
Оглавление

Введение	5
Часть I. Физические основы наноэлектроники	9
1. Фундаментальные явления	9
1.1. Квантовое ограничение.	9
1.2. Баллистический транспорт носителей заряда	15
1.3. Туннелирование носителей заряда	20
1.4. Спиновые эффекты	25
2. Элементы низкоразмерных структур	28
2.1. Свободная поверхность и межфазные границы	28
2.2. Сверхрешетки	31
2.3. Моделирование атомных конфигураций.	34
3. Структуры с квантовым ограничением за счет внутреннего электрического поля.	38
3.1. Квантовые колодцы	38
3.2. Модуляционно-легированные структуры	42
3.3. Дельта-легированные структуры	44
4. Структуры с квантовым ограничением за счет внешнего электрического поля	46
4.1. Структуры металл/диэлектрик/полупроводник	46
4.2. Структуры с расщепленным затвором	47
Литература к части I.	48
Часть II. Методы формирования наноэлектронных структур (нанотехнология)	50
1. Традиционные методы осаждения пленок	51
1.1. Химическое осаждение из газовой фазы	51
1.2. Молекулярно-лучевая эпитаксия	56
2. Методы, основанные на использовании сканирующих зондов.	58
2.1. Физические основы	58
2.2. Атомная инженерия	61
2.3. Локальное окисление металлов и полупроводников	67

2.4. Локальное химическое осаждение из газовой фазы	68
3. Нанолитография	69
3.1. Электронно-лучевая литография	70
3.2. Профилирование резистов сканирующими зондами	73
3.3. Нанопечать	74
3.4. Перьевая нанолитография	77
3.5. Сравнение нанолитографических методов	78
4. Саморегулирующиеся процессы	81
4.1. Самосборка	82
4.2. Самоорганизация в объемных материалах	84
4.3. Самоорганизация при эпитаксии	90
4.4. Осаждение пленок Лэнгмюра–Блоджетт	98
5. Формированиеnanoструктурированных материалов	105
5.1. Пористый кремний	105
5.2. Пористый оксид алюминия и структуры на его основе	110
5.3. Углеродные нанотрубки	122
Литература к части II	128
Часть III. Перенос носителей заряда в низкоразмерных структурах	129
1. Транспорт носителей заряда вдоль потенциальных барьеров	129
1.1. Интерференция электронных волн	130
1.2. Вольт-амперные характеристики низкоразмерных структур	134
1.3. Отрицательное сопротивление изгиба	137
1.4. Квантовый эффект Холла	139
1.5. Приборы на основе интерференционных эффектов	146
2. Туннелирование носителей заряда через потенциальные барьеры	152
2.1. Одноэлектронное туннелирование	153
2.2. Приборы на основе одноэлектронного туннелирования .	162
2.3. Резонансное туннелирование	178
2.4. Приборы на основе резонансного туннелирования	181
3. Спин-зависимый транспорт носителей заряда	189
3.1. Гигантское магнитосопротивление	190
3.2. Спин- зависимое туннелирование	196
3.3. Управление спинами носителей заряда в полупроводниках	202
3.4. Эффект Кондо	211
3.5. Спинтранные приборы	216
Литература к части III	223

Введение

Наноэлектроника (*nanoelectronics*)¹ — это область науки и техники, занимающаяся созданием, исследованием и применением электронных приборов с нанометровыми размерами элементов. В основе функционирования таких приборов лежат квантовые эффекты.

Приставка «нано» обозначает одну миллиардную долю той единицы измерения, перед которой она поставлена. Так, один нанометр равен одной миллиардной доле метра, т. е. $1 \text{ нм} = 10^{-9} \text{ м}$. Типичные размеры элементов, с которыми имеет дело наноэлектроника, составляют от единиц до сотен нанометров.

Следует различать понятия «наноразмерные структуры» и «низкоразмерные структуры». В первом случае главным признаком является линейный размер структуры, который хотя бы в одном направлении должен соответствовать нанометровому диапазону. **Низкоразмерными структурами** (*low-dimensional structures*) называют структуры, у которых по крайней мере один размер равен нулю. Такое определение носит, конечно, условный характер, поскольку реальный физический мир состоит из трехмерных объектов. Например, толщина плоскости, сконструированной из атомов, равна не нулю, а диаметру одного атома, что составляет около 10^{-10} м . Нужно понимать, что двумерные, одномерные и нульмерные структуры не являются таковыми в строгом геометрическом смысле, а называются так лишь потому, что их размер в одном, двух или трех направлениях меньше определенного «критического» значения, ниже которого физические свойства структуры в этом направлении (направлениях) становятся существенно отличными от свойств объемного (трехмерного) материала, из которого данная структура изготовлена. В твердотельных структурах так обычно происходит, когда один из линейных размеров

¹ Здесь и далее основные русскоязычные термины и понятия сопровождаются их англоязычными эквивалентами, что призвано облегчить последующее чтение и понимание оригинальных публикаций на английском языке.

уменьшается до величины порядка нанометра. В научной и технической литературе наноразмерные структуры часто называют *наноструктурами* (*nanostructures*). Квантово-механические явления в них являются доминирующими, что определяет их специфические электронные, оптические, магнитные и другие свойства.

В части I данного учебного пособия рассмотрены фундаментальные физические эффекты, имеющие место в наноструктурах и обусловленные их пониженней мерностью. Данна классификация элементарных и комбинированных низкоразмерных структур. Описаны подходы, позволяющие формировать такие структуры в полупроводниках.

Развитию нанотехнологии во многом способствовало постоянное совершенствование традиционных методов изготовления полупроводниковых приборов, таких как химическое осаждение из газовой фазы, молекулярно-лучевая эпитаксия и электронно-лучевая литография. Кроме того, применение сканирующих атомарно острых зондов и саморегулирующихся процессов значительно расширило возможности создания твердотельных наноструктур. Нанотехнологические приемы постоянно совершенствуются, появляются все новые и новые методы. В части II основное внимание уделено методам, уже прошедшем экспериментальную проверку и использующимся достаточно широко. Описаны и некоторые подходы, которые хоть и кажутся сегодня «экзотическими», завтра вполне могут превратиться в коммерческие.

Особенности транспорта (переноса) носителей заряда в наноразмерных структурах рассмотрены в части III. Главный акцент сделан на то, что размеры наноструктур сравнимы с длиной свободного пробега электронов или дырок. Это приводит к свободному перемещению носителей через наноструктуру, без рассеяния на дефектах, примесях и фонах. Так как фазы невзаимодействующих электронных волн сохраняются на всем пути их следования, то весьма естественно ожидать появления эффектов интерференции. Кроме того, в наноструктурах нарушается большинство предположений, на которых основано использование уравнения Больцмана для описания переноса носителей. Транспорт носителей заряда, обусловленный, с одной стороны, волновой природой электрона (волнам свойственна непрерывность), а с другой — дискретностью переносимого им заряда, рассматривается как квантовый режим переноса. Он принципиально отличается от классического режима, в котором длина свободного пробега электронов намного меньше размеров системы, а сами электроны считаются классическими частицами. На транспорт носителей накладывает свой отпечаток и специфическая для низкоразмерных

структур дискретность разрешенных энергетических состояний. Каждая глава части III включает примеры электронных приборов, основанных на рассматриваемых явлениях и эффектах.

При представлении материала использованы термины, основные понятия и определения из первого энциклопедического справочника по нанонауке и нанотехнологиям — V. E. Borisenko, S. Ossicini, *What is What in the Nanoworld* (Wiley-VCH, Weinheim, 2004, 2008).

Авторы надеются, что приведенные в книге материалы позволят читателю овладеть физическими и технологическими основами наноэлектроники и использовать их для дальнейшего развития этого перспективного направления.

Выражаем признательность А. Л. Данилюку и Л. А. Опенову за полезное критическое обсуждение материалов, вошедших в данное пособие, и за предложения по его совершенствованию.

Часть I

Физические основы nanoэлектроники

1. Фундаментальные явления

Поведение подвижных носителей заряда (электронов и дырок) в наноразмерных структурах определяют три группы фундаментальных явлений: квантовое ограничение, баллистический транспорт и квантовая интерференция, а также туннелирование. Все эти эффекты по своему происхождению представляют собой типичные квантовомеханические явления.

Квантовое ограничение возникает, когда свободное движение электронов в одном из направлений оказывается ограниченным потенциальными барьерами, образующими nanoструктуру, в которой эти электроны находятся. Оно изменяет спектр разрешенных энергетических состояний и влияет на перенос носителей заряда через nanoструктуры. Транспорт носителей заряда может, в принципе, осуществляться как параллельно, так и перпендикулярно потенциальным барьерам. В случае движения носителей вдоль потенциальных барьеров доминирующими эффектами оказываются баллистический транспорт и квантовая интерференция. Прохождение же носителей заряда через потенциальные барьеры имеет место исключительно посредством их туннелирования, что и обеспечивает перенос носителей из одной области nanoэлектронного прибора в другую. В данном разделе рассмотрена физическая природа и основные закономерности проявления перечисленных фундаментальных явлений.

1.1. Квантовое ограничение

Свободный электрон, движущийся в трехмерной системе (3D), имеет кинетическую энергию, величина которой, в соответствии с пространственными компонентами его импульса p_x, p_y, p_z , составляет

$$E = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2),$$

или, в волновом представлении,

$$E = \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2), \quad (1.1)$$

где m^* — эффективная масса электрона (в твердых телах она обычно меньше, чем масса покоя электрона m_0); \hbar — приведенная постоянная Планка ($\hbar = h/2\pi$); k_x, k_y, k_z — пространственные компоненты волнового вектора. Плотность электронных состояний при этом является непрерывной функцией энергии:

$$n_{3D}(E) = \frac{m^* \sqrt{2m^* E}}{\pi^2 \hbar^3}. \quad (1.2)$$

В низкоразмерной структуре свободное движение электрона ограничено по крайней мере в одном направлении. В данном направлении (пусть это будет направление вдоль оси x) потенциальная энергия электрона может быть представлена в виде бесконечно глубокой потенциальной ямы, как это показано на рис. 1.1. Если ширина ямы вдоль оси x равна a , то в области $0 < x < a$ электрон имеет нулевую потенциальную энергию. Бесконечно высокий потенциальный барьер делает невозможным нахождение электрона за границами этой области.

Таким образом, волновая функция электрона должна обращаться в нуль на границах потенциальной ямы, т. е. при $x = 0$ и $x = a$. Такому условию отвечает лишь ограниченный набор волновых функций. Это — стоячие волны с длиной λ , определяемой соотношением

$$\lambda_n = 2a/n \quad (n = 1, 2, \dots). \quad (1.3)$$

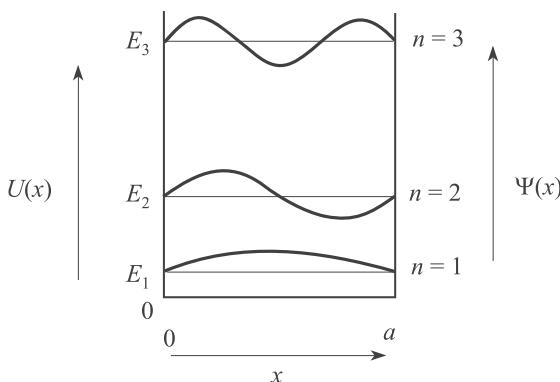


Рис. 1.1. Потенциальная яма и волновые функции электронов в ней

Соответствующие разрешенные значения волнового вектора дискретны и равны

$$k_n = 2\pi/\lambda_n = n\pi/a. \quad (1.4)$$

Как следствие, энергии разрешенных энергетических состояний электрона в яме тоже оказываются дискретными. Спектр этих состояний имеет вид

$$E_n = \frac{\hbar^2 k_n^2}{2m^*} = \frac{\hbar^2 \pi^2 n^2}{2m^* a^2}. \quad (1.5)$$

Целое число n является квантовым числом, обозначающим квантовое состояние. Таким образом, электрон, помещенный в ограниченную область пространства, может занимать только дискретные энергетические уровни. Самое низкое состояние имеет энергию

$$E_1 = \frac{\hbar^2 \pi^2}{2m^* a^2}, \quad (1.6)$$

которая всегда больше нуля. Ненулевая минимальная энергия отличает квантово-механическую систему от классической, для которой энергия частицы, находящейся на дне потенциальной ямы, тождественно равна нулю. Кроме того, разрешенные значения энергии для электрона оказываются квантованными и пропорциональны n^2 .

Для того чтобы удовлетворить принципу неопределенности $\Delta p \Delta x \geq \hbar/2$ (в нашем случае $\Delta x = a$), неопределенность импульса электрона должна быть $\Delta p \geq \hbar/2a$, что отвечает минимальному изменению энергии $\Delta E = (\Delta p)^2/2m^* = \hbar^2/8m^* a^2$, которое (с точностью до множителя $\pi^2/4$) соответствует приведенному выше выражению для E_1 . Таким образом, принцип неопределенности также приводит нас к выводу о ненулевом значении минимальной энергии электрона в потенциальной яме.

Ограничение движения электронов (дырок) в низкоразмерной структуре, приводящее (вследствие их квантово-волновой природы) к ненулевому минимальному значению их энергии и к дискретности энергий разрешенных состояний, называют **квантовым ограничением** (*quantum confinement*). В твердых телах квантовое ограничение может быть реализовано в трех пространственных направлениях. Количество направлений, в которых эффект квантового ограничения отсутствует, используется в качестве критерия для классификации элементарных низкоразмерных структур по трем группам: квантовые пленки, квантовые шнуры и квантовые точки. Схематически они показаны на рис. 1.2.

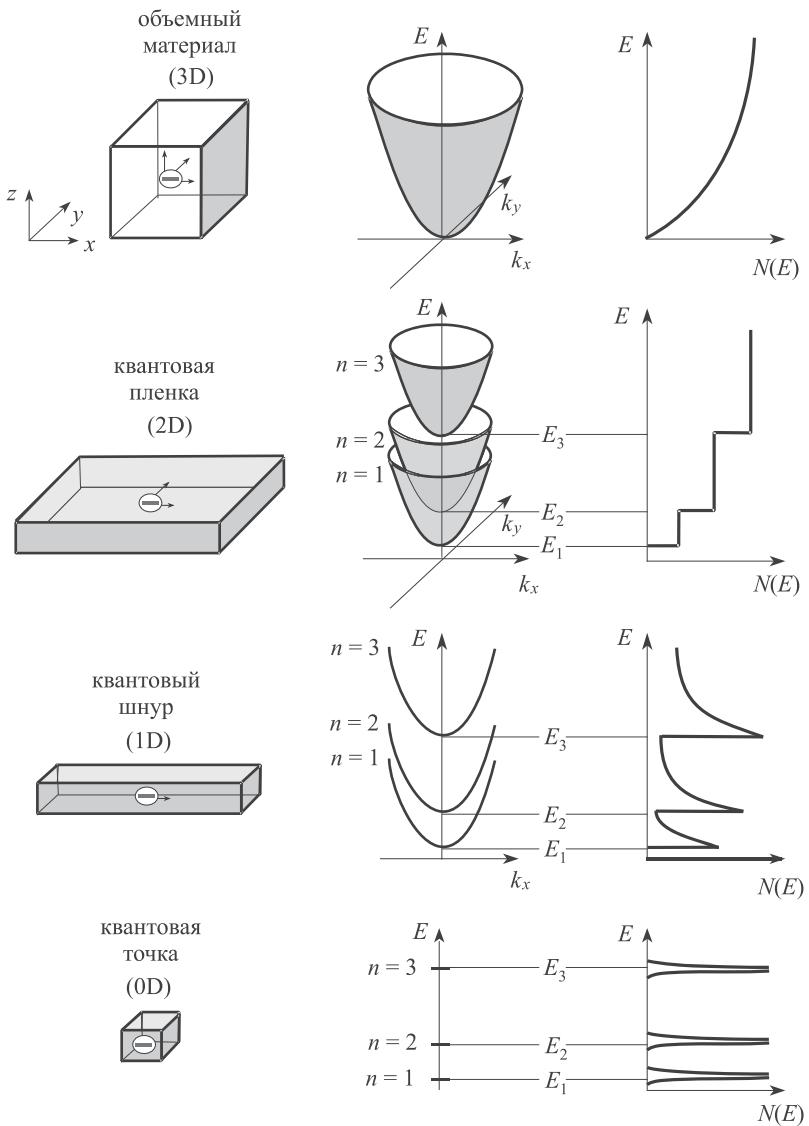


Рис. 1.2. Элементарные низкоразмерные структуры, их энергетические диаграммы и плотности состояний $N(E)$ в сравнении с трехмерной структурой

Квантовые пленки (quantum films) представляют собой двумерные (2D) структуры, в которых квантовое ограничение действует только в одном направлении — перпендикулярно пленке (направле-

ние z на рис. 1.2). Носители заряда в таких структурах могут свободно двигаться в плоскости xy . Их энергия складывается из квантованных значений, определяемых эффектом квантового ограничения в направлении z (в соответствии с толщиной пленки l_z), и непрерывных составляющих в направлениях x и y :

$$E = \frac{\hbar^2 \pi^2 n^2}{2m^* l_z^2} + \frac{\hbar^2 k_x^2}{2m^*} + \frac{\hbar^2 k_y^2}{2m^*}. \quad (1.7)$$

В k -пространстве энергетическая диаграмма квантовой пленки представляет собой семейство параболических зон, которые, перекрываясь, образуют подзоны. Минимальная энергия электрона в n -й подзоне задается соотношением (1.5). Электрон с такой энергией не-подвижен в плоскости пленки.

Зависимость плотности электронных состояний от энергии в квантовой пленке имеет ступенчатый вид (вместо параболической зависимости в трехмерных структурах):

$$n_{2D}(E) = \frac{m^*}{\pi \hbar^2 l_z} \sum_i \Theta(E - E_i), \quad i = 1, 2, \dots, \quad (1.8)$$

где $\Theta(E - E_i)$ — ступенчатая функция.

Электроны в квантовых пленках обычно называют **двумерным электронным газом** (*two-dimensional electron gas, 2DEG*).

Квантовые шнуры (*quantum wires*)¹ — это одномерные (1D) структуры. В отличие от квантовых пленок, они имеют не один, а два нанометровых размера, в направлении которых и действует эффект квантового ограничения. Носители заряда могут свободно двигаться только в одном направлении — вдоль оси шнура. Таким образом, вклад в энергию носителя заряда дают кинетическая составляющая вдоль одного направления и квантованные значения в двух других направлениях:

$$E = \frac{\hbar^2 \pi^2 n_1^2}{2m^* l_y^2} + \frac{\hbar^2 \pi^2 n_2^2}{2m^* l_z^2} + \frac{\hbar^2 k_x^2}{2m^*}, \quad n_1 = 1, 2, \dots; n_2 = 1, 2, \dots. \quad (1.9)$$

Для каждой пары дискретных уровней в направлениях квантового ограничения плотность электронных состояний в квантовом шнуре зависит от энергии по закону $E^{-1/2}$:

¹ В русскоязычной научной и учебной литературе оригиналный англоязычный термин *quantum wires* иногда переводится как «квантовые нити», «квантовые провода» и «квантовые проволоки», хотя наиболее точным и благозвучным, по нашему мнению, является его перевод как «квантовые шнуры».

$$n_{1D}(E) = \frac{(2m^*)^{1/2}}{\pi \hbar l_y l_z} \sum_{i,j} (E - E_{i,j})^{-1/2}, \quad i,j = 1, 2, \dots \quad (1.10)$$

Квантовые точки (*quantum dots*) — это нуль-мерные (0D) структуры, в которых движение носителей заряда ограничено во всех трех направлениях. В каждом из этих направлений энергия электрона оказывается квантованной в соответствии с формулой (1.5), а плотность состояний представляет собой набор острых пиков, описываемых δ -функциями:

$$n_{0D}(E) = \frac{2}{l_x l_y l_z} \sum_{i,j,k} \delta(E - E_{i,j,k}), \quad i,j,k = 1, 2, \dots \quad (1.11)$$

Из-за сходства энергетических характеристик атомов и квантовых точек последние иногда называют «искусственными атомами». Квантовые точки состоят из сравнительно небольшого количества атомов. В этом отношении к ним близки атомные кластеры и нанокристаллы (кристаллы нанометровых размеров), где также имеет место эффект квантового ограничения.

Рассмотренные элементарные низкоразмерные структуры в определенном смысле являются идеализированными объектами. Очевидно, что низкоразмерные структуры, представляющие практический интерес, должны располагаться на какой-либо подложке и иметь контакт с другими структурами и функциональными элементами. Более того, приборные применения требуют комбинации нескольких элементарных структур. Но, несмотря на появление в сложных комбинированных структурах новых квантово-механических эффектов, определяющую роль в них продолжает играть квантовое ограничение.

Для изготовления низкоразмерных структур используют два принципиальных подхода, которые можно охарактеризовать как «геометрический» и «электронный». Геометрический подход предполагает использование технологий, обеспечивающих формирование объектов с нанометровыми размерами. Для этого используются специальные нанотехнологические приемы, которые будут рассмотрены в части II данного учебного пособия. Электронный подход основан на возможности управления размерами областей с определенным типом и концентрацией носителей заряда в полупроводниках посредством электрического поля. При этом используются как традиционные структуры металл/диэлектрик/полупроводник и металл/полупроводник, так и полупроводниковые гетероструктуры. Способы создания в полупроводниках областей с квантовым ограничением за счет внутреннего встроенного электрического поля и путем приложения внешнего электрического смещения рассмотрены ниже.

1.2. Баллистический транспорт носителей заряда

Особенности транспорта носителей заряда в твердых телах во многом определяются процессами их рассеяния при движении из одной области в другую. Электрон, сталкиваясь с другим электроном или рассеиваясь на колебаниях решетки, дефектах либо границе раздела, неизбежно изменяет свое состояние. Среднее расстояние, которое электрон проходит между двумя последовательными актами рассеяния, называют *средней длиной свободного пробега* (*mean free path*).

В макроскопических системах средняя длина свободного пробега электронов всегда намного меньше размера этих систем. Для них справедливы следующие допущения: 1) процессы рассеяния носителей заряда локальны, т. е. имеют место в определенных точках пространства; 2) рассеяние происходит непрерывно во времени; 3) и рассеяние, и поля, инициирующие движение носителей заряда, настолько слабые, что вызванные ими отклонения электронной системы от равновесия можно рассматривать независимо друг от друга; 4) масштаб времени для наблюдения за системой выбран таким образом, что регистрируются только те события, которые происходят за времена, превышающие среднее время между двумя ближайшими актами рассеяния. Такие допущения позволяют использовать для описания транспорта носителей заряда в макроскопических системах кинетическое уравнение Больцмана.

Вnanoструктурах условия для транспорта носителей заряда существенно отличаются от таковых в макросистемах. В структурах с размером меньше длины свободного пробега перенос носителей заряда происходит без их рассеяния. Такой перенос называют *баллистическим транспортом* (*ballistic transport*). При этом допущения, перечисленные выше, теряют свою силу. Основные эффекты, относящиеся к баллистическому транспорту, определяются соотношением между размерами структуры, в которой имеет место перенос носителей заряда, и характерными длинами свободного пробега. Ключевыми являются длины свободного пробега, характеризующие упругое и неупругое рассеяние носителей.

Средняя длина свободного пробега при упругом рассеянии (*elastic mean free path*) — это среднее расстояние, которое проходит носитель заряда между двумя последовательными актами упругого рассеяния. Когда электронная система вырождена при низкой температуре, она определяется скоростью Ферми $v_F = (2E_F/m^*)^{1/2}$ (E_F — энергия Ферми) и временем рассеяния $\tau_{sc} = Dd/v_F^2$ (D — коэффициент диффузии носителей, d — мерность структуры) как $l_e = v_F \tau_{sc}$.

Средняя длина свободного пробега при неупругом рассеянии (*inelastic mean free path*) — это расстояние, на котором электронная волна изменяет свою фазу вследствие рассеяния. Она равна $l_{in} = v_F \tau_\phi$, где τ_ϕ — время релаксации фазы (или энергии). Имеется и другой параметр, характеризующий неупругое рассеяние носителей заряда. **Длина фазовой когерентности** (*phase coherence length*) $l_\phi = (D\tau_\phi)^{1/2}$ — это расстояние, на котором электронная волна сохраняет свою фазу (или, как еще говорят, подвижный носитель сохраняет свою «фазовую память»). Следует иметь в виду, что средняя длина свободного пробега при неупругом рассеянии и длина фазовой когерентности представляют собой различные характеристики. Длина фазовой когерентности меньше, чем средняя длина свободного пробега при неупругом рассеянии. Оба эти параметра важны при анализе условий интерференции электронных волн.

В твердых телах средняя длина свободного пробега для неупругого рассеяния больше, чем для упругого. Транспорт носителей заряда в структурах с размером больше l_e , но меньше l_{in} происходит «квазибаллистически», т. е. со слабым рассеянием.

Важной характеристикой наноструктур является **длина волны Ферми** $\lambda_F = 2\pi/k_F$, где k_F — волновой вектор Ферми. При $T = 0$ электроны находятся в состояниях, определяемых соотношением $|\mathbf{k}| \leq k_F$, что эквивалентно электронным волнам с длиной волны $\lambda \geq \lambda_F$.

Сравнивая размер наноструктуры со средней длиной свободного пробега электронов и с длиной волны Ферми (являющимися характеристиками материала, из которого данная структура изготовлена), можно предсказать основные особенности движения носителей заряда в этой наноструктуре. В металлах средняя длина свободного пробега электронов даже при низких температурах обычно не превышает 10 нм, что меньше или порядка размеров типичных наноструктур. По этой причине баллистический транспорт в металлических наноструктурах реализуется с трудом. Более того, длина волны Ферми в них тоже очень мала и обычно составляет 0,1–0,2 нм. Поэтому квантование энергетических уровней в металлах не является существенным фактором (за исключением очень низких температур, когда расстояние между двумя соседними энергетическими уровнями становится сравнимым с тепловой энергией $k_B T$), и энергетический спектр металлических квантовых точек определяется преимущественно кулоновским взаимодействием.

Транспорт носителей заряда в полупроводниках характеризуется большой, до нескольких микрометров, средней длиной свободного пробега электронов. Некоторые важные параметры для Si и GaAs при

низких температурах приведены в качестве примера в табл. 1.1. Следует иметь в виду, что эти параметры даны для носителей с энергией вблизи уровня Ферми и могут меняться в определенном диапазоне. Для горячих носителей характеристики существенно иные.

Таблица 1.1. Параметры, характеризующие транспорт электронов в Si и GaAs при низких температурах (≈ 4 К)

Параметр, единица измерения	Si	GaAs
Скорость Ферми, 10^7 см/с	0,97	2,76
Длина волны Ферми, нм	39	39
Время рассеяния, 10^{-12} с	1,1	3,8
Время релаксации фазы, 10^{-12} с	5,7	18
Коэффициент диффузии, 10^3 см 2 /с	0,52	1,45
Средняя длина свободного пробега при упругом рассеянии, нм	107	1050
Средняя длина свободного пробега при неупругом рассеянии, нм	500	5000
Длина фазовой когерентности, нм	540	1620
Эффективная масса, m_0	0,19	0,067
Время релаксации спина	14–60 мс*	5–80 пс

* В ^{28}Si , легированном фосфором.

При комнатной температуре средняя длина свободного пробега электронов при неупругом рассеянии достигает 50–100 нм в Si и около 120 нм в GaAs. Очевидно, что вnanoструктурах из полупроводников баллистический транспорт легко реализуется. Более того, длина волны Ферми в полупроводниках достигает 30–50 нм. Когда размер структуры сравнивается с этими длинами, квантование энергии вследствие квантового ограничения становится существенным фактором, определяющим ее электронные свойства и транспорт носителей заряда.

Идеальный баллистический транспорт носителей заряда в nanoструктурах характеризуется **универсальной баллистической проводимостью** (*universal ballistic conductance*), которая не зависит от типа материала и определяется лишь фундаментальными константами. Наиболее простым прибором для иллюстрации баллистического транспорта является проводник с двумя контактами. Такой проводник схематически изображен на рис. 1.3, где сужение между двумя резервуарами с электронами играет роль проводящего квантового шнура. Предполагается, что в этом канале нет никаких неоднородностей, при-

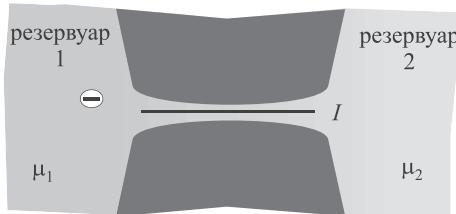


Рис. 1.3. Два резервуара с электронами, соединенные проводящим каналом

водящих к рассеянию носителей. Кроме того, предполагается, что связь проводящего канала с резервуарами электронов осуществляется посредством безотражательных проводников; это приводит к неизбежному попаданию в резервуар всех электронов, вышедших из канала.

Будем считать, что $T = 0$, а резервуары заполнены электронами вплоть до электрохимических потенциалов μ_1 и μ_2 , где $\mu_1 > \mu_2$. Если электронные состояния первого резервуара в интервале энергий от μ_1 до μ_2 полностью заняты, то между резервуарами протекает ток

$$I = (\mu_1 - \mu_2)ev(dn/d\mu), \quad (1.12)$$

где e — заряд электрона; v — составляющая скорости электронов на поверхности Ферми вдоль оси канала; $dn/d\mu$ — плотность электронных состояний в канале (с учетом вырождения по спину). В квантовом шнуре $dn/d\mu = 1/\pi\hbar v$. Учитывая, что $\mu_1 - \mu_2 = e(V_1 - V_2)$, где V_1 и V_2 — электрохимические потенциалы, приводящие к различию электрохимических потенциалов в резервуарах, для проводимости квантового шнура получим:

$$G = I/(V_1 - V_2) = e^2/\pi\hbar = 2e^2/h. \quad (1.13)$$

Это — проводимость идеального одномерного проводника в баллистическом режиме. Она определяется только фундаментальными константами — зарядом электрона и постоянной Планка. Величину $e^2/h = 38,740$ мкСм называют квантовой единицей проводимости (или квантом проводимости). Соответствующее сопротивление равно $h/e^2 = 25812,807$ Ом (квант сопротивления).

Приведенные рассуждения можно обобщить на случай, когда проводящий канал имеет более одного энергетического состояния ниже уровня Ферми. Для канала с N разрешенными состояниями (или, что то же самое, с N передающими модами) получаем

$$G = N(2e^2/h). \quad (1.14)$$

Таким образом, проводимость канала с несколькими передающими модами квантуется в единицах $2e^2/h$. Это наблюдается в **квантовых точечных контактах** (*quantum point contacts*), которые представляют собой узкие перемычки, соединяющие широкие резервуары с электронами. Пример такой структуры и ее проводимость, показаны на рис. 1.4.

Короткое сужение в «захороненном» двумерном электронном газе (2DEG) контролируется находящимися на поверхности затворами, имеющими форму направленных навстречу друг другу острых штырей. Чтобы вытеснить двумерный электронный газ из-под затворов и сформировать узкий проводящий канал, на затворы подается отрицательное смещение. Напряжение V_g определяет ширину канала. При увеличении отрицательного смещения ширина канала постепенно уменьшается, и канал исчезает. Если же ширина канала возрастет, то количество разрешенных (лежащих ниже уровня Ферми) состояний увеличивается. На кривой проводимости появляются ступеньки, соответствующие числу этих состояний N . Следует помнить, что в случае, когда проводящий канал не является идеальным, процессы рассеяния носителей заряда приводят к уменьшению тока, поскольку вероятность переноса носителей заряда становится меньше единицы. Это необходимо учитывать при расчете проводимости канала.

Как только движение электрона становится когерентным (в том смысле, что электрон проходит через всю структуру без рассеяния), его волновая функция сохраняет свою фазу. При этом возникают раз-

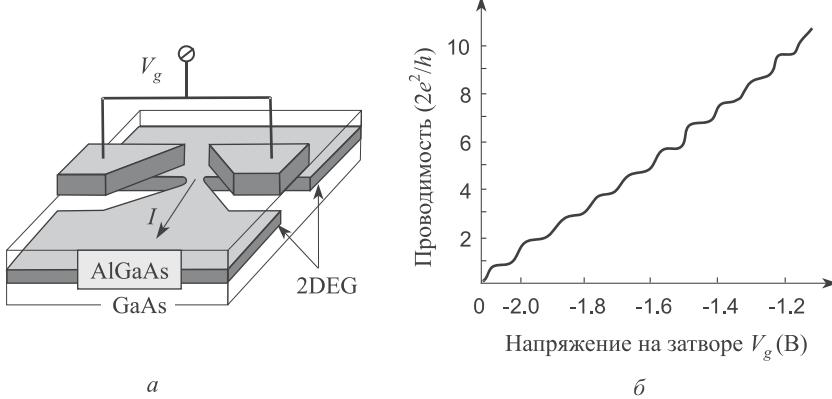


Рис. 1.4. Изменение проводимости квантового точечного контакта, сформированного в скрытом слое двумерного электронного газа поверхностью электродом: (а) схематическое изображение квантового контакта; (б) изменение его проводимости

нообразные интерференционные эффекты, которые (наряду с другими особенностями транспорта носителей заряда вдоль потенциальных барьеров) рассматриваются в одном из последующих разделов данного учебного пособия.

1.3. Туннелирование носителей заряда

Термин **туннелирование** (*tunneling*) означает перенос частицы через область, ограниченную потенциальным барьером, высота которого больше полной энергии данной частицы (или проникновение в эту область). Такой эффект невозможен с точки зрения классической механики, однако имеет место для квантовых частиц, которым, как известно, присущ корпускулярно-волновой дуализм. Волновые свойства квантовых частиц приводят и к другому, аномальному с точки зрения классической механики, эффекту — надбарьерному отражению. Основные закономерности этих явлений рассмотрены ниже.

Взаимодействие квантовых частиц с различными потенциальными барьерами иллюстрирует рис. 1.5.

Согласно классической механике, частица с энергией $E < U_0$, движущаяся слева направо, то есть приближающаяся к потенциальному

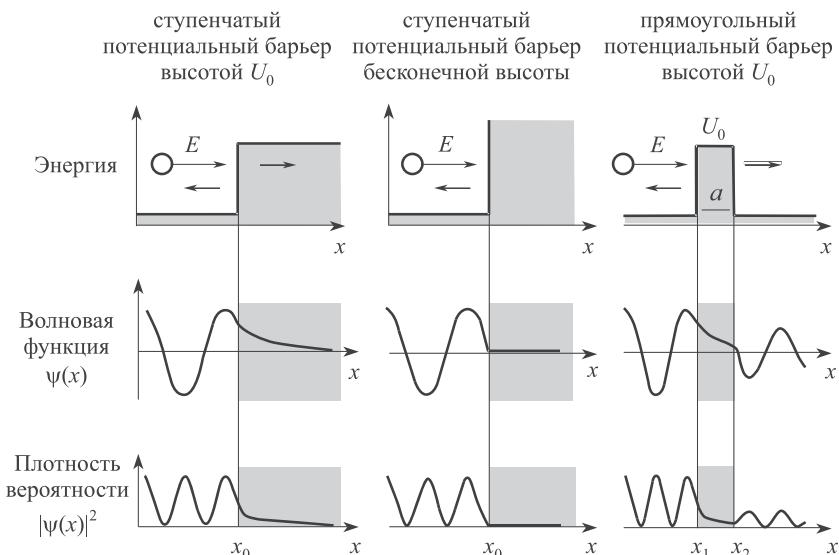


Рис. 1.5. Взаимодействие квантовой частицы с полной энергией E со ступенчатым барьером конечной высоты U_0 , с бесконечно высоким потенциальным барьером и с прямоугольным барьером высотой U_0 и шириной a

барьеру, отразится от него и начнет двигаться в обратном направлении. Если же $E > U_0$, то частица продолжит движение в прежнем направлении. В квантовой механике картина иная.

На языке квантовой механики движение частицы в одномерном потенциальном поле $U(x)$ описывается уравнением Шредингера:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + U(x)\psi(x) = E\psi(x), \quad (1.15)$$

где m — масса частицы, $\psi(x)$ — ее волновая функция.

Если энергия квантовой частицы больше высоты ступенчатого барьера ($E > U_0$), то в области перед потенциальным барьером, где $U(x) = 0$, решение уравнения Шредингера для частицы с импульсом $k_1 = (1/\hbar)\sqrt{2mE}$ имеет вид суперпозиции двух волн:

$$\psi = A'\exp(i k_1 x) + B\exp(-ik_1 x), \quad (1.16)$$

где A' , B — константы, i — мнимая единица. Волновую функцию обычно нормируют таким образом, что $A' = 1$.

Первый член в (1.16) соответствует падающей на барьер волне, движущейся вдоль оси x слева направо. Второй член описывает отраженную волну, движущуюся вдоль оси x в противоположном направлении.

При $x \rightarrow \infty$ волновая функция прошедшей над потенциальным барьером частицы имеет асимптотический вид:

$$\psi = A\exp(ik_2x), \quad (1.17)$$

где $k_2 = (1/\hbar)\sqrt{2m(E - U_0)}$. Плотность потока в падающей волне пропорциональна k_1 , в отраженной — $k_1|B|^2$, а в прошедшей — $k_2|A|^2$. Коэффициент прохождения частицы через границу потенциального барьера, определяемый как отношение плотности потока в прошедшей волне к плотности потока в падающей волне, равен:

$$T(E) = \frac{k_2}{k_1} |A|^2. \quad (1.18)$$

Коэффициент отражения частицы от потенциального барьера, определяемый отношением плотности отраженного потока к плотности падающего потока, равен:

$$R(E) = |B|^2 = 1 - \frac{k_2}{k_1} |A|^2. \quad (1.19)$$

Очевидно, что $R(E) = 1 - T(E)$.

Постоянныи A и B , определенные из условия непрерывности волновой функции и ее первой производной при $x = x_0$, равны:

$$A = \frac{2k_1}{k_1 + k_2}, \quad B = \frac{k_1 - k_2}{k_1 + k_2}. \quad (1.20)$$

В соответствии с (1.20) коэффициенты отражения и прохождения равны:

$$R(E) = \left(\frac{k_1 - k_2}{k_1 + k_2} \right)^2, \quad T(E) = \frac{4k_1 k_2}{(k_1 + k_2)^2}. \quad (1.21)$$

Из выражений (1.21) следует, что при $E = U_0$ ($k_2 = 0$) коэффициент отражения R обращается в единицу, а коэффициент прохождения — в нуль. С ростом энергии частицы коэффициент отражения уменьшается и стремится к нулю при $E \rightarrow \infty$, так как $R \approx (U_0/4E)^2$ при $E \gg U_0$.

Если квантовая частица движется над прямоугольным потенциальным барьером высотой U_0 и конечной толщиной $a = x_2 - x_1$, то решение уравнения Шредингера для каждой из трех областей имеет вид:

$$\begin{aligned} \psi_1 &= \exp(ik_1 x) + C \exp(-ik_1 x), \quad x < x_1, \\ \psi_2 &= D \exp(ik_2 x) + F \exp(-ik_2 x), \quad x_1 < x < x_2, \\ \psi_3 &= G \exp(ik_1 x), \quad x > x_2, \end{aligned} \quad (1.22)$$

где волновые векторы k_1 и k_2 определены выше, а C, D, F, G — константы. В выражениях (1.22) члены $\exp(ik_1 x)$ и $C \exp(-ik_1 x)$ описывают, соответственно, падающую и отраженную волны, а $G \exp(-ik_1 x)$ — прошедшую волну. Постоянныи C, D, F, G определяются из условий непрерывности волновой функции и ее первой производной в точках $x = x_1$ и $x = x_2$.

Коэффициент прохождения частицы в данном случае определяется как $T(E) = |G|^2$, что приводит к выражению:

$$T(E) = \frac{4k_1^2 k_2^2}{(k_1^2 - k_2^2)^2 \sin^2(ak_2) + 4k_1^2 k_2^2}. \quad (1.23)$$

Максимум коэффициента прохождения, $T(E) = 1$, достигается для частиц с энергиями:

$$E = U_0 + \frac{\pi^2 \hbar^2}{2ma^2} n^2, \quad \text{где } n = 0, 1, 2, 3 \dots. \quad (1.24)$$

При других значениях энергии наблюдается частичное отражение падающих на барьер частиц. Таким образом, из квантовой теории следует, что даже в случае, когда энергия падающей на потенциальный барьер частицы больше высоты этого барьера, коэффициент ее отражения может быть отличен от нуля. Этим квантовая частица отличается от классической, для которой никакого отражения в подобной ситуации быть не может.

Рассмотрим теперь другой практически важный случай, когда квантовая частица взаимодействует с прямоугольным потенциальным барьером шириной a , высота которого больше ее энергии ($E < U_0$). Классическая частица не может пройти через такой барьер. Она будет отражаться в так называемых классических точках поворота. *Точка поворота* (*turning point*) — это точка с координатой x , в которой кинетическая энергия частицы обращается в нуль, то есть ее полная энергия E равна $U(x)$. Для прямоугольного барьера точки поворота совпадают с координатами его границ (точки x_1 и x_2 на рис. 1.5). Достигнув точки поворота, частица меняет направление своего движения и начинает двигаться в обратном направлении.

Для квантовой частицы решение уравнения Шредингера в каждой из трех областей (перед, внутри и за барьером) имеет вид:

$$\begin{aligned}\psi_1 &= \exp(ik_1x) + B_1\exp(-ik_1x), \quad x < x_1, \\ \psi_2 &= A_2\exp(-\zeta x) + B_2\exp(\zeta x), \quad x_1 < x < x_2, \\ \psi_3 &= A_3\exp(ik_1x), \quad x > x_2,\end{aligned}\tag{1.25}$$

где волновой вектор k_1 определяется выражением, приведенным выше, $\zeta = \frac{1}{\hbar} \sqrt{2m(U_0 - E)}$, A_2 , A_3 , B_1 , B_2 — константы.

В выражениях (1.25) члены $\exp(ik_1x)$ и $B_1\exp(-ik_1x)$ описывают падающую и отраженную волну соответственно, а $A_3\exp(ik_1x)$ — волну, прошедшую через барьер. Существование прошедшей через барьер волны, отвечающей квантовой частице с энергией меньше высоты барьера, называют туннельным эффектом. Для его количественного описания используют коэффициент туннельной прозрачности барьера, под которым понимают модуль отношения плотности потока частиц, прошедших через барьер, к плотности потока падающих частиц. В нашем случае этот коэффициент равен:

$$T(E) = \frac{4k_1^2 \zeta^2}{(k_1^2 + \zeta^2)^2 \operatorname{sh}^2(a\zeta) + 4k_1^2 \zeta^2}.\tag{1.26}$$

Если выполняется условие $a\zeta \gg 1$, то выражение для коэффициента туннельной прозрачности упрощается:

$$T(E) \cong T_0 \exp\left(-\frac{2a}{\hbar} \sqrt{2m(U_0 - E)}\right), \quad (1.27)$$

где $T_0 = 16E(U_0 - E)/U_0^2$.

Прозрачность потенциального барьера произвольной формы можно оценить с помощью выражения:

$$T(E) \cong \exp\left(-\frac{2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m[U(x) - E]} dx\right), \quad (1.28)$$

где x_1 и x_2 — точки поворота, определяемые из условий $E = U(x_1) = U(x_2)$. Коэффициент отражения при этом определяется с использованием соотношения $R(E) = 1 - T(E)$.

Одним из практически важных случаев является взаимодействие квантовых частиц с потенциальным барьером, имеющим близкую к δ -функции форму. Такие барьеры характеризуются произведением aU_0 . Для них

$$T(E) = \left(1 + \frac{2\pi^2 ma^2 U_0^2}{\hbar^2 E}\right)^{-1}. \quad (1.29)$$

Коэффициент прохождения δ -образного барьера с ростом энергии частицы увеличивается квазилинейно в области относительно низких энергий, а потом выходит на константу, равную единице в области высоких энергий. Рис. 1.6 качественно иллюстрирует изменение коэффициента прохождения барьера прямоугольной формы и барьера, описываемого δ -функцией.

Туннелирование электронов является достаточно общим явлением для твердотельных структур. В низкоразмерных структурах это явление приобретает специфические особенности, отличающие его от эффектов в объемных системах. Одна из таких особенностей связана с дискретной природой переносимого электронами заряда и обнаруживает себя в явлении, которое получило название «одноэлектронное туннелирование». Другая особенность определяется дискретностью энергетических состояний носителей заряда в полупроводниковыхnanoструктурах с квантовыми колодцами, которая возникает из-за квантового ограничения. Туннельный перенос носителей заряда через потенциальный барьер с определенного уровня в эмиттирующей об-

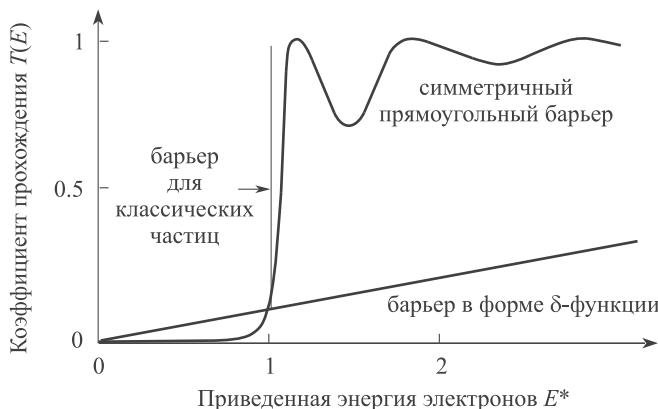


Рис. 1.6. Коэффициент прохождения электронами различных потенциальных барьеров как функция приведенной энергии E^* , которая для прямоугольного барьера равна E/U_0 , а для δ -образного барьера — E/U_0^* ,
где $U_0^* = ma^2U_0^2/h^2$

ласти на энергетически эквивалентный ему уровень в квантовом колодце происходит с сохранением энергии и импульса электрона. Такое совпадение уровней приводит к возрастанию туннельного тока, известному как эффект резонансного туннелирования. Более того, вnanoструктурах, содержащих магнитные и немагнитные материалы, определенная спиновая поляризация туннелирующих электронов приводит к дополнительным эффектам. Все эти явления находят широкое применение в наноэлектронных приборах. Их рассмотрение дано в последующих разделах данного пособия.

1.4. Спиновые эффекты

Спин, будучи одной из фундаментальных характеристик электрона, приводит к появлению новых особенностей транспорта носителей заряда в nanoструктурах. Спиновые эффекты возникают, когда в материале появляется спиновый дисбаланс заселенности уровня Ферми. Такой дисбаланс обычно присутствует в ферромагнитных материалах, у которых плотности вакантных состояний для электронов с различными спинами практически идентичны, однако эти состояния существенно различаются по энергии, как схематически показано на рис. 1.7 (здесь и далее под различными спинами электрона понимаются различные проекции его спина на ось квантования). Энергети-

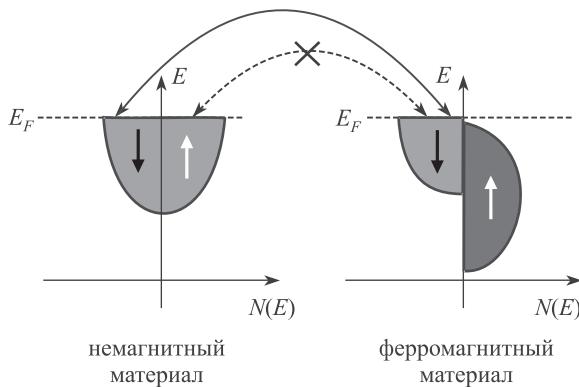


Рис. 1.7. Плотности состояний электронов с различными спинами в немагнитном и ферромагнитном материале и обмен электронами между ними

ческий сдвиг приводит к заполнению разрешенных зон электронами с одним определенным спином и к появлению у материала собственного магнитного момента (намагченности). Заселенность энергетических зон электронами с разным спином определяет как спиновую поляризацию инжектируемых из такого материала электронов, так и особенности транспорта носителей заряда через него.

Собственную спиновую поляризацию электронов в материале (P) определяют как отношение разности концентраций электронов с различными спинами (n_{\uparrow} и n_{\downarrow}) к их общей концентрации:

$$P = \frac{n_{\uparrow} - n_{\downarrow}}{n_{\uparrow} + n_{\downarrow}}. \quad (1.30)$$

Наиболее яркое проявление спиновых эффектов резонно ожидать в материалах с наибольшей спиновой поляризацией электронов. Это стимулирует поиск материалов со 100 %-й спиновой поляризацией. На практике же пока используются материалы с частичной спиновой поляризацией. Это — металлы и их сплавы, оксиды, магнитные полупроводники. Их примеры приведены в табл. 1.2. Анализируя экспериментальные данные по спиновой поляризации, следует иметь в виду, что этот параметр чувствителен к технологии изготовления материала, к его структуре и наличию в нем примесей.

Электрический ток в твердотельных структурах, составленных из материалов с различной спиновой поляризацией, зависит от спиновой поляризации носителей заряда и спиновой поляризации областей, через которые эти носители движутся. Электроны, инжектированные с

Таблица 1.2. Максимальная спиновая поляризация электронов проводимости в различных материалах

Материал	Co	Fe	Ni	Ni ₈₀ Fe ₂₀ *	CoFe	NiMnSb	CrO ₂
Поляризация, %	42	46	46	45	47	58	90

Материал	La _{0.7} Sr _{0.3} MnO ₃	Be _{0.07} Mn _{0.03} Zn _{0.9} Se	Zn _{0.94} Mn _{0.06} Se
Поляризация, %	78	90 (при $T \approx 5$ К, $H = 1.5$ Тл)	70 (при $T = 4.2$ К, $H = 4 - 5$ Тл)

* Сплав Ni₈₀Fe₂₀ называют пермаллоем.

определенным спином, могут занять в коллекторе только вакантные места с такой же ориентацией спина. Электрон, первоначально спин-поляризованный в инжектирующем электроде, по мере движения изменяет (из-за рассеяния) как свой импульс, так и свой спин. Для практических применений важно знать, как долго электрон «помнит» свою спиновую ориентацию. В качестве характеристики «спиновой памяти» используют среднее расстояние, проходимое электроном до изменения своего спина. Его называют *длиной спиновой релаксации* (*spin relaxation length*), l_s . В твердых телах ее величина превышает 100 нм и определяется спин-независимым средним свободным пробегом электронов, в качестве которого целесообразно рассматривать среднюю длину свободного пробега при неупругом рассеянии l_{in} , поскольку при изменении направления спина рассеянного электрона баланс по импульсу в системе взаимодействующих частиц не сохраняется. Тогда $l_s = (l_{in} v_F \tau_{\uparrow\downarrow})^{1/2}$, где v_F — скорость Ферми, $\tau_{\uparrow\downarrow}$ — время релаксации спина. Длина спиновой релаксации определяется, главным образом, процессами спин-орбитального и обменного рассеяния. При идентичном составе материала в кристаллах она больше, чем в аморфной фазе.

В спин-поляризованных материалах состояния с преобладающим спином контролируются намагниченностью этих материалов. Если намагниченность изменяется на противоположную, то преобладающая ориентация спинов также меняется на противоположную. При инъекции спин-поляризованных электронов в материал с отличной от нуля намагниченностью, а следовательно, и со спиновой поляризацией, контролируемой внешним магнитным полем, этот материал может вести себя как проводник или как изолятор — в зависимости от направления намагниченности и ориентации спинов инжектированных электронов.

При одинаковом направлении спинов инжектированных электронов и электронов материала обеспечивается наивысшая проводимость. Противоположное направление спинов препятствует прохождению электронов через материал.

Два основных эффекта, а именно — *гигантское магнитосопротивление* (*giant magnetoresistance*) и *туннельное магнитосопротивление* (*tunneling magnetoresistance*) — являются следствием особенностей транспорта носителей заряда, контролируемого спином электронов в наноразмерных структурах. Эти эффекты составляют основу нового направления в науке и технике, которое получило название *спинtronика* (*spintronics*). Целью спинtronики является создание элементов электронной обработки информации с использованием в качестве носителей информации как заряда электрона, так и его спина. Возможность контролировать и управлять спиновыми состояниями в твердых телах представляет также значительный интерес для практической реализации идей *квантовых вычислений* (*quantum computation*), которые обещают революционный прогресс в развитии информационных систем. Детальное рассмотрение упомянутых спиновых эффектов и примеры спинtronных приборов даны в последующих разделах.

2. Элементы низкоразмерных структур

Природа потенциальных барьеров, используемых для создания низкоразмерных структур, неизбежно влияет на свойства этих структур. Реальные границы раздела привносят дополнительные квантовые эффекты, которые следует учитывать при конструировании наноэлектронных, оптоэлектронных и оптических приборов. Рассмотрим электронные свойства границ раздела и комбинированных структур, включающих низкоразмерные элементы.

2.1. Свободная поверхность и межфазные границы

Практическая реализация квантового ограничения и связанных с ним эффектов требует пространственной локализации электронов. В твердых телах это может быть достигнуто путем использования свободных поверхностей и межфазных границ.

Свободная поверхность (*free surface*) любого твердого тела представляет собой естественный потенциальный барьер. Разрешенные энергетические состояния электронов при переходе через поверхность изменяются скачкообразно. Высота и пространственная конфигурация такого барьера определяются расположением ато-

мов твердого тела на поверхности и вблизи нее, а в особенности — чужеродными (примесными) атомами и молекулами, адсорбированными на поверхности. Поверхностные свойства кристаллов, находящихся в вакууме, определяются несколькими приповерхностными моноатомными слоями, которые по своей атомной конфигурации отличаются от таковых в объеме кристалла. Поверхностный слой кристалла, свободный от адсорбированных частиц, обычно подвержен структурным изменениям, которые называют *реконструкцией* (*reconstruction*). В некоторых редких случаях, однако, он может оставаться *нереконструированным* (*nonreconstructed*).

Реконструкция поверхности является следствием перестройки оборванных связей поверхностных атомов. Поверхностные атомы с разорванными ковалентными или ионными связями могут сгруппироваться в ряды с межатомными расстояниями больше или меньше, чем в объеме. Ближайшие соседние поверхностные атомы сближаются для образования связей за счет своих незадействованных валентных электронов и таким образом понижают энергию системы. При этом происходит реконструкция поверхности, характеризующаяся измененными (по отношению к объему) позициями атомов и образованием новой элементарной ячейки. Особенности реконструированной поверхности влияют на ее электронные свойства и на последующее эпитаксиальное осаждение на нее других материалов.

На нереконструированных поверхностях расположение атомов сохраняется таким же, как и в объеме, хотя расстояние между верхними моноатомными слоями изменяется. Это расстояние уменьшается, что находит объяснение в рамках представлений о двухатомной молекуле. Следуя этим представлениям, расстояния между атомами в приповерхностном слое характеризуются величинами, промежуточными между межатомными расстояниями в объеме кристалла и в свободной двухатомной молекуле, состоящей из тех же атомов, что и кристалл. Поскольку межатомное расстояние в двухатомной молекуле меньше, чем в кристалле, то существует движущая сила для релаксации атомов на поверхности. В процессе такой релаксации в плоскости поверхности атомы сохраняют свое взаимное расположение таким, как оно проецируется из объема, но расстояние между атомными плоскостями уменьшается.

Адсорбированные на поверхности атомы и молекулы образуют связи, нетипичные для объема твердого тела. В результате атомная структура и, соответственно, электронные свойства приповерхностного слоя приобретают существенные отличия от свойств, типичных как для объема твердого тела, так и для его реконструированной или

нереконструированной поверхности. Когда две свободные поверхности располагаются близко друг к другу (как, например, в квантовой пленке или квантовом шнуре), модифицированные приповерхностные слои могут перекрываться и кардинально изменять свойства таких структур.

Точный контроль потенциального барьера у свободной поверхности и управление им затруднены из-за неконтролируемой адсорбции примесей. Более того, достаточно сложно реализовать инжекцию носителей заряда через такой барьер, что необходимо для приборных применений. Как следствие этого, свободные поверхности следует рассматривать в качестве одного из важнейших элементов низкоразмерных структур, хотя их непосредственное использование в приборных структурах обычно ограничено пассивными функциями.

Межфазные границы (interfaces) образуются между материалами с различными физическими свойствами. В случае полупроводников, из всех возможных комбинаций монокристаллической, поликристаллической и аморфной фаз, граница между двумя монокристаллическими областями имеет наиболее управляемые и воспроизводимые характеристики. Для того чтобы получить потенциальный барьер на такой границе, должно быть удовлетворено одно из следующих требований. Если контактирующие полупроводники имеют одинаковый химический состав, они должны отличаться типом основных носителей заряда, а при одинаковом типе основных носителей заряда их концентрации должны быть существенно различны. Полупроводники же с разным химическим составом должны иметь близкие, а в идеальном случае — совпадающие параметры решеток.

Потенциальные барьеры, образованные на границе двух материалов с одинаковым химическим составом, обычно являются широкими и гладкими. Это — результат диффузационного перераспределения примесей, придающих определенные донорные или акцепторные черты контактирующим областям. Взаимная диффузия компонентов на границе материалов с различным химическим составом обычно ограничена несколькими монослоями (один моносвой — это самый тонкий слой, содержащий полный стехиометрический набор атомов данного материала). Для таких границ характерны резкие, ступенчатые потенциальные барьеры. Они образуются в гетероэпитаксиальных структурах полупроводников, создаваемых из бинарных, тройных и четверных соединений групп $A^{III}B^V$ и $A^{II}B^VI$. Подходящие пары материалов в данном случае определяют требование согласованности кристаллических решеток. Формируются такие структуры в виде сверхрешеток.

2.2. Сверхрешетки

Монокристаллическую пленку из одного материала, воспроизведенную постоянную решетки монокристаллической подложки из другого материала, называют *сверхрешеткой* (*superlattice*)². Когда оба материала имеют идентичные или очень близкие постоянные решеток, они образуют так называемые *псевдоморфные* (*pseudomorphic*) сверхрешетки. Среди полупроводников таких материалов очень мало. Между тем равенство постоянных решеток не является строго необходимым условием для псевдоморфного роста одного материала на другом.

В пределах некоторой ограниченной толщины наносимой пленки возможно «заставить» осаждаемые атомы занимать позиции, соответствующие расположению атомов в подложке, даже если это расположение отличается от равновесного расположения атомов в объемном материале пленки. При этом образуется *напряженная* (*strained*) сверхрешетка, структура которой, однако, совершенна. Формирование напряженной сверхрешетки в случае, когда постоянная решетки у материала подложки меньше, чем у материала пленки, схематически проиллюстрировано на рис. 1.8. Напряжения в такой пленке возрастают по мере увеличения ее толщины. По достижении некоторой критической толщины они релаксируют посредством образования дислокаций несоответствия, высвобождая накопленную в напряженном состоянии энергию и понижая полную энергию системы. Кристаллическая решетка наносимого материала приобретает свой естественный вид, и при дальнейшем поступлении материала на подложку пленка растет с уже *релаксированной* (*relaxed*) решеткой. Критическая толщина пленки зависит от величины рассогласования постоянных решеток и модулей упругости материалов пленки и подложки при температуре осаждения. В принципе, не превышая критической толщины, можно сформировать напряженную сверхрешетку из любого полупроводника на подложке с тем же типом кристаллической решетки.

² В Физической энциклопедии под редакцией академика А. М. Прохорова (Москва, 1988, Советская энциклопедия), а следом — и в некоторых других изданиях, *сверхрешетка* определяется как «*твёрдотельная периодическая структура, в которой на носители заряда (электроны), помимо обычного потенциала кристаллической решетки действует дополнительный, как правило одномерный, потенциал с периодом меньше длины свободного пробега электронов, но значительно большими периода основной решетки*». Данное определение скорее говорит об электронных свойствах многoperiodных сверхрешеток, нежели об их структуре, поскольку не дает ключа к пониманию того, почему сверхрешетки бывают псевдоморфными, напряженными и релаксированными (и не охватывает случая однопериодных сверхрешеток).

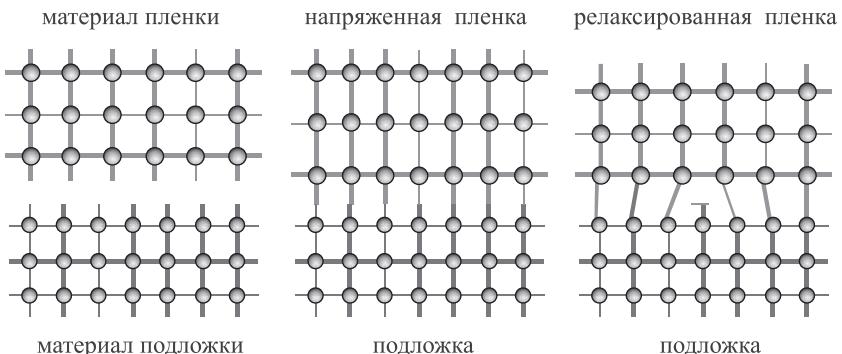


Рис. 1.8. Образование напряженной и релаксированной эпитаксиальной пленки

На практике для создания качественных потенциальных барьеров на сверхрешетках желательно наименьшее рассогласование параметров решеток пленки и подложки, а также определенное различие их электронных свойств. Среди полупроводников имеется достаточный выбор материалов, отвечающих этим требованиям. На рис. 1.9 представлены данные по ширине запрещенной зоны полупроводников со структурой алмаза и цинковой обманки и постоянные их решеток. Затененные вертикальные области объединяют группы полупроводников с близкими постоянными решетками. Материалы внутри этих групп обеспечивают создание псевдоморфных сверхрешеток и гетеропереходов на их основе; высота потенциального барьера на границе при этом определяется различием ширины запрещенной зоны. Полупроводники, соединенные сплошными линиями, образуют стабильные промежуточные соединения. Примером таких двойных соединений служит SiGe, тройных — AlGaAs, а четверых — GaInAsP.

Среди представленных материалов выделяются полупроводниковые нитриды. Они имеют другую решетку (гексагональную) и маленькие постоянные решетки, что требует использования специальных подложек для их эпитаксиального роста.

Постоянная решетки промежуточного соединения $a(x)$ линейно изменяется в интервале между постоянными решеток образующих его материалов a_1 и a_2 . Это — **правило Вегарда** (*Vegard's law*). Согласно ему

$$a(x) = x a_1 + (1 - x) a_2, \quad (1.31)$$

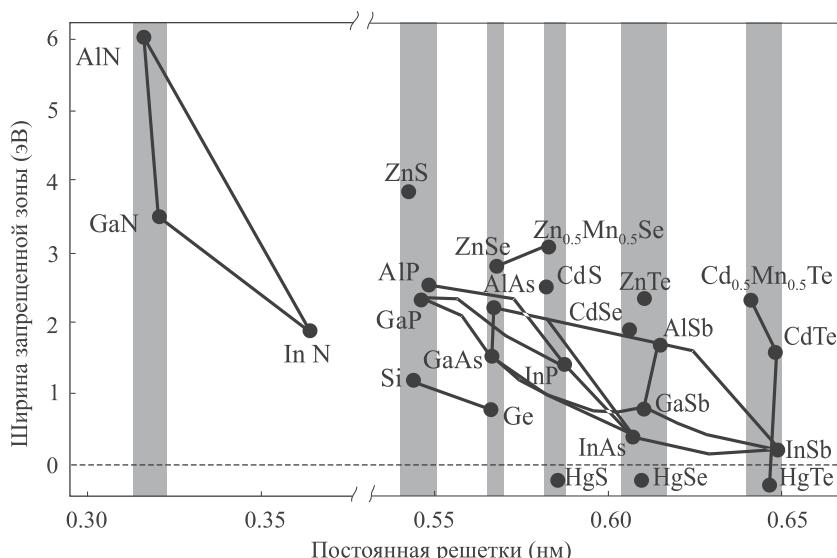


Рис. 1.9. Ширина запрещенной зоны при низкой температуре и постоянная решетки для полупроводников с кубической кристаллической структурой типа цинковой обманки и алмаза. Для сравнения представлены также гексагональные нитриды (приведены их постоянные решетки a)

где через x обозначена молярная доля материала 1 в материале 2. Промежуточные полупроводниковые соединения значительно увеличивают количество материалов для формирования согласованных сверхрешеток.

Так, например, постоянная решетки тройного соединения $\text{Al}_x\text{Ga}_{1-x}\text{As}$ равна $x a_{\text{AlAs}} + (1 - x)a_{\text{GaAs}}$. При изменении x от 0 до 1 она изменяется менее чем на 0,15%. Это позволяет выращивать из AlAs , GaAs или $\text{Al}_x\text{Ga}_{1-x}\text{As}$ сверхрешетки любого состава практически без напряжений. Кроме того, различие ширины запрещенной зоны материалов данной группы достигает 0,8 ЭВ, что (наряду со структурной совместимостью) и определяет их широкое применение для «зонной инженерии» при созданииnano- и оптоэлектронных приборов.

Как псевдоморфные, так и напряженные сверхрешетки, изготовленные путем многократного эпитаксиального осаждения различных по составу полупроводников, используют для формирования встроенных квантовых колодцев, в которых электроны и/или дырки испытывают квантовое ограничение.

2.3. Моделирование атомных конфигураций

Так как размеры наноструктур сопоставимы с межатомными расстояниями, то расположение атомов в них подвержено значительному влиянию поверхности, границ раздела и внутренних точечных дефектов. Экспериментальное определение связанный с этим релаксации кристаллической решетки крайне затруднено, поскольку требует использования методов анализа с атомным разрешением. Поэтому эффективным методом исследования и описания атомных конфигураций наноразмерных структур является их теоретическое моделирование с использованием современной вычислительной техники.

Поиск оптимального расположения атомов в наноразмерной структуре с учетом их взаимодействия друг с другом и влияния на них внешних сил осуществляется путем нахождения такого взаимного расположения, которое соответствует минимальной энергии всей структуры. Для решения этой задачи могут быть использованы два принципиально различных подхода. В их основе лежат, соответственно, представления о твердом теле как о механической и о квантово-химической системе. В первом случае используются методы *молекулярной динамики* (*molecular dynamics*) и *молекулярной механики* (*molecular mechanics*), во втором — методы квантовой химии. В данном разделе мы ограничимся рассмотрением теоретических основ лишь первой группы методов. Квантово-химические подходы будут описаны в последующих разделах, где мы рассмотрим фундаментальные электронные свойства наноразмерных структур. Именно для моделирования электронных свойств твердых тел они главным образом и разрабатывались.

В *методе молекулярной динамики* поиск оптимальной атомной конфигурации осуществляется путем пошагового движения атомов в положения, обеспечивающие минимум потенциальной энергии анализируемой структуры. Рассматривается конечная область кристалла — кристаллит. Движение атомов описывается уравнениями классической механики. Взаимодействие атомов с поверхностью или границей раздела имитируется с помощью направленных внутрь рассматриваемого объема постоянных сил, действующих только на атомы в этой граничной области. Расчет атомной конфигурации начинается с задания предполагаемых координат атомов. Затем, меняя положение атомов, находят конфигурацию, соответствующую минимуму потенциальной энергии. При этом решается система связанных уравнений движения для каждого атома,

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t), \quad (1.32)$$

$$\frac{d\mathbf{v}_i(t)}{dt} = \frac{1}{M} \mathbf{F}_i[\mathbf{r}_1(t), \dots, \mathbf{r}_N(t); \mathbf{v}_i(t)], \quad (1.33)$$

где $\mathbf{r}_i(t)$ — координата i -го атома; $\mathbf{v}_i(t)$ — его скорость; N — общее количество атомов; M — масса атома. Сила \mathbf{F}_i , действующая на i -й атом, является результатом взаимодействия этого атома с его соседями. Она определяется потенциалом многочастичного взаимодействия; однако, если последний неизвестен, используют потенциалы парного взаимодействия ϕ_{ij} между i -м атомом и всеми другими j -ми атомами:

$$\mathbf{F}_i = \sum_j \mathbf{F}_{ij}, \quad (1.34)$$

где сила, действующая на i -й атом со стороны j -го атома, равна

$$\mathbf{F}_{ij} = \frac{d\phi_{ij}}{dr_{ij}} \frac{\mathbf{r}_{ij}}{r_{ij}}. \quad (1.35)$$

Здесь $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. Потенциал парного взаимодействия ϕ_{ij} может быть задан в форме Ленарда–Джонса, Борна–Мейера, Томаса–Ферми, экранированного кулоновского потенциала или любой другой форме, адекватно описывающей особенности межатомного взаимодействия в моделируемой системе. Следует иметь в виду, что полный потенциал межатомного взаимодействия очень редко сводится к парным взаимодействиям, по крайней мере, в твердых телах. Поэтому использование потенциалов многочастичного взаимодействия дает более точный результат.

Решение системы уравнений (1.32), (1.33) с учетом (1.34) и (1.35) возможно только путем численного интегрирования. При этом следует учитывать следующее:

1. Метод интегрирования нужно выбирать с учетом размера моделируемой области, желательной точности расчетов и вычислительной мощности используемого компьютера.
2. Для получения объективного результата важно выбрать корректное описание потенциала межатомного взаимодействия в конкретной моделируемой структуре.
3. Учет тепловых колебаний атомов позволяет оценить динамическую стабильность найденных равновесных атомных конфигураций.
4. «Сшивка» границ выделенного кристаллита с остальным объемом должна быть корректной. Для описания гетерофазных границ необходимо использовать периодические граничные условия в соответствующих направлениях.
5. В процессе моделирования необходимо уменьшать скорости движения всех атомов на каждой итерации. В противном случае

атомы будут просто колебаться, но не релаксировать к положениям равновесия.

Процедура динамической релаксации моделируемого ансамбля атомов контролируется межатомными потенциалами и законами ньютонаской механики, что дает возможность отыскать такое расположение атомов, которое соответствует состоянию с наименьшей потенциальной энергией.

В основе *метода молекулярной механики* лежит процедура статической релаксации. Анализируется полная потенциальная энергия системы, которая рассчитывается для различных атомных конфигураций. Конфигурация с наименьшей потенциальной энергией считается равновесной.

Если вычислительные возможности используемого компьютера не позволяют моделировать поведение структуры в целом, то для анализа выделяется часть исследуемой структуры — кристаллит. Атомы внутри кристаллита рассматриваются как взаимодействующие друг с другом частицы. Между ними задается центральное парное взаимодействие, характеризуемое потенциалом ϕ_{ij} . При этом учитываются соседи из ближайших координационных сфер. Как отмечено выше, описание атомной структуры твердых тел с использованием парного потенциала является приближением к реальному многочастичному взаимодействию. Кроме этого, соседи только из ближайших координационных сфер учитываются «как правило», но не всегда.

Точность предсказания равновесной атомной конфигурации методом молекулярной механики во многом определяется интуицией и опытом исследователя, поскольку только выбор оптимального алгоритма изменения координат атомов позволяет за разумное время найти их расположение, отвечающее глобальному (а не локальному) минимуму потенциальной энергии. Расчет потенциальной энергии может быть осуществлен как в рамках классической механики, так и с использованием квантово-механических методов. При этом применение процедуры Монте-Карло делает результат статистически более надежным.

Одной из важнейших проблем практического применения методов молекулярной динамики и механики является ограничение на количество атомов в анализируемой структуре. Оно связано с ограниченной вычислительной мощностью используемого компьютера. Анализ взаимодействия N атомов со своими $N - 1$ соседями требует нахождения $N(N - 1)/2$ парных взаимодействий. Если не используется никакой специальный алгоритм, то время, необходимое для расчетов, пропорционально N^2 . Это не является большой проблемой для атомных

кластеров (нульмерных структур), включающих до нескольких тысяч атомов. Но в случае квантовых шнурков и квантовых пленок мощность используемого компьютера будет ограничивать максимальный размер структуры, для которой проводится моделирование.

Существуют приемы, позволяющие преодолевать указанные ограничения на размер моделируемых структур. Один из них заключается в использовании периодических граничных условий. При этом из структуры «вырезается» представительная область, состояния на границах которой описываются периодическими граничными условиями. Это эквивалентно рассмотрению бесконечного пространства, заполненного идентичными копиями моделируемой области. Такая периодичность имеет два важных следствия. Во-первых, любой атом, покидающий моделируемую область через границу, неизбежно входит в нее же с противоположной стороны. Во-вторых, атомы, расположенные на границе, взаимодействуют с атомами в соседней копии моделируемой области, что эквивалентно взаимодействию между атомами, находящимися у противоположных сторон самой моделируемой области — так называемый циклический эффект. Этот эффект необходимо учитывать при расчете межатомных взаимодействий и численном интегрировании уравнений движения.

Другой подход предполагает замену внешней (пассивной) области моделируемой структуры упругой средой с соответствующими интегральными свойствами. Деформация упругой среды и находящихся в ней атомов описывается сферически симметричным векторным полем

$$\mathbf{u} = C \frac{\mathbf{r}}{r^3}, \quad (1.36)$$

где C — параметр, характеризующий среду.

В таком приближении энергия кристаллита E равна

$$E = \frac{1}{2} \sum_i \sum_{j \neq i} \varphi_{ij} + aC + bC^2, \quad (1.37)$$

где член aC описывает работу против сил, удерживающих совершенную решетку в равновесии, а слагаемое bC^2 представляет собой энергию, запасенную упругим полем. Параметры C , a и b определяются из условия равенства нулю результирующей силы \mathbf{F}_Σ , действующей со стороны упругой среды на область внутри нее:

$$\mathbf{F}_\Sigma = -\frac{dE}{dC} = \sum_i \sum_k \mathbf{F}_{ik} \frac{d\mathbf{r}_{ik}}{dC} - a - 2bC = 0. \quad (1.38)$$

Индексы i и k относятся, соответственно, к атомам вне и внутри среды. После введения и описания внешней упругой среды к внутренней выделенной области можно применить процедуру динамической или статической релаксации. Рассмотренные приемы следует использовать с известной долей осторожности, поскольку они не полностью исключают эффекты, связанные с конечными размерами моделируемой структуры.

Сравнивая процедуры динамической и статической релаксации, следует подчеркнуть, что динамическая модель дает возможность атомам практически сразу занимать положения, отвечающие наименьшей потенциальной энергии. Процедура же статической релаксации позволяет определить, какая атомная конфигурация имеет наименьшую энергию (а следовательно, является наиболее устойчивой), только вычислив энергии различных возможных метастабильных конфигураций. Прогресс в использовании обоих подходов неразрывно связан с повышением вычислительной мощности (быстродействия и объема оперативной памяти) используемых для этих целей компьютеров.

3. Структуры с квантовым ограничением за счет внутреннего электрического поля

Донорные и акцепторные примеси в полупроводниках, так же как и гетеропереходы, образованные различными полупроводниками или полупроводником и диэлектриком, неизбежно индуцируют локальное перераспределение зарядов. Возникающее в результате этого внутреннее электрическое поле используют для создания потенциальных барьеров, ограничивающих движение электронов в наноразмерных областях. Данный подход может быть охарактеризован как «электронное» формированиеnanoструктур. Среди структур с квантовым ограничением, создаваемым внутренним электрическим полем, наибольшее распространение получили квантовые колодцы, модуляционно-легированные структуры и дельта-легированные структуры.

3.1. Квантовые колодцы

Структура, состоящая из полупроводников с различной шириной запрещенной зоны (или полупроводника и диэлектрика), в которой наноразмерная область из материала с меньшей шириной запрещенной зоны находится между областями из материала с большей шири-

ной запрещенной зоны, ведет себя как **квантовый колодец** (*quantum well*)³ для подвижных носителей заряда. Материал с меньшей шириной запрещенной зоны образует собственно колодец, а соседние области создают потенциальные барьеры, играющие для этого колодца роль стенок. Повторение такой структуры в пространстве дает **периодические квантовые колодцы** (*multiquantum wells*). Классическим примером твердотельных квантовых колодцев служат сверхрешетки, изготовленные из полупроводников с различными электронными характеристиками. Однако свойствами квантовых колодцев обладают и наноструктуры из полупроводников, встроенных в диэлектрические матрицы (даже при отсутствии согласования их кристаллических решеток). Примером таких структур являются наноразмерные кластеры и слои кремния, встроенные в диоксид кремния.

Для построения энергетической диаграммы квантового колодца необходимо соответствующим образом соединить зоны проводимости и валентные зоны материала колодца и материала барьера. Алгоритм соединения дает правило Андерсона, основанное на использовании такой характеристики материала, как **сродство к электрону** (*electron affinity*) χ — энергии, необходимой для переноса электрона со дна зоны проводимости E_c в вакуум, где он может покинуть материал. Сродство к электрону практически не зависит от положения уровня Ферми (в отличие от работы выхода, которая отсчитывается от энергии Ферми и поэтому существенно зависит от степени легирования материала). На рис. 1.10 показано соединение зон на границе между узкозонным материалом А со сродством к электрону χ_A и широкозонным материалом В со сродством к электрону χ_B (для случая $\chi_A > \chi_B$).

Правило Андерсона (*Anderson's rule*) устанавливает, что для двух материалов, образующих гетеропереход, энергии электронных состояний в вакууме одинаковы. Это приводит к тому, что сдвиг между зонами проводимости оказывается равным $\Delta E_c = E_{cB} - E_{cA} = \chi_A - \chi_B$. Соответственно, сдвиг между валентными зонами ΔE_v может быть определен на основе приведенной диаграммы с учетом сродства к электрону и ширины запрещенной зоны в каждом материале. В табл. 1.3 приведены параметры некоторых полупроводников и диэлектриков, позволяющие строить энергетические зонные диаграммы образуемых ими гетеропереходов. При температурах выше абсолютного нуля рассогласование уровней Ферми в контактирующих материалах (если таковое имеет место) устраняется за счет перераспределения

³ В русскоязычной научной и учебной литературе термин **quantum well** иногда переводится как «квантовая яма».

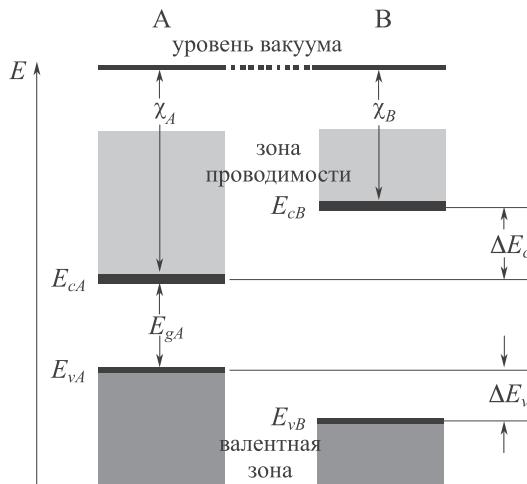


Рис. 1.10. Согласование энергетических зон на границе гетероперехода в соответствии с правилом Андерсона

ления свободных носителей заряда вблизи границы раздела между областью колодца и областью барьера.

В общем случае границы зон проводимости и валентных зон материалов, образующих квантовые колодцы, не совпадают. В зависимости от их взаимного расположения различают два основных типа периодических квантовых колодцев. Их энергетические диаграммы показаны на рис. 1.11, где предполагается, что материал А имеет меньшую запрещенную зону, чем материал Б.

Таблица 1.3. Ширина запрещенной зоны при комнатной температуре и сродство к электрону в некоторых полупроводниках и диэлектриках

Материал	Si	Ge	AlP	AlAs	AlSb	GaP	GaAs	GaSb	InP
$E_g, \text{ эВ}$	1,12	0,66	2,45	2,15	1,60	2,27	1,42	0,68	1,34
$\chi, \text{ эВ}$	4,05	4,13	3,44	3,51	3,65	3,08	4,07	4,06	4,38

Материал	InAs	InSb	ZnS	ZnSe	ZnTe	CdS	CdSe	CdTe
$E_g, \text{ эВ}$	0,35	0,18	3,58	2,67	2,26	2,42	1,7	1,44
$\chi, \text{ эВ}$	4,92	4,59	3,9	4,09	3,5	4,5	4,95	4,28

Материал	AlN	GaN	InN	SiO ₂	CaF ₂	Si ₃ N ₄	Al ₂ O ₃
$E_g, \text{ эВ}$	6,2	3,42	1,9	8,8	12,2	≈5	≈8
$\chi, \text{ эВ}$	0,6	3,3	5,8	1,1	1,5	1,8	1,35

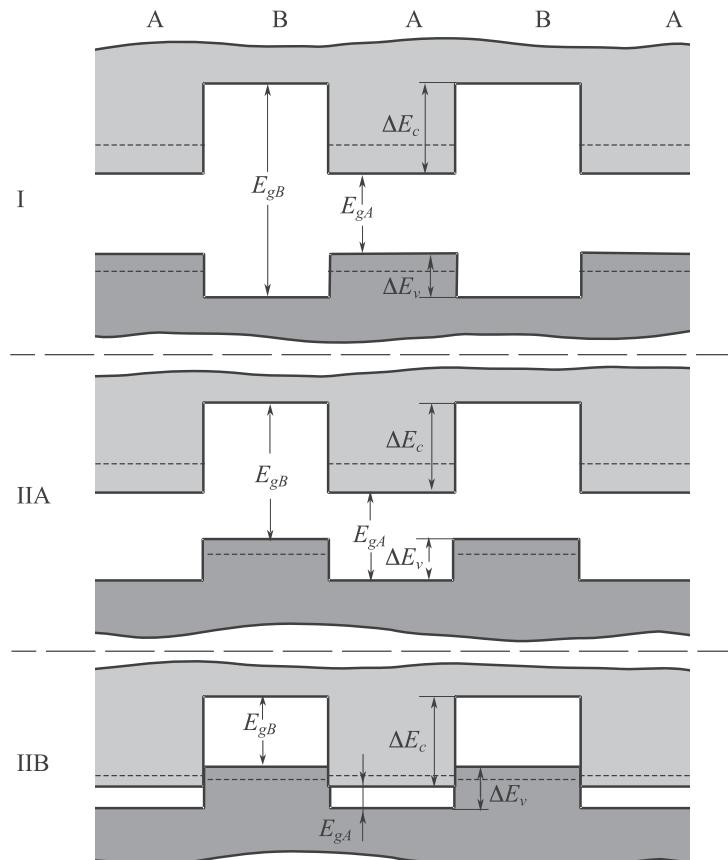


Рис. 1.11. Расположение энергетических зон в квантовых колодцах, образованных узкозонным материалом А и широкозонным материалом В. Пунктирными линиями обозначены минимальные уровни энергии электронов и дырок, разрешенные в соответствии с квантовым ограничением

В периодических квантовых колодцах типа I дно зоны проводимости широкозонного полупроводника располагается выше дна зоны проводимости узкозонного полупроводника. Для потолка валентной зоны эти соотношения обратны, т. е. потолок валентной зоны широкозонного полупроводника лежит ниже потолка валентной зоны узкозонного полупроводника. Такие энергетические соотношения приводят к тому, что и электроны, и дырки локализуются и испытывают квантовое ограничение геометрически в одной и той же области, а именно — в области узкозонного полупроводника, т. е. в колодце. Такую структуру можно назвать *пространственно прямозонной*.

Тип II периодических квантовых колодцев отличается от типа I тем, что при таких же энергетических соотношениях для дна зоны проводимости потолок валентной зоны широкозонного полупроводника находится выше потолка валентной зоны узкозонного полупроводника. Это приводит к тому, что электроны и дырки локализуются и испытывают квантовое ограничение в разных областях. Такой тип структур является *пространственно непрямозонным*.

В типе II выделяют также специфический случай, относящийся к узкозонным полупроводникам и полуметаллам. На рис. 1.11 он обозначен как тип IV. Для него характерно наличие очень малого энергетического зазора между минимальными уровнями энергии дырок и электронов в соседних областях. Встречаются также и квантовые колодцы типа III. Они образуются обычными и бесщелевыми полупроводниками. Однако это — достаточно редкий случай.

Квантовые колодцы являются одним из наиболее важных элементов большинства наноэлектронных и оптоэлектронных приборов.

3.2. Модуляционно-легированные структуры

Традиционный путь создания в полупроводниках областей с требуемым типом основных носителей заряда (электроны или дырки) предполагает легирование этих областей донорными или акцепторными примесями. Когда при температуре выше абсолютного нуля электроны или дырки покидают примесные атомы, они оставляют их в ионизированном состоянии. При этом по мере увеличения концентрации образующихся свободных носителей заряда их подвижность снижается из-за усиления рассеяния носителей на ионизированных примесях за счет кулоновского взаимодействия с ними. Поэтому необходимость достижения большой концентрации подвижных носителей заряда вступает в противоречие с возможностью обеспечить их высокую подвижность. Между тем высокочастотные полупроводниковые приборы требуют больших концентраций носителей заряда с максимально возможной подвижностью.

Эта проблема решается в *модуляционно-легированных структурах* (*modulation-doped structures*), в которых область полупроводника, где генерируются носители заряда, и область, где осуществляется их перенос, пространственно разделены. Для этого используются гетероструктуры, образованные полупроводниками с различной шириной запрещенной зоны. Электронные процессы в модуляционно-легированной структуре иллюстрируют энергетические диаграммы, приведенные на рис. 1.12.

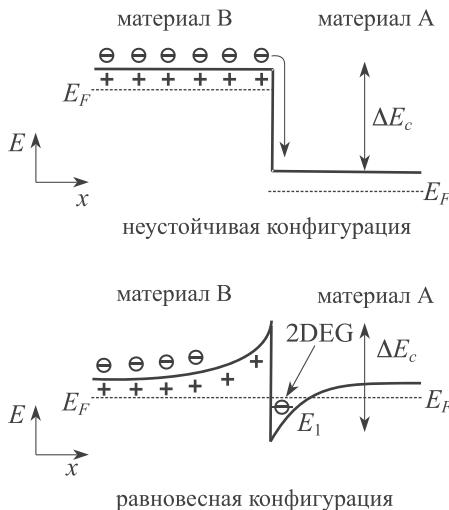


Рис. 1.12. Вид зоны проводимости в окрестности гетероперехода, образованного полупроводником А с малой шириной запрещенной зоны и полупроводником В с большой шириной запрещенной зоны

В модуляционно-легированной структуре донорная примесь обычно вводится в полупроводник с большей шириной запрещенной зоны. Структура сохраняет свою электрическую нейтральность до тех пор, пока электроны находятся у своих донорных атомов. Как только электроны покидают донорные атомы (вследствие тепловой активации при $T > 0$ К), они пересекают границу раздела и переходят в соседнюю область с более низкой потенциальной энергией. Там электроны теряют свою энергию и оказываются захваченными в приграничной области, поскольку не имеют возможности преодолеть потенциальный барьер ΔE_c и вернуться обратно. Эти электроны оказываются пространственно отделенными от сильно легированной донорной примесью области полупроводника, откуда они поступили. Скатившиеся в потенциальную яму электроны индуцируют электростатический потенциал, который (наряду с притяжением положительно заряженными ионами примеси) «прижимает» их к границе между материалами А и В. В результате у границы гетероперехода для электронов образуется квантовый колодец с примерно треугольным пространственным профилем распределения потенциала. Ширина этого колодца имеет порядок нескольких нанометров. В нем энергетические уровни для поперечного (вдоль оси x) движения электронов оказываются квантованными, как и в прямоу-

гольных квантовых колодцах. Заняты только нижние энергетические уровни. Однако электроны на этих уровнях сохраняют свободу для движения в двух других направлениях, т. е. в плоскости, параллельной плоскости гетероперехода. Каждый такой уровень представляет собой дно одной из двумерных подзон размерного квантования. Так в слаболегированном узкозонном полупроводнике у границы гетероперехода образуется двумерный электронный газ (2DEG). Электроны в нем заполняют локализованные в двух направлениях состояния. Для формирования двумерного электронного газа вполне достаточно одного локализованного в поперечном направлении уровня.

Гетероструктура, изготовленная как сверхрешетка из $n\text{-AlGaAs}$ (материал с большей шириной запрещенной зоны) и нелегированного GaAs, представляет классический пример модуляционно-легированной структуры. Подвижность электронов в ней выше на несколько порядков и при низких температурах достигает $2 \cdot 10^7 \text{ см}^2 \text{В}^{-1} \text{с}^{-1}$, что является рекордом для GaAs. Между тем слоевая концентрация электронов⁴ в двумерном электронном газе остается не очень высокой — менее $5 \cdot 10^{11} \text{ см}^{-2}$.

Модуляционное легирование дает два важных преимущества. Во-первых, электроны оказываются отделенными от донорных атомов, что ослабляет их рассеяние ионизированными примесями. Во-вторых, у границы гетероперехода образуется двумерный электронный газ.

3.3. Дельта-легированные структуры

Полупроводниковую структуру с существенно неравномерным профилем распределения примесей, характеризующимся локализацией примесных атомов в очень тонком внутреннем слое (в идеале — в пределах одного монослоя), называют **дельта-легированной (δ-легированной) структурой** (*delta-doped (δ-doped) structure*). Энергетическая диаграмма такой структуры представлена на рис. 1.13.

Носители заряда в сильно легированной области связаны кулоновским взаимодействием с порождающими их примесными атомами, поэтому они не уходят далеко от области локализации примесей. Электрическое поле ионизированных примесных атомов экранируется полем свободных носителей заряда. Для подвижных носителей образуется V-образный квантовый колодец с пространственным рас-

⁴ Концентрация электронов в расчете на единицу площади поверхности структуры.

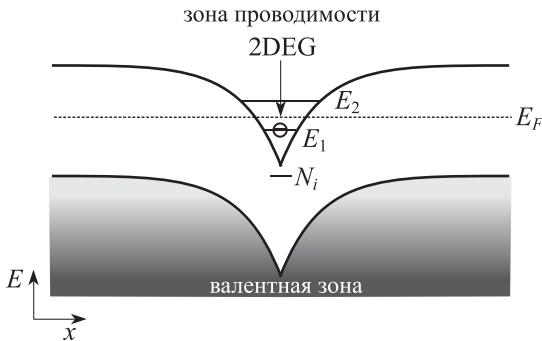


Рис. 1.13. Энергетические зоны в дельта-легированной полупроводниковой структуре

пределением потенциала

$$U(x) = \frac{me^4}{\varepsilon^2 \hbar^2} \left[\left(\frac{15\pi^3}{8\sqrt{2}} N_i^2 (a_B^*)^4 \right)^{-1/10} + \sqrt{\frac{2\sqrt{2}}{15\pi}} \frac{x}{a_B^*} \right]^{-4}, \quad (1.39)$$

где e — заряд электрона, ε — диэлектрическая проницаемость материала, N_i — слоевая концентрация примесей, $a_B^* = \varepsilon \hbar^2 / me^2$ — эффективный боровский радиус.

Энергетические состояния в колодце квантовутся в соответствии с эффектом квантового ограничения, что приводит к образованию двухмерных подзон, которые заполняются электронами вплоть до высоких концентраций.

Одним из примеров дельта-легированных структур является структура, содержащая моноатомный слой кремния, нанесенный на монокристалл GaAs и затем накрытый эпитаксиальным слоем GaAs. Донорные атомы кремния, хотя и перераспределяются вследствие диффузии, остаются в окрестности исходного моноатомного слоя. Область квантового ограничения для электронов составляет до 10 нм. Слоевая концентрация электронов в образовавшемся двухмерном электронном газе достигает высоких значений (вплоть до 10^{14} см $^{-2}$), но ценой уменьшения их подвижности.

Структуры, состоящие из периодически расположенных дельта-легированных слоев с n -типовом и p -типовом проводимости в материале с собственной проводимостью (i), известны как **n - i - p - i структуры**. Когда концентрация доноров в n -слое равна концентрации акцепторов в p -слое, свободные носители заряда не имеют возможности двигаться в равновесных условиях. Неравновесные же носители заряда,

генерируемые, например, светом, оказываются разделенными внутренним электрическим полем. Их заряд изменяет энергетическую зонную диаграмму точно так же, как это происходит в случае равновесных носителей заряда. Такой же эффект достигается за счет приложения внешнего смещения к n - и p -слоям. Все это дает возможность эффективно управлять энергетической диаграммой структуры, что важно для определенных приборных применений.

4. Структуры с квантовым ограничением за счет внешнего электрического поля

Электрическое смещение полупроводниковых структур внешним потенциалом, приложенным через металлический затвор, широко используется для управления потенциальными барьерами на границах металл/полупроводник и полупроводник/диэлектрик. Соответствующий выбор полярности внешнего потенциала позволяет формировать в полупроводнике области, обогащенные или обедненные носителями заряда вблизи этих границ. Размер таких областей существенно зависит от величины приложенного напряжения. Электроны или дырки в них могут испытывать квантовое ограничение. Эти структуры по сути являются электростатически индуцированными наноразмерными структурами. Примеры таких структур рассмотрены в данном разделе.

4.1. Структуры металл/диэлектрик/полупроводник

Движение электронов в одном направлении может быть легко ограничено в обычных структурах металл/окисел/полупроводник (МОП) полевого транзистора, как показано на рис. 1.14. Для этого наилучшим образом подходит кремниевый полевой транзистор.

В МОП-структуре монокристаллическая кремниевая подложка с p -типов проводимости, слой SiO_2 на ней и верхний металлический электрод, называемый затвором, образуют параллельные обкладки конденсатора. Положительное смещение, приложенное к затвору, отдаленному от полупроводника диэлектрическим слоем, притягивает электроны, обедняя таким образом приграничный слой полупроводника дырками.

По мере увеличения внешнего приложенного напряжения образуется тонкий инверсионный слой, обогащенный электронами, что приводит к искривлению энергетических зон в области полупроводника, граничащей с диэлектриком. Инверсионный слой играет роль квантового колодца для электронов. Потенциальные стенки этого колодца образованы с одной стороны границей диэлектрик/полупроводник, а

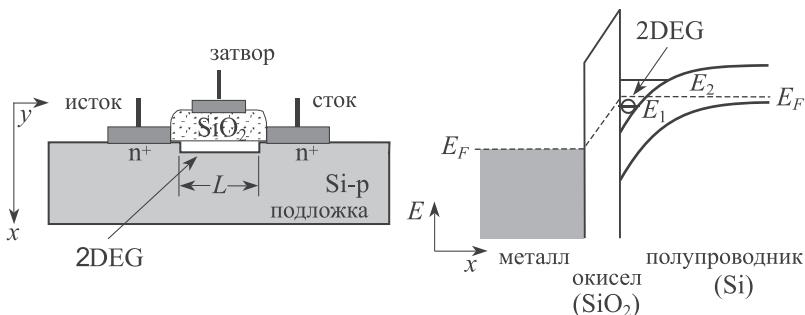


Рис. 1.14. Структура полевого транзистора металл/окисел/полупроводник и его энергетическая диаграмма, иллюстрирующая образование в нем двумерного электронного газа

с другой — электрическим полем, прижимающим электроны к этой границе. Из-за малой ширины колодца в нем формируется двумерный электронный газ. Поскольку общий индуцированный заряд определяется величиной напряжения на затворе, то количество электронов в двумерном газе легко регулируется внешним приложенным напряжением. В кремнии слоевая концентрация электронов достигает 10^{13} см^{-2} , а их подвижность находится в пределах $10^3\text{--}10^4 \text{ см}^2\text{В}^{-1}\text{с}^{-1}$.

Следует отметить две важных особенности двумерного электронного газа, формирующегося в МОП-структуратах полевого транзистора. Во-первых, ширина квантового колодца в таких структурах зависит от величины приложенного напряжения; как следствие, и энергетический зазор между квантованными уровнями в колодце определяется этим же напряжением. Во-вторых, только один тип носителей заряд может быть подвержен эффекту квантового ограничения. В рассмотренном примере это были электроны. Для другого типа носителей заряда потенциальный барьер отсутствует, и их энергетический спектр остается неизменным.

4.2. Структуры с расщепленным затвором

Электростатический подход используют также для того, чтобы «вырезать» квантовые шнуры и квантовые точки из двумерного электронного газа, образующегося, например, при модуляционном легировании и дельта-легировании. Для этого используют структуры с **расщепленным затвором** (split gate). Принцип их действия схематически показан на рис. 1.15.

Металлические электроды затвора наносят на поверхность полупроводника со скрытым слоем двумерного электронного газа. При по-

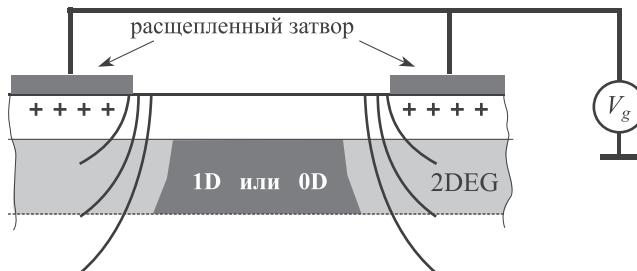


Рис. 1.15. Структура с расщепленным затвором, обеспечивающая формирование одномерных и нуль-мерных элементов в скрытом слое двумерного электронного газа

даче на них внешнего отрицательного смещения находящаяся под ними область полупроводника обедняется электронами вследствие эффекта Шоттки. Свободные электроны остаются только в узкой области, не подверженной воздействию электрического поля под затвором.

Увеличение обратного смещения приводит к расширению обедненной электронами области в окрестности расщепленных электродов, что дает возможность управлять геометрией индуцированной таким образом низкоразмерной структуры в одном и том же приборе. Конструируя расщепленные электроды соответствующим образом, можно сформировать одномерные (полосковыми электродами) или нульмерные (кольцевыми электродами) структуры. Применение расщепленных электродов наиболее эффективно в модуляционно-легированных структурах и менее эффективно в дельта-легированных образцах. Это связано с тем, что для обеднения областей, имеющих высокую слоевую концентрацию электронов (больше 10^{13} см^{-2}), необходимо прикладывать напряжения, которые превышают напряжение пробоя барьера Шоттки на границе металл/полупроводник.

Литература к части I

1. Borisenko V. E., Ossicini S. *What is What in the Nanoworld*. Berlin Weinheim: Wiley-VCH, 2008.
2. Mitin V. V., Kochelap V. A., Stroscio M. A. *Introduction to Nanoelectronics: Science, Nanotechnology, Engineering, and Applications*. Cambridge: Cambridge University Press, 2008.
3. *Concepts in Spin Electronics*, edited by S. Maekawa. New York: Oxford University Press, 2006.

4. Драгунов В. П., Неизвестный И. Г., Гридчин В. А. *Основы наноэлектроники*. Москва: Логос, 2006.
5. Tsu R. *Superlattice to Nanoelectronics*. Amsterdam: Elsevier Science, 2005.
6. Harrison P. *Quantum Wells, Wires and Dots: Theoretical and Computational Physics*. Chichester: John Wiley & Sons, 2000.
7. Yu P. Y., Cardona M. *Fundamentals of Semiconductors. Physics and Materials Properties*. Berlin: Springer, 1999.
8. Davies J. H. *The Physics of Low-Dimensional Semiconductors: An Introduction*. Cambridge: Cambridge University Press, 1998.
9. Imry Y. *Introduction to Mesoscopic Physics*. New York: Oxford University Press, 1997.
10. Shik A. *Quantum Wells: Physics and Electronics of Two-Dimensional Systems*. Singapore: World Scientific, 1997.
11. Ferry D. K., Goodnick S. M. *Transport in Nanostructures*. Cambridge: Cambridge University Press, 1997.
12. Frenkel D., Smit B. *Understanding Molecular Simulation*. San Diego: Academic Press, 1996.
13. Rapaport D. C. *The Art of Molecular Dynamics Simulation*. Cambridge: Cambridge University Press, 1995.