. 3D визуализация атомной структуры

Проиллюстрируем возможности визуализации атомной структуры на примере ГЦК решетки кремния, параметры самосогласованного расчета для которого приведены в разделе 3.1. Запустите p4vasp и откройте файл vasprun.xml (посредством кнопки Open в левой части окна программы или меню File/Load). После этого информация о модели появится в разделе System в верхней части окна p4vasp. Для отображения трехмерной атомной структуры необходимо нажать кнопку Show. При этом открывается дополнительное окно, пример которого приведен на рис. 4.2 вверху. Настройка параметров визуализации осуществляется посредством вкладки Control.

С помощью программы p4vasp можно смоделировать STM изображение модели (рис. 4.2, внизу). Графический интерфейс позволяет переключаться между двумя режимами (STM Mode): «constant height» и «constant current», менять ориентацию исследуемой плоскости (перпендикулярно осям x , y или z) и настраивать ряд других параметров.

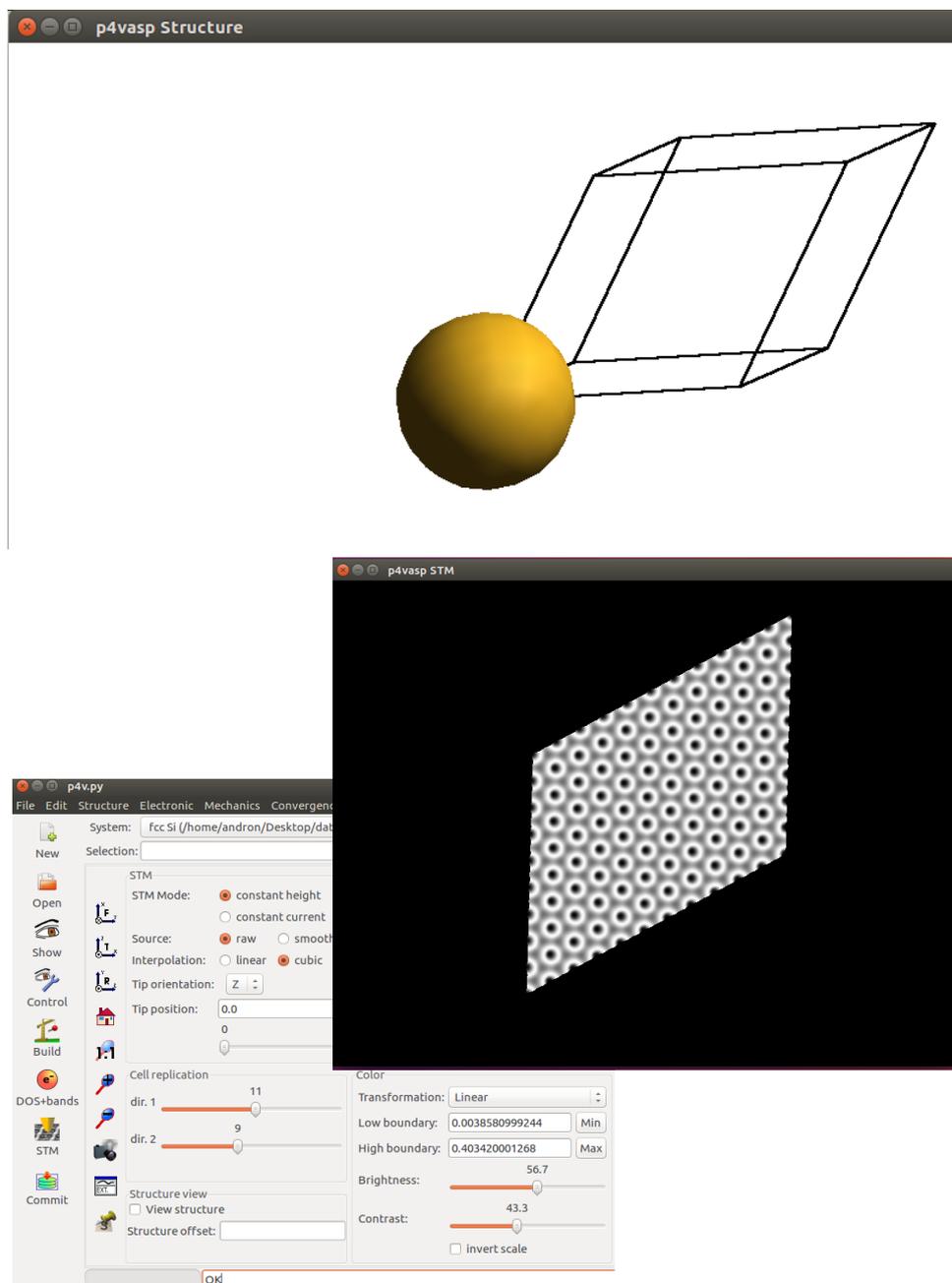


Рис. 4.2. Вверху — визуализация атомной структуры ГЦК–Si. Внизу — визуализация STM изображения модели и окно для его настройки

4.1.4. Визуализация плотности электронных состояний и зонной структуры

Для визуализации зонной структуры с помощью программы `p4vasp` необходимо провести соответствующий расчет пакетом VASP (см. раздел 3.3). Для построения плотности электронных состояний выберите пункт «DOS + bands» на панели инструментов (или «Electronic/DOS + bands» в меню программы). При этом откроется окно с графиком ПЭС (см. рис. 4.3), причем точка $E = 0$ на этом графике будет соответствовать энергии Ферми системы. Необходимо понимать, что ПЭС строится для тех k -точек, которые были указаны в файле `KPOINTS`. Плотность состояний для некоторых k -точек (выбранных для построения зонной структуры) будет выглядеть иначе, чем плотность состояний, построенная на всей сетке. Поэтому для того, чтобы построить и плотность электронных состояний, и зонную структуру (см. рис. 4.4), необходимо проводить два различных расчета.

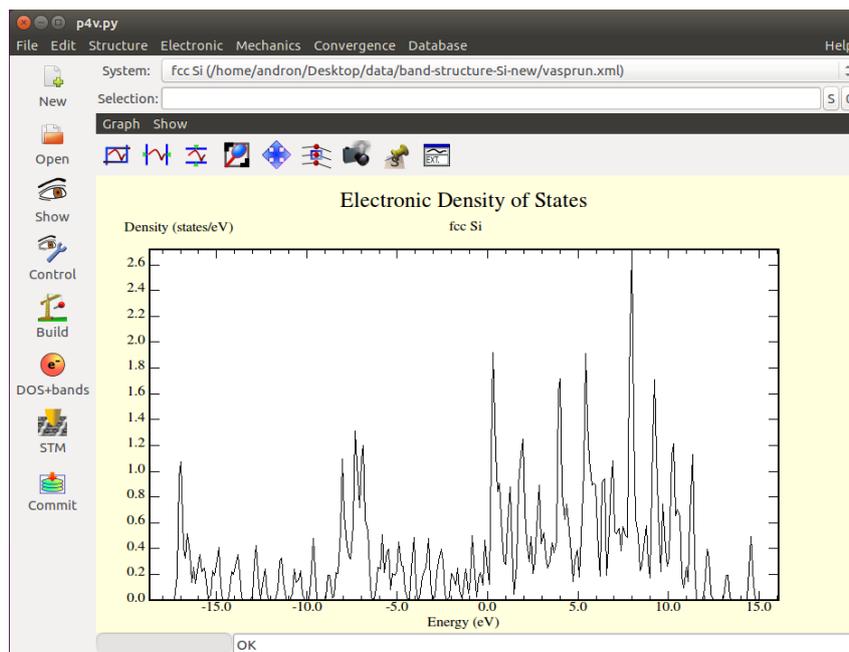


Рис. 4.3. Построение плотности электронных состояний (DOS)

Переключение режимов отображения (DOS / bandstructure) осуществляется с помощью пунктов меню `Show`:

- `show`: показать / скрыть DOS. Эта функция особенно полезна при анализе парциальной плотности состояний.
- `DOS`: отображать плотность состояний (по оси абсцисс — энергия, по оси ординат — DOS).
- `DOS (X-axis)`: отображать плотность состояний (по оси абсцисс — DOS, по оси ординат — энергия).

- Bands: отображать зонную структуру.
- DOS and Bands: отображать DOS и зонную структуру на одном графике.

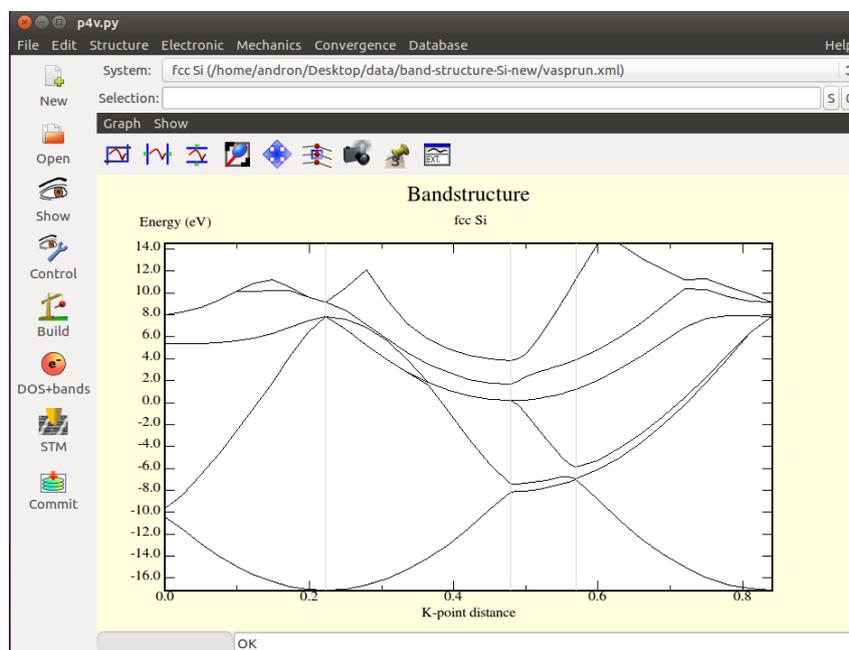


Рис. 4.4. Построение зонной структуры (Band)

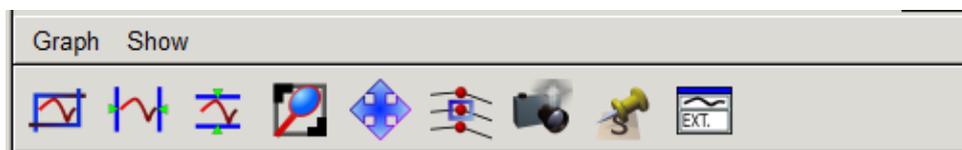


Рис. 4.5. Band

Все графики, отображаемые p4vasp, можно настроить и экспортировать в различных форматах с помощью дополнительной панели инструментов (см. рис. 4.5).

- Первые три кнопки позволяют автоматически выбрать масштаб графика: по двум осям, только по оси x, только по оси y.
- Иконка с лупой позволяет увеличить или уменьшить масштаб изображения:
 - клик левой кнопкой мыши — увеличить;
 - клик правой кнопкой мыши — уменьшить;
 - область для масштабирования можно задать удержанием любой кнопки мыши с одновременным перемещением мыши.
- выбор пятой кнопки (пересекающиеся стрелки) позволяет перемещать график.

- шестая кнопка позволяет увидеть значения, соответствующие точкам на графике.
- кнопка с изображением камеры позволяет экспортировать график в файл с автоматически заданным именем в множестве форматов (например, graph0001.agr, graph0001.ps, graph0001.dat):
 - raw data — необработанные данные;
 - XmGrace agr file — рекомендуемый разработчиками тип графического файла для дальнейшего редактирования с использованием программы XmGrace и окончательного представления результатов;
 - Postscript;
 - Encapsulated postscript;
 - XFig;
 - Scalable Vector Graphics — изображение в векторном формате.
- кнопка Pin/unpin позволяет фиксировать текущую систему, что может быть полезным при сравнении результатов ряда расчетов;
- кнопка Ext. открывает график в отдельном окне.

4.1.5. Сходимость

r4vasp позволяет визуализировать изменение энергии системы и сил, действующих в ней от шага к шагу моделирования. Для этого необходимо выбрать пункт меню Convergence->Forces/ Selected forces /Energy. Полученная при оптимизации структуры никеля (см. раздел 3.5) зависимость сил, действующих на атомы от номера шага моделирования приведена на рис. 4.6.

4.2. Phonopy

Phonopy — программный пакет с открытым исходным кодом, предназначенный для расчета фононных спектров в гармоническом и квазигармоническом приближении. Пакет был написан Ацсуши Того в университете Киото и расположен на странице <https://atztogo.github.io/>. Пакет написан на языке Python, и позволяет рассчитывать основные параметры фононных спектров на основании результатов расчета различных программ, реализующих теорию функционала электронной плотности, в том числе и VASP. Эти параметры включают в себя:

- зонную структуру, плотность состояний и ее проекцию на атомы;
- свободную энергию, теплоемкость при постоянном объеме, энтропию;
- групповую скорость;
- параметр Грюнайзена;

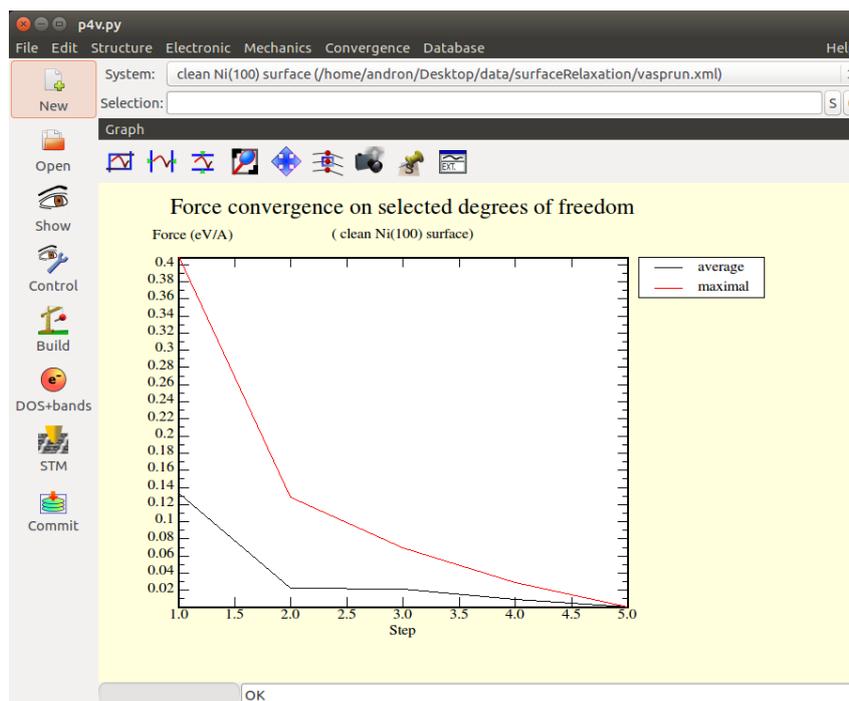


Рис. 4.6. Визуализация сходимости расчета

- и др.

Полный список можно найти на домашней странице Phonopy. Ниже кратко разберем порядок работы с пакетом.

4.2.1. Установка

Установка Phonopy может быть произведена различными способами: из открытых исходных кодов, с помощью менеджера пакетов языка Python `pip` или в рамках дистрибутива Anaconda от компании Continuum Analytics и его пакетного менеджера `conda`. Для работы программы, кроме самого Python и его заголовочных файлов (которые нужны для компиляции модулей Python, написанных на C) следует установить несколько зависимостей:

- математическая библиотека `numpy`;
- построитель графиков `matplotlib`;
- `python-yaml` — парсер языка сериализации данных YAML;
- `python-h5py` — интерфейс к формату данных HDF5.

Наиболее простой способ установки Phonopy — с помощью пакетного менеджера `conda`. Однако, для его работы требуется, чтобы на компьютере был установлен дистрибутив Anaconda, включающий интерпретатор

языка Python и необходимые для научной работы пакеты¹. Тогда Phonopy может быть установлен с помощью следующей команды:

```
conda install -c conda-forge phonopy h5py
```

Установка пакета другими способами описана на странице документации пакета (<https://atztogo.github.io/phonopy/install.html>). Для проверки правильности установки можно запустить команду `phonopy`

4.2.2. Теоретические основы

Фононы — квазичастицы, описывающие коллективные колебания в ансамбле атомов. Рассмотрим кристалл из L элементарных ячеек с S базисными атомами в каждой ячейке (общее число атомов в таком кристалле равно $N = LS$), находящийся при низкой температуре. Для каждой ячейки l можно определить положение s -го иона в ней как $\vec{R}_{ls} = \vec{R}_l + \vec{R}_s$. Ионы в таком кристалле претерпевают малые колебания относительно положения равновесия, причем смещения атомов будут зависеть от их положения в кристалле в целом: для каждого момента времени мы можем записать положение атома $\vec{R}_{ls}(t)$ как

$$\vec{R}_{ls}(t) = \vec{R}_{ls} + \vec{u}_{ls}(t), \quad (4.1)$$

где \vec{R}_{ls} — радиус-вектор положения равновесия s -го атома в l -й ячейке, и $\vec{u}_{ls}(t)$ — его малое смещение в зависимости от времени. Мы можем рассчитать потенциальную энергию такого кристалла, разложив ее в ряд Тейлора около равновесных позиций атомов:

$$U(\vec{R}_{ls}(t)) = U(\vec{R}_{ls}) + \sum_{ls\alpha} \frac{\partial U}{\partial u_{ls\alpha}} u_{ls\alpha}(t) + \frac{1}{2} \sum_{ls\alpha} \sum_{l't\beta} \frac{\partial^2 U}{\partial u_{ls\alpha} \partial u_{l't\beta}} u_{ls\alpha}(t) u_{l't\beta}(t) + \dots$$

В данном разложении α и β — декартовы компоненты вектора \vec{u} , и все производные взяты в равновесном положении кристалла. Первый член ряда — потенциальная энергия идеального кристалла, в котором все атомы находятся в равновесных позициях E_{eq} . Если атомы в кристалле близки к конфигурации с минимальной энергией, линейный член разложения (так как производные берутся в равновесном положении) будет равен нулю, и первый ненулевой член разложения будет квадратичным по смещениям атомов. Кроме этого, будем считать, что все члены разложения

¹С инструкциями по установке Anaconda можно ознакомиться в учебном пособии «Компьютерная физика».

высших порядков будут пренебрежимо малы (такое приближение называется *гармоническим*):

$$U_{harm} = U_{eq} + \frac{1}{2} \sum_{l\alpha, l't\beta} \Phi_{l\alpha, l't\beta} u_{l\alpha} u_{l't\beta}, \quad (4.2)$$

В последнем уравнении константы $\Phi_{l\alpha, l't\beta}$ называются *силовыми константами* в реальном пространстве и определяются как

$$\Phi_{l\alpha, l't\beta} = \frac{\partial^2 U}{\partial u_{l\alpha} \partial u_{l't\beta}}. \quad (4.3)$$

Эти константы должны удовлетворять ряду условий, следующих из изотропии пространства, симметрии кристалла и инвариантности кристалла относительно его переноса на вектор, кратный вектору трансляции кристалла — вторая производная энергии может зависеть только от разницы между l и l' :

$$\Phi_{l\alpha, l't\beta} = \Phi_{0s\alpha, (l'-l)t\beta}. \quad (4.4)$$

Найдем ускорение частиц. Для этого продифференцируем соотношение (4.1) два раза по времени:

$$\vec{a}_{ls} = \frac{d^2 \vec{R}_{ls}(t)}{dt^2} = \frac{d^2 \vec{u}_{ls}}{dt^2},$$

или при разложении на декартовы компоненты

$$a_{ls\alpha} = \frac{d^2 u_{ls\alpha}}{dt^2}. \quad (4.5)$$

Выразим силы через энергии (4.3) и подставим соотношение (4.5) во второй закон Ньютона. Получим систему из $3N$ связанных уравнений относительно малых смещений атомов:

$$M_s \frac{d^2 u_{ls\alpha}}{dt^2} = - \sum_{l't\beta} \Phi_{l\alpha, l't\beta} u_{l't\beta}. \quad (4.6)$$

Будем искать решения этой системы уравнений в виде коллективного колебания — *фонона*:

$$u_{ls\alpha}(t) = \eta_{m\vec{q}}(s\alpha) e^{i\vec{q}\cdot\vec{R}_l} e^{-i\omega_{m\vec{q}}t}, \quad (4.7)$$

где зависимость решения от времени является периодической функцией с периодом колебаний $\omega_{m\vec{q}}$, а в пространстве в каждый момент времени решение представляет собой плоскую волну с волновым вектором \vec{q} ,

определенную в точках, кратных \vec{R}_l . m в решении является его номером («модой» колебаний), а $\eta_{m\vec{q}}(s\alpha)$ — α -я компонента вектора, называемого *вектором поляризации* нормальной моды колебаний. Необходимо заметить, что такой подход очень похож на методику, применяемую для получения волновой функции электронов в рамках теоремы Блоха: колебания ионов отличаются волновым вектором \vec{q} , причем если кристалл бесконечно периодический, используются только разрешенные значения волнового вектора.

Подставим решения (4.7) в систему (4.6). Для этого возьмем вторую производную (4.7) по времени, сократим обе части на $e^{-i\omega_{m\vec{q}}t}$ и домножим обе части на $e^{-i\vec{q}\cdot\vec{R}_l}$:

$$M_s\omega_{m\vec{q}}^2\eta_{m\vec{q}}(s\alpha) = \sum_{t\beta} \left[\sum_{l'} \Phi_{l s\alpha, l' t\beta} e^{i\vec{q}\cdot(\vec{R}_{l'} - \vec{R}_l)} \right] \eta_{m\vec{q}}(t\beta).$$

Вследствие того, что суммирование по l' производится по всем элементарным ячейкам кристалла, мы можем использовать свойство (4.4) матрицы силовых констант зависеть только от разницы между ячейками l' и l кристалла и перенести начало координат в ячейку с номером l :

$$M_s\omega_{m\vec{q}}^2\eta_{m\vec{q}}(s\alpha) = \sum_{t\beta} \left[\sum_{l'} \Phi_{0 s\alpha, l' t\beta} e^{i\vec{q}\cdot\vec{R}_{l'}} \right] \eta_{m\vec{q}}(t\beta). \quad (4.8)$$

В скобках стоит дискретное преобразование Фурье матрицы силовых констант в реальном пространстве:

$$\sum_{l'} \Phi_{0 s\alpha, l' t\beta} e^{i\vec{q}\cdot\vec{R}_{l'}} = \tilde{\Phi}_{s\alpha, t\beta}(\vec{q}).$$

Таким образом, движение атомов может быть описано в виде системы следующих динамических уравнений:

$$M_s\omega_{m\vec{q}}^2\eta_{m\vec{q}}(s\alpha) = \sum_{t\beta} \tilde{\Phi}_{s\alpha, t\beta}(\vec{q})\eta_{m\vec{q}}(t\beta).$$

Можно записать эту же систему в матричной форме:

$$\tilde{\Phi}(\vec{q})\eta(\vec{q}) = \omega^2(\vec{q})M_s\delta_{st}\delta_{\alpha\beta}\eta(\vec{q}), \quad (4.9)$$

где $\tilde{\Phi}(\vec{q})$ — матрица размером $3S \times 3S$, в которой содержится результат преобразования Фурье силовых констант, $\eta(\vec{q})$ — вектор длиной $3S$ собственных смещений фононов с волновым вектором \vec{q} , $\omega(\vec{q})$ — частота

фононов, и $M_s \delta_{st} \delta_{\alpha\beta}$ — диагональная матрица размером $3S \times 3S$, содержащая массы атомов для трех декартовых координат смещения каждого атома.

Мы можем заметить, что система (4.9) очень похожа на задачу поиска собственных значений матрицы $\tilde{\Phi}(\vec{q})$ за исключением наличия матрицы масс атомов. Для исключения этой матрицы из уравнения мы можем переопределить смещение (4.7) как

$$u_{l s \alpha}(t) = \frac{1}{\sqrt{M_s}} \gamma_{m \vec{q}}(s \alpha) e^{i \vec{q} \cdot \vec{R}_l} e^{-i \omega_{m \vec{q}} t}, \quad (4.10)$$

Тогда, повторяя предыдущие вычисления, мы можем определить *динамическую матрицу* системы как

$$\tilde{D}_{s \alpha, t \beta}(\vec{q}) = \frac{1}{\sqrt{M_s M_t}} \tilde{\Phi}_{s \alpha, t \beta}(\vec{q})$$

и переписать систему (4.9) в виде задачи на поиск собственных значений:

$$\tilde{D}(\vec{q}) \gamma(\vec{q}) = \omega^2(\vec{q}) \gamma(\vec{q}), \quad (4.11)$$

где $\gamma(\vec{q})$ — собственные вектора фононов. Можно показать, что динамическая матрица системы является эрмитовой.

Для решения уравнения (4.11) вначале необходимо найти динамическую матрицу (либо, что то же самое, матрицу силовых констант). Для этого существует два метода: метод линейного отклика, также называемый теорией возмущений функционала электронной плотности (Density Functional Perturbation Theory, DFPT), и техника малых отклонений (finite displacement method). Основная идея метода линейного отклика, предложенного Барони и др. [27] — использование теоремы Хеллмана–Фейнмана для того, чтобы показать, что линейное изменение электронной плотности при внесении возмущения в кристалл приводит к изменению энергии этого кристалла во втором (на самом деле, третьем [28]) порядке теории возмущений. При использовании стандартной теории возмущений изменение первого порядка зарядовой плотности может быть рассчитано с использованием невозмущенных волновых функций, и, соответственно, требует расчета в рамках ТФП для кристалла в основном состоянии. Если возмущение — фонон с волновым вектором \vec{q} , определение изменения электронной плотности в первом порядке теории возмущений может помочь в определении динамической матрицы системы для волнового вектора \vec{q} . Это может быть проделано для любого волнового вектора \vec{q} без необходимости строить большую суперячейку.

Методика малых отклонений атомов определяет матрицу силовых констант на основе того, как изменяется сила, действующая на атом s , если

сдвинуть атом t на малое расстояние в некотором направлении. Действительно, согласно определению силовых констант (4.3),

$$\Phi_{0s\alpha,l't\beta} = \frac{\partial}{\partial u_{0s\alpha}} \left(\frac{\partial U}{\partial u_{l't\beta}} \right) = -\frac{\partial F_{l't\beta}}{\partial u_{0s\alpha}},$$

то есть по силам, действующей на некоторый атом t в ячейке l' при сдвиге атома s в ячейке 0 в некотором направлении α , можно определить столбец матрицы $\Phi_{0s\alpha,l't\beta}$ для всех $(l't\beta)$. При повторении расчета сил для всех $(ls\alpha)$ полностью определяется матрица силовых констант Φ . Ясно, что количество проводимых расчетов при этом может быть втрое больше количества атомов в элементарной ячейке; однако для многих материалов симметрия может существенно уменьшить это количество. Кроме этого, следует иметь в виду одно важное обстоятельство. Из-за того, что расчеты в Vasp проводятся с периодическими граничными условиями, суперячейка, в которой проводятся расчеты фононных спектров, должна быть достаточно большой, чтобы элементы матрицы Φ были пренебрежимо малы на границе ячейки для любых $s\alpha$ в ячейке 0 . Для большинства материалов (особенно металлов) так оно и есть. Однако, в ионных материалах сходимость фононного спектра с увеличением размера суперячейки может быть очень медленной. Это происходит вследствие того, что сдвиг зарядов в таких материалах образует диполи, взаимодействующие с дальнедействующими силами, уменьшающимися только как r^{-3} . Эти диполи также создают макроскопическое электрическое поле, которое ответственно за расщепление мод оптической ветви на компоненты, параллельную и перпендикулярную внешнему электрическому полю (ЛО-ТО расщепление).

Phonору для расчета фононной зонной структуры и плотности фононных состояний использует методику малых отклонений. Таким образом, расчет в Phonору состоит из трех этапов:

1. построение суперячейки для расчета колебаний;
2. получение матрицы силовых констант;
3. решение задачи на собственные значения и получение собственных векторов и частот фононов.

4.2.3. Работа с пакетом

Суперячейки в пакете Phonору строятся из элементарной ячейки, задаваемой в файле POSCAR. Отклонения атомов выбираются с учетом группы симметрии кристалла. Для того, чтобы построить суперячейки с отклонениями, необходимо в директории с файлом POSCAR запустить Phonору с ключами `-d` и `--dim`:

```
phonopy -d --dim="2 2 3"
```

После этого в директории появятся файлы SPOSCAR (структура идеальной суперячейки), disp.yaml, содержащий информацию об отклонениях, и POSCAR- n , каждый из которых является входным файлом VASP для n -го отклонения, описанного в disp.yaml. Для расчета силовых констант требуется точно рассчитать силы, действующие на атомы в *таких конфигурациях* (PREC = Accurate; EDIFF $\approx 10^{-8}$). Не следует релаксировать систему (IBRION должен быть равен -1)! Рассчитанные силы собираются в один файл FORCE_SETS командой

```
phonopy -f disp-{001..003}/vasprun.xml
```

После построения наборов сил Phonopy может рассчитать плотности фононных состояний, термодинамические свойства и др. Для этого необходимо задать размерность суперячейки и сетку k -точек в файле mesh.conf, например, так:

```
1  АТОМ_NAME = Si O
2  DIM = 2 2 3
3  MP = 8 8 8
```

Тогда плотность состояний может быть получена запуском

```
phonopy -p mesh.conf
```

Термодинамические характеристики рассчитываются запуском

```
phonopy -t mesh.conf
```

Вопросы и задания

1. Визуализируйте структуру (100)-поверхности ГЦК-железа.
2. Постройте плотность электронных состояний для объемного ГЦК-железа и указанной поверхности. Будут ли они похожи? В чем заключаются различия?
3. Визуализируйте сходимость проведенного расчета. За сколько итераций достигается точность по энергии в 10^{-4} эВ? 10^{-5} эВ? 10^{-7} эВ?

Некоторые параметры файла INCAR

ENAUG

Действительный параметр ENAUG задает кинетическую энергию cut-off for the augmentation charges. Этот параметр определяет параметры NGXf, NGYf and NGZf. Значением по умолчанию является параметр EAUG из файла POTCAR.

NBLOCK, KBLOCK

После NBLOCK (целочисленный параметр, значение по умолчанию: 1) шагов моделирования рассчитываются парная корреляционная функция и DOS, а информация об атомной структуре сохраняется в файл XDATCAR. Кроме этого параметр NBLOCK указывает как часто масштабируется кинетическая энергия при параметре SMASS=-1. Все перечисленные операции не представляют значительной вычислительной сложности, поэтому рекомендуется использовать NBLOCK=1. После KBLOCK·NBLOCK (KBLOCK – целочисленный параметр, значение по умолчанию NSW) шагов моделирования усредненные парная корреляционная функция и DOS записываются в файлы PCDAT и DOSCAR.

RWIGS

Действительный массив RWIGS радиусов Вигнера-Зейца для каждого типа атомов, указанных в файле POSCAR, например

RWIGS = 1.0 1.5

для системы, содержащей два типа атомов. Значение по умолчанию: значения, прочитанные из файла POTCAR. Если LORBIT \geq 10, RWIGS игнорируются. RWIGS необходимо задавать в расчетах с ограничением локальных магнитных моментов. Для моноатомной системы RWIGS можно однозначно определить. Сумма объема сфер вокруг каждого атома должна быть такой же, как и общий объем ячейки (в предположении, что вакуумной области внутри расчетной ячейки нет). VASP записывает в выходной файл строку

Volume of Typ 1: 98.5 %

Для бинарных систем нет однозначного способа определить RWIGS, и

возможны несколько вариантов. В любом случае значения RWIGS необходимо подбирать таким образом, чтобы сумма объемов всех сфер приблизительно составляла 100%.

- Один из возможных вариантов - установить RWIGS таким образом, чтобы минимизировать перекрытие между сферами.
- В большинстве случаев можно выбрать радиус каждой сферы таким образом, чтобы они были близки к ковалентному радиусу.

Пожалуйста, имейте в виду, что результаты являются качественными - т.е. нет однозначного способа определения локализации электрона. При текущей реализации, например, вряд ли можно определить перенос заряда. Из парциальной плотности состояний можно получить, например, пики в DOS. Количественные результаты могут быть получены только путем тщательного сравнения с эталонной системой (например, объемной или поверхностной).

LORBIT

Допустимые значения LORBIT = 0-1-2-10-11-12. Значение по умолчанию LORBIT = 0;

LORBIT	строка INCAR	RWIGS	в файле	записываемые файлы
0	required			DOSCAR и PROCAR
1	required			DOSCAR и подробный PROCAR
2	required			DOSCAR
10	not read			DOSCAR и PROCAR
11	not read			DOSCAR и PROCAR file with phase factors
12				не поддерживается

Этот флаг определяет вместе с соответствующими RWIGS, записывается ли файл PROCAR. Файл PROCAR содержит spd- and site projected wavefunction character of each band. The wavefunction character вычисляется либо путем проецирования орбиталей на сферические гармоники, которые не равны нулю внутри сфер радиуса RWIGS вокруг каждого иона (LORBIT = 1, 2), либо с использованием схемы быстрой проекции, применяющейся только для метода PAW (LORBIT = 10,11,12). Если флаг LORBIT не равен нулю, the site and l-projected density of states is also calculated.

Если используется the projector augmented wave method, LORBIT также может быть установлен равным 10, 11 или 12. Эта альтернативная установка выбирает быстрый метод определения the spd- and site projected wave function character и не требует задания радиусов Вигнера-Зейца в файле INCAR (в этом случае пренебрегают строкой RWIGS). Метод работает только для PAW файлов POTCAR и не для ультрамягких или сохраняющих норму псевдопотенциалов.

LREAL (и ROPT)

Допустимые значения:

- .FALSE. проектирование выполняется в обратном пространстве
- .TRUE. проектирование выполняется в координатном пространстве, оптимизация операторов не проводится
- On или O проектирование выполняется в координатном пространстве, операторы проектирования оптимизируются
- Auto или A проектирование выполняется в координатном пространстве, полностью автоматическая оптимизация операторов проектирования

Значение по умолчанию: LREAL = .FALSE.

Параметр LREAL определяет в координатном или обратном пространстве получают операторы проектирования: нелокальная часть псевдопотенциала требует вычисления выражения $\sum_{ij} D_{ij} |\beta_j\rangle \langle \beta_i| \phi_{nk}\rangle$. The “projected wavefunction character” определяется как:

$$C_{ink} = \langle \beta_i | \phi_{nk} \rangle = \frac{\Omega}{N_{FFT}} \sum_r \langle \beta_i | r \rangle \langle r | \phi_{nk} \rangle = \frac{\Omega}{N_{FFT}} \sum_r \langle \beta(r) \phi_{nk}(r) \rangle$$

$$C_{ink} = \sum_G \langle \beta_i | k + G \rangle \langle k + G | \phi_{nk} \rangle = \sum_G \bar{\beta}(k + G) C_{Gnk}.$$

Это выражение может быть вычислено как в обратном, так и в координатном пространстве: в обратном пространстве (вторая строка) число операций растёт с размером базового набора (т.е. с числом плоских волн). В координатном пространстве (первая строка) операторы проектирования ограничены сферами вокруг каждого атома. Поэтому количество операций, необходимых для оценки одного C_{ink} , не увеличивается с размером системы (обычно количество точек сетки в cut-off-sphere составляет от 500 до 2000). Одним из основных недостатков метода, работающего в реальном пространстве, является то, что проекционные операторы должны быть оптимизированы, т.е. все высокочастотные компоненты должны быть удалены из операторов проектирования. В противном случае может

произойти «сглаживание» (то есть высокочастотные компоненты операторов проектирования будут сглажены до низкочастотных компонент и возникнет случайный шум).

В настоящее время VASP поддерживает три различные схемы удаления высокочастотных компонент оператора проектирования. `LREAL = .TRUE.` является самым простым из них. При этом используются операторы проектирования в координатном пространстве, которые были созданы кодом генерации псевдопотенциала. При `LREAL = On` операторы проектирования оптимизируются с использованием алгоритма, предложенного King-Smith и др. [29]. Для `LREAL = Auto` используется улучшенная схема. Если используется `LREAL = Auto` или `LREAL = On` для точной настройки процедуры оптимизации может использоваться флаг `ROPT`.

Для систем, содержащих более 20 атомов, рекомендуем использовать схему проектирования в координатном пространстве. Для версий VASP начиная с VASP.4.4 рекомендуется использовать только `LREAL = Auto`. При этом обычно достигается наилучшая производительность, но если производительность не так важна, можно использовать `LREAL = .TRUE.`, который требует меньшего вмешательства пользователя.

Для `LREAL = O` и `LREAL = A` проекционные операторы оптимизируются при запуске. На оптимизацию влияет несколько флагов

- `ENCUT` (т.е. энергия обрезания), компоненты, соответствующие большим энергиям, «удаляются» из операторов проектирования
- параметр `PREC` указывает, насколько точны должны быть проекторы реального пространства, и устанавливает переменные `ROPT` соответственно следующим значениям:
 - для `LREAL = On`
 - * `PREC = Low`: 700 точек в сфере в координатном пространстве (`ROPT=0.67`)
 - * `PREC = Med`: 1000 точек в сфере в координатном пространстве (`ROPT=1.0`)
 - * `PREC = High`: 1500 точек в сфере в координатном пространстве (`ROPT=1.5`)
 - для `LREAL = Auto`
 - * `PREC = Low`: точность 10^{-2} (`ROPT=0.01`)
 - * `PREC = Med`: точность 2×10^{-3} (`ROPT=0.002`)
 - * `PREC = High`: точность 2×10^{-4} (`ROPT=2E-4`)

Эти значения по умолчанию могут быть заменены строкой `ROPT = one_number_for_each_species` в файле `INCAR`. Например, `ROPT = 0.7 1.5` задает количество точек координатного пространства в сфере `cutoff` для первого вида атомов примерно 700, а для второго вида - 1500.

Режим «точность» работает как для $LREAL = On$, так и для $LREAL = Auto$ (но для обеспечения совместимости со старыми версиями VASP он выбирается только в том случае, если $LREAL = Auto$). Прецизионный режим обычно включается, если значение для $ROPT$ меньше 0,1. Режим «точности» и обычный режим могут быть перемешаны, то есть можно указать

$ROPT = 0.7 \text{ 1E-3}$

в этом случае число точек координатного пространства в cutoff сфере для первого вида атомов будет приблизительно 700, тогда как функции проектирования для второго вида атомов оптимизированы с точностью приблизительно 1 мэВ. Рекомендуется использовать режим «точности» с целевой точностью около 10^{-3} эВ / атом.

При использовании режима, в котором указано количество точек сетки необходимо тщательно выбрать $ROPT$, особенно если *hard species is mixed with a soft species*. В этом случае необходимо проверить следующие строки в файле OUTCAR (для $LREAL = On$, для $LREAL = Auto$ очень похож)

```
Optimization of the real space projectors
maximal supplied Q-value = 12.85
optimization between [QCUT,QGAM] = [ 6.33, 25.32] Ry
Optimized for a Real-space Cutoff 2.30 Angstroem
1 X(QCUT) X(cont) X(QGAM) max X(q) W(q)/X(q) e(spline)
0 9.518 9.484 -.004 18.582 .11E-03 .16E-06
0 -2.149 -2.145 .001 3.059 .17E-03 .25E-06
1 8.957 8.942 .003 9.950 .14E-03 .34E-06
1 1.870 1.870 .001 1.837 .95E-03 .51E-06
2 3.874 3.866 .000 4.764 .15E-03 .68E-07
```

Наиболее важная информация приведена в столбце $W(q)/X(q)$ (столбец $W(\text{low})/X(q)$ для $LREAL = Auto$). Значения в этом столбце должны быть как можно меньше. Как правило, максимальное допустимое значение в этом столбце составляет 10^{-3} для $PREC = Med$. Для $PREC = Low$ ошибки могут быть около 10^{-2} , а для $PREC = High$ ошибки должны быть меньше 10^{-4} . Если $W(q)/X(q)$ больше 10^{-2} , ошибки могут быть существенными. В этом случае $ROPT$ больший значения по умолчанию должен быть указан в файле INCAR.

Оптимизация операторов проектирования в координатном пространстве ($LREAL = .TRUE.$, $LREAL = On$ или $LREAL = Auto$) всегда приводит к небольшой (не обязательно незначительной) ошибке. Если необходимо определить разности энергий в несколько мэВ, следует использовать вычисления с идентичными настройками ($ENCUT$, $PREC$, $LREAL$ и $ROPT$) для всех вычислений.

NWRITE

Параметр NWRITE определяет информацию, выводимую в файл OUTCAR. Допустимые значения:

- 0 — используется для молекулярно-динамических расчетов;
- 1 — как и 0, используется для молекулярно-динамических расчетов;
- 2 (значение по умолчанию) — для коротких расчетов;
- 3 — может дать информацию, если что-то пойдёт не так;
- 4 — используется исключительно для дебаггинга.

Что именно определяет то или иное значение параметра, указано в таблице.

NWRITE	0	1	2	3
вклад в электронную энергию на каждой электронной итерации	f	f	e	e
информация о сходимости собственных значения	f	f	e	e
DOS + зарядовая плотность	f+1	i	i	e
полная энергия и её разбиение по вкладам	f+1	i	i	e
тензор напряжений	i	i	i	i
базисные вектора	i	i	i	i
силы на атомах	f+1	i	i	i
информация о времени расчета			X	X

В таблице f — первый ионный шаг; f+1 — первый и последний ионные шаги, i — каждый ионный шаг, e — каждый электронный шаг, X — во всех местах, где применимо.

TEBEG и TEEND

Действительные параметры TEBEG и TEEND задают начальную и конечную температуры в методе молекулярной динамики (IBRION=0). Значения по умолчанию: TEBEG = 0; TEEND = TEBEG. Если начальные скорости не заданы в файле POSCAR, они устанавливаются в соответствии с распределением Максвелла для начальной температуры TEBEG. Температура в пакете VASP определяется следующим выражением

$$T = \frac{1}{2k_B T N_{ions}} \sum_n M_n |\vec{V}_n|^2.$$

Но, поскольку центр масс системы неподвижен, имеется только $3(N_{ions} - 1)$ степеней свободы (сумма всех скоростей равна нулю, если начальные

скорости задаются случайным образом). Это означает, что на самом деле температура моделируемой системы определяется как

$$T = \text{TEBEG} \times N_{ions} / (N_{ions} - 1).$$

Записываемые пакетом VASP температуры (например, в файл OUTCAR) некорректны. Обычно поправки незначительны, однако о них нужно помнить, если требуется очень высокая точность. Это означает, что начальная температура должна быть задана как

$$\text{TEBEG} = T_{requested} \times (N_{ions} - 1) / N_{ions},$$

в файле INCAR.

SMASS

Параметр SMASS влияет на скорости ионов в методе молекулярной динамики. Он может принимать значения: -3, -2, -1, ≥ 0 и параметр Но-зе массы. В ранних версиях пакета Vasp значением по умолчанию было значение 0. Начиная с версии 4.4 значение по умолчанию составляет -3.

При значении SMASS=-3 используется микроканонический ансамбль: несмотря на то, что силы Гельмана-Феймана придают ионам ускорение, полная свободная энергия ионов (свободная энергия электронов + потенциальная энергия ионов в электростатическом поле + кинетическая энергия ионов) поддерживается постоянной.

При значении SMASS=-2 скорости ионов поддерживаются постоянными. Этот подход позволяет рассчитать энергии для набора линейно связанных координат ионов (например, димеров с изменяющейся длиной связи). Обратим внимание на то, что при SMASS=-2 скорости умножаются на параметр POTIM. Для устранения неопределенности необходимо установить POTIM=1.

При значении SMASS=-1 скорости масштабируются на каждом NBLOCK шаге (начиная с первого шага) до температуры

$$TEMP = \text{TEBEG} + (\text{TEEND} - \text{TEBEG}) * \text{NSTEP} / \text{NSW},$$

где NSTEP – номер шага (нумерация начинается с единицы). Это позволяет непрерывно увеличивать или уменьшать кинетическую энергию. На промежуточных итерациях используется микроканонический ансамбль.

При SMASS ≥ 0 рассматривается канонический ансамбль с использованием алгоритма Нозе. Параметр массы алгоритм Нозе определяет колебаний температуры в процессе моделирования. Для SMASS=0 масса

Нозе соответствует периоду в 40 шагов по времени. Этот параметр должен быть задан таким образом, чтобы колебания температуры по частоте совпадали с характерными частотами фононов. Для жидкостей аналогом частот фононов может быть спектр автокорреляционной функции скоростей. Если частоты колебания ионов отличаются на порядок от частоты колебаний температуры, термостат Нозе и движение ионов может рассинхронизироваться, что приведет к неканоническому ансамблю.

Библиографический список

1. Allen MP, Tildesley DJ. Computer Simulation of Liquids Oxford Science Publications // New York. — 1987.
2. Zamalin Vadim Moiseevič, Norman GE, Filinov Vladimir Sergeevič. Metod Monte-Karlo v statističeskoj termodinamike. — Nauka, 1977.
3. Martin Richard M. Electronic structure: basic theory and practical methods. — Cambridge university press, 2004.
4. Car Richard, Parrinello Mark. Unified approach for molecular dynamics and density-functional theory // Physical review letters. — 1985. — Vol. 55, no. 22. — P. 2471.
5. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients / Mike C Payne, Michael P Teter, Douglas C Allan et al. // Reviews of modern physics. — 1992. — Vol. 64, no. 4. — P. 1045.
6. Ceperley David M, Mitas Lubos. Quantum Monte Carlo methods in chemistry // New methods in computational quantum mechanics. — 1995. — P. 1–38.
7. Kohn Walter, Sham Lu Jeu. Self-consistent equations including exchange and correlation effects // Physical review. — 1965. — Vol. 140, no. 4A. — P. A1133.
8. Hohenberg Pierre, Kohn Walter. Inhomogeneous electron gas // Physical review. — 1964. — Vol. 136, no. 3B. — P. B864.
9. Jones Robert O, Gunnarsson Olle. The density functional formalism, its applications and prospects // Reviews of Modern Physics. — 1989. — Vol. 61, no. 3. — P. 689.
10. Perdew John P, Wang Yue. Accurate and simple analytic representation of the electron-gas correlation energy // Physical Review B. — 1992. — Vol. 45, no. 23. — P. 13244.
11. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation / John P Perdew, John A Chevary, Sy H Vosko et al. // Physical Review B. — 1992. — Vol. 46, no. 11. — P. 6671.

12. Perdew John P, Burke Kieron, Ernzerhof Matthias. Generalized gradient approximation made simple // *Physical review letters*. — 1996. — Vol. 77, no. 18. — P. 3865.
13. Teter Michael P., Payne Michael C., Allan Douglas C. Solution of Schrödinger's equation for large systems // *Phys. Rev. B*. — 1989. — Dec. — Vol. 40. — P. 12255–12263.
14. Kresse Georg, Furthmüller Jürgen. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set // *Physical review B*. — 1996. — Vol. 54, no. 16. — P. 11169.
15. Hamann DR, Schlüter M, Chiang C. Norm-conserving pseudopotentials // *Physical Review Letters*. — 1979. — Vol. 43, no. 20. — P. 1494.
16. Bachelet GB, Hamann DR, Schlüter M. Pseudopotentials that work: From H to Pu // *Physical Review B*. — 1982. — Vol. 26, no. 8. — P. 4199.
17. Vanderbilt David. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism // *Physical Review B*. — 1990. — Vol. 41, no. 11. — P. 7892.
18. Kresse G, Hafner J. Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements // *Journal of Physics: Condensed Matter*. — 1994. — Vol. 6, no. 40. — P. 8245.
19. Blöchl PE. PE Blöchl, *Phys. Rev. B* 50, 17953 (1994). // *Phys. Rev. B*. — 1994. — Vol. 50. — P. 17953.
20. Kresse Georg, Joubert D. From ultrasoft pseudopotentials to the projector augmented-wave method // *Physical Review B*. — 1999. — Vol. 59, no. 3. — P. 1758.
21. Monkhorst H, Pack J. Special points for Brillouin zone integrations // *Physical Review B*. — 1976. — Vol. 13, no. 12. — P. 5188–5192.
22. Mermin N. David. Thermal Properties of the Inhomogeneous Electron Gas // *Phys. Rev.* — 1965. — Mar. — Vol. 137. — P. A1441–A1443.
23. Methfessel M., Paxton A. T. High-precision sampling for Brillouin-zone integration in metals // *Phys. Rev. B*. — 1989. — Aug. — Vol. 40. — P. 3616–3621.
24. Kresse Georg, Hafner Jürgen. Ab initio molecular dynamics for liquid metals // *Physical Review B*. — 1993. — Vol. 47, no. 1. — P. 558.
25. Pulay P. Convergence acceleration of iterative sequences. the case of scf iteration // *Chemical Physics Letters*. — 1980. — Vol. 73, no. 2. — P. 393 – 398.
26. Heyden Andreas, Bell Alexis T, Keil Frerich J. Efficient methods for finding transition states in chemical reactions: Comparison of improved

- dimer method and partitioned rational function optimization method // The Journal of chemical physics. — 2005. — Vol. 123, no. 22. — P. 224101.
27. Baroni Stefano, Giannozzi Paolo, Testa Andrea. Green's-function approach to linear response in solids // Physical Review Letters. — 1987. — Vol. 58, no. 18. — P. 1861.
28. Gonze Xavier, Vigneron J-P. Density-functional approach to nonlinear-response coefficients of solids // Physical Review B. — 1989. — Vol. 39, no. 18. — P. 13120.
29. King-Smith R. D., Payne M. C., Lin J. S. Real-space implementation of nonlocal pseudopotentials for first-principles total-energy calculations // Phys. Rev. B. — 1991. — Dec. — Vol. 44. — P. 13063–13066.