

САМОСОГЛАСОВАННОЕ РЕШЕНИЕ УРАВНЕНИЙ КОНА – ШЭМА ДЛЯ НЕОГРАНИЧЕННЫХ СИСТЕМ С НЕОДНОРОДНЫМ ЭЛЕКТРОННЫМ ГАЗОМ

*Д. В. Посвянский**, *А. Я. Шульман***

*Институт радиотехники и электроники им. В. А. Котельникова Российской академии наук
125009, Москва, Россия*

Поступила в редакцию 10 декабря 2008 г.

Метод функционала плотности в приближении Кона – Шэма широко используется при изучении свойств многоэлектронных систем. В силу нелинейности уравнений Кона – Шэма общим методом нахождения самосогласованного решения является итеративный способ, при котором надо поочередно решать уравнения Пуассона и Шредингера. Одной из проблем такого подхода является то, что полученное после решения уравнения Шредингера распределение заряда не согласуется с граничными условиями к уравнению Пуассона, которым должен удовлетворять кулоновский потенциал. Возникающая в таких случаях неустойчивость или даже расходимость итерационного процесса особенно существенно проявляется в случае неограниченных систем. Представленные в литературе методы преодоления этой трудности, по существу, сводятся к отказу от первоначально сформулированного итеративного способа решения и замене его какой-либо приближенной схемой расчета, которая обычно строится полуэмпирически и не позволяет оценить степень отклонения от точного решения. В настоящей работе реализована итеративная схема решения уравнений Кона – Шэма для протяженных систем с неоднородным электронным газом, которая основана на устранении кулоновского дальнего действия как причины жесткой связи между распределением заряда и граничными условиями. Предложенный алгоритм применяется для расчета энергетического спектра, самосогласованного потенциала и электростатической емкости полубесконечного вырожденного электронного газа, ограниченного бесконечно высоким барьером, для расчета работы выхода и поверхностной энергии простых металлов в модели с однородным распределением положительного фона. Исследуется различие самосогласованных решений, полученных в приближении Хартри и при учете обменно-корреляционного взаимодействия электронов. Проведено сравнение с результатами предшествующих работ. На примере туннельного контакта металл–полупроводник показано применение алгоритма для случая неограниченной системы, в которой возможно протекание тока.

PACS: 71.15.-m, 71.15.Mb, 65.40.gh, 85.30.Mn

1. ВВЕДЕНИЕ

Система уравнений метода функционала плотности, описывающая неоднородный электронный газ в приближении Кона – Шэма, должна удовлетворяться самосогласованными распределениями плотности электронов $N(\mathbf{r})$ и кулоновского потенциала $U(\mathbf{r})$ [1]. Ввиду существенной нелинейности этой системы единственным общим методом построения самосогласованного решения является итеративный способ, при котором надо поочередно

решать уравнение Пуассона для потенциала и систему уравнений Шредингера для одночастичных волновых функций в эффективном потенциале $U_{eff}(\mathbf{r}) = U(\mathbf{r}) + U_{xc}(\mathbf{r})$, где U_{xc} — обменно-корреляционный потенциал в приближении локальной плотности. Известной проблемой такого подхода является необходимость учитывать, что граничные условия, которым должно удовлетворять решение уравнения Пуассона, накладывают определенные требования на распределение заряда в правой части этого уравнения в виде интегральных соотношений [2]. С физической точки зрения эти интегральные соотношения означают, что заданы,

*E-mail: posvyanski@cplire.ru

**E-mail: ash@cplire.ru, ash@east.ru

например, полный заряд электронной системы, если граничные условия наложены на электрическое поле, либо ее полный дипольный момент, если заданы значения потенциала на границах бесконечной области. Но, пока не достигнуто самосогласование, на каждом шагу итерационного процесса распределение электронной плотности, получаемое после решения уравнения Шредингера, как правило, оказывается несовместным с заданными граничными условиями для уравнения Пуассона. В результате, как показано в работе [2], в случае бесконечно-протяженных систем решение следующего шага либо вообще не может быть построено, либо отсутствует сходимость итерационного процесса, отмечавшаяся, например, в работах [3, 4]. В ряде работ при этом прибегают к замене процедуры решения уравнения Пуассона некоторой вариационной схемой относительно числовых параметров, которые определяют функциональную форму, выбранную для аппроксимации распределения потенциала или электронов. Конкретнее это обсуждается в пп. 3.1 и 4.1.

Для преодоления этой трудности используются и другие искусственные приемы, общим свойством которых является изменение параметров рассматриваемой системы, например, плотности заряда компенсирующего фона [5]. Однако существование и единственность решения уравнений самосогласованного поля следуют из того, что они получены как уравнения Эйлера–Лагранжа для вариационной задачи на минимум полной энергии невырожденного основного состояния (см. например, [1]). При этом варьируются волновые функции одночастичных орбиталей и кулоновский потенциал при заданных граничных условиях, а параметры системы остаются постоянными. В случае варьирования параметров системы в процессе итераций необходимо иметь доказательство существования и единственности решения, которое на настоящий момент отсутствует.

Другой подход предложен в работах [4, 6], где для нахождения электростатического потенциала $U(\mathbf{r})$ используется линейное интегральное уравнение второго рода, к которому искусственно сведено уравнение Пуассона. Утверждается, что тем самым устраняется необходимость в специальных мерах для соблюдения совместности граничных условий к уравнению Пуассона с распределением заряда, полученного на каждом цикле итераций (см. [6, Appendix]). Однако ясно, что точного решения интегрального уравнения, равносильного исходному дифференциальному, также не существует, если не выполняется какое-либо из условий совместности.

Скорее всего, по этой причине указанный метод, даже в тех задачах, где он был успешно применен, для выполнения условия нейтральности при каждой итерации также дополнялся перед решением уравнения Пуассона перенормировкой (неясным образом) электронной плотности, полученной после решения уравнения Шредингера [4]. Еще одним недостатком этого подхода, по-видимому, является фактическое сохранение члена, описывающего экранирование, в области, где электронов нет, но потенциал не равен нулю, как это имеет место в задачах о поверхности.

В настоящей работе описывается реализация предложенного ранее и описанного в работе [2] алгоритма итеративного решения уравнений Кона–Шэма для равновесного неоднородного электронного газа, при котором на каждом шагу итеративной схемы обеспечивается удовлетворение граничных условий к уравнению Пуассона для любого распределения электронов, порожденного эффективным потенциалом предшествующего приближения (разд. 2). Отмечаются особенности алгоритма, возникающие из-за разрывного характера квазиклассического выражения для индуцированной плотности электронов при учете обменно-корреляционного потенциала (п. 2.2). Сходимость алгоритма показана на примере самосогласованных расчетов в модели с однородным нейтрализующим фоном при исследовании таких задач как: 1) существование самоиндуцированных связанных состояний вблизи поверхности электронного газа, ограниченного потенциальным барьером (п. 3.1), 2) влияние на емкость плоского конденсатора реального распределения экранирующего заряда в электродах с учетом фриделевских осцилляций плотности электронов (п. 3.2), 3) расчет работы выхода и поверхностной энергии простых металлов (разд. 4). На примере туннельного контакта металл–полупроводник с барьером Шоттки показана применимость алгоритма к расчету структур, в которых возможно протекание тока (разд. 5). Часть результатов была предварительно представлена в работе [7] и докладывалась на Российских конференциях по физике полупроводников в 2001–2007 гг.

2. ИСХОДНЫЕ УРАВНЕНИЯ И АЛГОРИТМ ИТЕРАТИВНОГО РЕШЕНИЯ

2.1. Общая постановка задачи

В приближении Кона–Шэма газ взаимодействующих электронов при температуре $T = 0$ описывается системой уравнений для одночастичных волновых функций Ψ_ϵ :

$$\frac{1}{2}\nabla^2\Psi_\varepsilon(\mathbf{r}) + (\varepsilon - U_{eff}(\mathbf{r}))\Psi_\varepsilon(\mathbf{r}) = 0 \quad (2.1)$$

с эффективным потенциалом

$$U_{eff}(\mathbf{r}) = U(\mathbf{r}) + U_{xc}(\mathbf{r}). \quad (2.2)$$

Здесь ε — собственное значение энергии одночастичного состояния, а через U обозначена кулоновская потенциальная энергия электрона, для которой уравнение Пуассона записывается в виде

$$\nabla^2 U = 4\pi(N_+(\mathbf{r}) - N(\mathbf{r})), \quad (2.3)$$

где $N_+(\mathbf{r})$ — плотность носителей положительного заряда, $N(\mathbf{r})$ — плотность электронов. Всюду, где не оговорено иное, используются атомные единицы $|e| = m = \hbar = 1$, основанные на заряде e и массе m свободного электрона, когда единица длины — боровский радиус a_B , единица энергии — Хартри $\text{Ha} = e^2/a_B$.

Обменно-корреляционный потенциал, входящий в уравнение (2.2), берется в локальном приближении функционала плотности:

$$U_{xc}(\mathbf{r}) = \left. \frac{d[N\varepsilon_{xc}(N)]}{dN} \right|_{N=N(\mathbf{r})}, \quad (2.4)$$

где $\varepsilon_{xc} = \varepsilon_x + \varepsilon_c$ — сумма обменной и корреляционной энергий на одну частицу.

Электронная плотность выражается через волновые функции как

$$N(\mathbf{r}) = 2 \int_{\varepsilon \leq \varepsilon_F} \mathcal{D}\{\varepsilon\} |\Psi_\varepsilon(\mathbf{r})|^2, \quad (2.5)$$

где множитель 2 учитывает спиновое вырождение одночастичных состояний и интегрирование по дифференциальной спектральной мере $\mathcal{D}\{\varepsilon\}$ гамильтониана

$$\hat{H} = -\frac{1}{2}\nabla^2 + U_{eff}(\mathbf{r})$$

ведется по всем занятым состояниям с энергией ε , не превышающей энергию Ферми ε_F .

Как отмечалось во Введении, граничные условия для уравнения Пуассона накладывают определенные требования на распределение электронной плотности в пространстве, которым результат интегрирования в формуле (2.5) не обязан удовлетворять до достижения самосогласования. Однако, если представить полную плотность электронов в виде суммы

$$N(\mathbf{r}) = N_{ind}(U(\mathbf{r})) + N_{qu}(\mathbf{r}), \quad (2.6)$$

где индуцированная плотность N_{ind} зависит явно от неизвестного потенциала, то проблемы несовместимости граничных условий с правой частью уравнения Пуассона не будет, так как кулоновское дальнее действие сменится экранированным взаимодействием конечного радиуса. В методе Кона – Шэма между плотностью электронов и эффективным потенциалом существует приближенное локальное соотношение, аналогичное тому, что имеет место в классическом идеальном равновесном ферми-газе,

$$N_{ind}(\mathbf{r}) = \frac{2^{3/2}}{3\pi^2} (\varepsilon_F - U_{eff}(\mathbf{r}))^{3/2}. \quad (2.7)$$

Формула (2.7) может быть получена из решений уравнения Шредингера (2.1) для волновых функций непрерывного спектра в квазиклассическом приближении, если подставить их в (2.5) и усреднить полученную плотность электронов по масштабам, большим фермиевской длины волны λ_F . Величина $N_{qu}(\mathbf{r})$ при использовании для $N_{ind}(\mathbf{r})$ формулы (2.7) представляет собой квантовую поправку к квазиклассическому распределению электронов. Отметим, что представление полной плотности формулой (2.6) всегда является точным, несмотря на использование приближенной формулы (2.7) для $N_{ind}(\mathbf{r})$.

В выражении (2.2) $U_{xc} = U_x + U_c$. Обменный потенциал для электронного газа в локальном приближении дается известной формулой

$$U_x(\mathbf{r}) = -\left(\frac{3}{\pi}\right)^{1/3} N^{1/3}(\mathbf{r}) = -\left(\frac{3}{2\pi}\right)^{2/3} \frac{1}{r_s(\mathbf{r})}. \quad (2.8)$$

Корреляционный потенциал согласно (2.4) определяется через корреляционную энергию на один электрон ε_c . Мы примем для последней следующее выражение:

$$\varepsilon_c = -\frac{0.44}{r_s + 11.5}. \quad (2.9)$$

Здесь r_s — локальный радиус Вигнера–Зейтца, определяемый соотношением

$$4\pi r_s^3(\mathbf{r})/3 = N^{-1}(\mathbf{r}). \quad (2.10)$$

Для дальнейшего удобно также обозначить через $R_s = r_s(\infty)$ значение этого параметра на бесконечности, где всегда будет предполагаться выполненным условие локальной нейтральности:

$$\lim_{|\mathbf{r}| \rightarrow \infty} (N - N_+) \rightarrow 0.$$

После вычисления производной (2.4) для случая корреляционной энергии, определенной формулой (2.9), выражение для корреляционного потенциала можно представить в виде

$$U_c(\mathbf{r}) = -\frac{22}{75} \frac{8r_s(\mathbf{r}) + 69}{[2r_s(\mathbf{r}) + 23]^2}. \quad (2.11)$$

Формула (2.9) для корреляционной энергии отличается от известной формулы Вигнера заменой слагаемого 7.8 в знаменателе на 11.5. Основания для задания корреляционной энергии формулой Вигнера с измененным числовым параметром подробнее требуют отдельного рассмотрения. Здесь лишь отметим, что при таком выборе ε_c , с одной стороны, лучше описывается зависимость от R_s энергии связи простых металлов, представленная в табл. 8 книги [8]. С другой стороны, принятое в литературе значение $R_s = 5.63$ для Cs оказывается меньше критического значения $R_{sc} = 5.64$, характеризующего устойчивость однородного состояния электронного газа в модели желе относительно пространственно-неоднородных возмущений. Поскольку для системы с $R_s \geq R_{sc}$ невозможность использования формулы (2.7) для индуцированного заряда делает неприменимым предлагаемый алгоритм, выбор корреляционной энергии в форме (2.9) обеспечивает существование металлического цезия в рамках разрабатываемого метода расчета. Для корреляционной энергии в локальном приближении в литературе предлагаются и другие выражения (см., например, обсуждение в [1, 5]), однако в настоящей работе, цель которой продемонстрировать в деталях применение нового алгоритма на ряде классических задач, конкретный вид корреляционного потенциала не столь важен, так как соответствующие изменения в выражении для U_c имеют следствием только количественные отличия в результатах вычислений, но не создают принципиальных трудностей при реализации предлагаемой итерационной процедуры.

Выделение индуцированного заряда приводит уравнение Пуассона (2.3) к виду

$$\nabla^2 U + 4\pi N_{ind}(U) = 4\pi(N_+(\mathbf{r}) - N_{qu}(\mathbf{r})). \quad (2.12)$$

Введенная здесь с использованием формул (2.5)–(2.7) квантовая поправка $N_{qu}(\mathbf{r})$ к электронной плотности определяется как

$$N_{qu}(\mathbf{r}) = 2 \int_{\varepsilon \leq \varepsilon_F} \mathcal{D}\{\varepsilon\} |\Psi_\varepsilon(\mathbf{r})|^2 - N_{ind}(\mathbf{r}). \quad (2.13)$$

Алгоритм последовательных шагов $i = 0, 1, 2, \dots$ при итеративном решении уравнений (2.1)–(2.5) с за-

меной уравнения (2.3) на (2.12) может быть теперь записан в следующем виде.

1. Пренебрегаем мелкомасштабными квантовыми вариациями плотности электронов

$$N_{qu}^0 = 0. \quad (2.14)$$

2. Решаем нелинейное уравнение Пуассона с известной правой частью:

$$\begin{aligned} \nabla^2 U^i(\mathbf{r}) + 4\pi N_{ind}(U^i, N_{qu}^i) = \\ = 4\pi(N_+(\mathbf{r}) - N_{qu}^i(\mathbf{r})). \end{aligned} \quad (2.15)$$

3. Определяем самосогласованную с U полную плотность электронов и эффективный потенциал:

$$\begin{aligned} N_s^i(\mathbf{r}) &= N_{ind}^i(\mathbf{r}) + N_{qu}^i(\mathbf{r}), \\ U_{eff}^i(\mathbf{r}) &= U^i(\mathbf{r}) + U_{xc}(N_s^i(\mathbf{r})). \end{aligned} \quad (2.16)$$

4. Находим собственные функции одночастичного гамильтониана Кона–Шэма

$$\frac{1}{2} \nabla^2 \Psi_\varepsilon^i(\mathbf{r}) + (\varepsilon - U_{eff}^i(\mathbf{r})) \Psi_\varepsilon^i(\mathbf{r}) = 0. \quad (2.17)$$

5. Переопределяем полную плотность электронов, используя полученный спектр собственных значений и найденные собственные функции:

$$N^i(\mathbf{r}) = 2 \int_{\varepsilon \leq \varepsilon_F} \mathcal{D}\{\varepsilon\} |\Psi_\varepsilon^i(\mathbf{r})|^2. \quad (2.18)$$

6. Определяем квантовую поправку для следующего шага итераций через новую полную плотность электронов:

$$N_{qu}^{i+1}(\mathbf{r}) = N^i(\mathbf{r}) - N_{ind}(U_{eff}^i(\mathbf{r})). \quad (2.19)$$

7. Возвращаемся к решению уравнения Пуассона с новой правой частью.

Индекс i здесь означает номер итерации. Важно отметить, что в процессе решения уравнения Пуассона (2.15) одновременно находятся кулоновский потенциал $U^i(\mathbf{r})$ и согласованная с ним индуцированная плотность электронов $N_{ind}^i(\mathbf{r})$ при неизменной квантовой плотности $N_{qu}^i(\mathbf{r})$.

Критерием сходимости итерационного процесса служит величина δ , ограничивающая максимальное отклонение плотности электронов $N_s^i(\mathbf{r})$, полученной после решения уравнения Пуассона и самосогласованной с потенциалом $U^i(\mathbf{r})$, от плотности $N^i(\mathbf{r})$, полученной после решения уравнения Шредингера,

$$\frac{\max_{\mathbf{r}} |N_s^i(\mathbf{r}) - N^i(\mathbf{r})|}{N(\infty)} \leq \delta. \quad (2.20)$$

Здесь предполагается, что в бесконечно удаленных областях плотность электронов задана условием локальной нейтральности и не меняется в процессе итераций.

В ходе решения самосогласованного уравнения Пуассона (2.15) индуцированная плотность

$$N_{ind}^i = N_{ind}(U^i, N_{qu}^i)$$

должна вычисляться как неявная функция неизвестного потенциала U^i при известной квантовой поправке N_{qu}^i согласно формулам (2.2) и (2.7). Неявность этой функции обусловлена тем, что входящий в (2.7) обменно-корреляционный потенциал зависит от полной плотности электронов $N_{ind}^i + N_{qu}^i$.

Следует отметить, что для представления индуцированного заряда в форме (2.7) необходимо выполнение условия

$$\frac{dN_{ind}}{dU} < 0. \quad (2.21)$$

В противном случае квадрат длины линейного экранирования в линеаризованном по U уравнении (2.15) становится отрицательной величиной и отклик на малые пространственные возмущения электронной плотности перестает быть локализованным из-за осциллирующего поведения решений самосогласованного уравнения Пуассона. Обычно неравенство (2.21) нарушается там, где плотность электронов мала и соответственно велика относительная роль обменно-корреляционного потенциала. В таких областях неустойчивости введение индуцированной плотности электронов формулой (2.7) не справедливо, и мы будем полагать там $N_{ind}^i(\mathbf{r}) \equiv 0$. Эта ситуация возникает, например, вблизи поверхности электронных систем в области «хвостов» электронного распределения, где плотность электронов всегда стремится к нулю.

В заключение этого раздела сделаем несколько замечаний об используемой в дальнейшем терминологии. Результаты самосогласованных вычислений полной системы уравнений (2.14)–(2.19) с формулой (2.7) для индуцированной плотности электронов, обменным потенциалом в виде (2.8) и корреляционным потенциалом в виде (2.11) будем называть точным или просто самосогласованным расчетом. Решение той же системы уравнений без учета U_{xc} назовем самосогласованным расчетом в приближении Хартри. Самосогласованное решение уравнения Пуассона (2.12) с индуцированной плотностью (2.7) без учета U_{xc} и при $N_{qu} = 0$ дает приближение Томаса–Ферми. Учет обменного потенциала U_x в этой схеме называется приближени-

ем Томаса–Ферми–Дирака. Использование в формуле (2.7) для индуцированного заряда полного эффективного потенциала в приближении локальной плотности иногда называют приближением Томаса–Ферми–Дирака–Гомбаша (см., например, [9, § 11]).

2.2. Особенности применения алгоритма

Конкретнее техника вычислений с использованием индуцированной плотности электронов будет пояснена в следующих разделах, где описана реализация предложенного итерационного алгоритма на примере расчета квантовой поправки к емкости барьерных структур и вычислений работы выхода и плотности поверхностной энергии простых металлов в модели с однородным положительным фоном. Однако однородность фона позволяет продвигаться дальше в аналитических преобразованиях общих формул. Поэтому рассмотрим полубесконечный электронный газ в предположении, что заряд ионов образует однородный положительный фон с плотностью N_+ в виде

$$N_+(z) = \begin{cases} N_+, & z \geq z_+, \\ 0, & z < z_+. \end{cases} \quad (2.22)$$

В этой модели эффективный потенциал в уравнении Шредингера (2.17) зависит только от координаты z . Такая система может представлять, например, поверхность металла с самосогласованным барьером либо полупроводниковую структуру, ограниченную потенциальным барьером.

Выражение для индуцированного заряда в форме (2.7) асимптотически правильно (по параметру квазиклассичности, который в рассматриваемых системах будет всегда мал при $z \rightarrow \infty$) описывает самосогласованную реакцию электронов на изменение длинноволновой части кулоновского потенциала, что и требуется для устранения бесконечного радиуса действия прямого кулоновского взаимодействия зарядов. Однако, как отмечалось выше, оно не справедливо в области, где электронная плотность мала и нарушается условие (2.21). С учетом сказанного самосогласованное уравнение Пуассона можно записать в виде

$$\begin{aligned} \frac{d^2 u^i}{d\zeta^2} + c_n n_{ind}(u^i, n_{qu}^i) &= \\ &= c_n (\theta(\zeta - \zeta_+) - n_{qu}^i(\zeta)), \quad \zeta > \zeta_c^i, \\ \frac{d^2 u^i}{d\zeta^2} &= c_n (\theta(\zeta - \zeta_+) - n^i(\zeta)), \quad \zeta \leq \zeta_c^i, \end{aligned} \quad (2.23)$$

где $c_n = (8/3\pi)(4/9\pi)^{1/3} R_s$ и $\theta(\zeta)$ — тета-функция Хевисайда. Здесь введены безразмерные переменные

$$\zeta = k_F z, \quad \zeta_+ = k_F z_+, \\ n(\zeta) = \frac{N(z)}{N_+}, \quad u(\zeta) = \frac{U(z)}{\varepsilon_F^0},$$

где $k_F = (3\pi^2 N_+)^{1/3}$, $\varepsilon_F^0 = k_F^2/2$. Граничные условия должны обеспечивать обращение в нуль на бесконечности электрического поля $du/d\zeta$, чтобы иметь там ограниченность потенциала, $u(\infty) = \text{const}$, и локальную нейтральность электронного газа, $n(\infty) = 1$.

В уравнении (2.23) предполагается наличие одной критической точки ζ_c , которая отделяет область, где существует индуцированная плотность n_{ind} в форме (2.7), от области, где ее введение недопустимо согласно критерию (2.21). При расчете систем, удовлетворяющих условию $R_s < R_{sc}$, всегда имеется возможность ввести n_{ind} в области $\zeta_c^i < \zeta < \infty$. Там же окажется определенной плотность n_{qu}^i . На самом деле задача всегда считается на конечном отрезке вещественной оси $0 \leq \zeta \leq \zeta_{max}$. Через $\zeta_+ \geq 0$ обозначен возможный сдвиг начала положительного фона от нуля. Величины ζ_+ и ζ_{max} предполагаются настолько большими, чтобы перенос граничных условий из точек $\zeta = -\infty$ и $\zeta = \infty$ в точки $\zeta = 0$ и ζ_{max} не оказывал влияния на результат в пределах выбранной точности. Это условие легко проверяется расчетом при нескольких последовательно увеличивающихся значениях ζ_+ и ζ_{max} .

В точке ζ_c правое и левое решения сшиваются так, чтобы потенциал u и его первая производная были непрерывны. Критическая точка ζ_c определяется как точка, в которой знак неравенства (2.21) меняется на противоположный, т. е. производная dn_{ind}/du становится положительной. Поскольку при смене знака производная многозначной неявной функции $n_{ind}(u)$ проходит через бесконечность, удобнее, как в методе функционала плотности, анализировать обратную к ней однозначную функцию $u(n_{ind})$, производная которой меняет знак, проходя через нуль. Эта функция может быть получена из формулы (2.7) и во введенных безразмерных переменных имеет вид

$$u(n_{ind}) = \mu - n_{ind}^{2/3} - u_{xc}(n_{ind} + n_{qu}), \quad (2.24)$$

где $u_{xc} = U_{xc}/\varepsilon_F^0$, а $\mu = 1 - u_{xc}(\infty)$ — безразмерная энергия Ферми электронов при учете обменно-корреляционного взаимодействия.

В отличие от анализа условий применимости формулы для индуцированной плотности в однородном электронном газе, проведенного в работе [10], ко-

торый с учетом (2.9) дает оценку критической плотности в виде $R_{sc} = 5.64$, в настоящем случае на зависимость $n_{ind}(u)$ влияет еще квантовая поправка к плотности n_{qu} . Это влияние обусловлено тем, что обменно-корреляционный потенциал в приближении локальной плотности зависит от полной плотности электронов. Поэтому необходимо провести рассмотрение заново, основываясь на выражении

$$\frac{du}{dn_{ind}} = -\frac{2}{3}n_{ind}^{-1/3} - \frac{d}{dn_{ind}}[u_{xc}(n_{ind} + n_{qu})]. \quad (2.25)$$

Необходимый для качественного понимания результат может быть выведен аналитически, если пренебречь корреляционной энергией. Ее учет не изменит полученной картины, так как даже при наименьшем допустимом значении плотности однородного электронного газа на бесконечности, отвечающем наибольшему допустимому значению $R_{sc} = 5.64$, отношение $u_c(R_{sc})/u_x(R_{sc}) = 0.26$, а отношение производных еще меньше: $(du_c/dn)/(du_x/dn) \approx 0.07$. С ростом плотности эти оценки только улучшаются. В то же время в ходе численной реализации алгоритма условие $du/dn_{ind} \leq 0$ легко проверяется без всяких упрощений по известной численно правой части формулы (2.25).

С учетом сказанного неравенство, определяющее область применимости квазиклассического выражения (2.7), удобно записать в виде

$$-\frac{2}{3}n_{ind}^{-1/3} + \frac{1}{3}c_x R_s (n_{ind} + n_{qu})^{-2/3} \leq 0, \quad (2.26) \\ c_x = 2(2/3\pi^2)^{2/3}.$$

Входящая сюда производная обменного потенциала (2.8) записана в введенных нами безразмерных переменных. Из формулы (2.26) можно получить уравнение для критических значений n_c полной плотности $n = n_{ind} + n_{qu}$, при которых происходит смена знака производной du/dn_{ind} :

$$n_c^2 - \beta n_c + \beta n_{qu} = 0, \quad \beta = (c_x R_s)^3/8. \quad (2.27)$$

Его решения должны удовлетворять условию неотрицательности полной плотности, $n \geq 0$, и индуцированной плотности, $n_{ind} \geq 0$, в силу их определения формулами (2.5) и (2.7). Значения квантовой поправки n_{qu} могут иметь любой знак, поскольку она определяется формулой (2.6) как разность двух выражений для плотности электронов, полученных из точного и приближенного решений уравнения Шредингера.

Корни уравнения (2.27) даются выражением

$$n_{c1,2} = \frac{\beta}{2} \left(1 \pm \sqrt{1 - \frac{4}{\beta} n_{qu}} \right). \quad (2.28)$$

Ясно, что при $n_{qu} \leq 0$ только один корень

$$n_{c1} = \frac{\beta}{2} \left(1 + \sqrt{1 - \frac{4}{\beta} n_{qu}} \right) \quad (2.29)$$

удовлетворяет сформулированному условию неотрицательности n_c и $n_{ind}^{(c)} = n_c - n_{qu}$. В случае $n_{qu} = 0$ формула (2.29) дает значение n_c для однородного электронного газа при учете только обменного потенциала, которое после перехода к атомным единицам отвечает значению $R_{sc} = 6.02$ [2]. Для $n_{qu} < 0$ критическая плотность будет только больше. Учет корреляционного потенциала также увеличивает оценку n_c , так как добавляет положительное слагаемое в левой части неравенства (2.26). В случае однородного газа с корреляционной энергией в виде (2.9) это смещает критическое значение к $R_{sc} = 5.64$.

Если отрицательные значения n_{qu} не меняют качественно зависимости индуцированной плотности n_{ind} от среднего кулоновского потенциала u по сравнению с однородным случаем, оставляя только одно значение критической плотности, то при $n_{qu} > 0$ картина меняется. Оба корня, даваемые формулой (2.28), имеют физический смысл и обеспечивают положительность как полной плотности n_c , что очевидно, так и индуцированной $n_c - n_{qu}$, что легко проверить. Видно также, что имеется ограничение на величину квантовой поправки $n_{qu} \leq \beta/4$, необходимое для существования корней. Если n_{qu} превысит это значение, кулоновский потенциал u становится монотонно убывающей функцией n_{ind} во всей области физически допустимых значений и условие (2.21) возможности использования формулы (2.7) для индуцированной плотности не нарушается при любых значениях u и n .

Подчеркнем, что условия для n_c , порожденные неравенством (2.26), не совпадают с условием неотрицательности подкоренного выражения в формуле (2.7), а выполняются раньше. Именно благодаря этому обстоятельству, вызванному отрицательностью u_{xc} и его зависимостью от n , действие формулы (2.7) с ростом потенциала u прекращается раньше, чем обращается в нуль n_{ind} . Поэтому при росте u или уменьшении n индуцированный заряд обращается в нуль скачком от конечного значения, формируя в процессе итераций разрывы первого рода в распределении полной плотности электронов и в эффективном потенциале, как это можно увидеть на рис. 1. По достижении самосогласованности этот разрыв оказывается, естественно, меньше заданной точности, т. е. критерия выхода (2.20) из итерационного цикла. Однако реализация обработки этого раз-

рыва при решении уравнений Шредингера и Пуассона оказалась существенным моментом, влияющим на эффективность всего алгоритма, о чем еще будет сказано дальше.

Оценка величины разрывов в плотности и эффективном потенциале может быть получена точно, если пренебречь n_{qu} . Из результатов такого анализа следует, что величина скачка в плотности равна $n_{ind}^{(c)}$ и на начальном цикле итераций принимает значение $(R_s/R_{sc})^3$. Видно, что с ростом неидеальности электронного газа и приближением R_s к R_{sc} скачок в безразмерной плотности стремится к единице и становится все более сильным возмущением.

Разрыв в эффективном потенциале равен значению $u_{xc} [n_{ind}^{(c)}]$, так как кулоновский потенциал непрерывен всюду. В атомных единицах этот разрыв имеет универсальное значение, зависящее только от выбранного выражения для корреляционной энергии, и меняется от 0.1 На при учете только обменного потенциала до 0.15 На при добавлении к нему корреляционного потенциала согласно стандартной формуле Вигнера [2]. Однако в безразмерных единицах разрыв эффективного потенциала также значительно увеличивается с ростом неидеальности электронов из-за уменьшения ε_F^0 .

Выпишем теперь алгоритм итеративного решения (2.14)–(2.19) применительно к рассматриваемому в этом разделе типу полубесконечных многоэлектронных систем с однородным положительным фоном, используя введенные безразмерные переменные и сопровождая каждый шаг необходимыми пояснениями.

Стартовое значение квантовой поправки к плотности электронов и сама полная плотность полагаются равными нулю, начальное положение критической точки выносится за границу положительного фона:

$$n_{qu}^0(\zeta) \equiv 0, \quad n^0(\zeta) \equiv 0, \quad \zeta^0 = -1. \quad (2.30)$$

Решив уравнение Пуассона (2.23), получаем кулоновский потенциал u^i , удовлетворяющий поставленным граничным условиям, и самосогласованную с ним новую полную плотность n_s^i электронов

$$\begin{aligned} n_s^i &= n_{ind}^i + n_{qu}^i, & \zeta > \zeta_c^i, \\ n_s^i &= n^i, & \zeta \leq \zeta_c^i, \quad i = 0, 1, \dots \end{aligned} \quad (2.31)$$

При этом проверка условия $dn_{ind}^i/du^i < 0$ может показать, что необходимо перенести положение критической точки на новое место, ζ_c^{i+1} . При известных численно функциях $n_s^i(\zeta)$, $u^i(\zeta)$ проще всего эту про-

верку сделать с помощью формулы (2.25), если заменить n_{ind}^i выражением через потенциал u^i и переписать определение (2.7) в безразмерных величинах:

$$n_{ind}^i = (\mu - u^i - u_{xc}(n_s^i))^{3/2}. \quad (2.32)$$

В случае смещения критической точки вправо следует переопределить n_s^i , заменив в процедуре (2.31) ζ_c^i на ζ_c^{i+1} . Сдвиг критической точки влево оставляет n_s^i без изменения, однако новое положение критической точки определит область формирования индуцированной плотности после решения уравнения Шредингера.

Несколько особняком стоит случай нулевого цикла, $i = 0$, когда можно обойти использование прямого численного метода решения дифференциального уравнения второго порядка. Положив $\zeta_c^0 = 0$, мы должны начать решение с уравнения (2.23). В отсутствие квантовой поправки к плотности, когда нелинейный по u индуцированный заряд не зависит явно от координаты ζ , первый интеграл этого уравнения может быть найден в явном виде. После этого решение находится численной квадратурой первого интеграла. В процессе этого решения определяются кулоновский потенциал u , самосогласованное с ним распределение электронов, которое сводится здесь только к n_{ind}^0 , и положение критической точки ζ_c^1 . Если для описания системы в целом достаточно квазиклассического приближения, такое решение может иметь самостоятельное значение (см., например, формулы (35.70)–(35.73) в [9] и расчет туннельного тока через барьер Шоттки в [10]).

Зная решение уравнения Пуассона, можно сформировать эффективный потенциал $u_{eff}^i = u^i + u_{xc}^i(n_s^i)$, который входит в уравнение Шредингера. В силу трансляционной инвариантности рассматриваемой системы в плоскости xy , полная волновая функция может быть представлена формулой

$$\Psi_{k, \mathbf{k}_{\parallel}}(\zeta, \mathbf{r}_{\parallel}) = \frac{1}{2\pi} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}} C \psi_k(\zeta), \quad (2.33)$$

где вектор \mathbf{r}_{\parallel} лежит в плоскости xy , а константа C определяется условием нормировки функции ψ_k . Уравнение Шредингера для ψ_k^i на i -м цикле итераций удобно записать в виде

$$\frac{d^2}{d\zeta^2} \psi_k^i(\zeta) + (k^2 - u_{eff}^i(\zeta) + u_{xc}(\infty)) \psi_k^i(\zeta) = 0, \quad (2.34)$$

где безразмерное квантовое число k нормировано на k_F и выбрана калибровка скалярного потенциала условием $u(\infty) = 0$. Собственное значение k

определяет поведение на бесконечности собственных функций непрерывного спектра, нормированных на $\delta(k - k')$ согласно [2] (см. формулы (24), (25)) и имеющих асимптотический вид

$$\psi_k^i(\zeta) \rightarrow \sqrt{\frac{2}{\pi}} \sin(k\zeta + \gamma_k), \quad \zeta \rightarrow \infty. \quad (2.35)$$

Найденное с точностью до произвольного постоянного множителя численное решение $\tilde{\psi}_k(\zeta)$ задачи Коши для однородного уравнения (2.34) будет правильно нормировано, если его асимптотика при $\zeta \rightarrow \infty$ имеет вид (2.35). Это достигается умножением полученного решения на постоянную A , которая определяется условием

$$A^2 (\tilde{\psi}_k^2(\zeta_{\infty}) + \tilde{\psi}_k'^2(\zeta_{\infty})) = 2/\pi. \quad (2.36)$$

Здесь через $\zeta_{\infty} \leq \zeta_{max}$ обозначена достаточно удаленная точка, где потенциал $u_{eff}(\zeta_{\infty})$ можно считать постоянным. Фаза γ_k может находиться из соотношения

$$\gamma_k = \arctg \left[k \frac{\tilde{\psi}_k(\zeta_{\infty})}{\tilde{\psi}_k'(\zeta_{\infty})} \right] - (k\zeta_{\infty} - l\pi), \quad (2.37)$$

где через $l = [k\zeta_{\infty}/\pi]$ обозначена целая часть выписанного отношения. При таком определении фаза будет инвариантна относительно выбора точки ζ_{∞} .

В завершение i -го цикла итераций вычисляется полная плотность электронов

$$n^i(\zeta) = 3 \int_0^1 dk (1 - k^2) |\psi_k^i(\zeta)|^2, \quad (2.38)$$

и с учетом (2.32) — квантовая поправка к плотности для следующего цикла:

$$n_{qu}^{i+1}(\zeta) = n^i(\zeta) - n_{ind}(u_{eff}^i(\zeta)). \quad (2.39)$$

Задача Коши для уравнения Шредингера решалась численно с использованием неявного метода четвертого порядка точности, который часто называют методом Нумерова. В алгоритм Нумерова напрямую не входит заданное в начальной точке значение первой производной искомого решения. Вместо этого надо знать значение искомой функции в первой от края точке. Для вычисления ее с точностью, соответствующей точности метода Нумерова, использовалось представление функции в виде отрезка ряда Тейлора с включением членов второго порядка. Необходимое для этого значение второй производной решения в краевой точке определялось из

самого уравнения Шредингера. Аналогичный прием использовался для прохождения точек разрыва эффективного потенциала. Было проверено, что такая процедура дает результат, совпадающий в пределах выбранной точности с результатом при использовании метода Рунге – Кутты для отхода на один шаг от точки, где заданы граничные условия задачи Коши.

Численное решение уравнения Пуассона проводилось методом установления [11]. Идея этого метода заключается в том, что вместо стационарной задачи решается соответствующее нестационарное уравнение

$$\frac{\partial u}{\partial t} - \left(\frac{\partial^2 u}{\partial \zeta^2} + c_n n_{ind}(u, n_{qu}) \right) = -c_n [\theta(\zeta - \zeta_+) - n_{qu}(\zeta)] \quad (2.40)$$

до тех пор, пока его решение не перестанет меняться со временем t в пределах заданной точности.

С формальной точки зрения уравнение (2.40) является нелинейным уравнением диффузии, для решения которого разработано большое количество численных реализаций. В данной работе применена схема, в которой после дискретизации (2.40) по времени с шагом τ и по пространству с шагом h задается начальное распределение потенциала, удовлетворяющее краевым условиям, и прослеживается эволюция решения во времени путем решения краевой задачи в каждой временной точке. При этом нет необходимости получать пространственное распределение потенциала u на каждом «временном» шаге с хорошей точностью, требуется лишь найти точное стационарное решение на больших временах, когда изменения u от шага к шагу становятся малы. Поскольку при дискретизации по времени была применена неявная схема первого порядка, для оценки нелинейного члена уравнения при формировании краевой задачи в новой точке по времени использовалась линеаризация индуцированной плотности $n_{ind}(u, n_{qu})$ относительно приращения потенциала на шаге τ . Выбор временного шага τ в ходе решения (2.40) проводится таким образом, чтобы, с одной стороны, обеспечить устойчивость численной схемы, а с другой стороны, увеличить скорость вычислений (скорость выхода на стационарное решение) за счет перехода ко все большим значениям τ . Согласно имеющимся в литературе рекомендациям, начальный шаг по времени выбирался из условия $\tau/h^2 = 0.1$.

Поскольку при переходе от одной временной точки к следующей в процессе решения использовалась линеаризация зависимости $n_{ind}(u)$, могла возникать

необходимость переопределить критическую точку в ходе решения уравнения Пуассона, чтобы исключить появление области, где $dn_{ind}(u, n_q)/du > 0$. Это проводилось таким же образом, как описано после формулы (2.32). Как отмечалось выше, при увеличении R_s амплитуда скачка в электронной плотности растет, что приводит ко все более сильным возмущениям численного решения уравнений. Было обнаружено (подробнее см. разд. 4), что при $R_s \lesssim R_{sc}$ сходимость решения системы уравнений (2.14)–(2.19) ухудшается. Однако представляется, что этот факт является, скорее, следствием конкретной реализации численного алгоритма, нежели дефектом предложенной итерационной схемы.

3. СВОЙСТВА ЭЛЕКТРОННОГО ГАЗА СТРУКТУР С ПОТЕНЦИАЛЬНЫМ БАРЬЕРОМ

В этом разделе описанный выше алгоритм решения уравнений Кона – Шэма применяется для нахождения самосогласованной электронной плотности и эффективного потенциала в барьерной структуре, представляющей собой два проводника с вырожденным электронным газом, разделенных слоем диэлектрика с диэлектрической проницаемостью κ_d . Будем считать, что ось z направлена перпендикулярно плоскости слоев структуры, проводники расположены соответственно при $z \leq -d/2$ и $z \geq d/2$, а слой диэлектрика занимает область $|z| < d/2$. В дальнейшем предполагается, что нейтрализующий фон в проводниках однороден и барьер не легирован, т. е.

$$N_+(z) = \begin{cases} 0, & |z| < d/2, \\ N_+, & |z| \geq d/2. \end{cases} \quad (3.1)$$

Высота потенциального барьера, образованного слоем диэлектрика, полагается равной бесконечности, так что волновые функции равны нулю на границах проводник–диэлектрик.

3.1. Самоиндуцированные уровни в полупроводниках

В случае бесконечно высокого потенциального барьера даже в отсутствие неоднородности положительного фона $N_+(z)$ и приложенного напряжения возможно появление локализованных по z электронных состояний и образование размерно-квантованных двумерных подзон. Впервые предположение о существовании таких состояний было выдвинуто в

работе [12], где соответствующие дискретные уровни были названы «плазменными уровнями». Появление локализованных состояний связывалось с обращением в нуль электронных волновых функций на границе с барьером. Благодаря непрерывности волновых функций, вблизи барьера возникает область с пониженной плотностью электронов, недостаточной для соблюдения локальной нейтральности. Нескомпенсированный положительный заряд фона создает около барьера самосогласованную потенциальную яму, в которой могут существовать связанные состояния. Согласно сделанному в работе [12] расчету, в случае вырожденного электронного газа уровни должны существовать при $R_s \leq 2.87$, а при $R_s \leq 0.15$ в прибарьерной потенциальной яме появляется второе связанное состояние.

Независимо от [12] в работе [13] также была высказана идея образования обедненной электронами области вблизи барьера и исследовалась зависимость энергии локализованных состояний и глубины самоиндуцированной потенциальной ямы от R_s . Однако в этих работах либо не достигалось самосогласования, как это было в [12], либо в уравнении Шредингера вместо точного решения уравнения Пуассона использовался некий модельный потенциал [13]. В обеих работах пренебрегалось обменно-корреляционным взаимодействием свободных носителей заряда. Самоиндуцированные уровни достаточно мелкие. Их энергии составляют максимум 0.1 от глубины потенциальной ямы (приблизительно 2 мэВ в случае полупроводника с параметрами $n\text{-GaAs}$), так что они находятся очень близко к состояниям непрерывного спектра. Поэтому незначительные изменения формы потенциальной ямы могут приводить к их исчезновению. В связи с этим представлял интерес вопрос, сохранятся ли эти состояния в самосогласованном решении.

Система уравнений Кона–Шэма решалась согласно итеративной схеме (2.14)–(2.19). Полная безразмерная плотность электронов n^i в области $z \geq d/2$, получаемая после решения уравнения Шредингера на i -й итерации, вычислялась по формуле

$$n^i(\zeta) = 3 \int_0^1 dk (1 - k^2) |\psi_k^i(\zeta)|^2 + \frac{3\pi}{2} \sum_j |\psi_j^i(\zeta)|^2 (1 - \varepsilon_j^i), \quad (3.2)$$

где $\zeta = k_F(z - d/2)$, j — номер дискретного уровня, ε_j — дискретные энергетические уровни, ψ_k , ψ_j — волновые функции соответственно непрерыв-

ной и дискретной частей спектра. Первое слагаемое в правой части (3.2) отвечает плотности электронов непрерывного спектра, второе — плотности электронов на локализованных состояниях.

Поскольку рассматриваемая задача с бесконечно высоким барьером близка к задаче о потенциальном рассеянии s -волны на короткодействующем потенциале, для определения наличия и числа уровней в потенциальной яме можно было бы воспользоваться теоремой Левинсона (см., например, [14]), которая выражает число m уровней в яме через фазу асимптотического поведения (2.35) волновых функций непрерывного спектра на бесконечности соотношением $\lim_{k \rightarrow 0} \gamma_k = m\pi$. Хотя эта формула используется, например в работе [15], при теоретическом анализе барьерной структуры, она неудобна при численных расчетах из-за избыточных вычислений волновых функций для незаполненных состояний. Вместо этого достаточно найти решение уравнения Шредингера (2.34) при $k = 0$ и подсчитать число нулей, лежащих внутри интервала $0 < \zeta < \infty$.

Однако при нахождении энергии и волновой функции связанного состояния мы использовали метод тригонометрической прогонки, который устраняет известную численную неустойчивость убывающего в классически запрещенной области решения, обусловленную наличием экспоненциально растущего второго решения [16]. Поскольку в этом методе волновая функция представляется в виде

$$\psi(\zeta) = a(\zeta) \sin \eta(\zeta), \quad \frac{d\psi}{d\zeta} = a(\zeta) \cos \eta(\zeta), \quad (3.3)$$

нетрудно показать, что введенная здесь полная фаза $\eta(\zeta)$ будет получать приращение π в нулях функции $\psi(\zeta)$, где логарифмическая производная последней обращается в бесконечность. Для фазы и амплитуды имеются два уравнения первого порядка:

$$\frac{d\eta}{d\zeta} = (\varepsilon - u_{eff}(\zeta) + u_{eff}(\infty)) \sin^2 \eta + \cos^2 \eta, \quad (3.4)$$

$$\frac{da}{d\zeta} = \frac{1}{2} a \sin(2\eta) (1 + u_{eff}(\zeta) - u_{eff}(\infty) - \varepsilon), \quad (3.5)$$

которые выводятся из (2.34) с учетом (3.3). Уравнение для $\eta(\zeta)$ решается с начальным условием $\eta(0) = 0$, начальное условие для амплитуды может быть выбрано любым числом, которое окончательно определяется условием нормировки волновой функции на 1. Энергии уровня $\varepsilon_j \leq 0$ соответствует предельное значение фазы $\eta_\varepsilon(\infty) = j\pi (j = 1, 2, \dots)$. Поэтому для подсчета числа уровней в яме достаточно было решить уравнение (3.4) для двух значений энергии, отвечающих дну потенциальной ямы и

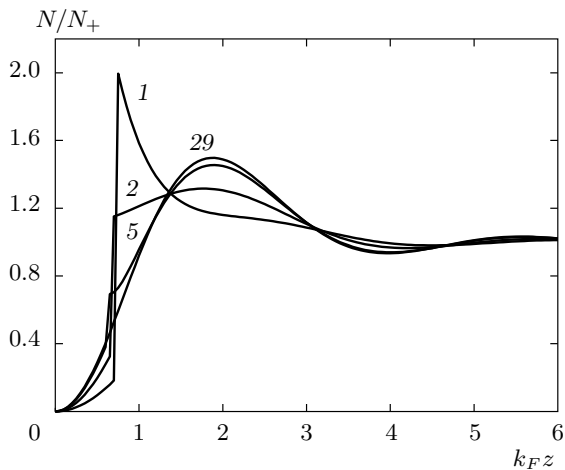


Рис. 1. Зависимости от координат плотности электронов, получаемой на различных циклах итерационного процесса. Числа около кривых соответствуют номерам итераций, после которых эти кривые были рассчитаны. В этом разделе координата z на рисунках отсчитывается от точки $d/2$

границе непрерывного спектра $\varepsilon = 0$, и определить кратность π разности соответствующих фаз на бесконечности.

Граничные условия к уравнению Пуассона брались в виде

$$\left. \frac{du}{d\zeta} \right|_{\zeta=0} = 0, \quad \left. \frac{du}{d\zeta} \right|_{\infty} = 0. \quad (3.6)$$

Отметим, что такие граничные условия вместе с заданием уровня Ферми посредством требования локальной нейтральности системы при $\zeta \rightarrow \infty$ и выбором $u(\infty) = 0$ однозначно определяют численное решение уравнения Пуассона (2.23) и обеспечивают его стремление к нулю на бесконечности.

На рис. 1 представлены распределения электронной плотности, полученные после различного числа итераций. На рисунке хорошо видно, что на начальных итерациях зависимость электронной плотности от координаты имеет разрывы, образующиеся на границе области применимости квазиклассического выражения для индуцированного заряда в форме (2.7). В ходе итерационного процесса эти разрывы «зарастают», и конечная самосогласованная плотность является гладкой функцией координаты. Существенно еще отметить, что заданные критерии сходимости итераций обычно достигались на 10–14 циклах (в зависимости от значения R_s). Продолжение вычислений не приводило к заметному увеличению точности при выбранных параметрах дискре-

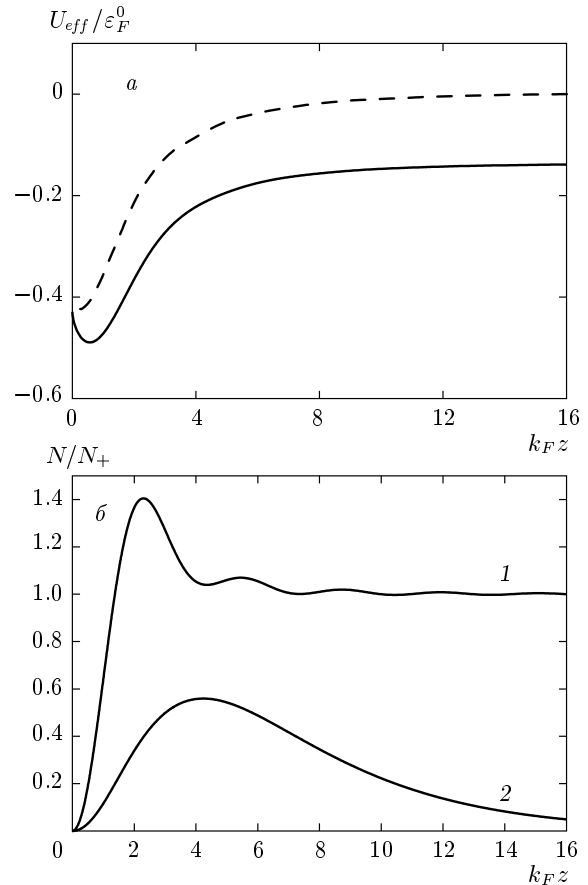


Рис. 2. Зависимости от координат самосогласованного потенциала (а) и электронной плотности (б) при $R_s = 0.4$. На рис. а — самосогласованный потенциал, полученный с учетом обменно-корреляционного взаимодействия (сплошная линия) и в приближении Хартри (штриховая линия). На рис. б — полная самосогласованная плотность электронов (кривая 1) и плотность электронов, локализованных в размерно-квантованной подзоне, (кривая 2)

тизации, но и не разрушало, как видно из рис. 1, полученного решения, что показывает устойчивость принятого алгоритма.

На рис. 2 представлены зависимости самосогласованного потенциала и электронной плотности от координат. Как видно на рис. 2а, учет U_{xc} меняет форму и глубину потенциальной ямы и, соответственно, энергию локализованного состояния (см. кривые 1, 2 на рис. 3). Полная самосогласованная электронная плотность на рис. 2б (кривая 1) равна сумме плотностей электронов, находящихся в размерно-квантованной подзоне, (кривая 2) и плотности электронов непрерывного спектра. Локализован-

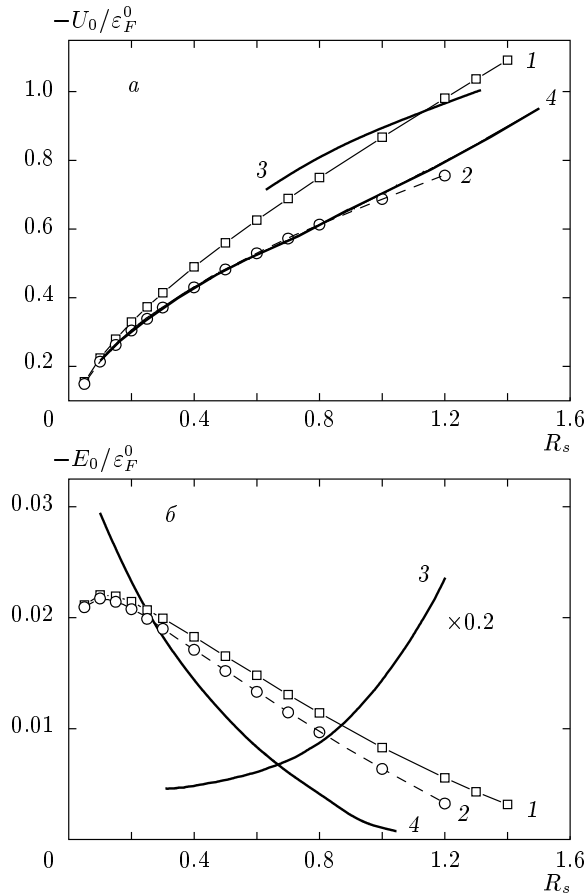


Рис. 3. Зависимости глубины потенциальной ямы (а) и энергии связанного состояния (б) от R_s . Кривые 1 — результаты расчета с учетом обменно-корреляционного взаимодействия, кривые 2 — самосогласованный расчет в приближении Хартри. Кривые 3 и 4 — соответственно результаты работ [12] и [13]

ные электроны в основном сосредоточены вблизи границы потенциального барьера, однако область локализации зависит от R_s .

Зависимости глубины самосогласованной ямы и энергии локализованного в ней состояния от R_s представлены соответственно на рис. 3. Видно, что результаты вычислений энергии уровня и ее зависимости от R_s в работе [12] (кривая 3 на рис. 3б) и наших самосогласованных вычислений расходятся как количественно, так и качественно. Кроме того, ни при каких значениях $R_s \geq 0.05$ мы не обнаружили появления в прибарьерной потенциальной яме двух локализованных уровней в отсутствие внешнего электрического поля. Второй уровень с безразмерной энергией связи порядка 0.0016 обнаружился только при значении $R_s = 0.005$, которое физически

не реализуемо в реальных структурах. Тем не менее следует отметить, что высказанная в работе [12] идея о существовании вблизи бесконечно высокого барьера самосогласованной потенциальной ямы и возможности появления в ней связанных состояний оказалась верной.

Результаты работы [13] для глубины потенциальной ямы неплохо согласуются с результатами самосогласованного расчета настоящей работы в приближении Хартри (см. кривые 2 и 4 на рис. 3а). Однако энергия связанного состояния (кривая 4 на рис. 3б), полученная в результате расчета в работе [13], отличается от точного решения как по величине, так и по зависимости от R_s , особенно в области малых значений этого параметра. Причем эти различия не связаны с пренебрежением в работе [13] обменно-корреляционным взаимодействием, что с очевидностью следует из относительной близости полученных нами кривых 1 и 2 на рис. 3б.

С большой вероятностью можно предположить, что причина значительной ошибки в расчете энергии самоиндуцированных уровней методом, использованным в работе [13], является неточное описание формы потенциальной ямы трехпараметрическим модельным потенциалом.

Учет конечной высоты барьера приводит к изменению граничного условия для волновых функций электронов на границе с диэлектриком, что, как показывает расчет, в свою очередь ведет к уменьшению верхней границы области R_s , в которой наблюдается существование локализованных состояний. При значительном уменьшении высоты барьера дискретный уровень исчезает вовсе.

Вдали от потенциального барьера, $k_F z \gg 1$, проявляются фриделевские осцилляции эффективного потенциала и электронной плотности (см. рис. 4). На рисунке видно, что амплитуда осцилляций эффективного потенциала, который содержит вклад U_{xc} , существенно меньше, чем амплитуда осцилляций кулоновского потенциала. Такое подавление осцилляций может быть понято следующим образом. Согласно уравнению Пуассона, осциллирующая компонента плотности электронов имеет максимумы в тех же точках, что и кулоновский потенциал. Обменно-корреляционный вклад в U_{eff} всегда имеет знак, обратный знаку прямого взаимодействия. Поэтому в силу отталкивательного характера межэлектронного взаимодействия потенциал U_{xc} отрицателен и согласно формулам (2.4), (2.8) и (2.11) растет монотонно по абсолютной величине с ростом плотности электронов. Следовательно, там, где кулоновский потенциал имеет максимум, вклад об-

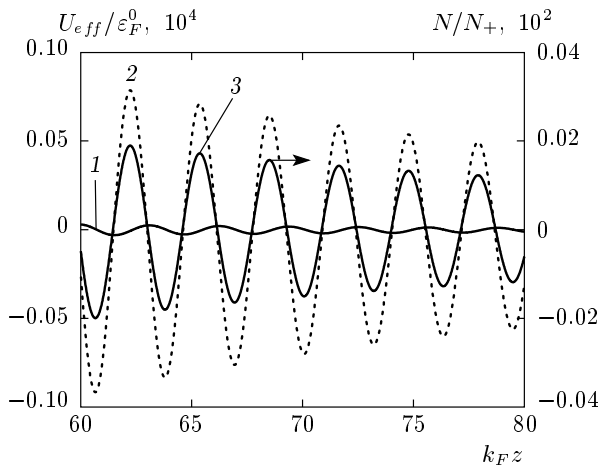


Рис. 4. Фриделевские осцилляции эффективного потенциала (1) и электронной плотности (3) в электронном газе, ограниченном бесконечно высоким барьером. Кривая 2 — кулоновская составляющая эффективного потенциала. $R_s = 0.4$

менно-корреляционного взаимодействия будет тоже максимален по абсолютной величине и отрицателен.

Аналогичные рассуждения показывают, что в минимумах эффективного потенциала кулоновский и обменно-корреляционный вклады также суммируются в противофазе. Поскольку относительная роль U_{xc} возрастает с уменьшением плотности электронов, это ведет к большему подавлению амплитуды осцилляций эффективного потенциала относительно осцилляций кулоновского с ростом R_s . Однако заметим, что влияние U_{xc} на зависимость амплитуды осцилляций в безразмерной плотности электронов от R_s имеет обратный характер. Причина этого более подробно обсуждается в разд. 5.

В заключение здесь стоит отметить, что рассмотренная система, в которой заполнены состояния дискретной и непрерывной областей спектра, имеет много общего с квазидвумерным электронным газом обогащенных слоев, образующихся вблизи границы полупроводник–диэлектрик при наложении или даже в отсутствие внешнего электрического поля. Хотя считается [17], что в работе [15] проблема самосоглаванного расчета спектра размерно-квантованных подзон обогащенного слоя решена, статьи, посвященные разработке методов расчета таких структур, продолжают появляться [18]. При этом по-прежнему не предлагается никакой обоснованной процедуры, которая позволила бы избежать рассогласованности между получающимся на промежуточных шагах итераций полным зарядом или

дипольным моментом квазидвумерного газа и наложенными на потенциал граничными условиями к уравнению Пуассона. Чтобы иметь возможность получать решение, проводят перенормировку найденной из уравнения Шредингера плотности электронов, а также варьируют положение уровня Ферми в объеме полупроводника. Результаты настоящего раздела нашей работы показывают, что предложенный алгоритм позволяет получить самосоглаванное решение для таких структур регулярным образом, оставаясь в рамках математически полностью корректного итерационного алгоритма.

3.2. Предельная емкость барьерной структуры

До сих пор рассматривался случай, когда к барьерной структуре не было приложено внешнее напряжение. В этом разделе находится самосоглаванное распределение электронной плотности и эффективного потенциала для случая приложенного к структуре постоянного смещения V и вычисляется емкость барьерной структуры. Помимо проверки работы алгоритма, эта задача представляет самостоятельный интерес в связи с проблемой предельной емкости как тонкопленочных конденсаторов вообще, которая была поставлена работой [19], так и частным случаем наноконденсаторов с тонкопленочным изолятором из сегнетоэлектрика в связи с их перспективностью для применений в электронике [20].

В работе Мида [19] было показано, что измеренная зависимость обратной емкости перехода металл–диэлектрик–металл от толщины d диэлектрического слоя не стремится к нулю при $d \rightarrow 0$ в противоречии с обычной формулой для геометрической емкости плоского конденсатора. Качественно это было объяснено конечной толщиной слоя пространственного заряда, экранирующего проникновение электрического поля в реальный металл. Количественный анализ этой картины в рамках приближения Томаса–Ферми для электронов металла был предпринят в работе [21] и дополнен учетом диэлектрической проницаемости металлического электрода в работе [22]. В последней статье было также отмечено, что использование диэлектрика с большой проницаемостью κ_d увеличивает роль неидеальности электродов даже при толщинах изолятора геометрически много больших, чем длина экранирования Томаса–Ферми в электродах, и введено понятие максимально достижимой с данными электродами емкости перехода как предела $\kappa_d \rightarrow \infty$.

Полученные в указанных работах формулы для

емкости показали качественное понимание экспериментальной зависимости последней от материала электродов и диэлектрика, а также позволили получить правильную по порядку величины оценку вклада области пространственного заряда в емкость. Однако отмечалось, что остается неясным влияние обменно-корреляционного взаимодействия электронов металла и заряда во фриделевских осцилляциях плотности, которое не может быть учтено в рамках приближения Томаса–Ферми. В то же время уже в работе Бардина [23] фактически было получено, что перераспределение электронов в пространстве, которое связано с порожденными барьером осцилляциями плотности, не сохраняет полную нейтральность системы из-за образующегося дефицита электронного заряда. В случае бесконечно высокого барьера без внешнего поля этот недостаток заряда равен (в приближении невзаимодействующих электронов) следующей величине:

$$\int_0^{\infty} dz (N(z) - N_+) = -\frac{k_F^2}{8\pi}, \quad (3.7)$$

что совпадает с избытком положительного заряда, который получается из формулы (26) работы [23] (после исправления опечатки ср. с (2) в [24]). Кроме того, из результатов предыдущего раздела следует, что самоиндуцированный поверхностный уровень также может захватить значительный электронный заряд (см. рис. 2б). При заметной зависимости от приложенного поля заряда в осцилляциях и на приборьерном уровне их вклад в емкость может оказаться сравнимым с вкладом заряда в области экранирования Томаса–Ферми.

Оценка всех перечисленных факторов может быть получена, если найти самосогласованное решение уравнений Кона–Шэма для барьерной структуры в модели желе. Рассмотрим систему из двух полубесконечных металлов, разделенных потенциальным барьером бесконечной высоты шириной d и с диэлектрической проницаемостью κ_d . Такой барьер предотвращает перенос заряда между двумя частями системы при наложении напряжения. Дифференциальная электростатическая емкость на единицу площади системы, в целом нейтральной, выражается через заряд Q на единицу площади правой половины структуры как

$$C(V) = -\frac{dQ}{dV} = \frac{d}{dV} \int_{z_0}^{\infty} dz (N(z, V) - N_+(z)), \quad (3.8)$$

где $V = \varphi(-\infty) - \varphi(\infty)$ — падение напряжения на

структуре, равно работе электрического поля по перемещению единицы положительного заряда из минуса бесконечности в плюс бесконечность. Здесь через $\varphi(z)$ обозначен размерный кулоновский потенциал. Точка z_0 определяет место смены знака плотности заряда в области, где электрическая индукция D максимальна, и задается условием $dD/dz|_{z_0} = 0$. В обеих бесконечностях электрическое поле равно нулю из-за требования ограниченности потенциала $\varphi(z)$, поэтому определение (3.8) обеспечивает равенство по абсолютной величине зарядов обеих частей структуры.

Положим, что металлические электроды одинаковы и распределение положительного фона описывается формулой (3.1). Пренебрежем также для упрощения расчета возможным отличием от нуля контактной разности потенциалов и связанными с ней встроенными полями и зарядом поверхностных состояний, как это было принято для рассматриваемой модели в работах [21, 22]. Поставленные условия означают отсутствие заряда внутри барьера, что позволяет выбрать в качестве z_0 любую точку внутри диэлектрика, включая граничные $z_0 = \pm d/2$. Если задана индукция D внутри барьера, то заряд правого электрода равен

$$Q(V) = -D/4\pi. \quad (3.9)$$

Представим падение напряжения на всей структуре в виде суммы падений напряжения на трех ее составляющих:

$$V = [\varphi(-\infty) - \varphi(-d/2)] + E_d d + [\varphi(d/2) - \varphi(\infty)], \quad (3.10)$$

где поле внутри барьера $E_d = D/\kappa_d$. Обозначим через $\varphi_{\pm}(z, E)$ решения уравнения Пуассона на участках $d/2 \leq z \leq \infty$ и $-d/2 \geq z \geq -\infty$ с краевыми условиями соответственно $d\varphi_{\pm}/dz|_{z=\pm d/2} = E$, $\varphi_{\pm}(\pm\infty) = 0$. Тогда равенство (3.10), учитывая непрерывность индукции на границе диэлектрика и металла, можно переписать в виде

$$V = -\varphi_-(-d/2, D/\kappa_m) + E_d d + \varphi_+(d/2, D/\kappa_m). \quad (3.11)$$

Введем обозначение $E_m = D/\kappa_m$ для поля в металле на границе с диэлектриком, где отличие диэлектрической поляризации κ_m металла от единицы учитывает возможную поляризуемость внутренних электронных оболочек ионных остовов. Поскольку в нашем случае $\varphi_-(-d/2, E_m) = \varphi_+(d/2, -E_m)$, в результате получаем формулу, позволяющую свести расчет емкости барьерной структуры к рассмотренной

в предыдущем разделе задаче о полубесконечной системе:

$$V = \varphi_+(d/2, E_m) - \varphi_+(d/2, -E_m) + E_d d. \quad (3.12)$$

Нас будет интересовать емкость при небольших напряжениях, когда дифференцирование в определении (3.8) можно заменить отношением малых приращений. Из-за аддитивной структуры V в (3.12) удобнее иметь дело с обратной емкостью. С учетом (3.9) получаем

$$C^{-1} = -\frac{V}{Q} = 4\pi \frac{\varphi_+(d/2, E_m) - \varphi_+(d/2, -E_m)}{\kappa_m E_m} + 4\pi \frac{d}{\kappa_d} \equiv 2C_i^{-1} + C_d^{-1}, \quad (3.13)$$

где использованы стандартные обозначения C_i и C_d для выделения вклада в полную емкость падения напряжения на интерфейсах (внутри металлических электродов) и на диэлектрике. Явное определение C_i и C_d легко увидеть из структуры выражения (3.11), если учесть, что $\varphi_+(d/2, 0) \neq 0$ (см. рис. 2а). Очевидно, что максимально возможная емкость C_{max} рассматриваемой структуры достигается в пределе малой электрической толщины диэлектрика $\lim d/\kappa_d \rightarrow 0$ и определяется формулой

$$C_{max}^{-1} = 4\pi \frac{\varphi_+(d/2, E_m) - \varphi_+(d/2, -E_m)}{\kappa_m E_m}. \quad (3.14)$$

В дальнейшем будем полагать $\kappa_m = 1$, так как при необходимости найденные численные значения емкости легко пересчитать, пользуясь выражением (3.14) и учитывая соответствующее переопределение атомных единиц.

Структура формулы (3.14) показывает, что для расчета предельной емкости барьерной структуры надо дважды решить задачу о полубесконечном металле, ограниченном бесконечно высоким барьером, с граничными условиями вида

$$\left. \frac{du}{d\zeta} \right|_{\zeta=0} = \pm E, \quad \left. \frac{du}{d\zeta} \right|_{\infty} = 0. \quad (3.15)$$

Здесь использованы введенные в п. 3.1 безразмерные величины для потенциальной энергии электрона и пространственной координаты. Безразмерное электрическое поле E нормировано на характерное поле $E_c = k_F \varepsilon_F^0 / e$. За исключением изменения граничного условия в нуле по сравнению с (3.6) никаких других изменений в программе расчета не потребовалось.

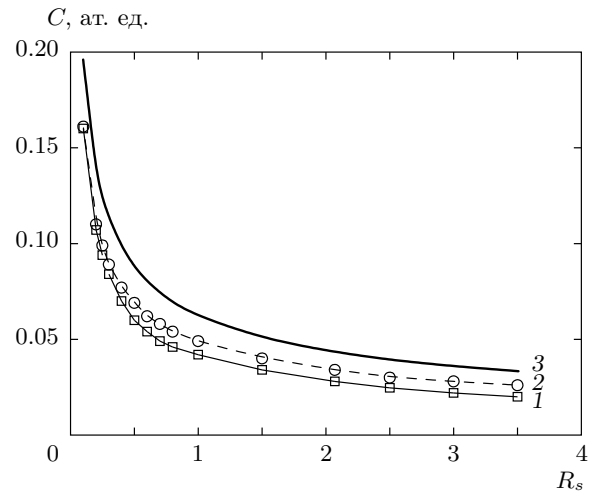


Рис. 5. Емкость барьерной структуры при различных значениях R_s : самосогласованный расчет с учетом обменно-корреляционного взаимодействия (1), в приближениях Хартри (2) и Томаса – Ферми (3)

Найденные значения потенциала на границе с диэлектриком при двух значениях полей использовались в формуле (3.14) для расчета емкости. Для повышения точности и проверки, что выбранные значения полей достаточно малы и не выводят за пределы линейной зависимости заряда от напряжения, вычисление потенциала проводилось для трех пар значений поля $\pm E$, а для расчета емкости использовался наклон полученной линейной зависимости.

На рис. 5 показаны зависимости предельной емкости структуры (обозначена C) от параметра R_s , полученные в различных приближениях. Кривые 1 и 2 отвечают самосогласованному расчету емкости с учетом обменно-корреляционного взаимодействия и в приближении Хартри. Кривая 3 получена в приближении Томаса – Ферми и может быть представлена простой явной формулой:

$$C = \frac{k_{TF}}{8\pi} = \frac{1}{8\pi} \left(\frac{12}{\pi} \right)^{1/3} \frac{1}{R_s^{1/2}}, \quad (3.16)$$

где k_{TF} — обратная длина Томаса – Ферми. Если в приближении Томаса – Ферми учесть еще обменно-корреляционный потенциал в соответствии с формулой (2.7), то для емкости получим формулу типа (3.16), только с заменой k_{TF} на перенормированную величину

$$\tilde{k}_{TF} = \frac{k_{TF}}{\left(1 + \frac{3}{2} \frac{du_{xc}}{dn} \right)^{1/2}}.$$

Соответствующая кривая не была построена на рис. 5, так как различие между ней и кривой 3 в рассматриваемой области значений R_s будет почти таким же незначительным, как и между кривыми 1 и 2. При точном расчете учет обменно-корреляционного потенциала приводит к уменьшению емкости, что, по-видимому, связано с удалением от барьера минимума самосогласованной потенциальной ямы (см. рис. 2а) и накопленного в ней пространственного заряда.

Сопоставление численных значений для интерфейсной емкости $C_i = 2C$, которые получаются из кривой 1 рис. 5, с ее оценками из экспериментальных данных, например, полученных в работе [25], дают совпадения по порядку величины. Так, без учета диэлектрической проницаемости металла электродов, которая заметно отличается от единицы для благородных [26] и переходных [27] металлов, диапазон изменения C_i согласно кривой 1 рис. 5 составляет от 275 до 85 фФ/мкм² при $0.5 \leq R_s \leq 3.5$. Если характеризовать емкость интерфейса эффективной электрической толщиной $d_{eff} = (4\pi C_i)^{-1}$, то получим диапазон изменения от 0.035 до 0.11 нм. Анализ современных теоретических и экспериментальных данных в работе [28] закончился выводом в пользу именно несовершенства экранирования в металлических контактах как основной причины обнаруживаемого экспериментально так называемого мертвого слоя («dead layer»). Последний определяет максимально достижимую емкость в структурах с сегнетоэлектрическим изолятором при заданном выборе металла для электродов.

Общий вывод, который следует из данных, представленных на рис. 5, состоит в том, что приближение Томаса – Ферми с его экспоненциальным распределением пространственного заряда действительно дает по порядку величины те же значения для предельной емкости, что и полный самосогласованный расчет. Ни обменно-корреляционное взаимодействие электронов, ни осцилляции плотности, ни учет самоиндуцированной потенциальной ямы и дискретного уровня в ней не оказывают существенными. Причем появление уровня при $R_s < 1.6$ не сказалось на плавности хода кривой $C(R_s)$, в согласии с результатом работы [29]. Аналогичное отсутствие резких изменений на экспериментальной кривой $C(V)$ при изменении числа уровней размерного квантования в обогащенном слое n -InAs отмечалось в работе [30].

В работе [15] сохранение непрерывной зависимости самосогласованного потенциала ямы от внешнего электрического поля независимо от появления/исчезновения связанного состояния объясня-

лось формированием «ортогональной дырки» в пространственном распределении электронов, заполняющих состояния непрерывного спектра. Полученные нами численные результаты показывают, что это явление обусловлено появлением или исчезновением дополнительного нуля у волновых функций непрерывного спектра при появлении или исчезновении связанного состояния в потенциальной яме. В результате заполнение связанного состояния сопровождается в некотором смысле перераспределением полной электронной плотности между двумя ее составляющими в формуле (3.2), а не добавлением вклада от уровня к имевшемуся от непрерывного спектра. Наоборот, неучет (или пропуск) вклада связанного состояния, даже достаточно мелкого, в процессе итераций приводит к формированию мелкой и широкой потенциальной ямы, в которой на следующем итерационном цикле появляется уже несколько уровней, так что итерации перестают сходиться.

4. РАБОТА ВЫХОДА И ПЛОТНОСТЬ ПОВЕРХНОСТНОЙ ЭНЕРГИИ МЕТАЛЛА В МОДЕЛИ ОДНОРОДНОГО ФОНА

В модели однородного положительного фона (модель желе) работа выхода из металла и плотность поверхностной энергии рассчитывались в многочисленных работах. Однако имеющиеся результаты обнаруживают подчас не только количественные, но и качественные различия. Нам представляется, что причина этих расхождений связана с тем, что в упомянутых работах не достигалось полного самосогласования при отыскании распределения электронной плотности и эффективного потенциала. К достаточно обстоятельным и содержащим описание методики расчета можно отнести статьи [3, 31, 32], с которыми мы и будем сопоставлять свои результаты. Отметим, что целью этой части работы было получить действительно самосогласованное решение для одной из классических задач теории неоднородного электронного газа и оценить важность самосогласованности путем сравнения с результатами, полученными другими методами в той же постановке задачи. Мы не ставили цели улучшить модель желе или теорию функционала плотности для расчета поверхностных свойств металлов.

Постановка задачи здесь в основном аналогична расчету свойств барьерной структуры, рассмотренному в предыдущем разделе. Однако в случае металлов поверхностный барьер формируется целиком самосогласованным полем и не содержит заданно-

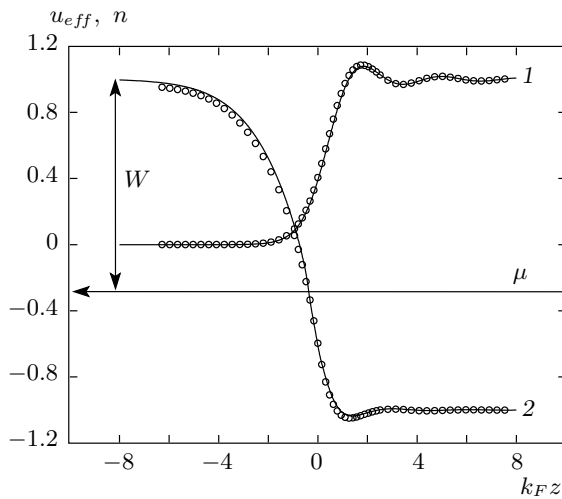


Рис. 6. Сравнение результата самосогласованного расчета настоящей работы для плотности электронов (кривая 1) и эффективного потенциала (кривая 2) с аналогичными данными из работы [3] (точки) при $R_s = 4$

го затравочного потенциала. Кроме того, для металлов в обычных условиях значения R_s , вычисленные для носителей тока с массой свободного электрона в среде с диэлектрической проницаемостью $\kappa_m = 1$, значительно превышают значения R_s , типичные для полупроводниковых структур, и лежат в диапазоне $2 < R_s < 5.7$. Все перечисленное обусловило интерес к расчету для выяснения характера сходимости к самосогласованному решению и для оценки устойчивости алгоритма при бóльших, чем в разд. 3, значениях R_s .

4.1. Работа выхода простых металлов

Пусть положительный фон занимает полупространство $z > z_+$, где z_+ — координата границы металл–вакуум. При расчете краевые условия к уравнению Пуассона (2.23) брались в виде $du/d\zeta|_{\zeta=\pm\infty} = 0$. Волновые функции $\psi_k(\zeta)$ одночастичных состояний с энергией, не превышающей уровень Ферми, находились как решения задачи Коши, экспоненциально затухающие при $\zeta \rightarrow -\infty$ и нормированные согласно (2.33) и (2.35). Безразмерная работа выхода $w = W/\varepsilon_F^0$ (см. рис. 6) равна высоте потенциального барьера, отсчитанной от значения объемного химического потенциала,

$$w = u_{eff}(-\infty) - \mu.$$

Самосогласованные плотность электронов и эффективный потенциал, полученные после сходимости итерационного процесса согласно описанной в разд. 2 схеме расчета, показаны на рис. 6. Для сравнения на этом же рисунке точками представлены зависимости n и u_{eff} , взятые из табл. 1 работы [3]. Следует отметить, что в работе [3] для корреляционной энергии была принята стандартная формула Вигнера

$$\varepsilon_c = -\frac{0.44}{R_s + 7.8}, \quad (4.1)$$

а не выражение (2.9), используемое в настоящей работе. С целью корректности сравнения представленный на рис. 6 результат нашего самосогласованного решения был вычислен с использованием формулы (4.1) для корреляционной энергии. Как видно из рис. 6, результаты расчетов неплохо согласуются между собой, хотя в работе [3] отыскание самосогласованного решения было заменено использованием 12-параметрической формулы для распределения электронов и определением этих параметров посредством некоторой вариационной процедуры (см. Приложение В в [3]).

Однако сходимость нашего итерационного решения и хорошее согласие с общепризнанным расчетом из работы Ланга и Кона [3] при частном значении R_s сами по себе еще не достаточны для полной оценки достигнутой степени самосогласованности. В работе [33] в рамках модели однородного фона было получено универсальное соотношение для электростатического потенциала:

$$U(z_+) - U(\infty) = N \frac{d\varepsilon}{dN} \equiv \tilde{\Delta}_{BV}, \quad (4.2)$$

где ε — полная энергия, приходящаяся на один электрон в однородном электронном газе. Для дальнейшего рассмотрения удобно перейти к безразмерному параметру самосогласованности $\Delta_{BV} = \tilde{\Delta}_{BV}/\varepsilon_F^0$ и записать его в виде

$$\Delta_{BV} = \frac{2}{5} + 2 \left(\frac{4}{9\pi} \right)^{2/3} R_s^2 [U_{xc} - \varepsilon_{xc}], \quad (4.3)$$

где выражение в квадратных скобках оставлено в атомных единицах. Соотношение (4.3) было использовано для проверки, является ли рассчитанный нами потенциал действительно самосогласованным с электронной плотностью. Хорошее совпадение $\Delta = u(\zeta_+) - u(\infty)$ и Δ_{BV} доказывает, что найдено действительно самосогласованное решение уравнений Кона–Шэма (см. [33, с. 1430]). В таблице

Таблица. Зависимости параметров Δ и Δ_{BV} , работы выхода W и плотности поверхностной энергии Σ от R_s , а также число итераций, потребовавшихся для достижения самосогласованности

R_s	Δ	Δ_{BV}	W , эВ	Σ , эрг/см ²	N_{iter}
0.3	0.3751	0.3751	3.12	—	12
0.5	0.358	0.358	3.21	—	8
1.0	0.317	0.317	3.40	—	6
1.3	0.291	0.291	3.51	—	6
1.5	0.274	0.274	3.60	—	7
1.65	0.261	0.261	3.63	—	7
1.8	0.248	0.248	3.64	—	7
2.07	0.224	0.224	3.60	-1340	9
2.3	0.204	0.204	3.53	-583	8
2.5	0.186	0.186	3.48	-321	12
3.28	0.115	0.115	3.12	106	9
3.99	0.051	0.051	2.87	71	17
4.96	-0.047	-0.048	2.50	65	28

приведены значения Δ и параметра Δ_{BV} , отвечающие результатам самосогласованного расчета для различных R_s с ε_c , определяемой формулой (2.9). Как видно, Δ и Δ_{BV} с точностью лучше 0.001 совпадают для всех $R_s \leq 3.99$. В отличие от работы [32], где критерий (4.2) явным образом использовался в процессе построения решения, в реализованном нами алгоритме это соотношение нигде не учитывается. Поэтому его выполнение доказывает, что полученные величины n и u_{eff} действительно являются самосогласованными решениями уравнений Кона–Шэма для принятой аппроксимации обменно-корреляционного потенциала электронов.

Стоит отметить чувствительность критерия самосогласованности к значениям потенциала на границе положительного фона. Представленная в таблице степень согласия значений Δ и Δ_{BV} достигается, только если в качестве значений потенциала $u(\zeta_+)$ брать значение, интерполированное на полшага дискретизации $\delta\zeta$ в сторону вакуума. Обычно при расчетах выбиралось $\delta\zeta$ в пределах 0.05–0.005.

Аналогичный тест на самосогласованность для плотности электронов и потенциала, полученных в работе Ланга и Кона [3], был выполнен в статье [33]. Представленные в таблице работы [33] данные показывают, что даже для уточненных (неопубликованных) результатов Ланга критерий самосогласованно-

сти (4.2) выполняется с точностью лучше третьего знака после запятой только до значений $R_s \leq 4$. Затем согласие прогрессивно ухудшается, и при $R_s = 6$ различие достигает 0.01. Это указывает, что при больших значениях R_s , близких к границе области существования реальных металлов, принятая в работе [3] схема параметризации электронной плотности и метод определения этих параметров недостаточно хорошо заменяют точное решение системы уравнений Кона–Шэма. Этот факт можно рассматривать как признак серьезности отличия поведения электронной плотности и потенциала при значениях R_s , близких к критическому R_{sc} , от соответствующего поведения решений модели желе при более высокой плотности электронного газа, что в какой-то мере обсуждалось в разд. 2.

Из таблицы следует, что и в наших расчетах при $R_s \approx 5$ критерий самосогласованности выполняется хуже, чем при меньших значениях. В области значений $R_s \leq 3.99$ итерационный процесс программно останавливался, если определенная в (2.20) невязка становилась меньше заданной величины $\delta = 10^{-5}$. Однако при $R_s = 4.96$ описанный в п. 2.2 алгоритм автоматического определения критической точки ζ_c не обеспечивал сходимости процесса решения уравнения Пуассона из-за монотонного сдвига критической точки ζ_c от поверхности вглубь металла. Для получения определенного результата в этом случае после некоторого числа начальных циклов итераций пришлось прибегнуть к отключению модуля, который переопределяет ζ_c . Однако после такого принудительного выбора положения критической точки скачок электронной плотности в ее окрестности не уменьшался в процессе итераций ниже $3 \cdot 10^{-4}$. При $R_s > 5$ добиться сходимости итерационного решения вообще не удалось. По-видимому, наиболее принципиальным с точки зрения анализа работы алгоритма здесь является то, что при $R_s \rightarrow R_{sc}$ эффективный потенциал (2.2) оказывается почти равным локальному обменно-корреляционному, и вклад действующего кулоновского потенциала в формирование поверхностного барьера становится практически несущественным.

В таблице приведены также вычисленные значения работы выхода W как функции R_s . На рис. 7 эта зависимость (кривая 3) сравнивается с аналогичными зависимостями, полученными в работах [31, 32]. Рассчитанные нами значения работы выхода находятся в достаточно разумном согласии с экспериментальными данными для металлов первой группы (Li ($R_s = 3.28$), Na ($R_s = 3.99$), K ($R_s = 4.96$)) и лежат к ним ближе, чем точки кривой 1, полученные

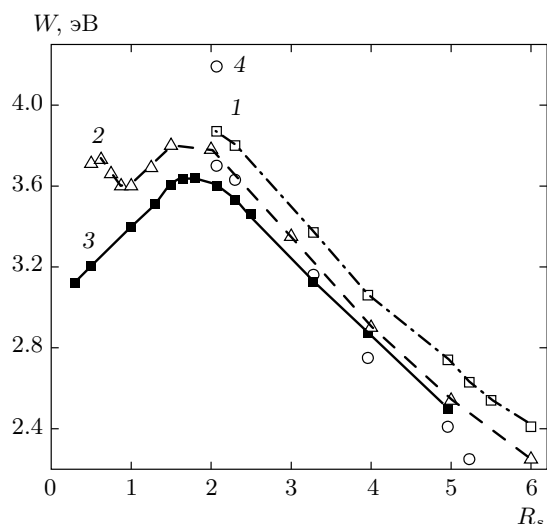


Рис. 7. Зависимости от R_s работы выхода в модели однородного фона: результаты расчетов [31] (1), [32] (2) и настоящей работы (3). Кружками (4) показаны экспериментальные значения работы выхода для поликристаллических образцов Al, Li, Na, K и Rb, взятые из табл. 3 обзора [34]. Меньшее из двух значений W при $R_s = 2.07$ (Al) измерено на поликристаллическом Al в работе [35]

в работе [31]. Скорее всего, это улучшение количественного согласия расчета с измерениями является следствием использования нами для корреляционной энергии измененной формулы Вигнера (2.9), так как при одинаковом выборе выражения для ϵ_c результаты расчетов потенциала в обеих работах близки, как показывает рис. 6, по крайней мере для $R_s = 4$.

При $R_s > 1$ имеется полуколичественное соответствие наших результатов для работы выхода с полученными в работе [32] (рис. 7, кривая 2), которое улучшается с ростом R_s . Как видно на рис. 7, обе кривые имеют максимум при $R_s \approx 2$, однако при малых R_s появляется качественное отличие. Согласно [32], работа выхода имеет локальный минимум при $R_s \approx 1$, который не подтвердился нашим расчетом. Этот локальный минимум в работе выхода, возможно, связан с несамосогласованностью полученного в [32] решения уравнений Кона–Шэма для малых значений R_s .

Использованный в работе [32] метод решения, разработанный и описанный детально в [36], включает в себя замену уравнения Пуассона интегральным уравнением для потенциала в некоторой конечной окрестности поверхности металла с граничными условиями, которые не фиксировали полный за-

ряд (см. [36, Appendix A]). Вместе с отмеченными в разд. 1 проблемами при замене уравнения Пуассона интегральным такие граничные условия не могли обеспечить автоматической самосогласованности окончательного результата итерационной процедуры, хотя это предполагается в [32] (см. текст после формулы (7)). Как указано в работе [32], при $R_s < 0.5$ метод вообще перестал сходиться. Причина этого, вероятно, в том, что с ростом плотности электронов возрастает роль кулоновского потенциала относительно обменно-корреляционного и чувствительность к несамосогласованности.

Отметим еще раз, что по физической постановке и формулировке исходных уравнений работа [32] и наша идентичны, за исключением различия в использованных формулах для корреляционной энергии. В то же время в наших вычислениях никаких принципиальных трудностей со сходимостью в области малых R_s не возникло. Ограничивали только рост требований к объему оперативной памяти и увеличение времени счета из-за расширения области вакуума, где плотность электронов отлична от нуля. Что касается выявившихся сложностей расчета в области больших R_s , они требуют дальнейшего анализа и, возможно, частично связаны с использованием приближения локальной плотности для обменно-корреляционного потенциала. С одной стороны, основания для такого предположения могут быть найдены в дискуссии о границах применимости локального приближения для обменно-корреляционного взаимодействия электронов (см., например, [37, 38]). С другой, попытка непосредственного применения итерационной процедуры к уравнению Кона–Шэма с потенциалом U_{xc} в отсутствие кулоновского вклада в эффективный потенциал показала, что сходимость итераций отсутствует.

4.2. Поверхностная энергия

Полученные нами самосогласованные распределения электронной плотности, потенциала и одночастичные волновые функции позволяют вычислить поверхностную энергию как функцию R_s . Однако имеющиеся в литературе формулы для расчета поверхностной энергии используют конечность размеров области, занятой металлом. С одной стороны, это позволяет определить поверхностную энергию как разность полных энергий двух состояний образца — до и после разделения его на две части. С другой — влечет за собой квазидискретность спектра собственных значений с необходимостью акку-

ратной асимптотической оценки зависимости различных составляющих полной энергии от размеров образца при стремлении последних к бесконечности (см., например, [39]). Для расчета свойств неограниченного образца надо иметь выражения для плотностей соответствующих составляющих полной энергии. В стандартной формуле теории функционала плотности для полной энергии ограниченного образца

$$E_{tot}(N(\mathbf{r})) = T_s(N(\mathbf{r})) + \int d\mathbf{r} \varepsilon_{xc}[N(\mathbf{r})]N(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{(N(\mathbf{r}) - N_+)(N(\mathbf{r}') - N_+)}{|\mathbf{r} - \mathbf{r}'|}, \quad (4.4)$$

где интегрирование понимается по его объему, второй и третий члены, описывающие обменно-корреляционную и электростатическую энергии многоэлектронной системы, уже содержат определение плотности соответствующих составляющих. Что касается полной кинетической энергии T_s газа невзаимодействующих электронов, то она таким свойством не обладает, так как обычно ее вычисляют, используя полученную в работе [24] формулу, которая выражает сумму одночастичных энергий, меньших энергии Ферми, через интеграл по энергетическому спектру от фазы волновых функций непрерывного спектра с асимптотикой (2.35). Следуя работе [1] (см. формулу (50)), определим плотность кинетической энергии для случая нулевой температуры как

$$t_s(\mathbf{r}) = 2 \int_{\varepsilon \leq \varepsilon_F} \mathcal{D}\{\varepsilon\} \varepsilon |\Psi_\varepsilon(\mathbf{r})|^2 - U_{eff}(\mathbf{r})N(\mathbf{r}). \quad (4.5)$$

Используя его, можно определить плотность полной энергии,

$$\epsilon_{tot}(\mathbf{r}) = t_s(\mathbf{r}) + \varepsilon_{xc}[N(\mathbf{r})]N(\mathbf{r}) + \frac{1}{2}U(\mathbf{r})(N(\mathbf{r}) - N_+(\mathbf{r})), \quad (4.6)$$

и соответственно получить выражение для полной энергии в объеме V , выделенном в бесконечном образце,

$$E_{tot}(V) = \int_V d\mathbf{r} \epsilon_{tot}(\mathbf{r}), \quad (4.7)$$

которое может использоваться для расчета поверхностной энергии.

Рассчитанная в работе [3] в модели желе поверхностная энергия становится отрицательной при $R_s < 2.5$. Это означает, что в такой модели металлы с большой плотностью электронов проводимости

($2 < R_s < 2.5$) не могли бы существовать. Наиболее разительно полученные в работе [3] значения поверхностной энергии отличаются от измеренных значений для многовалентных металлов. Но даже для одновалентных металлов первой группы наблюдается нарастающее ухудшение согласия расчета и измерений с уменьшением R_s . Основной причиной этого, как было показано, является пренебрежение в модели желе дискретной структурой положительного фона, который на самом деле представляет собой кристаллическую решетку ионных остовов (см. обсуждение, например, в [34]). Тем не менее представляло интерес выяснить, повлияет ли использование самосогласованного решения на зависимость поверхностной энергии от R_s .

Расчет поверхностной энергии проводился с использованием двух формул. Одна из них основана на строго термодинамическом определении поверхностной энергии как той части полной энергии однородной по объему системы, которая не зависит от ее объема V , но зависит от площади поверхности S . Согласно, например, определению (3.82) из работы [40], можно записать

$$E_{tot} = S\Sigma + V\epsilon_{tot}, \quad (4.8)$$

где Σ — плотность поверхностной энергии. В нашем случае полубесконечного образца можно изменять объем рассматриваемой части системы, удаляясь от поверхности вглубь. Определим энергию части системы единичной площади от поверхности до координаты z формулой

$$E_{tot}(z) = \int_{-\infty}^z dz \epsilon_{tot}(z). \quad (4.9)$$

Поскольку в глубине объема плотность энергии становится постоянной вместе со всеми определяющими ее величинами, при достаточно больших z величина $E_{tot}(z)$ должна стать линейной функцией от координаты. Продолжив эту линейную зависимость обратно до границы положительного фона, мы получим конечное, не зависящее от объема значение, которое по определению даст плотность поверхностной энергии. Сказанное можно выразить простой формулой:

$$\Sigma = \lim_{z \rightarrow \infty} [E_{tot}(z) - \epsilon_{tot}(\infty)z]. \quad (4.10)$$

Здесь предполагается, что граница положительного фона находится при $z = 0$.

Такое определение поверхностной энергии является строгим и, как полагается для термодинамической характеристики макроскопической системы,

не зависит от процесса образования поверхности, который иногда моделируют для определения поверхностной энергии. Однако оно требует повышенной точности счета при решении дифференциальных уравнений и вычислении интегралов, так как формула содержит разность двух больших чисел. Эта трудность устраняется, если переписать соотношение (4.10) в виде

$$\Sigma = \int_{-\infty}^0 dz \epsilon_{tot}(z) + \int_0^{\infty} dz (\epsilon_{tot}(z) - \epsilon_{tot}(\infty)). \quad (4.11)$$

При такой формуле для поверхностной энергии результат оказывается менее чувствительным к точности вычислений, так как интегрируются выражения, которые стремятся к нулю при удалении от поверхности.

В литературе обычно используется другое определение поверхностной энергии как разности полной энергии полубесконечного образца и полной энергии половины бесконечного образца (см., например, [34]). Для удобства сравнения представим, как в работе [3], полную плотность поверхностной энергии в виде суммы трех составляющих — кинетической, электростатической и обменно-корреляционной энергий:

$$\Sigma = \Sigma_s + \Sigma_{es} + \Sigma_{xc}. \quad (4.12)$$

В безразмерном виде после интегрирования в формуле (4.5) по волновому вектору, параллельному поверхности, компоненты плотности поверхностной энергии записываются как

$$\begin{aligned} \sigma_s = & \frac{1}{2\pi^2} \int_{-\infty}^{\infty} d\zeta \int_0^1 dk (1 - k^4) \left(|\psi_k(\zeta)|^2 - \frac{1}{2} \theta(\zeta) \right) - \\ & - \frac{1}{3\pi^2} \int_{-\infty}^{\infty} d\zeta \{ (u(\zeta) + u_{xc}[n(\zeta)]) n(\zeta) - \\ & - u_{xc}[1] \theta(\zeta) \}, \quad (4.13) \end{aligned}$$

$$\begin{aligned} \sigma_{xc} = & \frac{1}{3\pi^2} \times \\ & \times \int_{-\infty}^{\infty} d\zeta \{ \varepsilon_{xc}[n(\zeta)] n(\zeta) - \varepsilon_{xc}[1] \theta(\zeta) \}, \quad (4.14) \end{aligned}$$

$$\sigma_{es} = \frac{1}{6\pi^2} \int_{-\infty}^{\infty} d\zeta \{ n(\zeta) - 1 \cdot \theta(\zeta) \} u(\zeta). \quad (4.15)$$

Сравнение формул (4.13)–(4.15) с формулами (45а)–(45в) работы [34] показывает, что различаются только выражения для вклада кинетической энергии σ_s . Используя формулы (4.5) и (4.11), удалось избежать дополнительного усложнения расчета, исключив интегралы по энергии от фазы (2.37) волновых функций (2.35). Результаты расчета поверхностной энергии приведены в таблице. Видно, что отрицательные значения поверхностной энергии появились даже при чуть больших значениях R_s , чем в работе [3]. Основная причина этого — в меньшей величине корреляционной энергии, даваемой формулой (2.9). Общий вывод о недостаточности модели желе для описания поверхностной энергии в области существования реальных металлов остался в силе.

5. ВЫЧИСЛЕНИЕ САМОСОГЛАСОВАННОГО ПОТЕНЦИАЛА СТРУКТУРЫ С БАРЬЕРОМ ШОТКИ

До сих пор нами рассматривались полубесконечные системы, в которых волновые функции электронов были чисто вещественные, и, как следствие, не проводящие электрический ток. В данном разделе рассматривается структура, одночастичные волновые функции которой ограничены и имеют осциллирующее поведение в обеих бесконечно удаленных областях. Такие граничные условия для решений уравнения Шредингера допускают комплексные собственные функции и состояния системы с не равным нулю стационарным током. Чтобы избежать излишней трудоемкости в работе с модельной задачей, возьмем в качестве такой системы структуру типа контакт металл–полупроводник с барьером Шоттки. Из-за намного большей концентрации свободных носителей в металле практически весь потенциальный барьер формируется в полупроводнике. Это позволяет свести вычислительную часть задачи к расчету волновых функций и потенциала в полупространстве и применить к такой структуре без существенной модификации алгоритм итерационного поиска самосогласованных электронной плотности и эффективного потенциального барьера, описанный в п. 2.2.

Ограничимся равновесной ситуацией, когда на структуру не наложено напряжение смещения и ток равен нулю. Пусть металл занимает область $z < 0$, вырожденный полупроводник n -типа с концентрацией ионизованных доноров N_D расположен при $z > 0$. Для определенности будем иметь в виду случай, когда большая плотность состояний на границе разде-

ла металл/полупроводник фиксирует на интерфейсе положение дна зоны проводимости полупроводника при энергии Φ_s относительно уровня Ферми. Плотность электронов под барьером вблизи интерфейса в полупроводнике мала, поэтому эффективный потенциал совпадает с кулоновским, и тем самым определяется одно граничное условие для электростатического потенциала. Второе граничное условие для решения уравнения Пуассона в области полупроводника задается требованием ограниченности потенциала на бесконечности, что одновременно означает равенство там нулю электрического поля. Выбрав значение потенциала в глубине полупроводника в качестве нуля отсчета энергии, имеем следующие граничные условия для решения уравнения Пуассона в полупроводнике:

$$U(0) \equiv U_0 = \Phi_s + \varepsilon_F, \quad \left. \frac{dU}{dz} \right|_{z=\infty} = 0. \quad (5.1)$$

Как обычно принято, будем пренебрегать проникновением электрического поля в металл. Поэтому электростатический потенциал в металле принимается постоянным и равным его значению на интерфейсе.

При описании одночастичных состояний непрерывного спектра в такой одномерно неоднородной структуре воспользуемся результатами анализа гильбертова пространства, связанного с гамильтонианом неограниченной системы, который был проделан в работе [2]. Выберем в качестве базиса волновые функции правых Ψ^R и левых Ψ^L состояний:

$$\Psi_{k, \mathbf{k}_{\parallel}}^R(z, \mathbf{r}_{\parallel}) = \frac{1}{2\pi} C_k^R \psi_k^R(z) \exp(i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}), \quad (5.2)$$

$$\Psi_{q, \mathbf{k}_{\parallel}}^L(z, \mathbf{r}_{\parallel}) = \frac{1}{2\pi} C_q^L \psi_q^L(z) \exp(i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}), \quad (5.3)$$

которые однозначно определяются своим асимптотическим поведением на бесконечности:

$$\psi_k^R(z) = \begin{cases} e^{-ikz} + r_k^R e^{ikz}, & z \rightarrow \infty, \\ t_k^R e^{-iqz}, & z \rightarrow -\infty, \end{cases} \quad (5.4)$$

и

$$\psi_q^L(z) = \begin{cases} e^{iqz} + r_q^L e^{-iqz}, & z \rightarrow -\infty, \\ t_q^L e^{ikz}, & z \rightarrow \infty. \end{cases} \quad (5.5)$$

Использование двух квантовых чисел k и q для описания асимптотического поведения собственных функций в правой и левой бесконечных областях связано с несимметричностью рассматриваемой барьерной структуры даже в отсутствие смещения из-за различия в положении дна зоны проводимости относительно уровня Ферми в металле и полупроводнике. Впрочем, даже симметричная структура

становится несимметричной при наложении смещения, так что различие этих квантовых чисел скорее правило, чем исключение. Связь между k и q находится из условия, что оба характеризуют собственную функцию, принадлежащую одному собственному значению ε гамильтониана:

$$\varepsilon - U_{xc}(\infty) = \frac{q^2 + \mathbf{k}_{\parallel}^2}{2m_M} + U_M = \frac{k^2 + \mathbf{k}_{\parallel}^2}{2m_S}. \quad (5.6)$$

Здесь U_M характеризует положение дна зоны проводимости металла относительно дна зоны проводимости полупроводника, индексами «M» и «S» обозначены эффективные массы свободных носителей соответственно в металле и полупроводнике. Правое равенство в (5.6) определяет q как неявную функцию k (и наоборот). Нормировочные константы C_q^L и C_k^R равны $1/\sqrt{2\pi}$, если ψ_q^L нормированы на $\delta(q-q')$ и ψ_k^R нормированы на $\delta(k-k')$. Доказательство взаимной ортогональности пары функций $\Psi_{\varepsilon}^{R,L}$ при одном значении энергии ε , которое необходимо, чтобы иметь полный ортонормированный базис в гильбертовом пространстве задачи, приведено в работе [2].

Однако в рассматриваемой структуре все результаты могут быть выражены через величины, относящиеся к полупроводниковому электроду контакта металл-полупроводник, так что удобнее ψ_q^L также нормировать на $\delta(k-k')$. В этом случае для нормировочной константы левых состояний надо использовать формулу

$$|C_k^L|^2 = \frac{1}{2\pi} \frac{\partial q}{\partial k}. \quad (5.7)$$

В задачах, где не нужно самосогласование, обычно используют нормировку волновых функций непрерывного спектра на единичную амплитуду падающей волны. Этого достаточно, чтобы вычислить туннельную прозрачность барьера заданной формы. Однако для самосогласованного решения задачи при учете кулоновского взаимодействия носителей заряда надо находить распределение электронов в пространстве, что нельзя сделать, не зная собственных функций непрерывного спектра, нормированных на дельта-функцию от квантовых чисел.

Приведем сводку соотношений между коэффициентами отражения r и прозрачности t барьера, которые необходимы для проведения расчета:

$$\begin{aligned} \frac{1}{\partial q / \partial k} |t_k^R|^2 &= 1 - |r_k^R|^2, \\ \frac{1}{\partial k / \partial q} |t_q^L|^2 &= 1 - |r_q^L|^2, \end{aligned} \quad (5.8)$$

$$t_{q(k)}^L = \frac{1}{\partial q / \partial k} t_k^R, \quad r_{q(k)}^L = -r_k^{R*} \frac{t_k^R}{t_k^{R*}}. \quad (5.9)$$

Формулы (5.9) выписаны для случая нормировки левых состояний на $\delta(k - k')$, что отражено указанием явной зависимости $q(k)$ в аргументах.

Волновая функция электронов в полупроводнике может быть представлена как суперпозиция двух вещественных, линейно независимых решений ϕ_1 и ϕ_2 , которые при $z \rightarrow \infty$ имеют следующее асимптотическое поведение:

$$\begin{aligned} \phi_{k1}(z) &= \sin(kz + \gamma_k), \\ \phi_{k2}(z) &= \cos(kz + \gamma_k), \end{aligned} \quad (5.10)$$

где γ_k — фаза волновой функции k -го состояния. Из этих функций, используя граничные условия и нормировку, можно построить функцию $\psi_k^R(z)$, которой, в свою очередь, достаточно, чтобы определить $\psi_q^L(z)$.

Решения ϕ_1 и ϕ_2 находятся численным интегрированием задачи Коши для уравнения Шредингера (2.34). С изменением аргумента z от интерфейса вглубь полупроводника функции ϕ_1 экспоненциально растут в области под барьером, а ϕ_2 — убывают. При достаточно больших значениях высоты потенциального барьера нахождение экспоненциально убывающего решения сопряжено с вычислительными трудностями. Для их преодоления применялся следующий прием: функция ϕ_1 находится из решения задачи Коши с начальными условиями

$$\phi_{k1}(0) = 1, \quad \phi'_{k1}(0) = \sqrt{U_0 - k^2/2},$$

заданными на границе металл–полупроводник $z = 0$. Интегрирование уравнения (2.34) ведется от точки $z = 0$ до точки z_∞ , которая находится в глубине полупроводника, где с достаточной точностью эффективный потенциал уже постоянен. Полученное решение нормируется на единичную амплитуду формулой, аналогичной (2.36), и с помощью соотношения (2.37) находится фаза γ , входящая в выражения (5.10). Следует отметить, что выбор начальных условий для ϕ_1 достаточно произволен, так как процедура нормировки обеспечивает правильное асимптотическое поведение решений в бесконечности. Для нахождения функции ϕ_2 также решается задача Коши, но начальные условия ставятся уже в точке $z = z_\infty$:

$$\begin{aligned} \phi_{k2}(z_\infty) &= \cos(kz_\infty + \gamma_k), \\ \phi'_{k2}(z_\infty) &= -k\phi_{k1}(z_\infty), \end{aligned} \quad (5.11)$$

и интегрирование ведется из глубины полупроводника к границе металла в направлении роста $\phi_{k2}(z)$ в области барьера, что обеспечивает большую устойчивость процедуры. Критерием достаточной точности полученной пары решений является постоянство вронскиана $W(\phi_{k1}, \phi_{k2})$ на отрезке $[0, z_\infty]$.

В дальнейшем будем предполагать, что эффективные массы электронов в полупроводнике и металле равны. Тогда на границе двух сред будет непрерывна как сама волновая функция, так и ее первая производная. Безусловно, такое предположение является достаточно модельным, но сложный вопрос о выводе условий сопряжения для огибающих волновых функций на границе металл–полупроводник не является целью настоящей работы. Другой вид граничного условия на интерфейсе, конечно, может повлиять на конкретные численные результаты, однако вряд ли это потребует изменения предлагаемой итеративной схемы решения уравнений Кона – Шэма для структур с токовыми состояниями. Основания для такого мнения будут обсуждены в конце раздела. Отметим еще, что проблемы, близкие к рассматриваемым здесь вопросам, существуют также в теории резонансно-туннельных диодов (см., например, [41]).

5.1. Некоторые детали расчета

Итак, из двух функций (5.10) надо образовать волновую функцию $\psi_k^R(z)$ «правого» состояния, которая описывает туннелирование электронов из правого электрода (полупроводник) в левый (его по-прежнему для удобства будем называть «металл»). Асимптотические граничные условия (5.4), которыми определяется $\psi_k^R(z)$, отражают требование иметь в металле чисто уходящую волну и представлять в полупроводнике линейную комбинацию из падающей и отраженной волн. Эти условия могут быть использованы, чтобы найти коэффициенты в разложении

$$\psi_k^R(z) = \alpha_k \phi_{k1}(z) + \beta_k \phi_{k2}(z). \quad (5.12)$$

Если ψ_k^R найдена, то ψ_k^L в общем виде может быть представлена линейной комбинацией двух независимых решений ψ_k^R и ψ_k^{R*} . Но для расчета плотности электронов в полупроводнике, которая только и нужна в этой задаче, нам надо знать $\psi_k^L(z)$ только при $z \geq 0$. Поэтому для выражения ψ_k^L через $\phi_{k1,2}$ достаточно воспользоваться асимптотикой (5.10) левого состояния при $z \rightarrow \infty$. С учетом сказанного легко получаем, что

$$\psi_{q(k)}^L(z) = t_{q(k)}^L \exp(-i\gamma_k)(i\phi_{k1}(z) + \phi_{k2}(z)). \quad (5.13)$$

Задача определения $\psi_k^R(z)$ сводится к определению четырех комплексных коэффициентов $\alpha_k, \beta_k, r_k^R, t_k^R$. Соответствующие четыре уравнения получаются путем приравнивания асимптотических формул для ψ_k^R и ее производной при $z \rightarrow \infty$ асимптотике искомой линейной комбинации, что дает два уравнения, а также из двух условий сопряжения для функции и ее производной на границе при $z = 0$. При этом в металле принимаем

$$\psi_k^R(kz) = t_k^R \exp(-iq(k)z). \quad (5.14)$$

Расчет проводится при условии $U_M < 0$, что обеспечивает положение дна зоны проводимости в левом электроде ниже дна в правом и сохраняет неравенство $q(\varepsilon, \mathbf{k}_{\parallel}) \neq k(\varepsilon, \mathbf{k}_{\parallel})$ несмотря на принятое равенство эффективных масс носителей в обоих электродах. В то же время условие равенства эффективных масс обеспечивает независимость функции $q(k)$ от \mathbf{k}_{\parallel} , что сокращает объем вычислений.

После несложных преобразований искомая система линейных алгебраических уравнений принимает вид

$$\begin{aligned} \alpha\phi_{k1}(0) + \beta\phi_{k2}(0) - t_k &= 0, \\ \alpha\phi'_{k1}(0) + \beta\phi'_{k2}(0) + iqt_k &= 0, \\ i\alpha + \beta &= 2 \exp(i\gamma_k), \\ i\alpha - \beta + 2r_k \exp(-i\gamma_k) &= 0. \end{aligned} \quad (5.15)$$

Здесь для краткости везде опущены индекс «R» правого состояния, индекс «k» волнового числа у коэффициентов α, β , и подразумевается, что $q = q(k)$.

Решение системы (5.15) можно представить в виде

$$r_k = -\frac{(q\Phi_k + i\Phi'_k)^*}{q\Phi_k + i\Phi'_k}, \quad (5.16)$$

$$t_k = \Phi_k^* + r_k\Phi_k, \quad (5.17)$$

$$\alpha_k = -i(\exp(i\gamma_k) - r_k \exp(-i\gamma_k)), \quad (5.18)$$

$$\beta_k = \exp(i\gamma_k) + r_k \exp(-i\gamma_k), \quad (5.19)$$

где для компактности записи введены функции

$$\Phi_k = \exp(-i\gamma_k)(\phi_{k2}(0) + i\phi_{k1}(0)), \quad (5.20)$$

$$\Phi'_k = \exp(-i\gamma_k)(\phi'_{k2}(0) + i\phi'_{k1}(0)) \quad (5.21)$$

и ϕ'_k означает производную по z .

Изложенная схема расчета волновых функций описывает изменения в процедуре решения уравнения Шредингера (2.17) для неограниченной системы по сравнению со случаем полуограниченной

(см. п. 2.2). Соответствующую формулу для расчета плотности электронов барьерной структуры, заменяющую (2.38), получим из общего выражения, выведенного в работе [2]. Приняв во внимание, что нужно знать плотность электронов только при $z \geq 0$, и отсутствие смещения, имеем

$$\begin{aligned} N(z) &= 2 \int_0^{\infty} dk \int_{-\infty}^{\infty} d\mathbf{k}_{\parallel} f(\varepsilon(k, \mathbf{k}_{\parallel})) \times \\ &\times \left[\left| \Psi_{q(k)}^L(z) \right|^2 + \left| \Psi_k^R(z) \right|^2 \right], \end{aligned} \quad (5.22)$$

где $f(\varepsilon)$ — функция распределения Ферми, а собственные значения энергии $\varepsilon(k, \mathbf{k}_{\parallel})$ заданы формулой (5.6). Положение уровня Ферми определяется условием нейтральности объема полупроводника при $z \rightarrow \infty$. Используя определения (5.2), (5.3), (5.7), (5.12), (5.13) и соотношения (5.8), (5.9) между амплитудными коэффициентами отражения и прозрачности барьера, получим из (5.22) выражение для плотности электронов через найденные численно два вещественных решения $\phi_{k1,2}(z)$:

$$\begin{aligned} N(z) &= \frac{2}{(2\pi)^3} \int_0^{\infty} dk \int_{-\infty}^{\infty} d\mathbf{k}_{\parallel} f(\varepsilon(k, \mathbf{k}_{\parallel})) \times \\ &\times \left[|\alpha_k\phi_{k1}(z) + \beta_k\phi_{k2}(z)|^2 + \left(1 - |r_k^R|^2\right) \times \right. \\ &\left. \times |\phi_{k1}(z) - i\phi_{k2}(z)|^2 \right]. \end{aligned} \quad (5.23)$$

Входящие сюда коэффициенты r_k^R, α_k и β_k выписаны в формулах (5.16)–(5.19).

Выражение (5.23) для плотности электронов дает явное представление формулы (2.18) итеративной процедуры для рассматриваемого случая.

5.2. Результаты и обсуждение

Расчеты проводились при значениях параметров, характерных для переходов Al/n-GaAs ($m^* = 0.07m_e$, $\Phi_s = 0.9$ эВ, $\kappa = 12.5$) с концентрацией ионизованных доноров $N_D = 10^{18}$ см⁻³, что соответствует $R_s = 0.66$. Для q , ввиду большой разницы энергий Ферми полупроводника и алюминия, оказалось возможным принять постоянное значение, которое было выбрано равным радиусу ферми-поверхности электронного газа при $R_s = 2.07$, $\kappa_m = 1$. На рис. 8 представлены зависимости от координат самосогласованного эффективного потенциала, рассчитанного с учетом

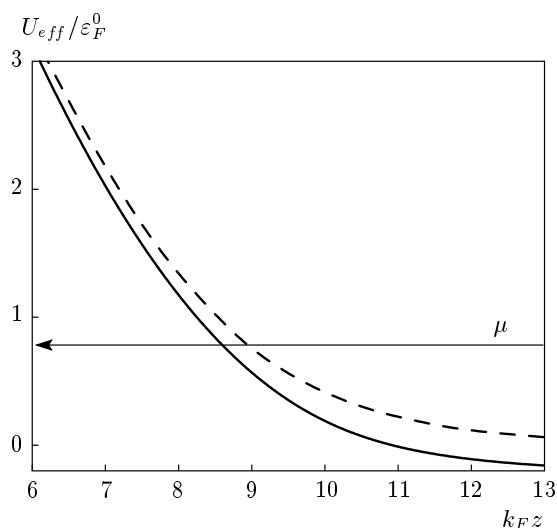


Рис. 8. Зависимости от координат самосогласованного потенциала для структуры с барьером Шоттки ($R_s = 0.66$), полученного с учетом обменно-корреляционного взаимодействия (сплошная линия) и в приближении Хартри (штриховая кривая). Отмечено положение уровня Ферми при учете U_{xc}

обменно-корреляционного взаимодействия электронов и в приближении Хартри, вблизи границы металл–полупроводник. Видно, что включение в расчет U_{xc} меняет форму потенциального барьера, делая изменение потенциала более крутым. Это различие заметно сказывается на виде теоретической вольт-амперной характеристики (ВАХ) туннельного контакта, особенно в области прямых смещений, когда электроны туннелируют из полупроводника в металл через область эффективного потенциального барьера, находящуюся ниже уровня Ферми (см. [10, 42, 43]). В этих же работах показано, что учет U_{xc} позволяет достичь более точного описания ВАХ реальных структур. Отсюда следует, что представление обменно-корреляционного взаимодействия электронов приближением локальной плотности остается вполне аккуратным для электронных состояний, расположенных достаточно глубоко под уровнем Ферми.

Кроме того, использование самосогласованного решения для потенциала в формуле для туннельного тока сделало возможным определение параметров Φ_s и N_D контакта металл–сильнолегированный n -GaAs по зависимости дифференциального сопротивления от напряжения смещения [44–46]. В указанных работах барьер Шоттки рассчитывался в приближении Томаса–Ферми и Томаса–Ферми–Дирака–Гомбаша. Хотя для таких структур хо-

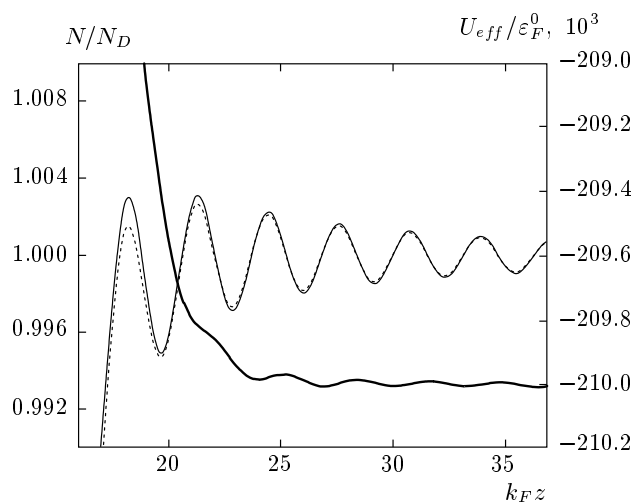


Рис. 9. Фриделевские осцилляции электронной плотности и эффективного потенциала в структуре с барьером Шоттки при $R_s = 0.66$. Электронная плотность с учетом обменно-корреляционного взаимодействия — сплошная кривая, в приближении Хартри — пунктирная; самосогласованный эффективный потенциал — жирная сплошная кривая

рошо выполняются критерии квазиклассичности барьера Шоттки $k_F L \gg 1$ и $k_F l_{TF} \gg 1$, где L и l_{TF} обозначают характерный масштаб изменения потенциала внутри и вне барьера [44], оставался вопрос, не произойдет ли заметного изменения формы барьера при выходе за рамки приближения Томаса–Ферми. Проведенный точный расчет показывает, что крупномасштабное поведение эффективного потенциала барьера Шоттки правильно передается потенциалом, найденным в приближении Томаса–Ферми, и согласие между ними улучшается с уменьшением R_s .

В глубине полупроводника на зависимости самосогласованного потенциала и плотности электронов от координаты проявляются осцилляции (рис. 9), которые отсутствуют в приближении Томаса–Ферми. В целом осцилляции, порожденные барьером Шоттки, менее выражены, чем в случае бесконечно высокой потенциальной стенки (ср. с рис. 4), что связано, очевидно, с квазиклассической плавностью первого. Особенно это относится к осцилляциям эффективного потенциала. Как видно из рис. 9, размах осцилляций эффективного потенциала очень мал, порядка и меньше 10^{-4} , тогда как осцилляции плотности электронов в точном решении, наоборот, более интенсивны, чем в приближении Хартри. Подавление осцилляций эффективного потенциала по сравнению с осцилляциями его кулоновской составляющей уже об-

суждалось в п. 3.1. Усиление осцилляций плотности в точном решении по сравнению с осцилляциями решения в приближении Хартри может быть понято из анализа линеаризованного выражения для индуцированной плотности.

Как известно, в приближении Хартри линеаризация формулы (2.7) при малых значениях потенциала U приводит в самосогласованном уравнении Пуассона (2.12) к выражению для приращения плотности (заряда), которое обычно записывают в виде $4\pi\delta N_{ind} = -k_{TF}^2 U$. Учет зависимости обменно-корреляционного потенциала от плотности электронов приводит к аналогичному соотношению $4\pi\delta N_{ind} = -k_{scr}^2 U$, где квадрат обратной длины экранирования имеет вид

$$k_{scr}^2 = \frac{k_{TF}^2}{1 + \frac{3}{2} \frac{du_{xc}(n)}{dn}}. \quad (5.24)$$

Поскольку безразмерный обменно-корреляционный потенциал растет по абсолютной величине с уменьшением концентрации электронов, производная $du_{xc}/dn < 0$ и с ростом R_s знаменатель в формуле (5.24) стремится к нулю, причем его равенство нулю является условием, определяющим критическое значение R_{sc} . Поэтому при сравнимой величине осцилляций самосогласованного кулоновского потенциала в точном решении и в приближении Хартри, что имеет место, осцилляции плотности электронов в точном решении будут усилены и тем больше, чем ближе R_s к R_{sc} . При этом будет сокращаться область, где возмущение уменьшается экспоненциально.

Суммируя результаты анализа осцилляций фриделевского типа, порождаемых нарушением однородности вырожденного электронного газа, отметим, что учет обменно-корреляционного взаимодействия электронов будет приводить к увеличению осцилляций плотности электронов и к уменьшению осцилляций эффективного потенциала относительно результата, полученного в приближении Хартри.

В заключение этого раздела обсудим, насколько полученные результаты зависят от принятого условия о равенстве эффективных масс электронов в обоих электродах. Вообще говоря, проблема граничных условий для огибающей волновой функции на интерфейсе двух твердых тел, которые могут различаться не только эффективными массами, но также типом и законом дисперсии носителей заряда, достаточно сложна. По сложившемуся в настоящее время мнению, эти условия не могут быть сведены к универсальному соотношению между значениями

волновой функции и ее производной с обеих сторон интерфейса, а должны быть выведены для каждого конкретного случая (см., например, критическое обсуждение вопроса применительно к гетеропереходам в работах [47, 48]).

Оставаясь в рамках феноменологического описания граничных условий, можно написать их наиболее общий вид с помощью матрицы перехода \mathbb{T} (transfer matrix), на элементы которой накладываются некоторые соотношения, выведенные из требования эрмитовости гамильтониана [48]. Если уравнения для огибающих в обоих электродах можно взять в однозонном приближении, то граничные условия описываются матрицей 2×2 . В тех случаях, когда матрица \mathbb{T} может быть диагональна [49], окончательные формулы (5.16)–(5.19) сохраняют свой вид при замене амплитудной прозрачности барьера t_k и волнового вектора q перенормированными величинами $\tilde{t} = T_{11}t_k$ и $\tilde{q} = (T_{22}/T_{11})q$.

Если матрица \mathbb{T} будет иметь отличные от нуля недиагональные элементы, это не внесет принципиальных трудностей в построение R - и L -решений уравнения Шредингера по их асимптотическому поведению при $z \rightarrow \pm\infty$ из частных вещественных решений. Однако рассмотрение самого общего случая в данной работе сделало бы изложение принципиальных сторон применения итерационного алгоритма неоправданно более громоздким. В то же время учет неравенства волновых векторов $q(\varepsilon) \neq k(\varepsilon)$, которое характерно для несимметричных барьерных структур, было необходимо включить в расчет.

6. ЗАКЛЮЧЕНИЕ

Выяснено, что реализованный алгоритм итерационного решения уравнений самосогласованного поля сходится во всех рассмотренных случаях, которые представляют типичные неоднородные электронные системы с непрерывным спектром. Сложности с его применением возникли только при рассмотрении модели желе для больших значений $R_s \geq 5$, которые не встречаются в случае большинства металлов и полупроводниковых структур. Показано также, что явный учет в алгоритме экранирующих свойств электронного газа позволяет при численном анализе неограниченных кулоновских систем рассматривать выделенный ограниченный объем такой системы, перенося на его поверхность граничные условия из бесконечности и считая энергетический спектр одноэлектронных состояний строго непрерывным. Такой подход устраняет

также трудности, возникающие при рассмотрении неоднородных неограниченных систем с непрерывным спектром как предела ограниченных с квазидискретным спектром.

Работа выполнена при частичной поддержке РФФИ (гранты №№ 07-02-01481, 06-02-16955).

ЛИТЕРАТУРА

1. В. Кон, П. Вашишта, в кн. *Теория неоднородного электронного газа*, пер. с англ. под ред. Д. А. Киржница, Е. Г. Максимова, Мир, Москва (1987), гл. 2. [W. Kohn and P. Vashishta, in *Theory of the Inhomogeneous Electron Gas*, ed. by S. Lundqvist and N. H. March, Plenum Press, New York (1983), Ch. 2.]
2. А. Я. Шул'ман, J. Phys.: Conf. Ser. **35**, 163 (2006).
3. N. D. Lang and W. Kohn, Phys. Rev. B **1**, 4555 (1970).
4. M. Manninen, R. Nieminen, P. Hautojarvi, and J. Arponen, Phys. Rev. B **12**, 4012 (1975).
5. A. Liebsch, *Electronic Excitations at Metal Surfaces*, Plenum Press, New York (1997), Sec. 2.3.2.
6. R. M. Nieminen, J. Phys. F **7**, 375 (1977).
7. А. Я. Шул'ман and Д. В. Посвянский, arXiv:cond-mat/0209335.
8. Д. Пайнс, *Элементарные возбуждения в твердых телах*, Мир, Москва (1965). [D. Pines, *Elementary Excitations in Solids*, Benjamin, New York (1963).]
9. П. Гомбаш, *Статистическая теория атома и ее применения*, Изд-во иностр. лит., Москва (1951). [P. Gombás, *Theorie des Atoms und Ihre Anwendungen*, Springer-Verlag, Wien (1949).]
10. А. Я. Шульман, В. В. Зайцев, Sol. St. Comm. **18**, 1623 (1976).
11. С. К. Годунов, В. С. Рябенький, *Разностные схемы*, Наука, Москва (1977).
12. О. В. Константинов, А. Я. Шик, ЖЭТФ **58**, 1662 (1970).
13. J. F. Appelbaum and G. A. Baraff, Phys. Rev. Lett. **26**, 1432 (1971).
14. Р. Ньютон, *Теория рассеяния волн и частиц*, пер. с англ. под ред. А. М. Бродского, В. В. Толмачева, Мир, Москва (1969). [R. G. Newton, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York (1966).]
15. G. A. Baraff and J. F. Appelbaum, Phys. Rev. B **5**, 475 (1972).
16. Р. П. Федоренко, *Введение в вычислительную физику*, Изд-во МФТИ, Москва (1994), § 15.
17. T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982); Е. Андо, А. Фаулер, Ф. Стерн, *Электронные свойства двумерных систем*, пер. с англ. под ред. Ю. В. Шмарцева, Мир, Москва (1985).
18. J. Suñé, P. Olivo, and B. Riccó, J. Appl. Phys. **70**, 337 (1991).
19. C. A. Mead, Phys. Rev. Lett. **6**, 545 (1961).
20. M. Stengel and N. A. Spaldin, Nature **443**, 679 (2006).
21. H. Y. Ku and F. G. Ullman, J. Appl. Phys. **35**, 265 (1964).
22. C. T. Black and J. J. Welser, IEEE Trans. ED **46**, 776 (1999).
23. J. Bardeen, Phys. Rev. **49**, 653 (1936).
24. H. B. Huntington, Phys. Rev. **81**, 1035 (1951).
25. L. J. Sinnamon, R. M. Bowman, and J. M. Gregg, Appl. Phys. Lett. **78**, 1724 (2001).
26. H. Ehrenreich and H. R. Philipp, Phys. Rev. **128**, 1622 (1962).
27. W. S. Choi, S. S. A. Seo, K. W. Kim et al., Phys. Rev. B **74**, 205117 (2006).
28. K. M. Rabe, Nature Nanotechnology **1**, 171 (2006).
29. W. Kohn and C. Majumdar, Phys. Rev. **138**, A1617 (1965).
30. D. C. Tsui, Phys. Rev. B **8**, 2657 (1973).
31. N. D. Lang and W. Kohn, Phys. Rev. B **3**, 1251 (1971).
32. J. P. Perdew and Y. Wang, Phys. Rev. B **38**, 12228 (1988).
33. H. F. Budd and J. Vannimenus, Phys. Rev. Lett. **31**, 1281, 1430 (Erratum) (1973).
34. Н. Лэнг, в кн. *Теория неоднородного электронного газа*, пер. с англ. под ред. Д. А. Киржница, Е. Г. Максимова, Мир, Москва (1987), гл. 5. [N. D. Lang, in *Theory of the Inhomogeneous Electron Gas*, ed. by S. Lundqvist and N. H. March, Plenum Press, New York (1983), Ch. 5.]
35. Б. Серочинска-Вояс, ФТТ **10**, 693 (1968).
36. R. Monnier and J. P. Perdew, Phys. Rev. B **17**, 2595 (1978).

37. R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).
38. W. Kohn, *Rev. Mod. Phys.* **71**, 1253 (1999).
39. Г. Пааш, М. Хитшольд, в кн. *Достижения электронной теории металлов*, пер. с нем. под ред. А. А. Абрикосова, Ю. Х. Векилова, Мир, Москва (1984), т. 2, гл. 10. [G. Raasch and M. Hietschold, in *Ergebnisse in der Elektronentheorie der Metalle*, ed. by P. Ziesche and G. Lehmann, Akademie-Verlag, Berlin (1983), Ch. 10.]
40. М. А. Леонтович, *Введение в термодинамику*, Наука, Москва (1983), гл. 3.
41. D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures*, Cambridge Uni Press, Cambridge, UK (1997)
42. I. N. Kotel'nikov and A. Ya. Shul'man, in *Proc. 19th Int. Conf. Phys. Semicond.*, Warsaw (1988), Vol. 1, p. 681.
43. А. Я. Шульман, И. Н. Котельников, Н. А. Варванин, Е. Н. Миргородская, *Письма в ЖЭТФ* **73**, 643 (2001).
44. И. Н. Котельников, И. Л. Бейнихес, А. Я. Шульман, *ФТТ* **27**, 401 (1985).
45. И. Н. Котельников, Д. К. Чепиков, Е. Г. Чиркова, А. Я. Шульман, *ФТП* **21**, 1854 (1987).
46. E. M. Dizhur, A. Ya. Shul'man, I. N. Kotel'nikov, and A. N. Voronovsky, *Phys. Stat. Sol. (b)* **223**, 129 (2001); arXiv:cond-mat/0010200.
47. Э. Е. Тахтамиров, В. А. Волков, *ЖЭТФ* **116**, 1843 (1999).
48. I. V. Tokatly, A. G. Tsibizov, and A. A. Gorbatsевич, *Phys. Rev. B* **65**, 165328 (2002).
49. T. Ando, S. Wakahara, and H. Akera, *Phys. Rev. B* **40**, 11609 (1989).