

ОБУХОВ Илья Андреевич

**Неравновесные эффекты как основа функционирования
твёрдых электронных приборов**

05.27.01 - Твёрдая электроника, радиоэлектронные компоненты, микро- и
наноэлектроника, приборы на квантовых эффектах

диссертация на соискание ученой степени
доктора физико-математических наук

г. Москва 2014

СОДЕРЖАНИЕ

ВВЕДЕНИЕ		5
ГЛАВА 1	Краевые задачи для уравнений переноса заряда в полупроводниках.....	14
	1.1 Дрейфово-диффузионная модель переноса заряда.....	15
	1.2. Уравнения дрейфово-диффузионной модели.....	16
	1.3. Некоторые особенности уравнений дрейфово-диффузионной модели.....	21
	1.4. Граничные условия для уравнений переноса заряда...	26
	1.5. Новый базис переменных.....	34
	1.6. Одномерные краевые задачи.....	44
ГЛАВА 2	Приближенные решения краевых задач переноса заряда в полупроводниках.....	47
	2.1. Асимптотика решения уравнения Пуассона.....	48
	2.1.1. Одномерная модель.....	49
	2.1.2. Двумерная краевая задача.....	51
	2.1.2.1 Область с омическим контактом.....	51
	2.1.2.2. Область с контактом к диэлектрику и омическим контактом.....	56
	2.2. Стационарная краевая задача для плотности тока.....	61
	2.3. Приближенные решения краевых задач для функции ψ в стационарном случае.....	66
	2.3.1. Одномерная краевая задача.....	67
	2.3.2. Двумерная краевая задача.....	71
ГЛАВА 3	Неравновесные эффекты и статические вольт - амперные характеристики полупроводниковых приборов.....	74
	3.1. Биполярный диод.....	74
	3.2. Биполярный транзистор.....	84
	3.3. Полупроводниковый тиристор.....	93
	3.4. Выводы.....	100
ГЛАВА 4	Модель переноса заряда в мезоскопических структурах.	102
	4.1. Уравнения многофазной модели переноса заряда.....	103
	4.1.1. Интерпретируемые состояния носителей заряда.....	103
	4.1.2. Уравнение для детерминированного потенциала.....	104
	4.1.3. Интерпретируемые состояния и фазы газа носителей заряда.....	105
	4.1.4. Уравнение для гидродинамической скорости носителей заряда.....	107
	4.1.5. Уравнения баланса числа носителей заряда.....	108

	4.1.6. Стационарные уравнения многофазной модели переноса заряда.....	110
	4.1.7. Длина релаксации к локальному химическому равновесию.....	111
	4.1.8. Закон Ома и формула Цу – Есаки.....	113
	4.2. Об обосновании многофазной модели переноса заряда..	116
	4.2.1. Квантовая динамика заряженных частиц, взаимодействующих со случайным потенциалом.....	116
	4.2.2. Термодинамика заряженных частиц, взаимодействующих со случайным потенциалом.....	119
	4.2.3. Уравнения баланса числа заряженных частиц, взаимодействующих со случайным потенциалом.....	123
	4.2.4. Время релаксации τ	125
	4.2.5. Необратимость уравнений многофазной модели.....	127
	4.3. Граничные условия для стационарных уравнений переноса заряда в мезоскопических структурах	131
ГЛАВА 5	Моделирование статических характеристик резонансно-туннельных диодов.....	137
	5.1. Перенос заряда в РТС.....	139
	5.1.1. Резонансное туннелирование электронов	139
	5.1.2. Одномерная стационарная двухфазная модель переноса заряда.....	145
	5.1.2.1. Интерпретируемые состояния.....	145
	5.1.2.2. Уравнения для химических потенциалов.....	152
	5.1.2.3. Выражения для плотностей числа частиц и потоков.	156
	5.1.2.4. Уравнение Пуассона в РТС.....	158
	5.2. Вольт - амперная характеристика РТС.....	159
	5.2.1. Соглашение о токах.....	160
	5.2.2. Сравнение теории и эксперимента, формула Цу – Есаки.....	161
	5.2.3. Дрейфово – диффузионный ток и нарушение химического равновесия.....	164
	5.2.4. Характерные участки ВАХ РТС.....	169
	5.3. Вольт – амперная характеристика РТД.....	175
	5.3.1. Двумерная гибридная модель РТД.....	178
	5.3.2. Простейший анализ гибридной модели РТД.....	180
ГЛАВА 6	Приборы на основе квантовых проводов и их статические характеристики.....	184
	6.1. Квантовые провода.....	185
	6.2. Модель переноса заряда в квантовом проводе.....	190

6.2.1. Интерпретируемые состояния электронов.....	190
6.2.2. Уравнения для химических потенциалов.....	201
6.2.3. Выражения для плотностей числа частиц и потоков...	205
6.2.4. Уравнение для самосогласованного электростатического потенциала.....	206
6.3. Статические характеристики квантового провода.....	208
6.3.1. Результаты моделирования.....	208
6.3.2. Механизмы проводимости в КП.....	219
6.3.2.1. Влияние контактов на ВАХ КП.....	223
6.3.2.2. Основные факторы, определяющие ВАХ КП.....	232
6.4. Температурные эффекты в квантовых проводах.....	235
6.4.1. Температурная зависимость сопротивления металлических квантовых проводов.....	235
6.4.2. Самоохлаждение эмиттерного контакта квантового провода.....	241
6.5. Полевой транзистор на основе квантового провода.....	245
6.6. Релаксационный квантовый диод.....	251
6.7. Релаксационный квантовый транзистор.....	259
6.8. Квантовый транзистор с инжекционным затвором.....	269
6.9. Выводы.....	275
ЗАКЛЮЧЕНИЕ	278
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ	284

ВВЕДЕНИЕ

Объект и предмет исследования. В диссертации исследуется влияние отклонений носителей заряда от состояния локального химического равновесия на электрические характеристики твердотельных электронных приборов. Предметом исследования является возможность использования указанных неравновесных явлений в качестве основы функционирования приборов с нанометровыми размерами активных областей.

Актуальность проблемы. Исследования в области наноэлектроники активно ведутся во всех крупных научно-исследовательских центрах мира. Ожидается, что применение электронных элементов нанометровых размеров приведет к:

- уменьшению размеров и массы электронных схем не менее, чем в 10 раз;
- увеличению плотности интеграции элементов в электронных схемах не менее, чем в 100 раз;
- уменьшению энергопотребления электронных схем не менее чем в 100 раз;
- увеличению рабочих частот связной аппаратуры и производительности процессоров не менее, чем в 100.

В абсолютных величинах для электронных схем это означает:

- плотность интеграции 10^{10} элементов на квадратный сантиметр и более;
- рабочие напряжения 10^{-4} В – 1,0 В;
- частоты функционирования от 100 ГГц до 10 ТГц;
- энергопотребление порядка 10^{-4} мкВт на элемент.

Применение приборов, схем и устройств с такими характеристиками произведет качественные изменения в военной и гражданской технике и обеспечит стратегические преимущества странам и компаниям, обладающим технологиями их производства [1-4].

Элементную базу нанoeлектроники не удастся создать простым масштабированием приборов микроэлектроники. Уменьшение их активных областей до величин, меньших некоторых пороговых, приводит к потере работоспособности [5]. В связи с этим актуальными являются:

- исследование физических процессов, которые можно положить в основу функционирования электронных приборов с нанометровыми размерами активных областей;
- поиск конструктивных решений таких приборов, имеющих требуемые электрические характеристики.

Степень разработанности темы исследования. Двадцать лет назад субмикронная технология считалась в микроэлектронике передовой. Нанотехнология была тогда предметом лабораторных экспериментов. Сегодня нанотехнология применяется для производства микросхем в коммерческих объемах [6,7].

В ближайшие годы ожидается, что норма проектирования для кремниевой технологии станет сопоставимой с длиной размерного квантования в этом материале, которая при комнатной температуре равна примерно 15 нм. Для полупроводников группы $A^{III}B^V$ длина размерного квантования электронов при температуре $T = 300^\circ\text{K}$ изменяется от 25 нм до 58 нм. Области с такими планарными размерами могут быть изготовлены при помощи современной нанолитографии [6-9].

Слои полупроводниковых материалов с толщиной в несколько атомов давно выращиваются методами молекулярно-лучевой эпитаксии, а приборы на основе многослойных гетероструктур нашли широкое применение в электронике [10-18].

Технологии, позволяющие управляемо создавать структуры с размерами, меньшими длины размерного квантования, открывают перспективы для использования квантовых явлений в качестве основы функционирования твердотельных электронных приборов. Спектр таких явлений достаточно широк. Это и собственно размерное квантование, и резонансное туннелирование электронов [1-3,12-16], и «одноэлектронное туннелирование» [2,19-22], и

интерференция электронных волн [2,23], и многие другие явления, активно изучаемые во всех ведущих мировых центрах электроники.

Перечисленные эффекты уже успешно используются в ряде приборных конструкций. Размерное квантование играет существенную роль в функционировании транзисторов с высокой подвижностью электронов [10,17]. Еще в конце 80-х годов прошлого века созданы и исследованы резонансно-туннельные диоды [14-16] и различные «одноэлектронные приборы» [2].

Серьезные надежды многие исследователи [2,23] возлагают на разработку квантовых приборов с новыми функциональными свойствами. Их использование сулит, например, создание совершенно новых и уникальных по своим возможностям вычислительных систем [2,24,25].

Однако практически без внимания остается вопрос: возможно ли полномасштабное использование достижений современной схемотехники, если электронные схемы будут строиться из квантовых приборов?

За десятилетия развития микроэлектроники разработаны стандартные схемотехнические решения. Они позволяют создавать электронные устройства широкой функциональности, используя при этом очень ограниченный набор элементов: резистор, емкость, диод, биполярный и полевой транзисторы. Можно ли построить квантовые приборы, имеющие электрические характеристики, аналогичные характеристикам этих элементов?

Цели исследования. Целями работы являются поиск и теоретический анализ:

- физических процессов, которые можно использовать в качестве основы функционирования приборов с нанометровыми размерами активных областей, имеющих электрические характеристики, аналогичные базовым элементам микроэлектроники;
- конструктивных решений, реализующих указанные приборы.

Задачи исследования. Для достижения целей в диссертации были решены следующие задачи:

- в рамках классической дрейфово-диффузионной модели исследовано влияние отклонений электронно-дырочного газа от состояния локального химического равновесия на электрические характеристики полупроводниковых приборов: биполярного диода, биполярного транзистора и тиристора;
- построена многофазная модель переноса заряда в мезоскопических структурах, позволяющая сохранить преемственность в формализме описании физических процессов с классической дрейфово-диффузионной моделью;
- на примере резонансно-туннельного диода показано, что многофазная модель позволяет описывать характеристики квантовых приборов адекватно эксперименту;
- в рамках многофазной модели исследована возможность использования в качестве основы функционирования приборов нанoeлектроники эффектов, обусловленных отклонениями электронного газа от состояния локального химического равновесия;
- предложены и исследованы конструкции планарных квантовых приборов на основе квантовых проводов с электрическими характеристиками, аналогичными характеристикам элементов микроэлектроники.

Методология и методы исследования. Поставленные задачи решались с применением аппарата современной математической физики, асимптотических методов и теории сингулярных возмущений. Для расчета характеристик приборов использовалось численное моделирование.

Научная новизна. Научная новизна работы заключается в следующем:

- получены новые приближенные решения дрейфово-диффузионных уравнений переноса заряда в полупроводниковых структурах, позволяющие проанализировать влияние отклонений электронов и дырок от состояния локального химического равновесия на вольт-амперные характеристики полупроводниковых приборов;

- построена новая многофазная модель переноса заряда в мезоскопических структурах, позволяющая в рамках единого формализма описывать классические и квантовые явления в твердотельных электронных приборах;
- в рамках многофазной модели дано теоретическое объяснение наблюдаемых в эксперименте особенностей вольт-амперных характеристик резонансно-туннельных диодов;
- впервые проведен анализ влияния классических контактных областей на вольт-амперные характеристики резонансно-туннельных диодов и показано что взаимодействие резонансно-туннельной структуры с подложкой не сводится к тривиальному эффекту нагрузочного сопротивления;
- проведен анализ переноса заряда в квантовом проводе с учетом взаимодействия одномерного проводящего канала с классическими контактными областями и неравновесных квантовых эффектов, возникающих при протекании тока через границы проводящего канала с контактами;
- предложены конструкции и рассчитаны ожидаемые статические электрические характеристики новых планарных квантовых приборов, функционирующих на основе эффектов, обусловленных отклонениями электронного газа от состояния локального химического равновесия: релаксационного квантового диода, релаксационного квантового транзистора, квантового транзистора с инжекционным затвором;
- показано, что эти приборы могут служить функциональными аналогами известных элементов микроэлектроники и использоваться в широком диапазоне частот, вплоть до терагерцового диапазона;
- исследован механизм охлаждения эмиттерного контакта квантового провода за счет инжекции электронов из контакта в проводящий канал;

- дано теоретическое объяснение изменения характера температурной зависимости сопротивления металлического квантового провода при уменьшении его поперечного сечения.

Теоретическая и практическая значимость работы. Теоретическая и практическая ценность полученных результатов заключается в следующем:

- приближенные решения дрейфово-диффузионных уравнений переноса заряда используются для быстрых оценочных расчетов характеристик полупроводниковых приборов, а также в качестве начального приближения при численном моделировании их работы;
- многофазная модель переноса заряда является эффективным инструментом проектирования квантовых приборов и моделирования их характеристик;
- новые квантовые приборы в случае их практической реализации позволят использовать в схемах и устройствах наноэлектроники хорошо известные в микроэлектронике схемотехнические решения;
- эффект охлаждения эмиттерного контакта квантового провода может быть использован в «нанохолодильниках» и для создания на основе матриц квантовых проводов «интеллектуальных» энергосберегающих покрытий.

Степень достоверности и апробация работы. Во всех случаях, где имеются экспериментальные результаты, проводилось сравнение расчетов и измерений. Для характеристик резонансно-туннельных диодов, квантовых проводов и полевого транзистора на квантовом проводе наблюдается хорошее соответствие теоретических и экспериментальных результатов. Это свидетельствует о достоверности и обоснованности применяемых методов и моделей.

Результаты, представленные в работе, на разных этапах ее выполнения докладывались и обсуждались на:

1. Всесоюзном научном совещании «Методы малого параметра». Нальчик. 1987;

2. III Всесоюзной конференции «Физические основы надежности и деградации полупроводниковых приборов». Кишинев. 1991;
3. Всесоюзной конференции «Асимптотические методы сингулярно-возмущенных уравнений и некорректно поставленных задач». Бишкек, 1991;
4. International Symposium "Nanostructures: Physics and Technology". St. Petersburg, 1993;
5. Second International Conference on Nanometer Scale Science and Technology (NANO-II), Moscow, 1993;
6. Nanomeeting'95, Minsk. 1995; Nanomeeting'97, Minsk. 1997; Nanomeeting'2011, Minsk. 2011;
7. WTEC Workshop: Russian Research and Development Activities on Nanoparticles and Nanostructured Materials, S. Petersburg, 1997;
8. Scanning Probe Microscopy – 2002, Nizhny Novgorod, 2002;
9. IX Научно-Технической Конференции «Твердотельная электроника, сложные функциональные блоки РЭА», Звенигород, 2010;
10. 6-й - 23-й Международной Крымской Микроволновой Конференции, Севастополь, 1996 – 2013 гг.

Публикации. Автором диссертации опубликовано 83 печатных работы, включая 3 монографии, 30 работ опубликованы в изданиях, входящих в перечень, рекомендованный ВАК при Минобрнауки России.

Основные результаты диссертации опубликованы в 63 печатных работах, из них 2 монографии, 15 работ опубликованы в изданиях, входящих в перечень, рекомендованный ВАК при Минобрнауки России. По тематике исследований получен 1 патент РФ.

Личный вклад. Лично автором получены все основные результаты диссертационной работы и выполнены необходимые расчеты. Приближенные решения уравнений Пуассона в полупроводниковых структурах получены совместно с Л.В.Калачевым. Постановка задачи о растекании тока в подложке резонансно-туннельного диода принадлежит О.Т. Гаврилову.

Положения диссертационной работы, выносимые на защиту:

1. Новый базис переменных для классических дрейфово-диффузионных уравнений переноса заряда. Позволяет явным образом связать характеристики полупроводниковых приборов с уровнем отклонения электронного и дырочного газов от состояния локального химического равновесия и выявить причины и источники этой неравновесности.

2. Приближенные решения дрейфово-диффузионных уравнений, описывающие перенос заряда в полупроводниковых структурах без введения условных внутренних границ в областях р-п-переходов. Применение этих решений для анализа влияния химической неравновесности электронного и дырочного газов на функционирование биполярных диода, транзистора и тиристора.

3. Многофазная модель переноса заряда в мезоскопических структурах. Позволяет в рамках единого формализма описывать квантовые и классические явления и, соответственно, квантовые и классические области электронных приборов. Сохраняет преемственность с классической дрейфово-диффузионной моделью. Основывается на:

- представлении об электронном газе в твердом теле, как о смеси фаз (компонент), соответствующих допустимым состояниям электронов и характеризующихся каждая своим химическим потенциалом;
- описании неравновесных явлений в электронном газе как отклонений от состояния локального химического равновесия;
- выражении для плотности тока, учитывающем как квантовую, так и классическую дрейфово-диффузионную составляющую.

4. Многофазная модель позволяет рассчитывать вольт-амперные характеристики резонансно-туннельных диодов на основе гетероструктур GaAs/AlGaAs адекватно измерениям и дать теоретическое объяснение особенностей этих характеристик:

- ступенька на участке ВАХ с отрицательной дифференциальной проводимостью объясняется нарушением локального химического

равновесия между «левыми» и «правыми» электронами и соответствующим накоплением неравновесных электронов в области потенциальной ямы;

- наблюдаемые экспериментально большие значения тока долины прибора обусловлены наличием дрейфово-диффузионного тока.

5. Квантовый прибор необходимо рассматривать как единый объект, состоящий из квантовых и классических областей. Пассивные классические области могут существенно влиять на характеристики квантового прибора:

- взаимодействие резонансно-туннельной структуры с подложкой изменяет параметры вольт - амперной характеристики резонансно-туннельного диода, а при некоторых условиях приводит к неединственности решения стационарных уравнений переноса заряда;
- инжекция электронов из эмиттерного контакта квантового провода в проводящий канал является одним из основных процессов, определяющих проводимость прибора, она может приводить к охлаждению эмиттерного контакта, а также к необычной зависимости сопротивления квантового провода от температуры.

6. Квантовые и неравновесные явления одинаково значимы для функционирования твердотельных электронных приборов с нанометровыми размерами активных областей. В мезоскопических структурах источники неравновесности электронного газа локализуются в малых окрестностях неоднородностей потенциального рельефа для электронов. Неоднородности порождаются гетерограницами, переходами между структурами с различной квантовой размерностью и между квантовыми проводами различного поперечного сечения, и т.п.

7. На основе эффектов, обусловленных отклонениями электронного газа от состояния локального химического равновесия, могут быть созданы квантовые приборы терагерцового диапазона частот, имеющие характеристики, аналогичные элементам микроэлектроники: релаксационный квантовый диод, релаксационный квантовый транзистор, квантовый транзистор с инжекционным затвором.

ГЛАВА 1. Краевые задачи для уравнений переноса заряда в полупроводниках

Задачи теоретического описания переноса заряда в полупроводниковых структурах не теряют своей актуальности уже долгие годы, со времен бурного развития микроэлектроники [26-28]. Их решения позволяют глубже понять физику функционирования полупроводниковых приборов, что необходимо как для создания новой элементной базы, так и для проектирования схем и устройств на основе уже известных компонент. Поэтому моделирование переноса заряда в полупроводниках находит применение и в научно-исследовательской, и в опытно-конструкторской деятельности, в частности, в системах автоматизированного проектирования (САПР) [28-32].

Одним из важных факторов, оказывающих существенное влияние на характеристики полупроводниковых приборов, является отклонение электронов и дырок от состояния локального химического равновесия [26,27]. В первых трех главах работы проведен последовательный теоретический анализ этого влияния на примере диода, биполярного транзистора и тиристора. Разработан соответствующий математический аппарат. Показано, что источником неравновесности электронно-дырочного газа является ток, протекающий перпендикулярно p-n-переходам. Получены новые приближенные решения дрейфово-диффузионных уравнений переноса заряда в полупроводниковых структурах описывающие влияние отклонений электронов и дырок от состояния локального химического равновесия на вольт – амперные характеристики (ВАХ) полупроводниковых приборов.

Процесс проектирования электронных схем и устройств включает в себя моделирование переноса заряда в полупроводниковых структурах [31-33]. Результатом такого моделирования являются статические и динамические электрические характеристики полупроводникового прибора, который может использоваться как самостоятельно, так и в качестве элемента интегральной схемы. При расчете этих характеристик учитываются геометрические и физические свойства структуры, реализующей прибор. Таким образом,

электрические характеристики полупроводникового прибора увязываются с его топологией и технологией изготовления.

Несмотря на длинную историю, решение задач переноса заряда в полупроводниках до сих пор связано с существенными трудностями. Это обусловлено сложностью моделируемого объекта и соответствующей сложностью уравнений, используемых для его описания. Это нелинейные дифференциальные уравнения в частных производных [17,26,28,31].

В приборах в качестве активных элементов используются р-n-переходы, контакты Шотки, гетеропереходы и другие конструктивные решения, подразумевающие резкую смену физических характеристик структуры в областях малых размеров. Этим обусловлены дополнительные математические проблем при моделировании работы полупроводниковых приборов.

Современная технология позволяет формировать структуры с нанометровыми размерами активных областей [6-9,34]. Эти возможности используются не только при создании приборов новых типов, но и при производстве стандартных элементов. Поэтому при моделировании приборов возникает потребность тщательного исследования свойств пограничных областей, влиянием которых ранее можно было пренебречь из-за относительно больших размеров структур. Это еще более усложняет решение задач переноса заряда в электронных приборах, которые остаются сегодня столь же актуальными, как и тридцать лет назад.

1.1. Дрейфово-диффузионная модель переноса заряда

Для описания переноса заряда в макроскопических полупроводниковых приборах используют различные физические модели:

- кинетическую, основанную на уравнении Больцмана [26,35,36];
- гидродинамическую, основанную на уравнениях для моментов функции распределения [30,35,36];
- дрейфово-диффузионную [26-32].

Простейшей является дрейфово-диффузионная модель, которая формально может быть получена из первых двух в качестве нулевого приближения. При этом возникает дополнительная задача вычисления коэффициентов переноса на основе модели более высокого уровня. Решить эту задачу зачастую бывает сложно, а иногда – практически невозможно.

Тем не менее, дрейфово-диффузионная модель наиболее часто используется в физике полупроводниковых приборов. Как правило, фигурирующие в уравнениях модели коэффициенты определяются феноменологически, на основе общих физических соотношений и анализа экспериментальных данных. Это позволяет наполнить дрейфово-диффузионную модель содержанием, достаточным для адекватного понимания физики функционирования приборов.

Более того, результаты расчетов электрических характеристик полупроводниковых приборов, проведенных на основе дрейфово-диффузионной модели, настолько хорошо согласуются с результатами измерения, что используются в системах автоматизированного проектирования электронных приборов и схем [29-33].

Сказанное позволяет считать дрейфово-диффузионную модель переноса заряда наиболее развитой для описания электрических характеристик макроскопических полупроводниковых приборов.

1.2. Уравнения дрейфово-диффузионной модели

Уравнения модели можно записать в следующем виде

$$\left\{ \begin{array}{l} (\nabla, \varepsilon_s \nabla \varphi) = e(n - p - N), \\ \partial n / \partial t - e^{-1} (\nabla, \mathbf{J}_n) = -U_n, \\ \partial p / \partial t + e^{-1} (\nabla, \mathbf{J}_p) = -U_p, \\ n = n_i \exp\{(e\varphi + F_n)/\Theta\}, p = n_i \exp\{-(e\varphi + F_p)/\Theta\}, \\ \mathbf{J}_n = \mu_n n \nabla F_n, \mathbf{J}_p = \mu_p p \nabla F_p, \\ \mathbf{J} = \mathbf{J}_n + \mathbf{J}_p - \varepsilon \partial (\nabla \varphi) / \partial t, \end{array} \right. \quad (1)$$

где ϵ_s – диэлектрическая проницаемость полупроводника, выраженная в единицах диэлектрической проницаемости вакуума ϵ_0 ; ϕ – электростатический потенциал; e – заряд электрона; n и p – концентрации электронов и дырок; N – концентрация легирующей примеси; $\mathbf{J}_{n,p}$ – плотности пространственных токов электронов и дырок; $U_{n,p}$ – скорости процессов рекомбинации и генерации электронов и дырок; $\mu_{n,p}$ – подвижности соответствующих носителей заряда; $F_{n,p}$ – их химические потенциалы; n_i – концентрация носителей заряда в собственном полупроводнике;

$$\Theta = k_B T,$$

k_B – постоянная Больцмана, а T – абсолютная температура, которую мы будем считать постоянной и равной температуре окружающей среды (кристаллической решетки); \mathbf{J} – плотность полного тока включающего в себя ток проводимости (два первых слагаемых в правой части последнего соотношения в (1)) и ток смещения (последнее слагаемое в правой части последнего соотношения в (1)). Эффекты, обусловленные магнитным полем, здесь рассматриваться не будут и в уравнениях (1) не учтены.

Первое уравнение в системе (1) представляет собой уравнение Пуассона. Его можно использовать для расчета электростатического потенциала и электрического поля в случаях, когда входящие в (1) величины достаточно медленно изменяются во времени. Если через ν_0 обозначить характерную частоту изменения потенциалов и плотностей токов, то должно быть справедливо неравенство

$$\nu_0 \ll \nu_c, \tag{2}$$

где

$$\nu_c = c/L,$$

c – скорость света, а L – характерный размер прибора. Если это неравенство не справедливо, то уравнение Пуассона следует заменить уравнением Даламбера и дополнить систему (1) соотношениями, позволяющими учитывать влияние магнитного поля на перенос заряда в приборе.

Для современных полупроводниковых приборов и элементов интегральных схем характерны субмикронные размеры $\sim 10^{-5} - 10^{-6}$ см. Для приборов

наноэлектроники эти размеры предполагаются порядка $10^{-6} - 10^{-7}$ см. В указанном диапазоне значений L неравенство (2) справедливо при частотах меньших, чем 10^{14} Гц, то есть применимо во всех практически интересных случаях.

Второе и третье уравнения в системе (1) описывают пространственное и временное поведение концентраций и плотностей потоков электронов и дырок. В этой связи следует иметь в виду, что хотя мы и не учитываем влияние магнитного поля на перенос заряда, уравнение Максвелла

$$\text{rot}(\mathbf{H}) = c^{-1}\mathbf{J}$$

однозначно разрешимо только в случае, когда сохраняется плотность суммарного тока

$$(\nabla, \mathbf{J}) = 0. \quad (3)$$

Поэтому уравнение (3) должно обязательно дополнять систему (1). Используя (3) из (1) получим

$$\partial N / \partial t = U_p - U_n. \quad (4)$$

В большинстве задач физики полупроводниковых приборов полагают, что $N = N_D - N_A$, где N_D и N_A – концентрации ионизированной донорной и акцепторной примесей. Как правило, считается, что величины $N_{D,A}$ не зависят от времени, то есть динамика ионизации примеси не учитывается. В этом случае, согласно (4) и (5), имеем

$$U_n = U_p = U. \quad (6)$$

Далее будем считать, что соотношение (6) выполнено. Это означает, в частности что рождение и уничтожение электронов и дырок происходит парами.

Часто в теории переноса заряда используют электрохимические потенциалы, которые связаны с химическими потенциалами соотношениями

$$\Phi_{n,p} = -F_{n,p}/e. \quad (7)$$

Химические потенциалы имеют размерность энергии, а электрохимические – потенциала. Электрохимические потенциалы удобны тем, что в контактных областях они равны приложенным напряжениям. Если считать справедливыми соотношения Эйнштейна

$$\mu_{n,p}/D_{n,p} = e/\Theta,$$

связывающие коэффициенты диффузии электронов и дырок $D_{n,p}$ с их подвижностями $\mu_{n,p}$, то плотности токов, выраженные в (1) через градиенты соответствующих химических потенциалов, можно представить в более привычном виде [26-28], как суммы дрейфовой и диффузионной составляющих

$$\mathbf{J}_n = -e\mu_n n \nabla \varphi + eD_n \nabla n, \quad \mathbf{J}_p = -e\mu_p p \nabla \varphi - eD_p \nabla p.$$

Подвижности $\mu_{n,p}$ и коэффициенты диффузии $D_{n,p}$ являются функциями концентраций и электрического поля

$$\mathbf{E} = -\nabla \varphi$$

Не будем приводить здесь используемые модельные зависимости. Их можно найти в работах [29,30]. Заметим только, что, как правило, $\mu_{n,p}$ и $D_{n,p}$ достаточно гладкие функции своих переменных.

Соотношения, связывающие концентрации с электростатическим и химическим потенциалами, записаны в системе (1) в приближении статистики Больцмана и справедливы для невырожденных полупроводников. Их можно обобщить и на вырожденный случай, когда справедлива статистика Ферми-Дирака, что обычно делается при численном решении уравнений переноса. В этой главе мы ограничимся невырожденным случаем, поскольку он позволяет продвинуться достаточно далеко по пути получения аналитических результатов.

В правой части последнего уравнения системы (1) два первых слагаемых представляют собой плотность суммарного тока проводимости, а последнее – плотность тока смещения.

В случае, когда концентрация примеси не зависит от времени, для U используются выражения, описывающие рекомбинацию - генерацию электронов и дырок посредством механизмов Шокли – Рида - Холла и Оже, а также генерацию носителей заряда за счет различных механизмов: лавинного умножения в электрическом поле [26-32], облучения светом, нейтронами и тому подобное. В общем случае для U имеем выражение

$$U = qY - G, \tag{8}$$

где

$$Y = (np - n_i^2)/n_i^2 = \exp\{(F_n - F_p)/\Theta\} - 1, \quad (9)$$

$q = q(N, n, p) > 0$ – положительно определенная функция, вид которой зависит от механизма рекомбинации; G – слагаемое, описывающее генерацию носителей заряда.

Введенная функция Y описывает отклонение электронов и дырок от состояния локального химического равновесия, которое, по определению, достигается при равенстве их химических потенциалов

$$F_n = F_p.$$

В этом случае

$$Y = 0.$$

Первое слагаемое в правой части выражения (8) описывает рекомбинацию электронно-дырочных пар при $Y > 0$ и их генерацию при $Y < 0$. В равновесии при $Y = 0$ это слагаемое также равно нулю и рекомбинации-генерации носителей заряда не происходит.

Для механизмов Шокли – Рида - Холла и Оже функция q имеет вид

$$\begin{cases} q = q^{SRH} + q^{Oge}, \\ q^{SRH} = n_i^2 / (\tau_p(n + n_1) + \tau_n(p + p_1)), \\ q^{Oge} = n_i^2 (C_n n + C_p p). \end{cases} \quad (10)$$

Величины $\tau_{n,p}$ представляют собой времена жизни неосновных носителей заряда (электронов или дырок, в зависимости от типа полупроводника) по отношению к процессу Шокли – Рида - Холла. Они могут изменяться в достаточно широком диапазоне в зависимости от уровня легирования и напряженности электрического поля. Для кремния $\tau_{n,p} \sim 10^{-5} - 10^{-8}$ с. Значения постоянных n_1 и p_1 определяются расположением в запрещенной зоне полупроводника эффективного уровня энергии, через который происходит рекомбинация [26,27].

Постоянные процесса Оже, как правило, малы. Для кремния $C_n = 2,8 \cdot 10^{-31}$ см⁶/с, $C_p = 9,9 \cdot 10^{-32}$ см⁶/с. Но поскольку q^{Oge} пропорциональна концентрациям носителей заряда, постольку начиная с концентраций порядка 10^{18} см⁻³ вклад q^{Oge} в q становится определяющим.

Уравнения системы (1) являются уравнениями классической физики. Квантовые процессы учитываются в них через коэффициенты: подвижности, собственную концентрацию n_i , параметры, входящих в выражения для U . Не отражены в (1) и явления, связанные с разогревом электронного газа [36,37].

Ясно, что такие уравнения могут быть справедливы лишь в том случае, когда характерные размеры неоднородностей полупроводниковых структур много больше длин волн Де Бройля для электронов и дырок. Проведенное в [5] детальное исследование области применимости уравнений (1) показывает, что ими можно пользоваться, если характерные размеры структур много больше, чем 10^{-6} см = 10 нм. Существует еще много дополнительных ограничений на применение уравнений (1), связанных, например, с междолинным рассеянием электронов в материалах группы $A^{III}B^V$, переносом заряда через примесные зоны и т.д. Тем не менее, область применимости системы (1) оказывается весьма широкой. Чаще всего эта система уравнений используется при моделировании кремниевых полупроводниковых приборов.

1.3. Некоторые особенности уравнений дрейфово-диффузионной модели

Для дальнейшего исследования уравнения системы (1) удобно привести к безразмерному виду. Будем считать, что величины $F_{n,p}$ и $e\phi$ выражены в единицах средней энергии тепловых флуктуаций Θ ; величины n , p , N – в единицах собственной концентрации n_i ; координаты x , y , z – в единицах характерного для полупроводниковой структуры линейного размера L ; время t и величины $\tau_{n,p}$ – в единицах характерного времени изменения сигналов

$$\tau_0 = 1/\nu_0;$$

плотности токов $\mathbf{J}_{n,p}$ и \mathbf{J} – в единицах

$$\mathbf{J}^* = -en_i D_0/L,$$

где $D_0 = 1 \text{ см}^2/\text{с}$; подвижности $\mu_{n,p}$ – в единицах eD_0/Θ ; скорости процессов рекомбинации и генерации носителей заряда U в единицах n_i/τ_0 ; постоянные Оже процесса $C_{n,p}$ – в единицах $(n_i^2\tau_0)^{-1}$.

Безразмерные плотности токов обозначим через \mathbf{j}_n , \mathbf{j}_p и \mathbf{j} . Для остальных безразмерных переменных сохраним те же обозначения, что и для размерных. В этом случае, с учетом сделанных замечаний и предположений, система уравнений (1) примет вид

$$\left\{ \begin{array}{l} (\nabla, \alpha^2 \nabla \varphi) = n - p - N, \\ \partial n / \partial t + \beta^2 (\nabla, \mathbf{j}_n) = -U, \\ \partial p / \partial t - \beta^2 (\nabla, \mathbf{j}_p) = -U, \\ n = \exp\{(\varphi + F_n)\}, p = \exp\{-(\varphi + F_p)\}, \\ \mathbf{j}_n = -\mu_n n \nabla F_n, \mathbf{j}_p = -\mu_p p \nabla F_p, \\ \mathbf{j} = \mathbf{j}_n + \mathbf{j}_p + (\alpha^2 / \beta^2) \partial (\nabla \varphi) / \partial t, \\ (\nabla, \mathbf{j}) = 0, \end{array} \right. \quad (11)$$

где

$$\left\{ \begin{array}{l} \alpha = L_D / L, L_D = (\epsilon_s \Theta / (e^2 n_i))^{1/2}, \beta = \alpha (\tau_0 / \tau_D)^{1/2}, \tau_D = L_D^2 / D_0, \\ U = Yq - G, \\ Q^{SRH} = 1 / (\tau_p (n + n_1) + \tau_n (p + p_1)), \\ Q^{Oguc} = C_n n + C_p p. \end{array} \right. \quad (12)$$

Здесь и далее считается, что $\epsilon_s = \text{const}$.

Величину L_D называют длиной Дебая в собственном полупроводнике [26,27]. Согласно первому уравнению в (11) она имеет смысл длины экранировки электрического поля. В легированном полупроводнике длина Дебая будет в $|N|^{1/2}$ меньше, чем L_D (концентрация легирующей примеси N выражена в единицах n_i). Для кремния при комнатной температуре $T = 300^0\text{К}$ длина $L_D = 3,4 * 10^{-3} \text{ см}$. Размеры областей современных полупроводниковых приборов могут варьироваться в широких пределах от 10^{-7} см до 10^{-3} см , соответственно, и величина α изменяется от 10^4 до 1.

Обозначим через N_{ch} абсолютную величину характерного для некоторого фрагмента полупроводникового прибора значения концентрации примеси, выраженной в единицах n_i . Для большинства кремниевых приборов, за исключением структур с нелегированными i -областями, N_{ch} порядка $10^6 - 10^8$. В i -областях за счет наличия неконтролируемой примеси величина N_{ch} находится в пределах $10^4 - 10^5$. Если уравнение Пуассона (первое уравнение системы (11)) поделить на N_{ch} , то правая часть будет порядка единицы, а при старшей производной электростатического потенциала появится параметр

$$\alpha_{ch}^2 = \alpha^2 / N_{ch} \quad (13)$$

Согласно сделанным оценкам, для легированных областей приборов α_{ch} варьируется от 10^{-4} до 10. Значения α_{ch} порядка единиц и десятков возможны для очень малых областей с размерами 10^{-7} см – 10^{-6} см (1 нм – 10 нм) и низким уровнем легирования $N_{ch} \sim 10^6$. В достаточно протяженных областях даже при низком уровне легирования α_{ch} оказывается малой величиной.

Нелегированные i -области кремниевых приборов, таких как p - i - n -диоды, имеют характерные размеры от 10^{-4} см до 10^{-2} см. Это означает, что α_{ch} для них находится в пределах от 10^{-1} до 10^{-3} .

Таким образом, для кремниевых полупроводниковых приборов в уравнении Пуассона при старшей производной электростатического потенциала почти всегда присутствует малый параметр, то есть оно является сингулярно возмущенным почти всюду, за исключением областей малых размеров по сравнению с общим характерным размером прибора. Отмеченными областями малых размеров являются, в частности, p - n -переходы.

Из сингулярной возмущенности уравнения Пуассона следует, что нулевое приближение его решения, которое обозначим как φ_{en} , следует искать из трансцендентного уравнения

$$n - p - N = 0. \quad (14)$$

Это уравнение является точным в случае, когда отсутствует пространственный заряд. Например, для однородно легированного полупроводника. Его часто называют условием (или уравнением)

электронейтральности. Учитывая выражения для n и p , а также очевидное неравенство

$$1 + Y = \exp\{F_n - F_p\} > 0, \quad (15)$$

получим два представления решения уравнения (14) относительно φ

$$\begin{cases} \varphi_{en} = -F_n + \ln((N^2/4 + 1 + Y)^{1/2} + N/2), \\ \varphi_{en} = -F_p - \ln((N^2/4 + 1 + Y)^{1/2} - N/2). \end{cases} \quad (16)$$

Легко видеть, что оба соотношения (16) эквивалентны. Первое из выражений (16) удобно применять в случае полупроводника, легированного донорной примесью, когда $N > 0$, а второе – в случае полупроводника, легированного акцепторной примесью, когда $N < 0$. Из соотношений (16) можно получить еще одно полезное представление для φ_{en}

$$\varphi_{en} = -(F_n + F_p)/2 + (1/2)\ln[((N^2/4 + 1 + Y)^{1/2} + N/2)/((N^2/4 + 1 + Y)^{1/2} - N/2)]. \quad (17)$$

Если подставить выражения (16) в формулы для n и p системы (11), то получим нулевое приближение для концентраций носителей заряда

$$\begin{cases} n_{en} = (N^2/4 + 1 + Y)^{1/2} + N/2, \\ p_{en} = (N^2/4 + 1 + Y)^{1/2} - N/2. \end{cases} \quad (18)$$

Выражения (18) являются точными в областях, где отсутствует пространственный заряд, в частности, в однородно легированном полупроводнике. Из (18) следует, что в состоянии локального химического равновесия при условии

$$|N| \gg 1$$

$$n \approx N/2 + |N|/2 + 1/2N^2, \quad p \approx N/2 - |N|/2 + 1/2N^2,$$

то есть

$$n \approx N, \quad p \approx 1/2N^2$$

в областях, легированных донорной примесью, когда $N > 0$, и

$$p \approx N, \quad n \approx 1/2N^2$$

в областях, легированных акцепторной примесью, когда $N < 0$.

Эти выражения показывают, что в состоянии локального химического равновесия в областях с $N > 0$, которые называют n -областями, основными носителями заряда являются электроны, а неосновными – дырки, концентрация

которых много меньше концентрации электронов. Соответственно, в областях с $N < 0$, которые называют р-областями, основными носителями заряда являются дырки, а неосновными – электроны, концентрация которых много меньше концентрации дырок. Однако, как видно из формул (18), вдали от химического равновесия, когда $Y \gg 1$, ситуация может измениться.

В уравнениях баланса электронов и дырок при старшей пространственной производной стоит параметр β^2 . В соответствии с оценками, сделанными для α , β^2 изменяется в пределах от τ_0/τ_D до $10^8\tau_0/\tau_D$. В современных полупроводниковых приборах характерные времена изменения сигналов занимают широкий диапазон от 10^{-11} с до 1 с. Это означает, что β^2 также изменяется в широкой области значений от 10^{-3} до 10^{12} . Учитывая полученные результаты для концентраций основных носителей и формулы (12) для скоростей рекомбинации, можно утверждать, что уравнения для концентраций и токов в системе (11) не являются, вообще говоря, сингулярно возмущенными.

Важнейшей особенностью уравнений (1) и, соответственно, (11) в полупроводниковых приборах является то, что их необходимо решать при наличии больших градиентов N . В областях р-п-переходов, они достигают значений 10^{23} см⁻⁴ или 10^9 в безразмерных величинах. Это приводит к возникновению соответствующих градиентов для концентраций носителей заряда и электростатического потенциала. Кроме того, дрейфово-диффузионные уравнения переноса заряда оказываются нелинейными. В результате, численное решение системы (1) с соответствующими граничными и начальными условиями наталкивается на значительные трудности [28,30,38]. В подобной ситуации очень полезными оказываются приближенные решения, которые позволяют качественно проанализировать функционирование приборов и получить количественные оценки ряда характеристик. Эти же решения могут использоваться и в качестве начального приближения при численном моделировании.

Часть границы γ^j области S_j , общую для диэлектрика и полупроводника, обозначим через γ_1^j , границу S_j с металлом – через Γ_{1-0}^j , оставшуюся часть γ^j – через γ_2^j , так что: $\gamma^j = \gamma_1^j \cap \gamma_2^j \cap \Gamma_{1-0}^j$. Будем использовать обозначения γ_{1-0}^j и γ_{1+0}^j для сторон границы $\gamma_1^j = \gamma_{1-0}^j \cap \gamma_{1+0}^j$, которые относятся к полупроводнику и диэлектрику соответственно. Внешнюю единичную нормаль к Γ_{1-0}^j , $\gamma_{1,2}^j$ будем обозначать через \mathbf{v} также, как и нормали к Γ_0 .

Наиболее простой вид имеют граничные условия для химических потенциалов. Предполагаем, что токи проводимости через границу Γ_0 не протекают и поэтому на этой границе

$$\partial F_{n,p} / \partial \mathbf{v} = 0. \quad (19)$$

Кроме того, считаем, что проводимость металлических контактов много больше проводимости полупроводника, и падением напряжения в металлических контактах можно пренебречь. Перенос заряда в контактах осуществляется электронами, безразмерный химический потенциал которых F_i (i – номер контакта) равен в каждом контакте приложенному к нему напряжению со знаком минус

$$F_i = -V_i.$$

(Если восстановить размерность, то $F_i = -eV_i$).

Требование локального химического равновесия электронов металла с электронами и дырками на границах полупроводник – металл приводят к условиям на Γ_{1-0}^i

$$F_n = F_p = -V_i. \quad (20)$$

Условия (20) предполагают очень малую, практически нулевую, длину релаксации к локальному химическому равновесию в металле.

Если интересоваться только переносом заряда, то различие между полупроводником и диэлектриком сводится к более широкой запрещенной зоне в диэлектрике (> 3 эВ) и, как следствие, малой концентрации электронов проводимости и дырок. Это означает, что условия (19) справедливы на границах γ_2^j , а условия (20) – на границах Γ_{1-0}^j . Как правило, переносом заряда через диэлектрик пренебрегают и считают, что условия (19) выполнены и на границах

γ_1^j . Физически это означает, что возможный градиент химических потенциалов $F_{n,p}$ на границе полупроводник-диэлектрик заменяется ступенчатой функцией, то есть разрывом $F_{n,p}$. Считая диэлектрик однородным, будем полагать, что

$$F_D = F_n|_{S_j} = F_p|_{S_j}, \quad (21)$$

где F_D – химический потенциал носителей заряда в диэлектрике.

Условия (21) могут быть нарушены при пробое диэлектрика, облучении его фотонами или нейтронами, и т.п. Эти и другие возможные ситуации, когда в областях S_j возникает неравновесные носители заряда и необходимо учитывать их перенос, мы рассматривать не будем. Из выражений (20) и (21) следует, что для каждой области S_j справедливы условия

$$F_D^j = -V_j. \quad (22)$$

Если справедлива статистика Больцмана и легирующая примесь в диэлектрике отсутствует ($N = 0$), уравнение Пуассона примет вид

$$(\varepsilon_D n_{iD} / \varepsilon n_i) \alpha^2 \Delta \varphi = 2 \operatorname{sh}(\varphi + F_D), \quad (23)$$

где ε_D и n_{iD} – диэлектрическая проницаемость диэлектрика и собственная концентрация носителей заряда в нем.

Сделаем ряд предположений относительно области S . Будем считать, что она заполнена диэлектриком с диэлектрической проницаемостью ε_S и практически нулевой собственной концентрацией электронов и дырок так, что справедливо уравнение

$$(\nabla, \varepsilon_S \mathbf{E}) = 0 \quad (24)$$

Вектор \mathbf{E} описывает внешнее по отношению к полупроводнику электрическое поле. Оно может быть создано как сторонними источниками, так и самой полупроводниковой структурой. Будем предполагать, что сторонние источники, если они существуют, расположены достаточно далеко от полупроводниковой структуры и задают на бесконечности поле \mathbf{E}_{ext} . Если $\mathbf{E}_{\text{ext}} = 0$, то поле \mathbf{E} , создается самим полупроводником. В ряде важных для практических приложений задач, в частности, в задачах об электромагнитной совместимости компонент электронных устройств, поле \mathbf{E}_{ext} не является статическим, а изменяется во времени с циклической частотой ω . Будем считать, что для

$$v = \omega/2\pi$$

выполнено условие (2) и ограничимся случаем, когда вклад токов, индуцированных магнитной составляющей как внешнего, так и самосогласованного полей мал по сравнению с токами, вызванными электрическим полем [27,28,29]. Тогда граничные условия для φ и поля \mathbf{E} можно записать в следующем виде

$$\left\{ \begin{array}{l} \mathbf{E}(|\mathbf{r}| = \infty) = \mathbf{E}_{\text{ext}}, \\ (\nabla\varphi, \boldsymbol{\tau}) = -(\mathbf{E}, \boldsymbol{\tau}), \text{ при } \mathbf{r} \in \Gamma_0, \gamma_2^j, \\ (\mathbf{E}, \boldsymbol{\tau}_1) = 0, \text{ при } \mathbf{r} \in \Gamma_i, \Gamma_j, \\ \varepsilon_S(\mathbf{E}, \mathbf{v}_1) = \sigma_{i,j}, \text{ при } \mathbf{r} \in \Gamma_i, \Gamma_j, \\ \varepsilon_S(\mathbf{v}, \nabla\varphi) + \varepsilon_S(\mathbf{E}, \mathbf{v}) = Q_S, \text{ при } \mathbf{r} \in \Gamma_0, \\ \varepsilon_D(\mathbf{v}, \nabla\varphi) + \varepsilon_S(\mathbf{E}, \mathbf{v}) = 0, \text{ при } \mathbf{r} \in \gamma_2^j, \\ \varepsilon_S(\mathbf{v}, \nabla\varphi(\mathbf{r} \in \gamma_{1-0}^j)) - \varepsilon_D(\mathbf{v}, \nabla\varphi(\mathbf{r} \in \gamma_{1+0}^j)) = Q_S, \\ \varphi(\mathbf{r} \in \gamma_{1-0}^j) = \varphi(\mathbf{r} \in \gamma_{1+0}^j). \end{array} \right. \quad (25)$$

Здесь Q_S – встроенный поверхностный заряд, который может существовать на границе полупроводника; σ_i и σ_j – поверхностные заряды, индуцированные на поверхностях проводников Ω_i и Ω_j .

Условия (25) являются стандартными для границ раздела различных сред [39,40]. Они фиксируют требование непрерывности тангенциальной составляющей и определяют условия разрыва нормальной составляющей электрического поля на границах. Однако условия (25) не определяют граничную задачу полностью. Необходимо еще фиксировать условия для потенциала φ на границах Γ_{1-0}^{ij} . Для этого необходимо сделать ряд дополнительных предположений [41].

В отсутствии внешнего электрического поля для получения условий на φ на границах Γ_{1-0}^{ij} обычно используют требование локальной электронейтральности $\Delta\varphi = 0$.

При этом пренебрегают не только полем удаленных источников, но и полем, создаваемым соседними с границей Γ_{1-0}^{ij} областями полупроводника. Учитывая

(26)

требования равновесия (20), на основании условия (26) из первого уравнения системы (11) получим граничное условие на Γ_{1-0}^i , следующее из соотношения (17) при $Y = 0$

$$\varphi_i = V_i + (1/2)\ln[((N^2/4 + 1)^{1/2} + N/2)/((N^2/4 + 1)^{1/2} - N/2)]. \quad (27)$$

Поскольку на границе металл-диэлектрик $N = 0$, то с учетом (22) имеем на Γ_{1-0}^j

$$\varphi_j = V_j. \quad (28)$$

Таким образом, учитывая (21), (22) и (23), приходим к выводу, что в отсутствии внешнего электрического поля уравнение Лапласа (26) выполняется не только на границе, но и в объеме диэлектрика (областях S_j). Вместе с условиями (25) и (28) это уравнение определяет граничную задачу для электростатического потенциала φ .

Однако во внешнем электрическом поле ситуация изменяется. Внешнее поле индуцирует на поверхности металла дополнительные поверхностные заряды, что, в свою очередь, приводит к появлению дополнительного электрического поля под контактами. Нормальные составляющие электрического поля на границах $\Gamma_{1-0}^{i,j}$ и $\Gamma_{1+0}^{i,j}$ будем считать одинаковыми

$$(\mathbf{E}, \mathbf{v})|_{\Gamma_{1-0}^{i,j}} = (\mathbf{E}, \mathbf{v})|_{\Gamma_{1+0}^{i,j}}. \quad (29)$$

Представим потенциал φ в виде двух слагаемых

$$\varphi = \varphi_0 + \varphi_1 \quad (30)$$

и потребуем, чтобы для φ_0 на границах $\Gamma_{1-0}^{i,j}$ были выполнены условия (27) и (28).

Кроме того, потребуем выполнения соотношений

$$\begin{cases} (\mathbf{v}, \nabla \varphi_0)|_{\Gamma_0, \gamma_2^j} = 0, \\ \varepsilon_2(\mathbf{v}, \nabla \varphi)|_{\Gamma_{1-0}^{i,j}} = -\varepsilon_s(\mathbf{E}, \mathbf{v})|_{\Gamma_{1-0}^{i,j}} + Q_s|_{\Gamma_{1-0}^{i,j}}. \end{cases} \quad (31)$$

Здесь и далее используется обозначение

$$\varepsilon_2 = \begin{cases} \varepsilon_s - \text{в полупроводнике и на его границе;} \\ \varepsilon_D - \text{в диэлектрике и на его границе.} \end{cases}$$

Как видно из выражений (25) и (31) для потенциала φ_1 на всех границах должны выполняться условия второго рода. Поскольку концентрации электронов

и дырок связаны с φ соотношениями (11), то должно быть выполнено условие устойчивости для уравнения Пуассона

$$\delta(n - p - N)/\delta\varphi > 0.$$

Этого достаточно, чтобы гарантировать единственность решения краевой задачи для φ_1 . На основании (25) и (31) получим соотношение, связывающее плотность поверхностного заряда на границах Γ_{1+0}^{ij} с нормальной производной потенциала φ на границах Γ_{1-0}^{ij}

$$\sigma_{i,j} = \varepsilon_2(\mathbf{v}, \nabla\varphi)|_{\Gamma_{1-0}^{ij}} - Q_s|_{\Gamma_{1-0}^{ij}}. \quad (32)$$

Формальное решение уравнения (24) с граничными условиями (25) можно записать в виде [30]

$$\varepsilon_S \mathbf{E}(\mathbf{r}) = \varepsilon_S \mathbf{E}_{\text{ext}}(\mathbf{r}) + \int_{\Gamma_0, \gamma_2^j} P_S(\mathbf{r}') \nabla_r K_0(\mathbf{r}', \mathbf{r}) ds(\mathbf{r}') + \sum_{i,j} \int_{\Gamma_{1-0}^{ij}} P_S^{ij}(\mathbf{r}') \nabla_r K_0(\mathbf{r}', \mathbf{r}) ds(\mathbf{r}') + \sum_{i,j} \int_{\Gamma^{ij}} \sigma_{i,j}(\mathbf{r}') \nabla_r K_0(\mathbf{r}', \mathbf{r}) ds(\mathbf{r}'). \quad (33)$$

Здесь P_S – плотность простого электрического слоя на границах Γ_0 и γ_2^j ; P_S^{ij} – плотность простого электрического слоя на границах Γ_{1-0}^{ij} ;

$$\mathbf{r}' \in \Gamma = \Gamma_0 + \sum_{i,j} \Gamma^{ij} + \sum_{i,j} \Gamma_{1-0}^{ij}; \quad \mathbf{r} \in S \cap \Gamma;$$

$$K_0(\mathbf{r}', \mathbf{r}) = \begin{cases} (4\pi|\mathbf{r}' - \mathbf{r}|)^{-1} - \text{в трехмерном случае;} \\ -(1/2\pi)\ln|\mathbf{r}' - \mathbf{r}| - \text{в двумерном случае.} \end{cases}$$

Отметим, что вообще говоря

$$P_S \neq Q_S \text{ и } P_S^{ij} \neq Q_S^{ij},$$

поскольку в область Ω распределен ненулевой объемный заряд. В областях металлических контактов Ω_j такой заряд отсутствует, и плотность простого электрического слоя на их границах совпадает с плотностью поверхностного заряда $\sigma_{i,j}$.

Если для E_{ext} справедливо уравнение

$$(\nabla, \varepsilon_S \mathbf{E}_{\text{ext}}) = 0,$$

то (33) является решением уравнения (24), удовлетворяющим условиям на бесконечности и границах Γ^{ij} . Наличие неизвестной функции P_S позволяет удовлетворить условиям на границах Γ_0 и γ_2^j .

Приравнивая согласно (31) нормальные составляющие поля \mathbf{E} на Γ_{1-0}^{ij} и Γ_{1+0}^{ij} , найдем

$$P_{i,j}^{ij}|\Gamma_{1-0}^{ij} = -\sigma_{i,j}|\Gamma_{1+0}^{ij}.$$

Введем обозначения

$$\begin{cases} \varepsilon_S \mathbf{E}_0(\mathbf{r}) = \int_{\Gamma_0, \gamma_2^j} P_S(\mathbf{r}') \nabla_r K_0(\mathbf{r}', \mathbf{r}) ds(\mathbf{r}'), \\ \varepsilon_S \mathbf{E}_2^k(\mathbf{r}) = \sum_{i,j \neq k} \int_{\Gamma_2^{ij}} \sigma_k(\mathbf{r}') \nabla_r K_0(\mathbf{r}', \mathbf{r}) ds(\mathbf{r}'), \\ \mathbf{E}_{\text{ext}}^k = \mathbf{E}_{\text{ext}} + \mathbf{E}_0 + \mathbf{E}_2^k, \end{cases} \quad (34)$$

Индекс k нумерует металлические контакты и пробегает все возможные значения индексов i и j . Тогда выражение (33) можно представить в следующем виде

$$\varepsilon_S \mathbf{E}(\mathbf{r}) = \varepsilon_S \mathbf{E}_{\text{ext}}^k(\mathbf{r}) + \int_{\Gamma_2^k} \sigma_k(\mathbf{r}') \nabla_r K_0(\mathbf{r}', \mathbf{r}) ds(\mathbf{r}') \quad (35)$$

Выражения (34) и (35) позволяют интерпретировать поле $\mathbf{E}_{\text{ext}}^k$ как внешнее к k -тому контакту. Если считать, что $\mathbf{r} \in \Gamma^k$ (границе k -того контакта), то из (34) и (35) видно, что поле на этой границе создается за счет влияния границ Γ_0 и γ_2^{ij} , других контактов, внешнего поля и собственного поверхностного заряда.

Воспользовавшись известными свойствами потенциала простого слоя [42] и условиями (25), получим интегральное уравнение

$$\int_{\Gamma_2^k} \sigma_k(\mathbf{r}') (\mathbf{v}_1, \nabla_r K_0(\mathbf{r}', \mathbf{r})) ds(\mathbf{r}') |_{(\mathbf{r}=\mathbf{r}_k)} - (1/2) \sigma_k(\mathbf{r}_k) + \varepsilon_S (\mathbf{v}_1, \mathbf{E}_{\text{ext}}^k(\mathbf{r})) |_{\Gamma^k} = 0, \quad \mathbf{r}_k \in \Gamma^k, \quad (36)$$

которое необходимо еще дополнить условием (32). Согласно (36) и (32) поверхностный заряд на Γ^k индуцируется полями \mathbf{E}_{ext} , \mathbf{E}_0 , \mathbf{E}_2^k и полем, создаваемым полупроводником на границе металл-полупроводник.

Как уже отмечалось, в отсутствии внешнего поля граничные условия для потенциала ϕ на контактах определяются выражениями (27) и (28). При этом не учитывается влияние на потенциал поля удаленных источников и полей, которые создают свободные границы полупроводника и другие контакты. Такая идеализация приемлема, если поверхностный заряд

$$\sigma_{i,j}^0 |\Gamma_{1+0}^{ij} = (\varepsilon_2(\mathbf{v}, \partial \phi_0 / \partial \mathbf{v}) - Q_S) |\Gamma_{1-0}^{ij}$$

является решением уравнения (36). Таким образом, краевая задача (36), (32) оказывается тесно связанной с граничными условиями для уравнений переноса заряда.

Представим $\sigma_{i,j}$ в виде

$$\sigma_k = \sigma_k^0 + f_k(\mathbf{E}_{\text{ext}}^k, \mathbf{v}_1)|\Gamma^k. \quad (37)$$

Для функции f_k из (30), (32) и (36) получим выражения

$$\begin{cases} \int_{\Gamma_2^k} f_k(\mathbf{r}')(\mathbf{E}_{\text{ext}}^k, \mathbf{v}_1)(\mathbf{v}_1, \nabla_{\mathbf{r}} K_0(\mathbf{r}', \mathbf{r})) ds(\mathbf{r}')|_{(\mathbf{r}=\mathbf{r}_k)} + ((\mathbf{E}_{\text{ext}}^k, \mathbf{v}_1)|\Gamma^k)(\varepsilon_S - (1/2)f_k(\mathbf{r}))|_{(\mathbf{r}=\mathbf{r}_k)} = 0, \\ \varepsilon_2(\mathbf{v}_1, \nabla\varphi_1)|\Gamma_{1-0}^k = -f_k(\mathbf{E}_{\text{ext}}^k, \mathbf{v}_1)|\Gamma_{1+0}^k. \end{cases} \quad (38)$$

С учетом (38) граничные условия для φ_0 и φ_1 примут вид

$$\begin{cases} \varphi_0|\Gamma_{1-0}^k = V_k + (1/2)\ln[((N^2/4 + 1)^{1/2} + N/2)/((N^2/4 + 1)^{1/2} - N/2)]|\Gamma_{1-0}^k, \\ (\mathbf{v}, \nabla\varphi_0)|\Gamma_0 = 0, \\ \varepsilon_2(\mathbf{v}, \nabla\varphi_1)|\Gamma_0 = -\varepsilon_S(\mathbf{E}, \mathbf{v})|\Gamma_0 + Q_S, \\ \varepsilon_2(\mathbf{v}, \nabla\varphi_1)|\Gamma_{1-0}^k = -f_k(\mathbf{E}, \mathbf{v})|\Gamma_{1+0}^k, \end{cases} \quad (39)$$

где предполагается, что в диэлектрике ($k=j$) $N=0$.

При моделировании функционирования полупроводниковых приборов в отсутствии внешнего поля ограничиваются двумя первыми условиями (39). Пренебрежение полем, создаваемым полупроводниковой структурой во внешней области S , можно оправдать ее малой диэлектрической проницаемостью, то есть неравенством

$$\varepsilon_S \ll \varepsilon_s.$$

Однако в задачах об электромагнитной совместимости элементов интегральных схем учет полей, создаваемых соседними для данного элемента схемы, может оказаться существенным.

Сформулированные в этом параграфе граничные условия используются для постановки задач моделирования переноса заряда в полупроводниковых структурах. При исследовании нестационарных процессов как правило считают, что

$$V_i = V_i(t). \quad (40)$$

То есть приложенные напряжения являются заданными функциями координат. Особняком стоят задачи о возникновении автоколебаний в полупроводниковых структурах [43].

При анализе работы электронных приборов мы ограничимся далее ситуациями, когда влияние внешнего электрического поля на перенос заряда можно не учитывать. Однако его учет при постановке граничных условий необходим для ясного понимания их физического смысла.

1.5. Новый базис переменных

Система уравнений (1) (или (11)) достаточно сложна для аналитического исследования. Это обусловлено, в частности, большими градиентами концентрации легирующей примеси N , сосредоточенными в малых областях p - n -переходов. Аналогичные большие градиенты возникают в p - n -переходах для потенциала φ и концентраций n и p .

Если пытаться построить приближенные решения системы (1) в базисе переменных $\{n, p, \varphi\}$, то для учета p - n -переходов приходится вводить дополнительные внутренние границы [28,44], соответствующие этим переходам. Поскольку, несмотря на высокие градиенты N , p - n -переходы являются все-таки протяженными, постольку введение внутренних границ на них является искусственной и неоднозначной процедурой. Преимуществом нового базиса переменных являются:

- возможность построить приближенные решения уравнений системы (1) (или (11)) без введения внутренних границ;
- возможность провести анализ функционирования полупроводниковых приборов уже в низших порядках асимптотического разложения решения уравнений системы (1) (или (11)).

В качестве новых переменных выберем следующие функции

$$\begin{cases} Y = (np - n_i^2)/n_i^2 = \exp\{(F_n - F_p)/\Theta\} - 1, \\ F = - (F_n + F_p), \\ \varphi_{ch} = \varphi - \varphi_{en}. \end{cases} \quad (41)$$

Заметим, что по определению Y для этой величины справедливо неравенство

$$Y > -1.$$

Потенциал φ_{en} является решением уравнения электронейтральности (14) и согласно (17) и (41) может быть представлен в виде

$$\varphi_{en} = (1/2e)\{F + \Theta \ln[(N^2 + 4n_i^2(1 + Y))^{1/2} + N]/((N^2 + 4n_i^2(1 + Y))^{1/2} - N)\}. \quad (42)$$

Для большинства полупроводниковых приборов концентрация заряда примеси N является функцией координат. Поэтому лапласиан от потенциала φ_{en} , вообще говоря, не равен нулю, то есть область, характеризуемая этим потенциалом, не обязательно является электронейтральной. Более корректно называть ее квазиэлектронейтральной.

Используя (1), (41) и (42), не составляет труда выразить переменные $\{n, p, \varphi\}$ через переменные $\{Y, F, \varphi_{ch}\}$

$$\begin{cases} n = (1/2)[(N^2 + 4n_i^2(1 + Y))^{1/2} + N]\exp\{e\varphi_{ch}/\Theta\}, \\ p = (1/2)[(N^2 + 4n_i^2(1 + Y))^{1/2} - N]\exp\{-e\varphi_{ch}/\Theta\}, \\ \varphi = (1/2e)\{F + \Theta \ln[(N^2 + 4n_i^2(1 + Y))^{1/2} + N]/((N^2 + 4n_i^2(1 + Y))^{1/2} - N)\} + \varphi_{ch}. \end{cases} \quad (43)$$

Для наглядности соотношения (41), (42) и (43) записаны в размерном виде. Далее, по-прежнему будут использоваться безразмерные величины, получаемые из размерных согласно правилам, описанным в параграфе 1.3.

Как уже отмечалось, переменная Y описывает отклонение электронов и дырок от состояния химического равновесия. Если $Y > 0$ в некоторой части полупроводниковой структуры, то эта часть обогащена неравновесными носителями заряда. Действительно, согласно (43)

$$n(Y > 0, \varphi_{ch}) > n(Y = 0, \varphi_{ch}),$$

$$p(Y > 0, \varphi_{ch}) > p(Y = 0, \varphi_{ch}).$$

В обогащенных электронами и дырками областях в соответствии с (11), (12) наблюдается их рекомбинация, интенсивность которой возрастает с ростом Y . В областях, где

$$-1 < Y < 0$$

наблюдается обеднение по сравнению с равновесным состоянием. В этом случае согласно (11), (12) происходит генерация носителей заряда.

Как следует из этих простых рассуждений, в случае однородной полупроводниковой структуры и в отсутствии лавинной генерации система уравнений (11) (или (1)) обеспечивает устойчивость состояния локального химического равновесия

$$Y = 0.$$

Однако могут возникать ситуации, когда эта устойчивость нарушается. Некоторые из таких возможностей рассмотрены в [43]. В книге [28] показано, что для системы уравнений (1) всегда найдется такая окрестность значений приложенных напряжений вблизи нулевых значений, в которой система (1) устойчива относительно малых возмущений n , p и ϕ .

Согласно теории сингулярных возмущений [44-46] потенциал ϕ_{en} представляет собой главный член разложения потенциала ϕ в ряд по малому параметру α_{ch} . В отсутствие внешнего электрического поля функция ϕ_{en} удовлетворяет граничным условиям на омических контактах. Поэтому существенных отличий потенциала ϕ от ϕ_{en} следует ожидать в окрестностях p-n-переходов.

Функция F удобна в том отношении, что для нее легко формулируются граничные условия. Кроме того, совместное использование переменных F и Y позволяет выделить в выражениях для плотностей токов часть, связанную с неравновесностью электронов и дырок.

С учетом (41), (42) и (12) система уравнений (11) примет вид

$$\left\{ \begin{array}{l}
(\nabla, \alpha^2 \nabla \varphi_{ch}) = wsh\varphi_{ch} + Nch\varphi_{ch} - (N + (\nabla, \alpha^2 \nabla \varphi_{en})), \\
- (\partial Y / \partial t) [(1 - Q / Q_+) \exp\{\varphi_{ch}\} + (1 + Q / Q_+) \exp\{-\varphi_{ch}\}] / w + \beta^2 (\nabla, (\mu_n \mu_p / Q_+) \nabla Y) = \\
= qY - G - (\partial \varphi_{ch} / \partial t) (1 + Y) (\mu_n - \mu_p) / Q_+ + (\beta^2 / 2) (\mathbf{i}, \nabla (Q / Q_+)), \\
(\nabla, \mathbf{j}) = 0, \\
\mathbf{j} = \mathbf{i} + (\alpha / \beta)^2 \partial (\nabla \varphi_{ch} + \nabla \varphi_{en}) / \partial t, \mathbf{i} = \mathbf{j}_n + \mathbf{j}_p, \\
\mathbf{j}_n = \mu_n n \nabla (F - \ln(1 + Y)) / 2, \mathbf{j}_p = \mu_p p \nabla (F + \ln(1 + Y)) / 2, \\
n = (1/2)(w + N) \exp\{\varphi_{ch}\}, p = (1/2)(w - N) \exp\{-\varphi_{ch}\}.
\end{array} \right. \quad (44)$$

Здесь введены обозначения

$$\left\{ \begin{array}{l}
w = (N^2 + 4(1 + Y))^{1/2}, \\
Q_{\pm} = \mu_n n \pm \mu_p p.
\end{array} \right. \quad (45)$$

При анализе уравнений переноса заряда часто бывает необходимо восстановить размерности входящих в них величин. Поэтому дополнительно к (44) приведем уравнения в новых переменных с сохранением размерностей.

$$\left\{ \begin{array}{l}
(\nabla, \varepsilon_s \nabla \varphi_{ch}) = e \{ wsh(e\varphi_{ch} / \Theta) + Nch(e\varphi_{ch} / \Theta) - (N + (\nabla, \varepsilon_s \nabla \varphi_{en})) \}, \\
- (\partial Y / \partial t) [(1 - Q / Q_+) \exp\{e\varphi_{ch} / \Theta\} + (1 + Q / Q_+) \exp\{-e\varphi_{ch} / \Theta\}] (n_i^2 / w) + \\
(\Theta / e) (\nabla, (\mu_n \mu_p n_i^2 / Q_+) \nabla Y) = \\
= qY - G - (e / \Theta) n_i^2 (1 + Y) ((\mu_n - \mu_p) / Q_+) (\partial \varphi_{ch} / \partial t) - e^{-1} (\mathbf{I}, \nabla (Q / Q_+)), \\
(\nabla, \mathbf{J}) = 0, \\
\mathbf{J} = \mathbf{I} - \varepsilon \partial (\nabla (\varphi_{ch} + \varphi_{en})) / \partial t, \\
\mathbf{I} = \mathbf{J}_n + \mathbf{J}_p, \\
\mathbf{J}_n = -(1/2) \mu_n n \nabla (F - \Theta \ln(1 + Y)), \mathbf{J}_p = -(1/2) \mu_p p \nabla (F + \Theta \ln(1 + Y)), \\
n = (1/2)(w + N) \exp\{e\varphi_{ch} / \Theta\}, p = (1/2)(w - N) \exp\{-\varphi_{ch} / \Theta\}.
\end{array} \right. \quad (44.P)$$

Для величин Q_{\pm} в (44.P) справедливы формулы (45), где концентрации и подвижности носителей заряда выражены в размерных единицах. Величина w в уравнениях (44.P) вычисляется по формуле

$$w = (N^2 + 4n_i^2(1 + Y))^{1/2},$$

то есть также как в (45) только с восстановленной размерностью концентрации примеси.

Рассмотрим сначала уравнения для плотностей токов проводимости электронов \mathbf{j}_n и дырок \mathbf{j}_p , фигурирующие в системе уравнений (44). Они являются обобщением дифференциального закона Ома на случай двухкомпонентной неравновесной системы. Легко получить

$$\nabla F = 2\mathbf{i}/Q_+ + (Q_-/Q_+)\nabla \ln(1 + Y). \quad (46)$$

Откуда видно, что вклад неравновесности в падение напряжения (также как и в плотности токов, согласно (44)) определяется градиентом Y . При $Y = \text{const}$

$$\nabla F = 2\mathbf{i}/(\mu_n n + \mu_p p),$$

то есть падение напряжения в полупроводнике является чисто омическим.

Как и для системы уравнений (11) граничные задачи для системы (44) можно формулировать, задавая приложенные напряжения на металлических (омических) контактах. В переменных Y и F условия (19) и (20) примут вид

$$\begin{cases} (\nabla F, \mathbf{v})|_{\Gamma_0} = (\nabla Y, \mathbf{v})|_{\Gamma_0} = 0, \\ Y|\Gamma_{1-0}^i = 0, \\ F|\Gamma_{1-0}^i = 2V_i. \end{cases} \quad (47)$$

Граничные условия для φ_{ch} без труда получаются из (25) и (39). Если представить φ_{ch} в виде суммы

$$\varphi_{ch} = \varphi_{ch}^0 + \varphi_1, \quad (48)$$

то для φ_{ch}^0 получим

$$\begin{cases} \varphi_{ch}^0|\Gamma_{1-0}^i = 0, \\ \partial \varphi_{ch}^0 / \partial \mathbf{v}|_{\Gamma_0} = -\partial \varphi_{en} / \partial \mathbf{v}|_{\Gamma_0}. \end{cases} \quad (49)$$

Для φ_1 , очевидно, справедливы граничные условия (39). Из выражений (48) и (49) вытекает, что

$$\varphi_0 = \varphi_{ch}^0 + \varphi_{en}. \quad (50)$$

Как следует из (46), граничные условия для задачи о переносе заряда в полупроводнике можно переформулировать в терминах не напряжений, а токов, протекающих через контакты. При этом приложенные напряжения вычисляются в

результате решения задачи (46), (47). Действительно, если известны \mathbf{i} , Y и φ_{ch} , то правая часть (46) полностью определена и справедливо выражение

$$F(\mathbf{r}) = F(\mathbf{r}_0) + \int_{\mathbf{r}_0}^{\mathbf{r}} d\mathbf{l} (2\mathbf{i}/Q_+ + (Q/Q_+) \nabla \ln(1 + Y)). \quad (51)$$

Здесь \mathbf{r}_0 некоторая точка, которую удобно выбрать на одном из омических контактов, к которому приложено напряжение V_0 , так что

$$F(\mathbf{r}_0) = 2V_0. \quad (52)$$

Постоянная V_0 определяет начало отсчета потенциала. Для всех $\mathbf{r}_i \in \Gamma_{1-0}^i$ получим

$$V_i = V_0 + \int_{\mathbf{r}_0}^{\mathbf{r}} d\mathbf{l} (\mathbf{i}/Q_+ + (Q/2Q_+) \nabla \ln(1 + Y)). \quad (53)$$

Заметим, что путь интегрирования в выражениях (52) и (53) может быть произвольным. Для того чтобы результат интегрирования не зависел от пути необходимо, чтобы ротор от правой части выражения (46) равнялся нулю. Это требование однозначной определенности потенциала F . Вместе с третьим, четвертым и седьмым уравнениями системы (44) оно определяет систему уравнений для плотности тока \mathbf{j}

$$\begin{cases} \mathbf{j} = \mathbf{i} + (\alpha^2/\beta^2) \partial(\nabla\varphi)/\partial t, \\ (\nabla, \mathbf{j}) = 0, \\ \text{rot}\{2\mathbf{i}/Q_+ + (Q/Q_+) \nabla \ln(1 + Y)\} = 0. \end{cases} \quad (54)$$

Граничные условия для системы уравнений (54) следуют из требований постоянства химических потенциалов F_n и F_p на контактах и отсутствия тока проводимости i через границы Γ_0 .

$$\begin{cases} (\mathbf{j}, \boldsymbol{\tau})|_{\Gamma_{1-0}^i} = (\alpha/\beta)^2 \partial(\nabla\varphi, \boldsymbol{\tau})/\partial t|_{\Gamma_{1-0}^i}, \\ (\mathbf{j}, \mathbf{v})|_{\Gamma_0} = (\alpha/\beta)^2 \partial(\nabla\varphi, \mathbf{v})/\partial t|_{\Gamma_0}, \\ J_i = \int_{\Gamma_{1-0}^i} (\mathbf{j}, \mathbf{v}) ds. \end{cases} \quad (55)$$

Постоянная J_i имеет смысл суммарного тока, протекающего через i -й контакт. Условием разрешимости задачи (54), (55) является требование сохранения суммарного тока

$$\int_{\Gamma} (\mathbf{j}, \mathbf{v}) ds = 0. \quad (56)$$

Откуда с учетом (55) найдем

$$\sum_{i=1}^k J_i = -(\alpha/\beta)^2 \partial(\int_{\Gamma_0} (\nabla\varphi, \mathbf{v}) ds) / \partial t, \quad (57)$$

где k – полное число контактов. Таким образом, можно независимо задать $k-1$ постоянную J_i и еще V_0 . Получается то же количество заданных постоянных, что и при постановке задачи по напряжениям на контактах. В стационарном случае, когда плотности токов \mathbf{i} и \mathbf{j} совпадают, формулировка задачи упрощается.

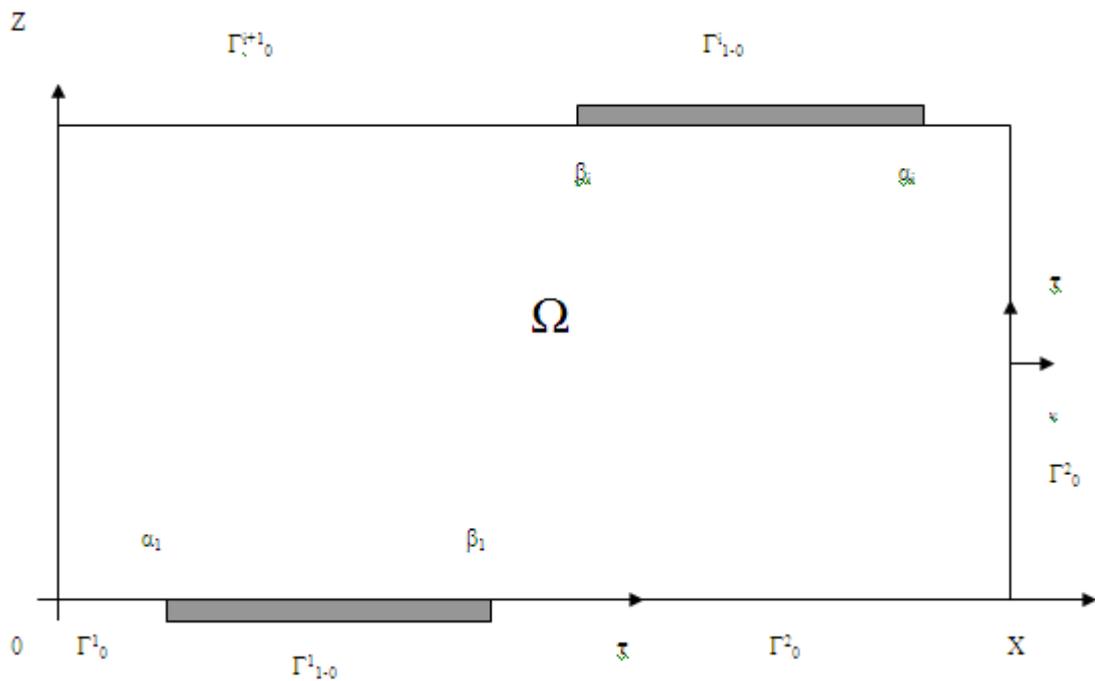


Рисунок 2. Схематичное изображение двумерной области.

Если решать задачу в двумерной области (см. Рисунок 2), чего часто бывает достаточно на практике, то удобно переформулировать ее в терминах функции тока χ

$$j_x = -\partial\chi/\partial z, j_z = \partial\chi/\partial x. \quad (58)$$

При такой подстановке второе из уравнений (54) будет удовлетворено автоматически, а из третьего найдем

$$\begin{cases} \partial(Q_+^{-1}\partial\chi/\partial x)/\partial x + \partial(Q_+^{-1}\partial\chi/\partial z)/\partial z = \mathbf{L}, \\ \mathbf{L} = (\alpha/\beta)^2[(\partial Q_+^{-1}/\partial x)(\partial^2\varphi/\partial t\partial z) - (\partial Q_+^{-1}/\partial z)(\partial^2\varphi/\partial t\partial x)] \\ \quad + (1/2)[(\partial(Q_-/Q_+)/\partial z)(\partial\ln(1+Y)/\partial x) - (\partial(Q_-/Q_+)/\partial x)(\partial\ln(1+Y)/\partial z)]. \end{cases} \quad (59)$$

Граница Γ_0, Γ_0' будет в двумерном случае разбита областями Γ_{1-0}^i на части, которые мы обозначим через Γ_0^i в порядке, указанном на Рисунке 2. Контактные области нумеруются при обходе области Ω вдоль границы Γ против часовой стрелки. Координаты начала контакта обозначим через α_i , а координаты конца контакта – через β_i . Введем точку

$$(\alpha_0, \beta_0) = (0, 0).$$

Граничные условия (55) примут вид

$$\begin{cases} (\mathbf{v}, \nabla\chi)|_{\Gamma_{1-0}^i} = (\alpha/\beta)^2 \partial(\nabla\varphi, \boldsymbol{\tau})/\partial t|_{\Gamma_{1-0}^i}, \\ \chi(s)|_{s \in \Gamma_0^{i+1}} = \chi(\alpha_i) - J_i - (\alpha/\beta)^2 \partial(\int_{\beta_i}^s (\nabla\varphi, \mathbf{v}) dl)/\partial t. \end{cases} \quad (60)$$

Положим

$$J_0 = 0$$

и произвольную постоянную $\chi(\alpha_0)$ также приравняем нулю. Таким образом, система уравнений (54) свелась в двумерном случае к смешанной краевой задаче для эллиптического уравнения (59).

Постановка задачи в терминах заданных токов через контакты удобна для ряда приборов, имеющих так называемую S-образную вольт – амперную характеристику. Решение уравнений переноса заряда для них оказывается единственным при заданных токах через контакты [47]. Естественно, что для приборов, обладающих N-образной ВАХ, предпочтительной является формулировка в терминах заданных на контактах напряжений.

В заключение этого параграфа отметим еще некоторые важные свойства уравнений для φ_{ch} и Y . Пусть функция φ_a – некоторое решение уравнения

Пуассона из системы уравнений (11). Если искать другое решение φ этого же уравнения в виде

$$\varphi = \varphi_a + \varphi_b, \quad (61)$$

то для φ_a и φ_b справедливы соотношения

$$\begin{cases} \varphi_a = (1/2)\{F + \ln[(w + Q)/(w - Q)]\}, \\ Q = N + \alpha^2 \Delta \varphi_a, \\ w_a = (Q^2 + 4(1 + Y))^{1/2}, \\ \alpha^2 \Delta \varphi_b = w_a \operatorname{sh} \varphi_b + Q(\operatorname{ch} \varphi_b - 1). \end{cases} \quad (62)$$

При таком выборе представления для φ все уравнения системы (44), кроме первого, сохраняют свой вид. В них необходимо только произвести замены

$$\varphi \rightarrow \varphi_a, \quad w \rightarrow w_a.$$

Ясно, что если

$$\alpha^2 |\Delta \varphi_a| \ll |N|,$$

то

$$\varphi_a \approx \varphi_{\text{ен}}.$$

Таким образом, решения уравнения Пуассона в полупроводнике могут быть представлены в некотором инвариантном виде (первое уравнение системы (62)). Если найдено решение φ_a , не удовлетворяющее граничным условиям, то им можно удовлетворить при помощи функции φ_b . Заметим, что для функции

$$f(\varphi_b) = w_a \operatorname{sh} \varphi_b + Q(\operatorname{ch} \varphi_b - 1)$$

при всех φ_b справедливо

$$\delta f(\varphi_b) / \delta \varphi_b = w_a \operatorname{ch} \varphi_b + Q \operatorname{sh} \varphi_b > 0.$$

Поэтому точка покоя уравнения для φ_b является седлом и решения граничных задач с условиями первого и второго рода для этого уравнения единственны.

Уравнение для Y в системе (44) представляет собой нелинейное параболическое уравнение. Правая часть его существенно зависит от Y . Источниками неравновесности электронов и дырок, согласно (44), являются:

- протекание тока через области, где присутствуют градиенты заряда примеси и подвижностей;

- генерация носителей заряда внешним источником, описываемым слагаемым G ;
- изменение во времени электростатического потенциала φ .

Последний фактор может стать также и причиной потери устойчивости уравнения для Y . Действительно, из второго уравнения системы (44) следует, что роль коэффициента затухания для Y играет величина

$$\Omega = w(q - ((\mu_n - \mu_p)/Q_+) \partial \varphi_{ch} / \partial t) / [(1 - Q/Q_+) \exp\{\varphi_{ch}\} + (1 + Q/Q_+) \exp\{-\varphi_{ch}\}]. \quad (63)$$

В случае, когда

$$((\mu_n - \mu_p)/Q_+) \partial \varphi_{ch} / \partial t > q, \quad (64)$$

Ω отрицательна и возможны экспоненциально возрастающие по времени решения уравнения для Y .

В большинстве задач моделирования элементов интегральных схем и полупроводниковых приборов основным источником неравновесности является функция

$$S = \beta^2 w(\mathbf{i}, \nabla(Q/Q_+)) / [(1 - Q/Q_+) \exp\{\varphi_{ch}\} + (1 + Q/Q_+) \exp\{-\varphi_{ch}\}]. \quad (65)$$

Уровень неравновесности, как правило, не очень высок и справедливо неравенство

$$-1 < Y \ll N_{ch}^2, \quad (66)$$

где N_{ch} – характерное значение абсолютной величины заряда примеси в единицах собственной концентрации n_i , и

$$N_{ch} \gg 1.$$

В этом случае

$$Q/Q_+ \approx \begin{cases} 1 & \text{в n-областях, когда } N > 0, \\ -1 & \text{в p-областях, когда } N < 0. \end{cases} \quad (67)$$

Соотношения (67) справедливы вплоть до членов порядка

$$(1 + Y)/(N_{ch})^2,$$

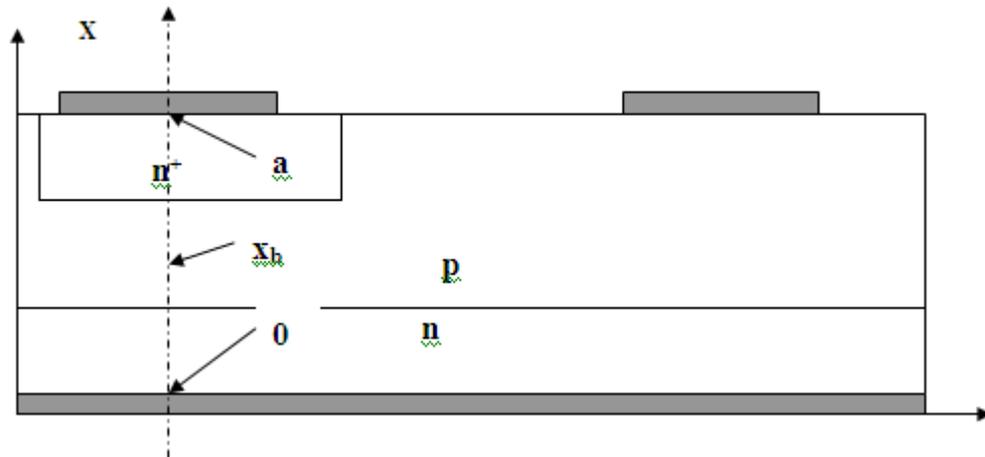
малых, если выполнено неравенство (66). Таким образом, величина S будет существенно отличаться от нуля только вблизи p-n-переходов, где отношение

Q_-/Q_+ меняет знак. То есть именно вблизи р-n-переходов при условии протекания тока в перпендикулярном к ним направлении сосредоточены источники неравновесности электронов и дырок.

1.6. Одномерные краевые задачи

В одномерном приближении постановка краевых задач, описывающих функционирование полупроводниковых приборов, требует дополнительных предположений. Поэтому они выделены в отдельный параграф.

Сложности возникают при рассмотрении структур с более чем двумя контактами [29]. Поскольку для каждой из переменных порядок уравнений в системе, описывающей перенос заряда, не превышает второго, то и количество краевых условий в одномерном случае не может превышать двух. Покажем, как обычно выходят из этого затруднительного положения.



Будем рассматривать задачу на отрезке $x \in [0, a]$. В точках $x = 0$ и $x = a$ расположены омические контакты. Выбор этого отрезка в реальной трехмерной структуре может осуществляться по-разному. Например, это может быть линия, вдоль которой плотность тока максимальна. На Рисунке 3 изображены такие линии для двух возможных вариантов транзисторных структур. Ясно, что расположение указанных линий не определяется полностью геометрией, но может зависеть и от режимов работы приборов. По этой причине адекватность одномерных моделей в ряде случаев сомнительна. Однако в качестве начального

приближения, позволяющего понять и оценить основные особенности работы приборов, одномерные модели вполне пригодны и часто используются.

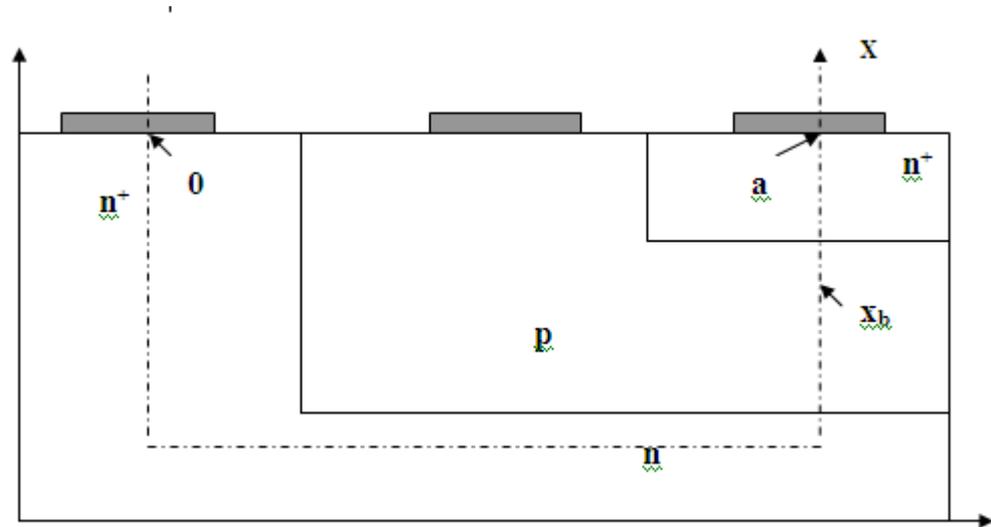


Рисунок 3. Выбор сечения при переходе к одномерной модели транзистора для двух различных конфигураций прибора.

Граничные условия для переменных φ_{ch} и Y логично вытекают из трехмерных и задаются в точках $x = 0$ и $x = a$

$$\begin{cases} Y(0) = Y(a) = 0, \\ \varphi_{ch0}(0) = \varphi_{ch0}(a) = 0. \end{cases} \quad (68)$$

В случаях, когда существенно влияние внешнего поля, согласно (39) и (41) получим

$$\begin{cases} \varphi_{ch} = \varphi_{ch0} + \varphi_{ch1}, \\ \partial\varphi_{ch1}/\partial x|_{x=0,a} = f(E_{ext}), \end{cases} \quad (69)$$

где f – известная функция. При $E_{ext} = 0$

$$f(E_{ext}) = 0 \text{ и } \varphi_{ch1} = 0,$$

то есть

$$\varphi_{ch} = \varphi_{cp0}.$$

Предполагается, что F связана с приложенными напряжениями на контактах в точках $x = 0$ и $x = a$

$$F(0) = 2V_0, F(a) = 2V_a. \quad (70)$$

При наличии третьего контакта, не принадлежащего линии токоперевода, но не отделенного от этой линии р-п-переходами, считают, что химический потенциал

основных носителей не меняется в области одного типа легирования под этим контактом. Будем называть такой контакт базовым. На линии Ox выбирают некоторую точку x_b и считают, что она соответствует базовому контакту. Предполагается, что в этой точке происходит разрыв плотности суммарного тока

$$j = \begin{cases} j_1, & 0 \leq x < x_b, \\ j_2, & x_b < x \leq a \end{cases} \quad (71)$$

и

$$\begin{cases} F_n = -V_b, & \text{если } x_b \in n \text{ – области,} \\ F_p = -V_b, & \text{если } x_b \in p \text{ – области.} \end{cases} \quad (72)$$

В терминах F и Y из (72) получим

$$\begin{cases} F = 2V_b + \ln(1 + Y(x_b)), & \text{если } x_b \in n \text{ – области,} \\ F = 2V_b - \ln(1 + Y(x_b)), & \text{если } x_b \in p \text{ – области.} \end{cases} \quad (73)$$

В принципе, количество базовых контактов может быть произвольным. Сделанные предположения физически обоснованы при не очень больших токах, когда омическим падением напряжения в области под базовым контактом можно пренебречь.

ГЛАВА 2. Приближенные решения краевых задач переноса заряда в полупроводниках

Аналитические методы исследования, основанные на асимптотических разложениях и теории сингулярных возмущений [44-46,48], позволяют при помощи относительно простых и легко интерпретируемых формул описать основные электрические характеристики исследуемых приборов и понять тенденции их изменения при варьировании электрофизических параметров структур. Полученные результаты могут служить рабочим инструментом для исследователей и инженеров, проектирующих новые приборы.

Кроме того, приближенные решения оказываются полезными при использовании в качестве начального приближения для численного решения сложных задач переноса заряда в полупроводниках [28,29,38,48-54]. Их применение не только многократно ускоряет сходимость численных алгоритмов, но и является дополнительной страховкой от получения физически бессмысленного «численного решения».

Здесь мы построим приближенные решения некоторых краевых задач физики переноса заряда в полупроводниках в одномерном и двумерном случаях. Полученные результаты мы обсудим применительно к функционированию полупроводниковых приборов.

В основу построения приближенных решений положен метод последовательных приближений. Уравнения переноса заряда решаются по очереди. Сначала находится приближенное решение уравнения Пуассона, учитывающее основные физические особенности задачи. Затем, во «внешнем поле», определяемым этим решением, строятся приближенные решения остальных уравнений. Наибольший интерес представляют низшие члены асимптотических разложений решений, содержащие основную физическую информацию. Уточнение решений эффективно производить на основе численных методов.

2.1. Асимптотика решения уравнения Пуассона

Приближенное решение уравнения Пуассона будем строить как асимптотическое разложение по малому параметру

$$\varepsilon = \alpha/m^{1/2} \ll 1, \quad (74)$$

где

$$m = \max(|N|).$$

Здесь рассматривается случай, когда внешнее поле отсутствует. Асимптотика решения уравнения Пуассона в полупроводнике с учетом внешнего поля построена в работе [51].

Для потенциала φ воспользуемся представлением (41), (42). С учетом определения (74) уравнение для φ_{ch} из системы (44) примет вид

$$\varepsilon^2 \Delta \varphi_{ch} = A \varphi_{ch} + B(ch\varphi_{ch} - 1) - g, \quad (75)$$

где

$$A = w/m, \quad B = N/m, \quad g = \varepsilon^2 \Delta \varphi_{en}.$$

Будем считать, что $A = O(1)$ и $B = O(1)$ всюду в Ω , а $g = O(1)$ внутри Ω и $g = O(\varepsilon)$ в некоторой окрестности границы и на самой границе Γ области Ω .

Требование гладкости для функции g накладывает согласно (42) соответствующие ограничения на N . Величина φ_{en} не может по порядку величины быть больше, чем x^2/ε^2 внутри Ω и x/ε на границе Γ . Соответственно, как следует из (42), функция N может быть порядка $(x/\varepsilon)^s$, где s – любое число в Ω и даже порядка $\exp\{a(x/\varepsilon)^2\}$ внутри Ω и порядка $\exp\{b(x/\varepsilon)\}$ вблизи границы Γ (a и $b \sim O(1)$). Возможна и любая регулярная зависимость N от ε . Таким образом, в область нашего анализа попадает достаточно широкий класс возможных функций N .

Основываясь на сделанных предположениях, исследуем далее одномерную и двумерную краевые задачи для уравнения (75).

2.1.1. Одномерная модель

Краевая задача для φ_{ch} на отрезке $x \in [0, a]$ в этом случае будет выглядеть так

$$\begin{cases} \varepsilon^2 d^2 \varphi_{ch} / dx^2 = A(x) \operatorname{sh} \varphi_{ch} + B(x) (\operatorname{ch} \varphi_{ch} - 1) - g(x, \varepsilon), \\ \varphi_{ch}(0) = 0, \\ \varphi_{ch}(a) = 0. \end{cases} \quad (76)$$

Асимптотическое решение будем строить в виде

$$\varphi_{ch}(x, \varepsilon) = \Phi(x, \varepsilon) + \Pi_1(\xi, \varepsilon) + \Pi_2(\xi^*, \varepsilon), \quad (77)$$

где $\Phi(x, \varepsilon)$ – регулярная часть асимптотики; $\Pi_{1,2}$ – пограничные функции;

$$\xi = x/\varepsilon, \quad \xi^* = (a - x)/\varepsilon \quad (78)$$

- погранслоиные переменные. Каждое из слагаемых в (77) представляет собой ряд по степеням ε

$$\begin{cases} \Phi(x, \varepsilon) = \sum_{l=0}^{\infty} \varepsilon^l \Phi_{1,l}(x), \\ \Pi_1(\xi, \varepsilon) = \sum_{l=0}^{\infty} \varepsilon^l \Pi_{1,l}(\xi), \\ \Pi_2(\xi^*, \varepsilon) = \sum_{l=0}^{\infty} \varepsilon^l \Pi_{2,l}(\xi^*). \end{cases} \quad (79)$$

Пограничные функции стремятся к нулю при стремлении погранслоиных переменных к бесконечности

$$\Pi_1(\xi, \varepsilon)|_{\xi \rightarrow \infty} \rightarrow 0, \quad \Pi_2(\xi^*, \varepsilon)|_{\xi^* \rightarrow \infty} \rightarrow 0.$$

Функцию

$$f(\varphi_{ch}, x) = A(x) \operatorname{sh} \varphi_{ch} + B(x) (\operatorname{ch} \varphi_{ch} - 1)$$

представим в виде

$$f(\varphi_{ch}, x) = f_0(x, \varepsilon) + \Pi_1 f(\xi, \varepsilon) + \Pi_2 f(\xi^*, \varepsilon), \quad (80)$$

где

$$\begin{cases} f_0(x, \varepsilon) = f(\Phi(x, \varepsilon), x), \\ \Pi_1 f(\xi, \varepsilon) = f(\Phi(\varepsilon \xi, \varepsilon) + \Pi_1(\xi, \varepsilon), \varepsilon \xi) - f_0(\varepsilon \xi, \varepsilon), \\ \Pi_2 f(\xi^*, \varepsilon) = f(\Phi(a - \varepsilon \xi^*, \varepsilon) + \Pi_2(\xi^*, \varepsilon), a - \varepsilon \xi^*) - f_0(a - \varepsilon \xi^*, \varepsilon). \end{cases} \quad (81)$$

Необходимо отметить, что равенство (80) не является тождеством. Оно справедливо в каждом порядке по ε с точностью до экспоненциально малых членов [49].

Подставив (79), (80) и (81) в (76) и приравняв однотипные слагаемые при одинаковых степенях ε , получим уравнения и дополнительные условия для членов асимптотики. В нулевом приближении для $\Phi_0(x, \varepsilon)$ имеем

$$A(x)\text{sh}\Phi_0 + B(x)(\text{ch}\Phi_0 - 1) = g(x, 0).$$

Так как $A > B$ при всех $x \in [0, a]$, то это уравнение имеет единственное решение

$$\begin{aligned} \Phi_0 &= \ln\{(B + g + ((B + g)^2 + A^2 - B^2)^{1/2})/(A + B)\} = \\ &= (1/2)\ln\{(((N + g)^2 + 4(1+Y))^{1/2} + (N + g))/(((N + g)^2 + 4(1+Y))^{1/2} - (N + g))\} - \\ &\quad -(1/2)\ln\{((N^2 + 4(1+Y))^{1/2} + N)/((N^2 + 4(1+Y))^{1/2} - N)\}. \end{aligned} \quad (82)$$

Нетрудно убедиться, что суммирование φ_0 с Φ_0 реализует первый шаг алгоритма итерационного решения уравнения Пуассона. Поскольку полагается, что функция $g(x, 0)$ в точках $x = 0$ и $x = a$ равна нулю, то и функции Π_1 и Π_2 равны нулю в нулевом приближении. В первом приближении по ε получим

$$\Phi_1 = g'_\varepsilon(x, 0)/F(x), \quad (83)$$

где обозначено

$$F = A\text{ch}\Phi_0 + B\text{sh}\Phi_0.$$

С точностью до экспоненциально малых слагаемых порядка

$$\exp\{-a/\varepsilon\}$$

для $\Pi_{1,1}$ и $\Pi_{2,1}$ справедливы задачи

$$\begin{cases} \partial^2 \Pi_{1,1}(\xi)/\partial \xi^2 = F(0)\Pi_{1,1}(\xi), \\ \Pi_{1,1}(0) = -\Phi_1(0), \Pi_{1,1}(\infty) = 0, \\ \partial^2 \Pi_{2,1}(\xi^*)/\partial \xi^{*2} = F(a)\Pi_{2,1}(\xi^*), \\ \Pi_{2,1}(0) = -\Phi_1(a), \Pi_{2,1}(\infty) = 0, \end{cases} \quad (84)$$

решение которых имеет вид

$$\begin{cases} \Pi_{1,1}(\xi) = -\Phi_1(0)\exp\{-F(0)^{1/2}\xi\}, \\ \Pi_{2,1}(\xi^*) = -\Phi_1(a)\exp\{-F(a)^{1/2}\xi^*\}. \end{cases} \quad (85)$$

Решения (85) можно «улучшить» таким образом, чтобы они обеспечивали точное выполнение граничных условий для уравнения Пуассона в точках $x = 0$ и $x = a$ [48,49,52]. Для этого граничные условия в краевых задачах (84) необходимо переопределить следующим образом

$$\begin{cases} \Pi_{1,1}(0) + \Pi_{2,1}(a/\varepsilon) = -\Phi_1(0), \\ \Pi_{1,1}(a/\varepsilon) + \Pi_{2,1}(0) = -\Phi_1(a). \end{cases} \quad (86)$$

Тогда получим

$$\begin{cases} \Pi_{1,1}(\xi) = -(\Phi_1(0) - \Phi_1(a)\exp\{-F(a)^{1/2}a/\varepsilon\})\exp\{-F(0)^{1/2}\xi\}/(1-\exp\{-(F(0)^{1/2}+F(a)^{1/2})a/\varepsilon\}), \\ \Pi_{2,1}(\xi) = -(\Phi_1(a) - \Phi_1(0)\exp\{-F(0)^{1/2}a/\varepsilon\})\exp\{-F(a)^{1/2}\xi^*\}/(1-\exp\{-(F(0)^{1/2}+F(a)^{1/2})a/\varepsilon\}). \end{cases} \quad (87)$$

Легко видеть, что формулы (85) и (87) отличаются друг от друга экспоненциально малыми слагаемыми, а для Π -функций справедливы оценки

$$|\Pi_{1,1}(\xi)| \ll C\exp\{-\varkappa\xi\}, |\Pi_{2,1}(\xi^*)| \ll C\exp\{-\varkappa\xi^*\}, \quad (88)$$

где C и \varkappa – некоторые положительные постоянные.

Построение асимптотики может быть продолжено до любого порядка по ε . При этом Π -функции будут определяться из линейных дифференциальных уравнений типа (84). Через

$$\sum_k \varphi_{ch}$$

обозначим частичную сумму k -того порядка ряда (77). Согласно [49] можно доказать, что при достаточно малых ε существует единственное решение φ_{ch} задачи (76) и функция $\sum_k \varphi_{ch}$ является равномерным на отрезке $x \in [0, a]$ асимптотическим приближением этого решения с точностью $O(\varepsilon^{k+1})$.

2.1.2. Двумерная краевая задача

2.1.2.1. Область с омическим контактом

Во избежание слишком громоздких выражений здесь будет построена асимптотика решения уравнения (75) в модельной двумерной области Ω (см. Рисунок 4), содержащей всего один омический контакт (граница Γ_1 на Рисунке 4). В случае большего числа контактов решение строится совершенно аналогично, появляются лишь дополнительные пограничные функции. Кроме участка Γ_1

граница Γ области Ω содержит еще участки Γ_0^i ($i = 0, 1, \dots, 5$) – границы с внешней средой. Граничная задача для φ_{ch} имеет вид

$$\begin{cases} \varepsilon^2(\partial^2\varphi_{ch}/\partial x^2 + \partial^2\varphi_{ch}/\partial z^2) = A(x,z)\text{sh}\varphi_{ch} + B(x,z)(\text{ch}\varphi_{ch} - 1) - g(x,z,\varepsilon), \\ \varphi_{ch}|_{\Gamma_1} = 0, \\ (\nabla\varphi_{ch}, \mathbf{v})|_{\Gamma_0^i} = \chi_i, \end{cases} \quad (89)$$

где

$$\chi_i = (\nabla\varphi_0, \mathbf{v})|_{\Gamma_0^i}.$$

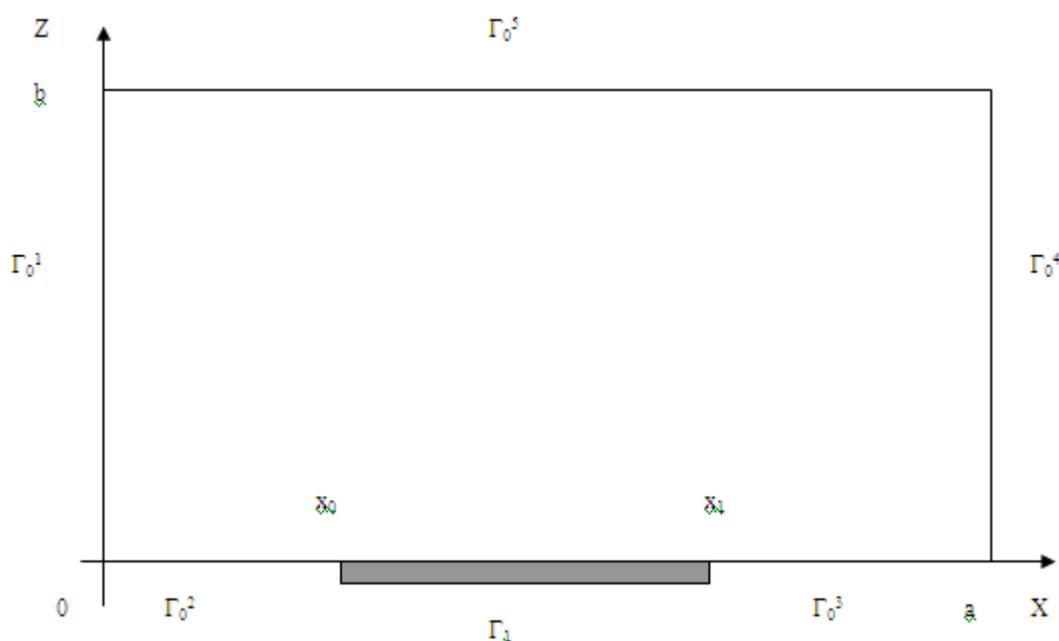


Рисунок 4. Модельная двумерная область Ω с одним омическим контактом.

Как упоминалось выше, возможен еще один тип границ – границы с диэлектриком, на которых возникают условия второго рода для φ_{ch} . Мы рассмотрим такую задачу в конце этого параграфа.

Будем считать, что функции

$$\chi_i = O(1),$$

что эквивалентно условию

$$(\nabla \ln|N|, \mathbf{v})|_{\Gamma_0^i} = O(1),$$

которое почти всегда выполняется на практике. Наложим на χ_i и g условия согласования в точках, где меняется тип граничных условий

$$\begin{cases} \partial^n \chi_2 / \partial x^n |_{x=x_0} = 0, \\ \partial^n \chi_3 / \partial x^n |_{x=x_1} = 0, \\ \partial^{n+k} g / \partial \varepsilon^k x^n |_{x=x_0, x_1} = 0, \quad (n, k = 0, 1, 2, \dots), \end{cases} \quad (90)$$

а также естественные условия согласования граничных данных в угловых точках области Ω .

Асимптотическое разложение решения задачи (89) будем искать в виде

$$\varphi_{ch}(x, z, \varepsilon) = \Phi(x, z, \varepsilon) + \Pi(x, \eta, \varepsilon) + Q_1(\xi, z, \varepsilon) + Q_2(x, \eta, \varepsilon) + Q_3(x, \eta, \varepsilon) + Q_4(\xi^*, z, \varepsilon) + \\ + Q_5(x, \eta^*, \varepsilon) + P_1(\xi, \eta, \varepsilon) + P_2(\xi^*, \eta, \varepsilon) + P_3(\xi^*, \eta^*, \varepsilon) + P_4(\xi, \eta^*, \varepsilon), \quad (91)$$

где

$$\eta = z/\varepsilon, \quad \eta^* = (b - z)/\varepsilon, \quad \xi = x/\varepsilon, \quad \xi^* = (a - x)/\varepsilon \quad (92)$$

- погранслоиные переменные; Π и Q_i – обыкновенные пограничные функции; P_j – угловые пограничные функции. Каждое из слагаемых в (91) представляет собой ряд по степеням ε . Знаки суммирования в (91) являются формальными в том смысле, что в некоторой окрестности каждого отдельно взятого участка границы за исключением угловых точек присутствует только одна из пограничных функций. Следуя изложенному в предыдущем параграфе алгоритму, представим $f(\varphi_{ch}, x, z)$ в виде, аналогичном (80)

$$f = f_0 + \Pi f + \sum_{i=1,5} Q_i f + \sum_{j=1,4} P_j f, \quad (93)$$

где

$$\begin{cases} f_0(x, z, \varepsilon) = f(\Phi(x, z, \varepsilon), x, z), \\ \Pi f(x, \eta, \varepsilon) = f(\Phi(x, \varepsilon \eta, \varepsilon) + \Pi(x, \eta, \varepsilon), x, \varepsilon \eta) - f_0(x, \varepsilon \eta, \varepsilon), \\ Q_1 f(\xi, z, \varepsilon) = f(\Phi(\varepsilon \xi, z, \varepsilon) + Q_1(\xi, z, \varepsilon), \varepsilon \xi, z) - f_0(\varepsilon \xi, z, \varepsilon), \\ Q_2 f(x, \eta, \varepsilon) = f(\Phi(x, \varepsilon \eta, \varepsilon) + Q_2(x, \eta, \varepsilon), x, \varepsilon \eta) - f_0(x, \varepsilon \eta, \varepsilon), \\ P_1 f(\xi, \eta, \varepsilon) = f(\Phi(\varepsilon \xi, \varepsilon \eta, \varepsilon) + Q_1(\xi, \varepsilon \eta, \varepsilon) + Q_2(\varepsilon \xi, \eta, \varepsilon) + P_1(\xi, \eta, \varepsilon), \varepsilon \xi, \varepsilon \eta) - \\ - Q_1 f(\xi, \varepsilon \eta, \varepsilon) - Q_2 f(\varepsilon \xi, \eta, \varepsilon) - f_0(\varepsilon \xi, \varepsilon \eta, \varepsilon). \end{cases} \quad (94)$$

Для других Q и P – функций имеют место аналогичные соотношения.

Действуя также как и в одномерном случае для Φ_0 и Φ_1 , получим выражения (82) и (83), в которых g, g_ε', A и B – функции двух переменных (x, z) . Поскольку $g(x, z, 0)$ на Γ равны нулю, то в нулевом приближении все Π, Q и P – функции также равны нулю. Для пограничных функций первого порядка по ε имеем задачи

$$\begin{cases} \partial^2 \Pi_1 / \partial \eta^2 = F(x, 0) \Pi_1, \\ \Pi_1(x, 0) |_{x \in [x_0, x_1]} = -\Phi_1(x, 0), \\ \Pi_1(x, 0) |_{x \in [0, x_0] \cap [x_1, a]} = 0, \\ \Pi_1(x, \infty) = 0, \end{cases} \quad (95)$$

$$\begin{cases} \partial^2 Q_{1,1} / \partial \xi^2 = F(0, z) Q_{1,1}, \\ Q_{1,1}(\infty, z) = 0, \\ \partial Q_{1,1}(0, z) / \partial \xi = \chi_1(z) - \partial \Phi_0(0, z) / \partial x = \gamma_1(z), \end{cases} \quad (96)$$

$$\begin{cases} \partial^2 Q_{2,1} / \partial \eta^2 = F(x, 0) Q_{2,1}, \\ Q_{2,1}(x, 0) |_{x \in [0, a]} = 0, \\ Q_{2,1}(x, \infty) = 0, \\ \partial Q_{2,1}(x, 0) / \partial \eta |_{x \in [0, x_0]} = \chi_2(x) - \partial \Phi_0(x, 0) / \partial z = \gamma_2(x). \end{cases} \quad (97)$$

Для остальных Q – функций имеем аналогичные задачи. Решения задач (95) – (97) имеют вид

$$\Pi_1(x, \eta) = \delta(x) \exp\{-F(x, 0)^{1/2} \eta\},$$

$$\delta(x) = \begin{cases} -\Phi_1(x, 0), & x \in \Gamma_1, \\ 0, & x \in [0, x_0] \cap [x_1, a], \end{cases} \quad (98)$$

$$Q_{1,1}(\xi, z) = -(\gamma_1(z) / F(0, z)^{1/2}) \exp\{-F(0, z)^{1/2} \xi\}, \quad (99)$$

$$Q_{2,1}(x, \eta) = \begin{cases} -(\gamma_2(x) / F(x, 0)^{1/2}) \exp\{-F(x, 0)^{1/2} \eta\}, & x \in [0, x_0], \\ 0, & x \in [x_0, a]. \end{cases} \quad (100)$$

Похожие выражения имеют место и для остальных Q – функций. Для всех Π и Q – функций справедливы оценки, аналогичные (88). В силу условий (90) все погранфункции являются сколь угодно гладкими в Ω и на границе Γ .

Нетрудно убедиться, что в первом приближении по ε все P – функции равны нулю. Ненулевые выражения для них имеют место, начиная со второго порядка

по ε . Выпишем, например, для угловой пограничной функции $P_{1,2}(\xi, \eta)$, существенной вблизи точки $(0,0)$ области Ω . $P_{2,1}$ компенсирует вблизи точки $(0,0)$ невязки, вносимые функциями $Q_{1,1}$ и $Q_{2,1}$ в граничные условия.

$$\left\{ \begin{array}{l} \partial^2 P_{1,2} / \partial \xi^2 + \partial^2 P_{1,2} / \partial \eta^2 = F(0,0) P_{1,2}, \\ P_{1,2} \rightarrow 0 \text{ при } (\xi + \eta) \rightarrow \infty, \\ \partial P_{1,2}(\xi, 0) / \partial \eta = -\partial Q_{1,1}(\xi, 0) / \partial z = \chi'_1(\xi) \exp\{-F(0,0)\xi\}, \\ \partial P_{1,2}(0, \eta) / \partial \xi = -\partial Q_{2,1}(0, \eta) / \partial x = \chi'_2(\eta) \exp\{-F(0,0)\eta\}. \end{array} \right. \quad (101)$$

Явные выражения для $\chi'_{1,2}$ легко получить из формул (99), (100).

Выше предполагалось выполнение некоторых условий согласования граничных данных в угловых точках области Ω . Запишем условия в точке $(0,0)$, при которых определяется функция $P_{1,2}$

$$\chi'_1(0) = \chi'_2(0), F_x(0,0) = F_z(0,0).$$

При определении других P -функций во втором порядке по ε необходимо наложить аналогичные условия согласования в точках $(0,b)$, (a,b) , $(a,0)$. В следующих порядках накладываются некоторые другие условия согласования в угловых точках области Ω . Все они могут быть выписаны в явном виде.

Решение задачи для $P_{1,2}$ запишем с помощью функции Грина [53]

$$P_{1,2}(\xi, \eta) = \int_0^{\infty} \int_0^{\infty} G(\xi, \eta; \xi_0, \eta_0) \sigma_1(\xi_0, \eta_0) \exp\{-F(0,0)^{1/2}(\xi_0 + \eta_0)\} d\xi_0 d\eta_0 + \\ + \sigma_2(\xi, \eta) \exp\{-F(0,0)^{1/2}(\xi + \eta)\},$$

где σ_1 и σ_2 – известные линейные функции своих аргументов,

$$G(\xi, \eta; \xi_0, \eta_0) = (1/2\pi)[K_0(R_1) + K_0(R_2) + K_0(R_3) + K_0(R_4)],$$

$K_0(z)$ – цилиндрическая функция мнимого аргумента,

$$R_1 = [(\xi - \xi_0)^2 + (\eta - \eta_0)^2]^{1/2}, R_2 = [(\xi + \xi_0)^2 + (\eta + \eta_0)^2]^{1/2},$$

$$R_3 = [(\xi - \xi_0)^2 + (\eta + \eta_0)^2]^{1/2}, R_4 = [(\xi + \xi_0)^2 + (\eta - \eta_0)^2]^{1/2}.$$

Аналогично [53] можно показать, что $P_{1,2}$ имеет оценку

$$|P_{1,2}(\xi, \eta)| \leq C \exp\{-\alpha(\xi + \eta)\}, \xi \geq 0, \eta \geq 0.$$

Такие же оценки справедливы и для других P – функций.

Умножая $\Pi_1, Q_{1,1}, Q_{2,1}$ стандартным образом на сглаживающую функцию, получим функции (обозначаемые теми же символами), которые не вносят никакой невязки в граничное условие при $Y = b$. Аналогичную процедуру проделаем и с другими функциями Π, Q, P .

Обозначив как и ранее $\sum_k \varphi_{ch}$ частичную сумму k -того порядка ряда (91), можно доказать теорему [50,53] о том, что при выполнении указанных выше условий и достаточно малых ε существует единственное решение φ_{ch} задачи (89) и функция $\sum_k \varphi_{ch}$ является равномерным в Ω асимптотическим приближением для этого решения с точностью $O(\varepsilon^{r+1})$.

2.1.2.2. Область с контактом к диэлектрику и омическим контактом

В случае, когда в задаче присутствует граница с подзатворным диэлектриком (граница типа γ_1^j на Рисунке 1) постановку задачи удобно немного видоизменить. Обозначим на Рисунке 5 эту границу через Γ_2 . Согласно (48) представим потенциал φ_{ch} в виде

$$\varphi_{ch} = \varphi_{ch}^0 + \varphi_1,$$

где под φ_{ch}^0 понимается решение только что рассмотренной задачи (89). Для φ_1 получим краевую задачу

$$\begin{cases} \varepsilon^2(\partial^2 \varphi_1 / \partial x^2 + \partial^2 \varphi_1 / \partial z^2) = A_1(x,z) \operatorname{sh} \varphi_1 + B_1(x,z)(\operatorname{ch} \varphi_1 - 1), \\ (\nabla \varphi_1, \nu)|_{\Gamma_2} = \lambda, \\ (\nabla \varphi_1, \nu)|_{\Gamma \setminus \Gamma_2} = 0. \end{cases} \quad (102)$$

В такой постановке φ_1 может внести невязку только в граничные условия для φ_{ch} на границе Γ_1 . Если Γ_1 расположена на противоположной Γ_2 стороне прямоугольника Ω , то эта невязка экспоненциально мала. Если же Γ_1 и Γ_2 расположены на одной стороне прямоугольника Ω , то невязка не появится, если считать выполненными условия согласования для λ в точках $x = x_2$ и $x = x_3$

$$\partial^n \mathcal{N} / \partial x^n |_{x=x_2, x_3} = 0, \quad n = 0, 1, 2, \dots \quad (103)$$

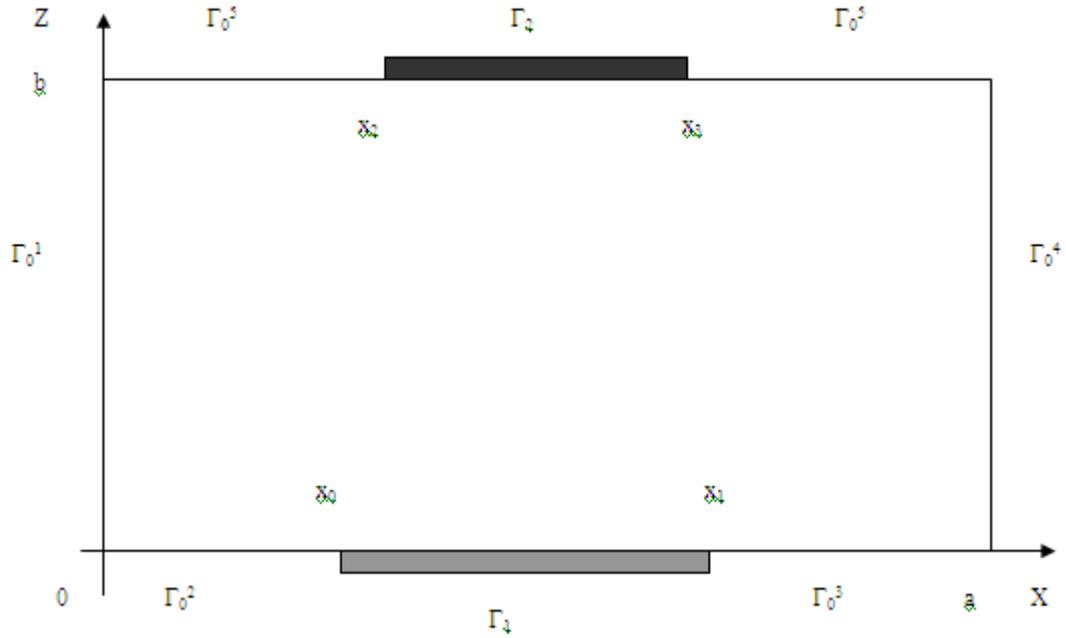


Рисунок 5. Модельная двумерная область Ω с одним омическим контактом и контактом к подзатворному диэлектрику.

В отличие от функций χ_i функция λ может быть и большой, например, порядка $O(\varepsilon^{-1})$. Такая ситуация возможна под затвором МДП - прибора при открытом канале проводимости [27]. Процедуру построения асимптотики в этом случае следует видоизменить, так как пограничные слои теперь будут возникать и в нулевом порядке по ε . Будем искать φ_1 в виде

$$\varphi_1(x, \eta^*, \varepsilon) = \sum_{l=0}^{\infty} \varepsilon^l \varphi_{1,l}(x, \eta^*). \quad (103)$$

Определим функцию h на границе Γ

$$h = \begin{cases} \varepsilon \lambda, & \text{на } \Gamma_2, \\ 0, & \text{на } \Gamma \setminus \Gamma_2. \end{cases}$$

Положим, что $h = O(1)$. Для $\varphi_{1,0}$ имеем задачу

$$\begin{cases} \partial^2 \varphi_{1,0} / \partial \eta^{*2} = A_1(x, b) \text{sh} \varphi_{1,0} + B_1(x, b) (c \text{h} \varphi_{1,0} - 1) = \gamma(\varphi_{1,0}, x), \\ \partial \varphi_{1,0} / \partial \eta^* |_{\Gamma_2 \cap \Gamma_0^6 \cap \Gamma_0^5} = -h, \\ (\nabla \varphi_{1,0}, \mathbf{v}) |_{\Gamma \setminus \Gamma_2 \setminus \Gamma_0^6 \setminus \Gamma_0^5} = 0, \\ \varphi_{1,0}(x, \infty) = 0. \end{cases} \quad (104)$$

На отрезке $x \in [x_2, x_3]$ обозначим

$$v = \varphi_{1,0}(x, \eta^*), \quad \omega = \partial \varphi_{1,0} / \partial \eta^*, \quad v_0 = \varphi_{1,0}(x, 0).$$

Тогда соотношения (104) на этом отрезке можно переписать так

$$\begin{cases} \partial v / \partial \eta^* = \omega, \\ \partial \omega / \partial \eta^* = \gamma(v, x), \\ \omega(x, 0) = -h, \\ v(x, \infty) = 0. \end{cases} \quad (105)$$

Функции v и ω зависят от x как от параметра.

Из явного вида γ и условия

$$A_1 > B_1$$

при всех $(x, z) \in \Omega$ следует, что

$$\gamma(0, x) = 0, \quad \gamma'(0, x) > 0$$

и, более того

$$\gamma(s, x) > 0 \text{ для всех } s \in (-\infty, \infty).$$

Точка покоя

$$v = 0, \quad \omega = 0$$

системы (105) является седлом. Корни соответствующего характеристического уравнения равны

$$\pm(\gamma'(0, x))^{1/2}.$$

Аналогично [53] получим решение задачи (105) в неявном виде

$$\begin{matrix} v & s \\ \int_{v_0}^v (\int_{\tau}^v \gamma(\tau, x) d\tau)^{1/2} ds = & \begin{cases} \eta^*, & v_0 < 0, \\ -\eta^*, & v_0 > 0, \end{cases} \end{matrix} \quad (106)$$

где v_0 – решение уравнения

$$h^2(x) = 2 \int_0^{v_0} \gamma(s, x) ds \quad (107)$$

и для каждого x выбирается

$$v_0 > 0 \text{ при } h > 0, \text{ и } v_0 < 0 \text{ при } h < 0.$$

Можно показать, что $v(x, \eta^*)$, а значит и $\varphi_{1,0}(x, \eta^*)$, имеет оценку

$$|v(x, \eta^*)| \leq C \exp\{-k\eta^*\}, \quad \eta^* \geq 0.$$

Таким образом, функция $\varphi_{1,0}(x, \eta^*)$ построена. В силу условий (103) она является сколь угодно гладкой в Ω . Не составляет особого труда построить и последующие приближения $\varphi_{1,l}$ как это сделано в работе [51]. Не будем воспроизводить здесь эти результаты.

В случае, когда $\lambda = O(1)$ и, следовательно, $h = O(\varepsilon)$ из (106) и (107) можно получить явные выражения для $\varphi_{1,0}(x, \eta^*)$

$$\varphi_{1,0}(x, \eta^*) \approx (h/(A_1(x, b)))^{1/2} \exp\{-(A_1(x, b))^{1/2} \eta^*\}. \quad (108)$$

На основе этого выражения можно оценить напряженности электрических полей на границе Γ_2 , при которых этот результат справедлив. Из условия

$$\varphi_{1,0} \ll 1$$

следует, что

$$|(\nabla \varphi_{1,0}, \nu)|_{\Gamma_2} \ll E_{cr} = (k_B T/e) L_D^{-1} (|N|/n_i)^{1/2} |\Gamma_2|. \quad (109)$$

Для кремния при комнатной температуре ($T = 300^0$ К) из (109) следует, что

$$E_{cr} = 10^3 \text{ В/см} - 10^6 \text{ В/см при } |N| = 10^{14} \text{ см}^{-3} - 10^{20} \text{ см}^{-3}.$$

То есть поля оказываются достаточно большими и используемый для их обозначения термин «слабые» является условным.

Соотношения (106), (107) и (108) указывают на то, что в полупроводнике под затвором МДП – структуры образуется экспоненциальный пограничный слой электростатического потенциала с характерной областью локализации, равной длине Дебая в легированном материале

$$L_{ch} = L_D (n_i/|N|)^{1/2}.$$

В кремнии при $T = 300^0$ К

$$L_{ch} = 10^{-5} \text{ см} - 10^{-8} \text{ см при } |N| = 10^{14} \text{ см}^{-3} - 10^{20} \text{ см}^{-3}.$$

В «слабых» полях, когда справедливо неравенство (109), φ_1 растет линейно с ростом поля (с ростом λ или h). Однако в «сильных» полях, когда справедливо обратное (109) неравенство, ситуация меняется. Из (107) и (104) получим

$$h^2(x) = 2A_1(x, b)(chv_0 - 1) + 2B_1(x, b)shv_0 - 2B_1(x, b)v_0. \quad (110)$$

В случае, когда $v_0 \gg 1$ из (110) найдем

$$v_0 \approx \ln(h^2(x)/(A_1(x, b) + B_1(x, b))) = \ln\{h^2(x)/((Q^2 + 4(1 + Y))^{1/2} + Q)\} |\Gamma_2, \quad (111)$$

а если $v_0 \ll -1$, то имеем

$$v_0 \approx -\ln(h^2(x)/(A_1(x,b) - B_1(x,b))) = -\ln\{h^2(x)/((Q^2 + 4(1 + Y))^{1/2} - Q)\}|\Gamma_2. \quad (112)$$

Таким образом, рост φ_1 в «сильных» полях замедляется и становится логарифмическим.

Потенциалом φ_1 определяется количество носителей заряда в полупроводнике под затвором (вблизи границы Γ_2). Согласно (43) и (62)

$$\begin{cases} n = (1/2)[(Q^2 + 4(1 + Y))^{1/2} + Q]\exp\{\varphi_1\}, \\ p = (1/2)[(Q^2 + 4(1 + Y))^{1/2} - Q]\exp\{-\varphi_1\}, \end{cases} \quad (113)$$

где

$$Q = N + \alpha^2 \Delta\varphi_0, \quad \varphi_0 = \varphi_{en} + \varphi_{ch}^0.$$

Если $\varphi_1 > 0$, то происходит обогащение подзатворной области полупроводника электронами и обеднение ее дырками по отношению к случаю, когда $\varphi_1 = 0$. При $\varphi_1 < 0$ наблюдается обратная картина.

Если знак φ_1 таков, что область полупроводника вблизи Γ_2 обедняется основными носителями заряда, то с ростом $|\varphi_1|$ происходит инверсия типа проводимости [27], то есть, начиная с некоторого порогового значения

$$\varphi_1 = \varphi_1^{inv},$$

количество неосновных носителей заряда начинает превышать количество основных. Значение φ_1^{inv} нетрудно найти из соотношения

$$n(\varphi_1^{inv}) = p(\varphi_1^{inv}),$$

откуда

$$\varphi_1^{inv} = (1/2)\ln[((Q^2 + 4(1 + Y))^{1/2} - Q)/((Q^2 + 4(1 + Y))^{1/2} + Q)]. \quad (114)$$

Явление инверсии типа проводимости на границе полупроводника и подзатворного диэлектрика лежит в основе работы МДП – структур [27]. Многие важные характеристики приборов на основе этих структур определяются распределением потенциала φ в инвертированной области, которую часто называют каналом проводимости. Поскольку глубина этой области весьма мала (порядка L_{chr}), то для численных расчетов приходится выбирать сетку с очень мелким шагом. В этой ситуации использование приведенных здесь приближенных решений оказывается весьма эффективным. Часто они

оказываются ближе к точным решениям, чем решения соответствующих разностных уравнений.

Заметим еще, что как следует из (111), (112) и (113), в случае, когда знак φ_1 таков, что происходит обогащение подзатворной области основными носителями заряда, рост их числа на границе Γ_2 становится не экспоненциальным, а квадратичным

$$n|_{\Gamma_2} \approx (1/2)h^2(x), \text{ при } \varphi_1|_{\Gamma_2} \gg 1,$$

$$p|_{\Gamma_2} \approx (1/2)h^2(x), \text{ при } \varphi_1|_{\Gamma_2} \ll -1.$$

Если восстановить размерности в выражениях для концентраций, то получим

$$n, p = (1/2)E^2(\epsilon_s/k_B T), \quad (115)$$

где E – напряженность электрического поля на границе Γ_2 .

2.2. Стационарная краевая задача для плотности тока

В одномерном стационарном случае определение плотности суммарного тока j не представляет труда. Согласно уравнениям (54)

$$j = \text{const}$$

в этом случае. Как уже обсуждалось, при одномерном моделировании структур с тремя и более контактами j может быть кусочно-постоянной.

В двумерном и трехмерном случаях задачу (54), (55) приходится решать численно. Для двумерной модели это не вызывает особых затруднений. Введение функции тока χ согласно (51) приводит к эллиптической краевой задаче (59), (60), которая хорошо решается разностными методами.

В некоторых случаях задачу (59), (60) можно упростить и решить аналитически. Это можно сделать в областях с постоянным уровнем легирования, когда

$$Q_+ = \mu_n n + \mu_p p = \text{const},$$

а правая часть в (59) равна нулю. С большой точностью такие предположения справедливы для резисторов. Однако и для более сложных структур часто можно пренебречь правой частью в (59) и считать, что [38]

$$|\nabla Q_+/Q_+| \ll 1.$$

Если эти допущения справедливы, то вместо (59) получим для χ уравнение Лапласа

$$\partial^2 \chi / \partial x^2 + \partial^2 \chi / \partial z^2 = 0. \quad (116)$$

Граничные условия (60) в стационарном случае будут выглядеть так

$$\left\{ \begin{array}{l} (\nabla \chi, \mathbf{v})|_{\Gamma_{1-0}^i} = 0, \\ \chi(s) = \chi(\alpha_i) - J_i, s \in \Gamma_0^{i+1}, \\ \chi(\alpha_1) = 0, \alpha_{k+1} = \alpha_1, \\ \sum_{i=1}^k J_i = 0. \end{array} \right. \quad (117)$$

Напомним, что J_i это ток, протекающий через i -й контакт с учетом направления нормали \mathbf{v} (см. выражение (55)), а k – полное число контактов.

Поскольку любая односвязная область конформно инвариантна полуплоскости [55], отобразим модельную область Ω (см. Рисунок 2) в верхнюю полуплоскость $v \geq 0$ таким образом, чтобы граница

$$\Gamma = \bigcup_i \Gamma_{1-0}^i \bigcup_j \Gamma_0^j$$

перешла в ось Ou , а точки начала и конца контактов (смены типа граничных условий) α_i и β_i в точки a_i и b_i на оси Ou соответственно. Всегда можно выбрать отображение так, чтобы точки (a_i, b_i) располагались справа от начала координат.

На рис. 8 показано, как такое отображение можно осуществить в случае прямоугольной области полупроводника ω . Введем комплексную переменную

$$\omega = x + iz$$

и определим другую комплексную переменную

$$t = u + iv$$

соотношением

$$\omega = A \int_0^t d\tau / (\tau(\tau - 1)(k^2\tau - 1))^{1/2} \quad (118)$$

Или, что тоже самое

$$t = \text{sn}^2(-\omega/2A, k), k < 1, \quad (119)$$

где $\text{sn}(\varphi, k)$ – эллиптический синус [56]. Постоянные A и k находятся из соотношений

$$a = -2AF(\pi/2, k), \quad b = -2AF(\pi/2, k_1), \quad k_1 = (1 - k^2)^{1/2},$$

а $F(\pi/2, k)$ – полный эллиптический интеграл I-го рода.

Для конкретных вычислений удобно использовать соотношения

$$q = \exp\{-\pi F(\pi/2, k_1)/F(\pi/2, k)\} = \exp\{-\pi b/a\},$$

$$q_1 = \exp\{-\pi F(\pi/2, k)/F(\pi/2, k_1)\} = \exp\{-\pi a/b\},$$

$$F(\pi/2, k) = (\pi/2)(1 + 2\sum_{n=1}^{\infty} q^{s^2}), \quad s = n^2,$$

$$F(\pi/2, k_1) = (\pi/2)(1 + 2\sum_{n=1}^{\infty} q_1^{s^2}), \quad s = n^2,$$

$$k = (2\pi/F(\pi/2, k))(\sum_{n=1}^{\infty} q^{s^2}), \quad s = (n - 1/2)^2,$$

$$k_1 = (2\pi/F(\pi/2, k_1))(\sum_{n=1}^{\infty} q_1^{s^2}), \quad s = (n - 1/2)^2,$$

$$u(x, z) = \Delta^{-2} \{ \text{sn}^2(x', k) \text{dn}(z', k_1) - \text{sn}^2(z', k_1) \text{cn}^2(z', k_1) \text{cn}^2(x', k) \text{dn}^2(x', k) \},$$

$$v(x, z) = 2\Delta^{-2} \text{sn}(x', k) \text{sn}(z', k_1) \text{cn}(x', k) \text{cn}(z', k_1) \text{dn}(x', k) \text{dn}(z', k_1),$$

$$\Delta = 1 - \text{sn}^2(z', k_1) \text{dn}^2(x', k),$$

$$x' = -x/(2A), \quad z' = -z/(2A).$$

Здесь $\text{sn}(\varphi, k)$, $\text{cn}(\varphi, k)$, $\text{dn}(\varphi, k)$ – эллиптические функции.

В силу конформной инвариантности уравнения Лапласа из (116) имеем

$$\partial^2 \chi / \partial u^2 + \partial^2 \chi / \partial v^2 = 0. \quad (120)$$

Введем функцию $g(u)$ переменной u , такую что

$$g(u) = \begin{cases} 0, & u \in (a_i, b_i), \\ g(u), & u \in (b_i, a_{i+1}). \end{cases} \quad (121)$$

Потребуем, чтобы были выполнены граничные условия

$$\partial \chi / \partial v|_{v \rightarrow 0} = g(u). \quad (122)$$

Функция $g(u)$ пока что не определена полностью и мы будем искать $g(u)$ такую, чтобы краевые задачи (120), (122) и (116), (117) были эквивалентны. Задача Неймана (120), (122) разрешима при условии [44]

$$\int_{-\infty}^{\infty} g(u) du = 0. \quad (123)$$

Ее решение имеет вид

$$\chi(u,v) = \pi^{-1} \int_{-\infty}^{\infty} g(s) \ln((s-u)^2 + v^2)^{1/2} ds + \text{const.} \quad (124)$$

Для потоков j_u и j_v получим выражения

$$\begin{cases} j_u = \partial\chi/\partial u = -\pi^{-1} \int_{-\infty}^{\infty} ds g(s) (s-u) / ((s-u)^2 + v^2) = -(2\pi)^{-1} \int_{-\infty}^{\infty} ds g(s) \{ (s-t)^{-1} + (s-t^*)^{-1} \}, \\ j_v = \partial\chi/\partial v = \pi^{-1} \int_{-\infty}^{\infty} ds g(s) v / ((s-u)^2 + v^2) = -i(2\pi)^{-1} \int_{-\infty}^{\infty} ds g(s) \{ (s-t)^{-1} - (s-t^*)^{-1} \}, \\ t = u + iv, t^* = u - iv. \end{cases} \quad (125)$$

Из известного соотношения

$$(\varepsilon/\pi)(x^2 + \varepsilon^2)^{-1} \Big|_{\varepsilon \rightarrow 0} = \delta(x)$$

следует, что

$$j_v|_{v \rightarrow 0} = g(u),$$

то есть граничное условие (122) выполнено.

Построим аналитическую в верхней полуплоскости функцию $q(t)$

$$q(t) = j_u + ij_v = \pi^{-1} \int_{-\infty}^{\infty} ds g(s) / (s - t - i\varepsilon). \quad (126)$$

Ясно, что

$$\text{Im}(q(t))|_{v \rightarrow 0} = g(u).$$

Интеграл Коши (126) осуществляет аналитическое продолжение функции $g(u)$ в верхнюю полуплоскость.

Чтобы найти функцию $g(u)$ воспользуемся важным следствием условий (117)

$$(\tau, \nabla\chi)|_{\Gamma_0^j} = 0,$$

которое в полуплоскости t эквивалентно соотношению

$$\partial\chi/\partial u|_{v=0, u \in (b_i, a_{i+1})} = 0. \quad (127)$$

Сформулируем задачу: найти функцию $f(t)$, аналитическую в верхней полуплоскости, такую, что на действительной оси

$$\text{Re}(f(t)) = -j_u(u), \text{Im}(f(t)) = j_v(u)$$

и известно, что $j_u(u)$ определена всюду при $u \in (b_i, a_{i+1})$, а $j_v(u)$ определена всюду при $u \in (a_i, b_i)$. Кроме того, потребуем, чтобы предел

$$f(t)|_{|t| \rightarrow \infty} = f(\infty)$$

существовал и был действительным.

Это задача Келдыша-Седова [44], решение которой можно представить в виде

$$f(t) = -i(\pi\sigma(t))^{-1} \left\{ -\sum_{i=1}^k \int_{b_i}^{a_{i+1}} ds j_u(s) \sigma(s)/(s-t) + \sum_{i=1}^k \int_{a_i}^{b_i} ds j_v(s) \sigma(s)/(s-t) \right\} + h(t) + f(\infty)/\sigma(t), \quad (128)$$

где

$$\begin{aligned} \sigma(t) &= \prod_{i=1}^k ((t - a_{i+1})/(t - b_i))^{1/2}, \\ h(t) &= (\gamma_0 + \gamma_1 t + \dots + \gamma_{k-2} t^{k-2}) / (\prod_{i=1}^k (t - b_i)(t - a_{i+1}))^{1/2}, \\ a_{k+1} &= a_1. \end{aligned}$$

Согласно (121), (122) и (127)

$$j_v(u) = 0 \text{ при } u \in (a_i, b_i),$$

а

$$j_u(u) = 0 \text{ при } u \in (b_i, a_{i+1}).$$

Наложим на $f(t)$ дополнительное требование убывания на бесконечности быстрее, чем $|t|^{-1}$. Тогда из (128) получим

$$f(t) = h(t) = h(t) = (\gamma_0 + \gamma_1 t + \dots + \gamma_{k-2} t^{k-2}) / (\prod_{i=1}^k (t - b_i)(t - a_{i+1}))^{1/2}. \quad (129)$$

Используя формулу Коши, найдем

$$f(t) = (2\pi i)^{-1} \int_{-\infty}^{\infty} ds 2i \operatorname{Im}(h(s)) / (s - t) = \pi^{-1} \int_{-\infty}^{\infty} ds \operatorname{Im}(h(s)) / (s - t - i\epsilon). \quad (130)$$

Сравнивая выражения (129) и (126) приходим к выводу

$$g(u) = \operatorname{Im}(h(u)) = (\gamma_0 + \gamma_1 u + \dots + \gamma_{k-2} u^{k-2}) / (\prod_{i=1}^k (u - b_i)(u - a_{i+1}))^{1/2}. \quad (131)$$

Можно показать, что условие (123) выполнено для $g(u)$ при любых γ_i ($i = 0, 1, 2, \dots, k-2$). Величины γ_i , очевидно, определяются токами J_i .

Используя выражения (129), (130) и (131) из (125), получим

$$\begin{cases} j_v(u, v) = \operatorname{Im}(h(t)), \\ j_u(u, v) = -\operatorname{Re}(h(t)), \\ t = u + iv. \end{cases} \quad (132)$$

Теперь не составляет труда найти плотности токов j_x и j_z

$$\begin{cases} j_x = -(\partial v / \partial z) \operatorname{Im}(h(t)) + (\partial u / \partial z) \operatorname{Re}(h(t)), \\ j_z = (\partial v / \partial x) \operatorname{Im}(h(t)) - (\partial u / \partial x) \operatorname{Re}(h(t)), \\ u = u(x, z), v = v(x, z), t = u + iv. \end{cases} \quad (133)$$

Если известно конформное отображение области полупроводника в полуплоскость, то выражения (133) и (131) полностью определяют j_x и j_z . Функцию тока χ легко найти из формулы (123).

Полученные здесь результаты оказываются полезными и при численном моделировании переноса заряда в полупроводниковых приборах. В случае прямоугольных областей полученные $j_{x,z}$ можно использовать для расчета начального приближения функции Y .

2.3. Приближенные решения краевых задач для функции Y в стационарном случае

В стационарном случае из выражений (44), (47) для

$$Y = (np - n_i^2) / n_i^2$$

получим следующую краевую задачу

$$\begin{cases} (\nabla, g \nabla Y) = qY + f, \\ Y|_{\Gamma_{1-0}^i} = 0, (\nabla Y, \nu)|_{\Gamma \setminus U_i \Gamma_{1-0}^i} = 0, \end{cases} \quad (134)$$

где

$$\begin{cases} g = \beta^2 \mu_n \mu_p / Q_+, \\ f = (\beta^2 / 2) (\mathbf{j}, \nabla(Q / Q_+)) - G. \end{cases} \quad (135)$$

Впервые уравнение для Y было получено, по-видимому, в работе [57]. Там же отмечалась важность этого уравнения для понимания основ функционирования полупроводниковых приборов. Последовательный анализ решений уравнений для функции Y применительно к полупроводниковым структурам был предпринят в работах [38,41,47,54,58].

Коэффициенты и правая часть в уравнении (134) зависят от Y . В общем случае - это нелинейное уравнение, решение которого следует искать численно, используя итерационные методы. Однако при некоторых предположениях,

которые на практике часто выполняются, можно построить приближенное решение задачи (134). Эти решения оказываются полезными и тогда, когда сделанные предположения не выполнены. Они могут играть роль начального приближения при итерационном решении задачи.

Будем считать, что неравенство (66)

$$-1 < Y \ll N_{ch}^2$$

справедливо почти всюду в полупроводнике за исключением, быть может, областей, размеры которых малы по сравнению с характерными размерами структуры. В этом случае при построении главного члена асимптотики решения задачи (134) можно приближенно считать, что коэффициенты g , q и правая часть f не зависят от Y , то есть использовать их значения при $Y = 0$. При известных \mathbf{j} , φ и G получаем линейную задачу, которую можно исследовать стандартными методами.

Отметим, что согласно (134) и (135) источником возникновения неравновесности (неравенства Y нулю) является функция f , аналог функции S , возникающей в нестационарной задаче (65). Отличие f от нуля может быть обусловлено:

- протеканием тока перпендикулярно неоднородностям типа p - n , n^+ - n , p^+ - p – переходам;
- умножением (генерацией) носителей заряда за счет лавинного эффекта, облучения полупроводника светом, нейтронами и т.п.

2.3.1. Одномерная краевая задача

В одномерном случае из (134) получим

$$\begin{cases} d(gdY/dx)/dx - qY = f, \\ Y(0) = Y(a) = 0, x \in [0, a]. \end{cases} \quad (136)$$

Функции g и q положительно определены при всех значениях своих аргументов.

При выполнении условия (66) справедливо неравенство

$$\delta(qY + f)/\delta Y = q_0 > 0,$$

где

$$q_0 = q|_{Y=0}.$$

Поэтому решение линейризованной задачи (136), когда

$$Y \rightarrow Y_0, g \rightarrow g_0 = g|_{Y=0}, q \rightarrow q_0 = q|_{Y=0}, f \rightarrow f_0 = f|_{Y=0}, \\ \lambda = (q/g)^{1/2} \rightarrow \lambda_0 = \lambda|_{Y=0},$$

единственно. Если $f = 0$ то и $Y = 0$, то есть электроны и дырки находятся в состоянии локального химического равновесия. Источниками неравновесности (отличия f от нуля) могут быть либо внешняя или лавинная генерация G , либо ток, протекающий перпендикулярно p - n -переходу.

Построение приближенного решения задачи (136) даже в линейризованном варианте не всегда возможно, так как коэффициенты зависят от координат.

Однако при выполнении условия гладкости

$$(3\|\lambda_0'/\lambda_0\|^2 + \|\lambda_0''/\lambda_0\|)/8\lambda_0^{\min}(1 - \exp\{-2\lambda_0^{\text{int}}\}) \ll 1, \quad (137)$$

приближенное решение можно построить методом ВКБ [46,59,60]. В выражении (137)

$$\|(\dots)\| = \left(\int_0^a (\dots)^2(x)dx\right)^{1/2}, \lambda_0 = (q_0/g_0)^{1/2}, \lambda_0^{\min} = \min\{\lambda_0(x)\}, \lambda_0^{\text{int}} = \int_0^a \lambda_0(\xi)d\xi.$$

Если условие (137) выполнено, то ВКБ - приближение для решения линейризованной задачи (136) Y_0 имеет вид

$$Y_0(x) \approx \int_0^a G_0(x,x')f_0(x')dx', \quad (138)$$

а для функция Грина $G(x,x')$ справедливо выражение

$$G_0(x,x') = Z(x,x') \begin{cases} \exp\left\{-\int_x^{x'} \lambda_0(\xi)d\xi\right\} (1 - \exp\left\{-2\int_0^x \lambda_0(\xi)d\xi\right\}) (1 - \exp\left\{-2\int_{x'}^a \lambda_0(\xi)d\xi\right\}), & 0 \leq x \leq x' \leq a, \\ \exp\left\{-\int_{x'}^x \lambda_0(\xi)d\xi\right\} (1 - \exp\left\{-2\int_0^a \lambda_0(\xi)d\xi\right\}) (1 - \exp\left\{-2\int_x^{x'} \lambda_0(\xi)d\xi\right\}), & 0 \leq x' \leq x \leq a, \end{cases}$$

$$Z(x,x') = -(2(\lambda_0(x)g_0(x)\lambda_0(x')g_0(x')))^{1/2} (1 - \exp\{-2\lambda_0^{\text{int}}\})^{-1}. \quad (139)$$

Как уже отмечалось, источник неравновесности электронов и дырок функция f в большинстве случаев отлична от нуля в малых областях вблизи p - n -

переходов. Распространение Y в области, где $f = 0$, описывается решением (138), (139), согласно которому неравновесные носители заряда локализованы вблизи источника неравновесности на расстоянии порядка λ^1 . Эта величина имеет смысл длины неравновесности L_{neq} или, что то же самое, длины релаксации электронов и дырок к состоянию локального химического равновесия. Если восстановить размерности, воспользовавшись уравнениями (44.P), то получим

$$L_{\text{neq}}^2 = (k_B T/e)(\mu_n \mu_p n_i^2)/(q(\mu_n n + \mu_p p)). \quad (140)$$

Когда преобладающим процессом является рекомбинация Шокли – Рида - Холла, из (140) в n-области найдем

$$L_{\text{neq}} \approx L_p = (\tau_p \mu_p k_B T/e)^{1/2} = (\tau_p D_p)^{1/2}, \quad (140.1)$$

а в p-области имеем

$$L_{\text{neq}} \approx L_n = (\tau_n \mu_n k_B T/e)^{1/2} = (\tau_n D_n)^{1/2}. \quad (140.2)$$

Величины L_n и L_p являются диффузионными длинами для неосновных носителей заряда в p и n областях полупроводниковой структуры, соответственно. Они определяют размер области полупроводника, в которой происходит релаксация к локальному химическому равновесию.

Времена $\tau_{n,p}$ и подвижности $\mu_{n,p}$ [26,27,29] зависят от температуры, концентрации примеси и напряженности электрического поля. Соответственно, от этих величин зависят и длины релаксации L_{neq} , $L_{n,p}$. Для кремния при комнатной температуре и напряженностях электрического поля меньших, чем 10^5 В/см, в зависимости от концентрации примеси значения времен $\tau_{n,p}$ находятся в пределах от 10^{-8} с в сильнолегированных областях до 10^{-5} с в нелегированном материале. Максимальные значения подвижностей в кремнии при комнатной температуре

$$\mu_n = 1500 \text{ см}^2/\text{Вс}, \mu_p = 600 \text{ см}^2/\text{Вс}.$$

Используя эти данные, получим оценку

$$L_{\text{neq}} \sim 10^{-4} - 10^{-2} \text{ см}.$$

Это означает, что длина релаксации электронов и дырок к состоянию локального химического равновесия в кремнии при комнатной температуре сопоставима с размерами традиционных кремниевых элементов интегральных схем.

В соответствии с формулами (138) и (139) при размерах элементов меньших или сопоставимых с длиной релаксации $L_{\text{неq}}$ функция Y будет иметь пограничные слои вблизи контактов. Их наличие существенно влияет на величину Y в приборе и, как следствие, на его выходные характеристики.

Заметим, что получение наглядного и легко интерпретируемого результата (138), (139) связано с выделением в уравнении (134) члена

$$qY,$$

описывающего релаксацию неравновесности электронов и дырок в полупроводниковой структуре.

Условие (137) выполняется в случае, когда $\lambda_0 = O(1)$, а производные λ_0'' и λ_0' малы. Такая ситуация характерна для большинства полупроводниковых структур. Но в ряде случаев, в частности для полупроводниковых приборов больших размеров возможно, что

$$\lambda_0 \gg 1.$$

Физически это неравенство означает, что

$$L \gg L_{\text{неq}},$$

где L – характерный размер структуры. В этом случае задача (136) становится сингулярно возмущенной и для выполнения неравенства (137) достаточно требований

$$\|\lambda_0'\| = O(1) \text{ и } \|\lambda_0''\| = O(1).$$

Введем погранслойные переменные

$$\eta(x) = \int_0^x \lambda_0(\xi) d\xi, \quad \eta^*(x) = \int_x^a \lambda_0(\xi) d\xi.$$

Произведя в выражении для Y интегрирование по частям, отбрасывая члены меньшие и порядка λ_0^{-1} , получим

$$Y_0 \approx -(f_0(x)/q_0(x)) + (f_0(0)/q_0(0))\exp\{-\eta(x)\} + (f_0(a)/q_0(a))\exp\{-\eta^*(x)\}. \quad (141)$$

Выражение (141) представляет собой стандартный результат теории сингулярных возмущений в применении к задаче (136). При этом следует иметь в виду, что функция Y , записанная как (138), (139), удовлетворяет граничным условиям точно, а записанная в виде (141) – лишь с точностью до экспоненциально малых

слагаемых. Решение (138), (139) содержит главные члены разложения теории сингулярных возмущений в качестве предельного случая $\lambda_0 \gg 1$. Область применимости самого решения (138), (139) гораздо шире и определяется неравенством (137).

2.3.2. Двумерная краевая задача

В двумерном случае приближенное решение задачи (134) удастся построить, если область Ω является прямоугольником. Примеры таких областей показаны на Рисунке 3 и Рисунке 6. Количество омических контактов (областей Γ_{1-0}^i , $i = 1, 2, 3 \dots$) и p-n-переходов может быть произвольным.

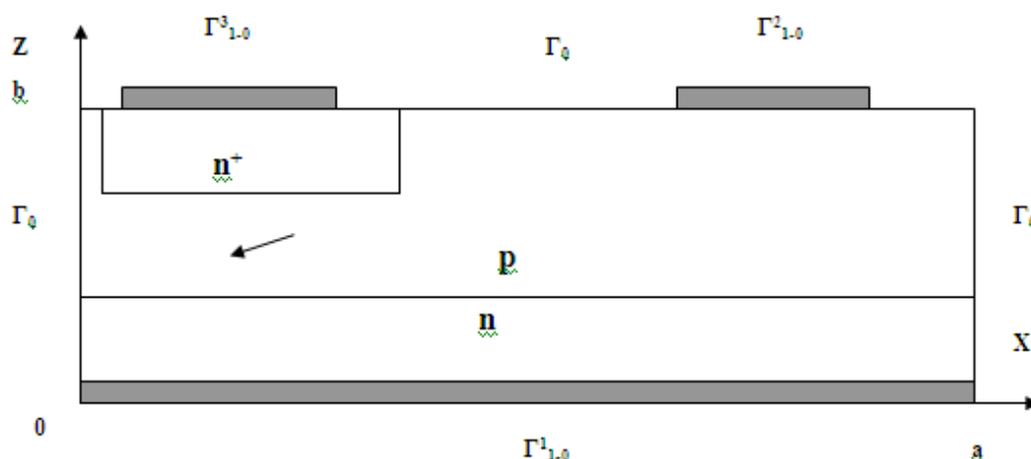


Рисунок 6. Прямоугольное сечение полупроводниковой структуры.

Введем функции $\gamma(x,0)$ и $\gamma(x,b)$, удовлетворяющие условиям

$$\gamma(x,0 \text{ или } b) = \begin{cases} 0, & x \in \Gamma_{1-0}^i, \\ 1, & x \in \Gamma_0, \end{cases}$$

и будем считать их достаточно гладкими. Граничные условия задачи (142) примут вид

$$\begin{cases} \partial Y / \partial x|_{x=0} = 0, \quad \partial Y / \partial x|_{x=a} = 0, \\ [\gamma(x,0) \partial Y / \partial z + (1 - \gamma(x,0)) Y]_{z=0} = 0, \\ [\gamma(x,b) \partial Y / \partial z + (1 - \gamma(x,b)) Y]_{z=b} = 0. \end{cases} \quad (142)$$

Уравнение для Y в двумерном случае можно записать следующим образом

$$\partial^2 Y / \partial x^2 + \partial^2 Y / \partial z^2 - g_z \partial Y / \partial z - g_x \partial Y / \partial x = \lambda^2 Y + h(x, z), \quad (143)$$

где

$$g_z = -g^{-1} \partial g / \partial z, \quad g_x = -g^{-1} \partial g / \partial x, \quad \lambda^2 = q/g, \quad h = f/g.$$

Потребуем выполнение условий

$$g_x(0, z) = g_x(a, z) = 0.$$

Тогда для линейризованной задачи (142), (143) для функции Y_0 , когда

$Y \rightarrow Y_0$, $g \rightarrow g_0 = g|_{Y=0}$, $q \rightarrow q_0 = q|_{Y=0}$, $f \rightarrow f_0 = f|_{Y=0}$, $h \rightarrow h_0 = h|_{Y=0}$, $\lambda \rightarrow \lambda_0 = \lambda|_{Y=0}$, при выполнении условий гладкости (137) можно получить приближенное решение, имеющее следующий вид [54]

$$Y_0 = \sum_{n=0}^{\infty} v_n(x, z) u_n(x, z), \quad (144)$$

$$v_n(x, z) = \cos(\pi n \int_0^x \lambda_0(x', z) dx' / \lambda_0^{\text{int}}(z)) \exp\left\{ \int_0^x g_x(x', z) dx' / 2 \right\}, \quad (145)$$

$$u_n(x, z) = \int_0^b G_n(x, z; z') h_n(z') dz', \quad (146)$$

$$G_n(x, z; z') = Z(x, z; z') \begin{cases} R_+(z, z; z') & 0 \leq z \leq z' \leq b, \\ R_-(x, z; z') & 0 \leq z' \leq z \leq b, \end{cases} \quad (147)$$

где

$$R_+(x, z; z') = \exp\left\{ -\int_z^{z'} p_{2n}(x, \xi) d\xi \right\} (1 - g_0(x) \exp\left\{ -\int_0^z k_n(x, \xi) d\xi \right\}) (1 - g_b(x) \exp\left\{ -\int_{z'}^b k_n(x, \xi) d\xi \right\}),$$

$$R_-(x, z; z') = \exp\left\{ -\int_{z'}^z p_{1n}(x, \xi) d\xi \right\} (1 - g_0(x) \exp\left\{ -\int_0^z k_n(x, \xi) d\xi \right\}) (1 - g_b(x) \exp\left\{ -\int_z^b k_n(x, \xi) d\xi \right\}),$$

$$\lambda_0^{\text{int}}(z) = \int_0^a \lambda_0(x, z) dx,$$

$$Z(x, z; z') = -((k_n(x, z') k_n(x, z))^{1/2} (1 - g_0(x) g_b(x) \exp\left\{ -\int_0^b k_n(x, \xi) d\xi \right\}))^{-1},$$

$$g_0(x) = (1 - \gamma(x,0) + p_{2n}(x,0)\gamma(x,0))/(1 - \gamma(x,0) - p_{1n}(x,0)\gamma(x,0)),$$

$$g_b(x) = (1 - \gamma(x,b) + p_{1n}(x,b)\gamma(x,b))/(1 - \gamma(x,b) - p_{2n}(x,b)\gamma(x,b)),$$

$$k_n = p_{1n} + p_{2n}, p_{2n,1n} = (\lambda_0^2(1 + (\pi n / \lambda_0^{\text{int}})^2) + (g_x^2 + g_z^2)/4)^{1/2} \pm g_z/2$$

$$h_n(z) = (\alpha_n / \lambda_0^{\text{int}}(z)) \int_0^a \lambda_0(x,z) h(x,z) v_n(x,z) \exp\left\{-\int_0^x g_x(x',z) dx'\right\}, \alpha_0 = 1, \alpha_{n \neq 0} = 2.$$

По своей структуре решение (144)-(147) аналогично результату (138), (139), полученному в одномерном случае. Точно также оно описывает появление химической неравновесности электронов и дырок вблизи p-n-переходов и пространственную релаксацию неравновесности при удалении от них.

ГЛАВА 3. Неравновесные эффекты и статические вольт - амперные характеристики полупроводниковых приборов

Полученные результаты позволяют провести аналитическое исследование переноса заряда в полупроводниковых приборах в одномерном и двумерном приближениях [41,47,54,58], а также оценить влияние на их характеристики реальной трехмерной геометрии [50]. При численном моделировании работы полупроводниковых приборов и элементов интегральных схем использование приведенных выше формул в качестве начального приближения на порядок сокращает время расчетов [38,49,54].

Основной задачей настоящей работы является выявление среди множества эффектов, влияющих на характеристики полупроводниковых приборов, тех, которые могли бы быть использованы при конструировании элементной базы наноэлектроники. Для ее решения мы ограничимся простейшим одномерным анализом явлений, определяющих выходные характеристики диодов, транзисторов и тиристоров.

3.1. Биполярный диод

Биполярный полупроводниковый диод [27] состоит из полупроводниковой структуры и двух омических контактов к ней. Полупроводниковая структура представляет собой области полупроводника n и p типов, разделенные p-n-переходом. Один из контактов – это контакт к n-области, а второй – к p-области. Возможные варианты двумерных сечений прибора показаны на Рисунках 7 и 8.

Будем считать p-n-переход достаточно резким, так что

$$|N| \gg 1$$

всюду внутри структуры, за исключением малой окрестности p-n-перехода, который в одномерном сечении $[0, X]$ определяется точкой x_j . В этой точке

происходит смена знака функции $N(x)$. Справедливо приближенное выражение, аналогичное (67)

$$Q/Q_+ \approx \begin{cases} 1 & \text{при } x_e \leq x < x_j \text{ (в п-области),} \\ -1 & \text{при } x_j < x \leq x_c \text{ (в р-области)} \end{cases} \quad (148).$$

Для определенности здесь полагается, что точка x_e принадлежит границе п-области с омическим контактом к ней, а точка x_c – границе р-области с другим омическим контактом.

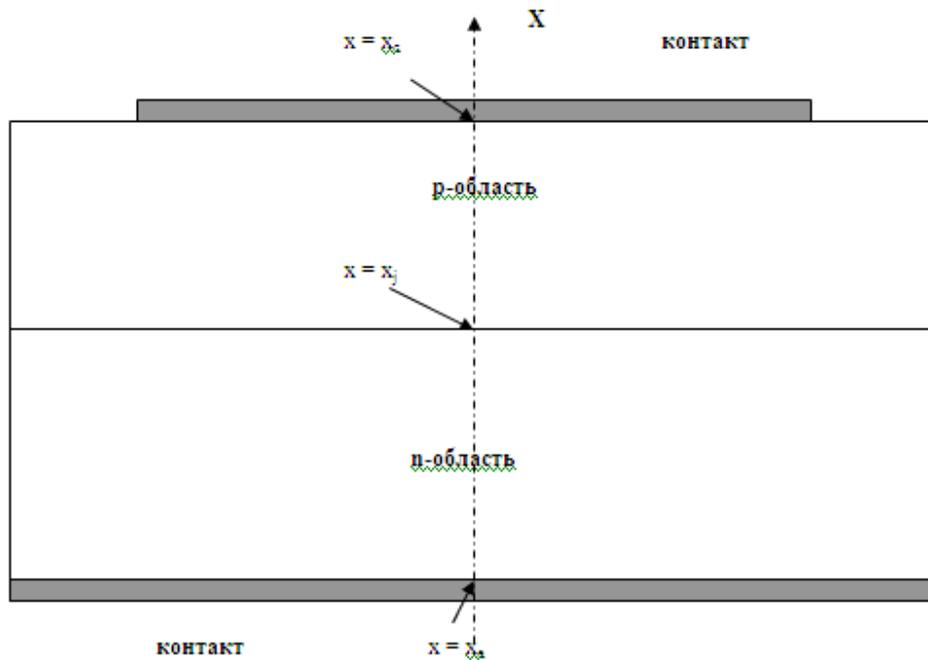


Рисунок 7. Диод с вертикальным расположением п и р областей.

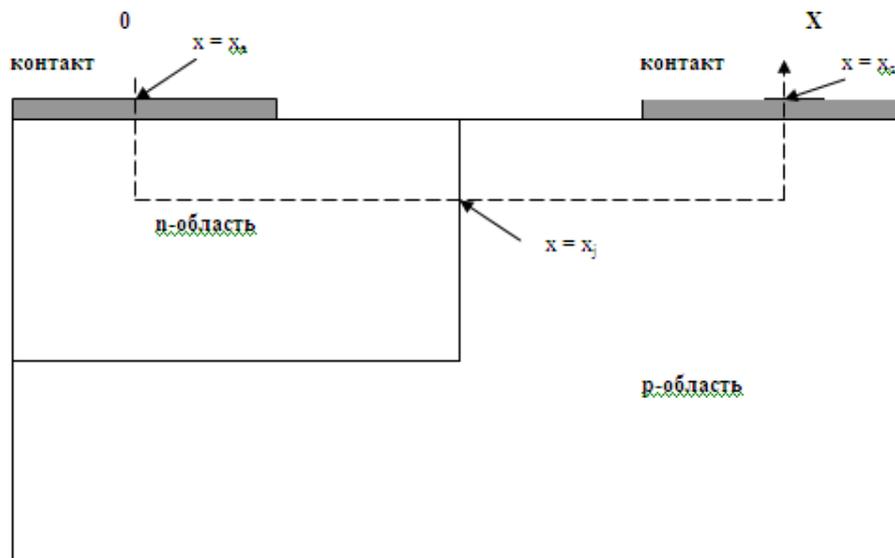


Рисунок 8. Диод с планарным расположением п и р областей.

Согласно (47) в точках x_e и x_c справедливы граничные условия

$$Y(x_e) = Y(x_c) = 0, F(x_e) = 2V_e, F(x_c) = 2V_c. \quad (149)$$

По определению величина

$$V = V_c - V_e$$

является напряжением, приложенным к диоду. Поскольку мы рассматриваем одномерную стационарную задачу вдоль сечения $[0, X]$, то в соответствии с уравнениями (44) плотность тока через диод постоянна

$$j = \text{const.}$$

Применим формулу (53) и с учетом приближения (148) для отношения Q/Q_+ получим

$$V \approx jr + \ln(1 + Y(x_j)), \quad (150)$$

где удельное сопротивление r определено выражением

$$r = \int_{x_e}^{x_c} dx / Q_+ \quad (151)$$

Чтобы найти соотношение, определяющее ВАХ прибора, нужно связать уровень неравновесности $Y(x_j)$ электронов и дырок в области р-п-перехода с плотностью тока в приборе. Для этого воспользуемся результатами решения задачи (136). Согласно (138) имеем

$$Y(x_j) \approx \int_{x_e}^{x_c} G_0(x_j, x') f_0(x') dx', \quad (151)$$

где в соответствии с формулой (135)

$$f_0 = j(\beta^2/2)(d(Q/Q_+)/dx) - G, \quad (152)$$

и действует ограничение (66)

$$-1 < Y \ll N_{ch}^2.$$

Второе слагаемое в правой части выражения (152) определяет зависимость уровня неравновесности электронов и дырок от интенсивности лавинной генерации носителей заряда, либо их генерации внешним воздействием. Первое слагаемое в (152) связывает f_0 , а, следовательно, и Y (см. формулу (151)) с плотностью тока и градиентом примеси в области р-п-перехода. Если лавинная генерация либо внешняя генерация отсутствуют, функция f_0 будет существенно

отличаться от нуля только при отличном от нуля градиенте отношения Q_-/Q_+ , то есть вблизи p-n-перехода.

Далее будут рассматриваться полупроводниковые структуры, содержащие более чем один p-n-переход, поэтому важно определиться с используемой терминологией. Будем называть p-n-переход или, возможно, некоторую его часть при двумерном или трехмерном моделировании, прямо смещенным, если справедливо неравенство

$$(\mathbf{j}, \nabla Q_-/Q_+) < 0. \quad (153.1)$$

Если справедливо обратное неравенство

$$(\mathbf{j}, \nabla Q_-/Q_+) > 0, \quad (153.2)$$

то такой переход будем называть обратным смещенным.

Эта терминология имеет простое физическое обоснование. Прямо смещенный p-n-переход дает положительный вклад в функцию Y , то есть является источником обогащения полупроводниковой структуры неравновесными носителями заряда, а обратный смещенный p-n-переход способствует обеднению полупроводниковой структуры носителями заряда. В простейшем случае диода, то есть прибора, содержащего один p-n-переход, разделяющий области с различным знаком заряда легирующей примеси, снабженные омическими контактами, в режиме прямого смещения согласно (150) $V > 0$, а в режиме обратного смещения $V < 0$. В более сложных ситуациях прямое или обратное смещение p-n-перехода или его части может быть не связано так просто со знаком суммарного падения напряжения на приборе, однако локальный вклад p-n-перехода в это падение будет иметь тот же знак: положительный, для прямо смещенного, и отрицательный, для обратного смещенного, p-n-переходов.

При рассмотрении даже очень простых полупроводниковых структур могут встречаться ситуации, когда одномерные модели неприменимы. Одной из них является планарный резистор с «изолирующим» p-n-переходом. Это прибор, типа изображенного на Рисунке 7, но только в случае, когда оба омических контакта находятся сверху, то есть являются в этой ситуации контактами к n-области. Простейший анализ распределения тока в таком резисторе показывает, что при

приложении разности потенциалов к омическим контактам часть «изолирующего» p-n-перехода будет прямо смещенной, а часть – обратно смещенной.

На Рисунке 9 показаны характерные распределения функции Y в диоде при $G = 0$ и различных положительных значениях плотности тока, соответствующих согласно (148) и (152) режиму прямого смещения p-n-перехода, или, что в случае диода то же самое, прибора в целом. Хорошо видно, что при прямом смещении диод обогащен неравновесными носителями заряда. Максимального значения функция Y достигает в p-n-переходе. Чем больше значение плотности тока, тем большее количество неравновесных носителей заряда появляется в приборе.

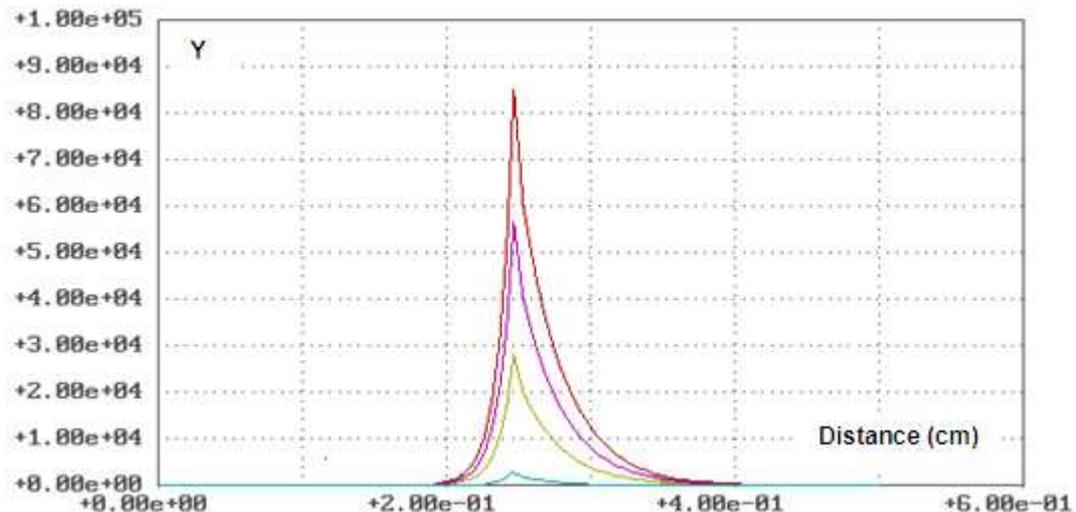


Рисунок 9. Уровень неравновесности электронов и дырок Y в одномерном диоде при прямом смещении и плотностях тока: $j = 1 \cdot 10^{-10}$ А/см²; $j = 1 \cdot 10^{-09}$ А/см²; $j = 2 \cdot 10^{-09}$ А/см²; $j = 3 \cdot 10^{-9}$ А/см² в порядке возрастания максимального значения функции Y .

При отрицательных плотностях тока, когда p-n-переход обратно смещен, величина Y приближается со стороны больших значений к -1 . Хотя приближенные выражения для Y , полученные выше (см. (138) и (139)), могут приводить и к величинам $Y \leq -1$, эти значения находятся за пределами области определения Y и их появление является следствием некорректности приближения. Для корректного вычисления функции Y и ВАХ прибора в режиме обратного смещения важным становится учет лавинного умножения носителей заряда.

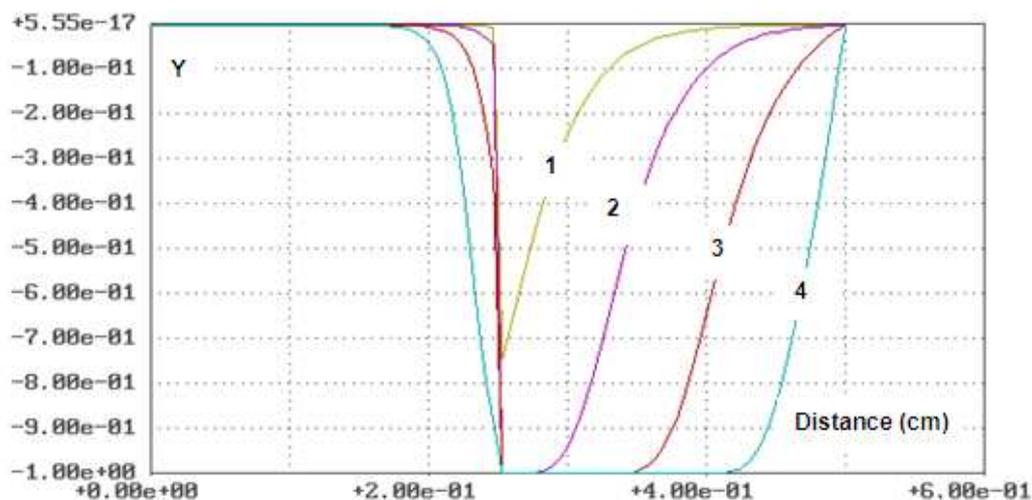


Рисунок 10. Уровень неравновесности электронов и дырок Y в одномерном диоде при обратном смещении при различных значениях плотности тока: $-1 \cdot 10^{-16}$ А/см² (кривая 1); $-1 \cdot 10^{-15}$ А/см² (кривая 2); $-1 \cdot 10^{-14}$ А/см² (кривая 3); $-1 \cdot 10^{-13}$ А/см² (кривая 4).

На Рисунке 10 показаны результаты расчетов Y при отрицательных значениях плотности тока. Видно, что при обратном смещении р-п-перехода прибор обедняется носителями заряда ($Y < 0$). Сначала функция Y достигает значений, близких к -1 в небольшой области обеднения, расположенной вблизи р-п-перехода, а затем с ростом абсолютного значения плотности тока до некоторой величины j_s , эта область увеличивается до значительных размеров. Величина j_s называется плотностью тока насыщения. Далее, уменьшение j уже не сопровождается уменьшением Y . Согласно (150) уменьшение функции Y и приближение ее к значению -1 сопровождается быстрым ростом абсолютного значения напряжения V ($V < 0$). При этом растет напряженность электрического поля в приборе. При некотором пороговом напряжении V_G существенным становится вклад в f_0 , а значит и в Y , эффекта лавинного умножения носителей заряда. При этом величина Y фиксируется вблизи некоторого значения $Y_G \sim -1$ и практически перестает изменяться с ростом тока. Увеличивается только область, в которой Y близка к -1. Таким образом, лавинное умножение компенсирует обеднение прибора носителями заряда, вызванное рекомбинацией. При дальнейшем увеличении абсолютного значения плотности тока в приборе наблюдаются явления типа теплового пробоя, которые мы здесь рассматривать не будем.

С учетом соотношения (148) для функции f_0 в отсутствии генерации ($G=0$) справедливо приближенное выражение

$$f_0(x) \approx -j\beta^2\delta(x - x_j).$$

Это означает, что согласно (138), (139)

$$Y(x) \approx -j\beta^2 G_0(x, x_j), \quad (154)$$

а для функции Грина $G(x, x_j)$ справедливо выражение

$$G_0(x, x_j) = Z(x, x_j) \begin{cases} \frac{x_j}{x} \exp\{-\int_{x_j}^x \lambda_0(\xi) d\xi\} (1 - \exp\{-2\eta(x)\}) (1 - \exp\{-2\eta^*(x_j)\}), & x_e \leq x \leq x_j \leq x_c, \\ \frac{x}{x_j} \exp\{-\int_{x_j}^x \lambda_0(\xi) d\xi\} (1 - \exp\{-2\eta^*(x)\}) (1 - \exp\{-2\eta(x_j)\}), & x_e \leq x_j \leq x \leq x_c, \end{cases}$$

$$Z(x, x_j) = -[2(\lambda_0(x)g_0(x)\lambda_0(x_j)g_0(x_j))^{1/2}(1 - \exp\{-2\lambda_0^{int}\})]^{-1}, \quad (155)$$

где

$$\eta(x) = \int_{x_e}^x \lambda_0(\xi) d\xi, \quad \eta^*(x) = \int_x^{x_c} \lambda_0(\xi) d\xi.$$

Из выражений (154) и (155) явным образом следует, что функция Y , характеризующая уровень неравновесности электронов и дырок, экспоненциально затухает при удалении от р-п – перехода, являющегося при протекании через него ненулевого тока источником этой неравновесности. Характерная длина затухания

$$L_{neq} \sim \lambda_0^{-1},$$

определяется выражением (140). Из выражений (154) и (155) следует, что при размерах прибора

$$L = x_c - x_e < L_{neq}$$

в областях вблизи контактов (точки x_e и x_c) возникают экспоненциальные пограничные слои функции Y (см. рис. 9). Как уже отмечалось, в кремнии

$$L_{neq} \sim 10^{-4} - 10^{-2} \text{ см}$$

и сопоставима с размерами элементов интегральных схем.

Поскольку в соответствии с формулой (150) значением $Y(x_j)$ определяется ВАХ диода, то важным оказывается влияние на $Y(x_j)$ пограничных слоев уровня

неравновесности. На рисунке 11 показано распределение функции Y при прямом смещении в диодах различной длины при одинаковой плотности тока. Хорошо видно, что в «коротких» приборах значение $Y(x_j)$ меньше чем в длинных. То есть заданная плотность тока достигается при меньших приложенных напряжениях. Это подтверждается показанными на Рисунке 12 ВАХ (нижний рисунок) диодов с одинаковыми уровнями профилями легирования, но с разной длиной.

С учетом выражений (154), (155) уравнение (150), определяющее ВАХ диода, перепишем в виде:

$$V = jr + \ln(1 + j/j_s), \quad (156)$$

где

$$j_s = -(\beta^2 G_0(x_j, x_j))^{-1} = \beta^{-2} [2\lambda_0(x_j) g_0(x_j) (1 - \exp\{-2\lambda_0^{int}\})] / [(1 - \exp\{-2\eta(x_j)\})(1 - \exp\{-2\eta^*(x_j)\})] \quad (157)$$

- плотность тока насыщения диода. Как мы уже упоминали выше, величиной j_s определяется предельная плотность тока в приборе при обратном смещении, когда отсутствуют лавинное умножение носителей заряда и их генерации внешним воздействием. Согласно (154) и (155) при сделанных предположениях

$$Y(x_j) = -1 \text{ при } j = -j_s.$$

Если воспользоваться уравнениями (44.Р) и восстановить размерности, то для плотности тока насыщения получим

$$J_s = -J^* j_s = J_{s0} (1 - \exp\{-2\lambda_0^{int}\}) / [(1 - \exp\{-2\eta(x_j)\})(1 - \exp\{-2\eta^*(x_j)\})]. \quad (158)$$

где

$$J_{s0} = -J^* j_{s0} = 2L_{neq}^{-1} \Theta n_i^2 \mu_n \mu_p / (\mu_n n + \mu_p p) |_{x=x_j}. \quad (159)$$

В n-области

$$J_{s0} = J_{s0n} \approx 2eD_p(n_i^2/n)/L_p$$

В p-области

$$J_{s0} = J_{s0p} \approx 2eD_n(n_i^2/p)/L_p$$

В случае резкого p-n-перехода в точке x_j

$$J_{s0} = (J_{s0n} + J_{s0p})/2 \approx eD_p(n_i^2/n)/L_p + eD_n(n_i^2/p)/L_p. \quad (160)$$

Формула (160) совпадает с известным выражением Шокли [27,61]. Соотношение (158) обобщает это выражение на случай «коротких» приборов, когда влияние пограничных эффектов на ВАХ диода может быть существенно.

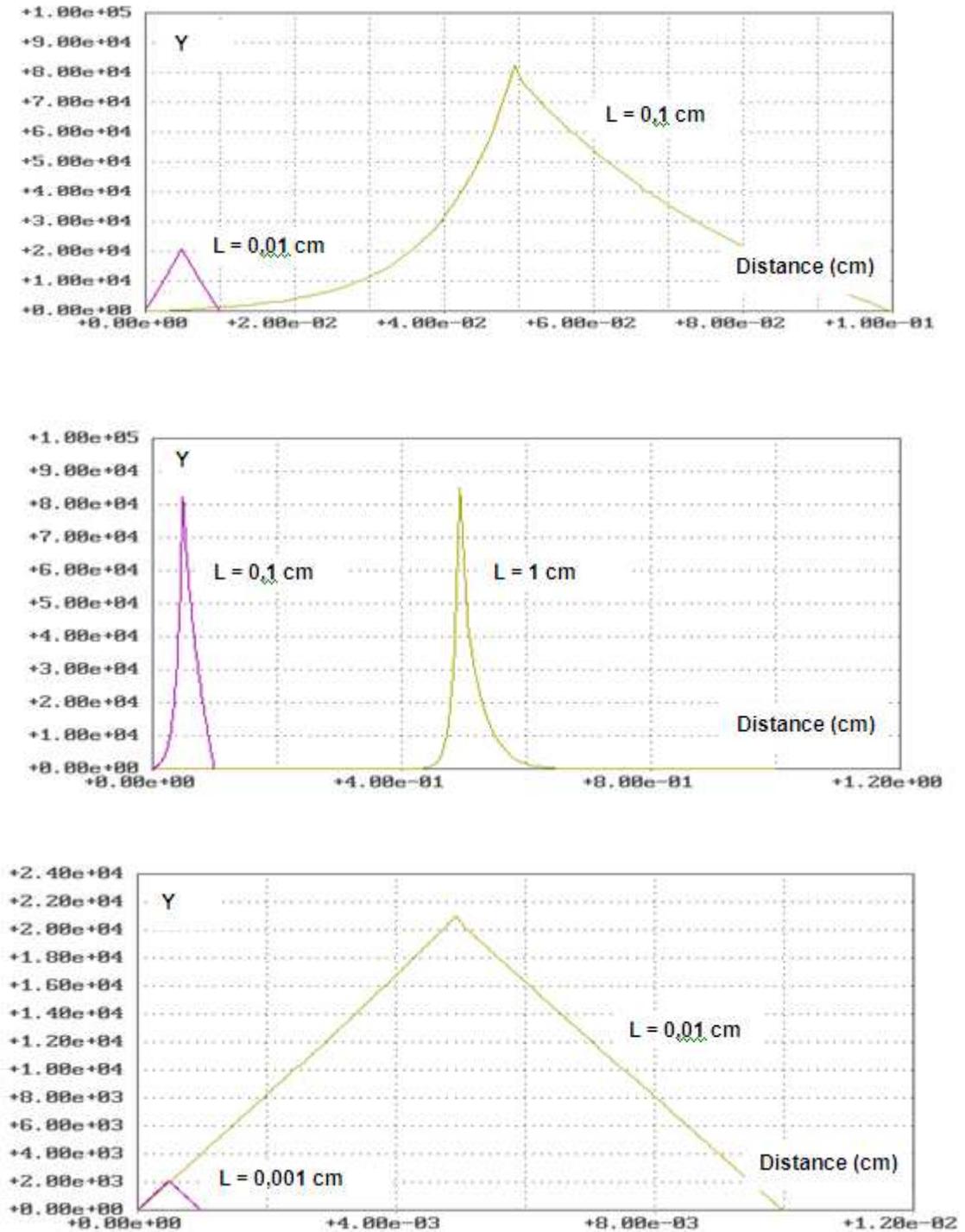


Рисунок 11. Уровень неравновесности электронов и дырок Y в одномерных диодах при одинаковой плотности тока и различной длине прибора L : 1 см и 0,1 см на верхнем рисунке; 0,1 см и 0,01 см на среднем рисунке; 0,01 см и 0,001 см на нижнем рисунке.

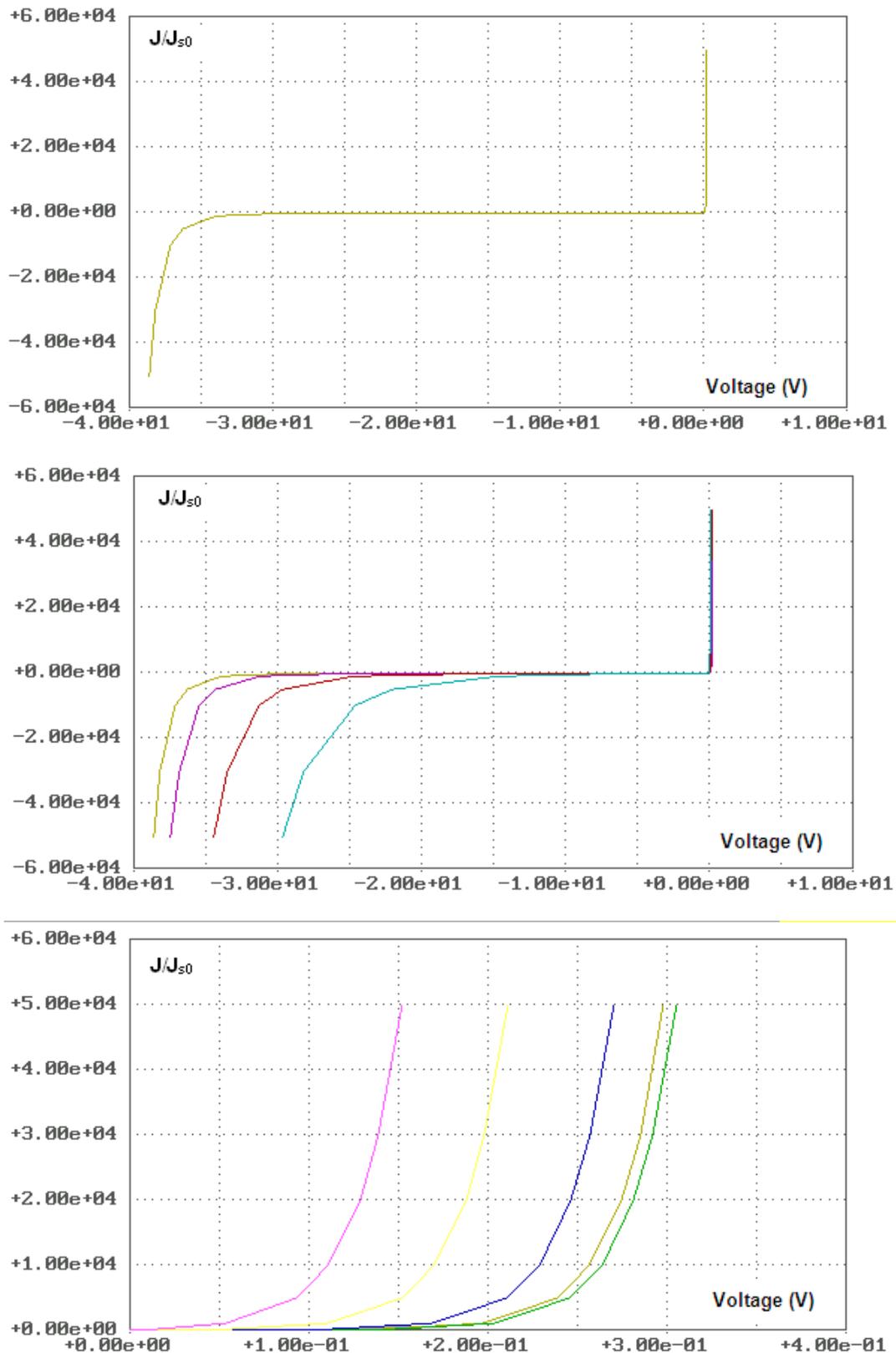


Рисунок 12. Характеристики диодов с одинаковыми профилями легирования, но разной длиной: ВАХ диода с длиной 1 см – верхний рисунок; ВАХ диодов с длинами 1 см, 0,01 см, 0,001 см, 0,0001 см (кривые слева направо) – средний рисунок; прямые ветви ВАХ диодов с длинами 1 см, 0,1 см, 0,01 см, 0,001 см, 0,0001 см (кривые справа налево)– нижний рисунок.

Формула (156) удобна для расчета ВАХ прибора при условии заданной плотности тока. Из нее легко получить известное уравнение Шокли для тока диода, которое обычно используют в инженерных расчетах [27,61]. Если сопротивление r достаточно мало и справедливо неравенство

$$r \ll j_s^{-1}, \quad (161)$$

то из (156) найдем

$$j = j_s(\exp\{V\} - 1). \quad (162)$$

Если восстановить размерности и воспользоваться результатом (160) для оценки j_s , то из (161) получим

$$L/L_{\text{neq}} \ll 1 + Y. \quad (163)$$

Неравенство (163) показывает, что применение формулы Шокли (162) корректно для достаточно «коротких» диодов. Она хорошо описывает ВАХ прибора в режиме прямого смещения, но ее применение проблематично для обратно смещенного прибора. Использование приближенного выражения (156) является более корректным.

В заключение этого параграфа отметим, что применение биполярного диода как элемента электронной схемы основано на особенностях его ВАХ:

- очень малые токи в широком диапазоне напряжений обратного смещения ($V < 0$);
- экспоненциальный рост тока как функции напряжения в режиме прямого смещения ($V > 0$).

Как показывает проведенный здесь анализ, обе эти особенности обусловлены отклонениями газа электронов и дырок от состояния локального химического равновесия.

3.2. Биполярный транзистор

Биполярный полупроводниковый транзистор [27] представляет собой полупроводниковую структуру с тремя омическими контактами к ней. Полупроводниковая структура состоит из трех областей полупроводника с

чередующимся типом проводимости, разделенных р-п-переходами. Эти области называют эмиттером, базой и коллектором транзистора. Геометрически база располагается между эмиттером и коллектором, так что носители заряда перемещающиеся между этими областями не могут базу миновать. Эмиттер и коллектор представляют собой области с одинаковым типом проводимости, а база – с типом проводимости, противоположным эмиттеру и коллектору. Таким образом, эмиттер и коллектор отделены от базы р-п-переходами. Различают два типа биполярных транзисторов п-р-п и р-п-р, которые отличаются типом легирования областей: в первом случае эмиттер и коллектор легированы донорной примесью, а база – акцепторной; во втором случае эмиттер и коллектор легированы акцепторной примесью, а база – донорной. Каждая область снабжена своим омическим контактом, которые мы будем называть эмиттерным, базовым и коллекторным, соответственно. Пример двумерного сечения п-р-п-транзистора показан на Рисунке 13.

Как и в случае с диодом для простоты и наглядности мы ограничимся рассмотрением простейшей одномерной модели биполярного транзистора. Проблемы одномерного моделирования такого рода приборов были подробно обсуждены в параграфе 1.6. Там же были сформулированы граничные условия для уравнений переноса заряда в этом случае.

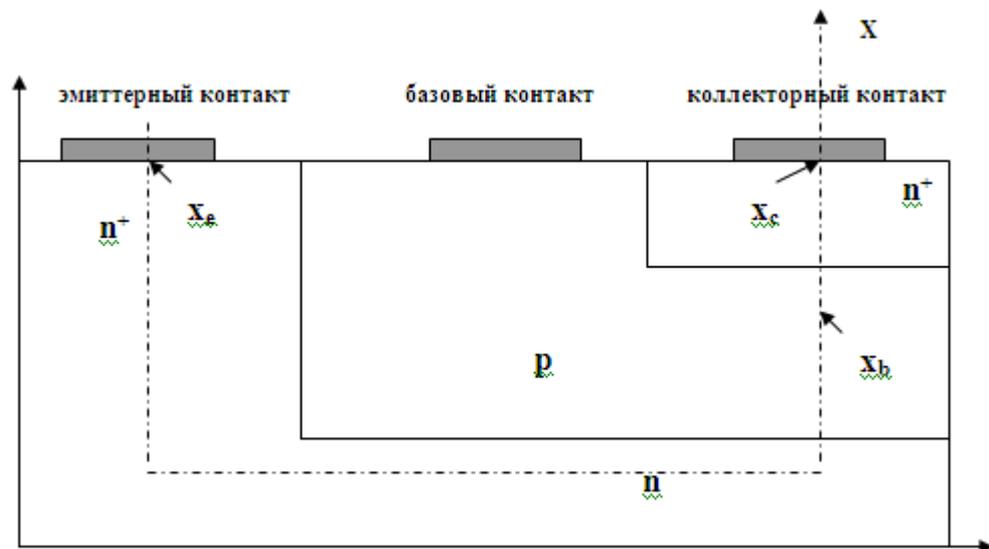


Рисунок 13. Один из возможных вариантов реализации биполярного п-р-п-транзистора.

В рамках одномерного приближения область переноса заряда в транзисторе представляет собой отрезок $[x_e, x_c]$. В точках x_e и x_c располагаются омические контакты. Граничные условия для функций Y и F в этих точках такие же, как и для диода и описываются выражением (149). В точке, x_b , соответствующей базовому контакту, граничные условия ставятся на основании выражения (73). В результате имеем

$$\begin{cases} Y(x_e) = Y(x_c) = 0, \\ F(x_e) = 2V_e, F(x_c) = 2V_c, \\ F(x_b) = 2V_b - \ln(1 + Y(x_b)), \text{ для n-p-n-транзистора,} \\ F(x_b) = 2V_b + \ln(1 + Y(x_b)), \text{ для p-n-p-транзистора.} \end{cases} \quad (164)$$

Как отмечалось в параграфе 1.6, в точке x_b предполагается разрыв плотности тока. В соответствии с формулой (71) имеем

$$j = \begin{cases} j_e, & x_e \leq x < x_b, \\ j_c, & x_b < x \leq x_c \end{cases} \quad (165)$$

где j_e и j_c плотности тока эмиттера и коллектора соответственно. Разница между этими величинами

$$j_b = j_e - j_c, \quad (166)$$

обеспечивающая сохранение плотности тока в приборе, называется плотностью тока базы.

Далее для определенности будем анализировать работу n-p-n-транзистора. Полученные результаты без труда обобщаются на случай p-n-p-приборов путем инверсии соответствующих плотностей токов и напряжений.

Положим, что эмиттерный и коллекторный p-n-переходы достаточно резкие и для них справедливы приближенные выражения, аналогичные (148)

$$Q/Q_+ \approx \begin{cases} 1 & \text{при } x_e \leq x < x_{je} \text{ и } x_{jc} < x \leq x_c \text{ (в n-областях),} \\ -1 & \text{при } x_{je} < x < x_{jc} \text{ (в p-области).} \end{cases} \quad (167).$$

Здесь x_{je} и x_{jc} точки, в которых происходит смена знака концентрации примеси (функции $N(x)$), и которые мы будем называть точками эмиттерного и коллекторного p-n-переходов соответственно.

Используя формулу (51) и граничные условия (164) найдем

$$F(x) = 2V_e + \int_{x_e}^x dx (2j/Q_+ + (Q_-/Q_+) d \ln(1 + Y)/dx). \quad (168)$$

Из (168) с учетом (164), (165) и (167) получим

$$\begin{cases} V_{be} \approx j_e r_{eb} + \ln(1 + Y(x_{je})), \\ V_{cb} \approx j_c r_{bc} - \ln(1 + Y(x_{jc})), \end{cases} \quad (169)$$

где удельные сопротивления r_{eb} и r_{bc} определяются формулами

$$r_{eb} = \int_{x_e}^{x_b} dx/Q_+, \quad r_{bc} = \int_{x_b}^{x_c} dx/Q_+, \quad (170)$$

и приняты обозначения

$$V_{be} = V_b - V_e, \quad V_{cb} = V_c - V_b.$$

Каждая из формул (169) аналогична соотношению (150) для диода. Такое представление для напряжений V_{be} и V_{cb} получено благодаря «специфическому» граничному условию для F в точке x_b . При этом существенно, что формулы (169) описывают не два встречно включенных диода, а единый прибор, поскольку фигурирующие в них величины $Y(x_{je})$ и $Y(x_{jc})$ вычисляются из одной краевой задачи типа (136).

Согласно (138) для уровня неравновесности электронов и дырок в приборе имеем выражение

$$Y(x) \approx \int_{x_e}^{x_c} G_0(x, x') f_0(x') dx'. \quad (171)$$

В отсутствии лавинного умножения носителей заряда и их генерации из внешних источников в соответствии с формулами (135), (165) и (167) найдем

$$f_0 = j(\beta^2/2)(d(Q_-/Q_+)/dx) \approx -j_e \beta^2 \delta(x - x_{je}) + j_c \beta^2 \delta(x - x_{jc}). \quad (172)$$

Это означает, что для $Y(x)$ справедливо приближенное выражение

$$Y(x) \approx -j_e \beta^2 G_0(x, x_{je}) + j_c \beta^2 G_0(x, x_{jc}). \quad (173)$$

Для функций Грина $G(x, x_{je,c})$ справедлива формула (155), в которой x_j заменяется на x_{je} или x_{jc} , а для Y действует ограничение (66). В соответствии с (173) для фигурирующих в формулах (169) значений $Y(x_{je,c})$ получим выражения

$$\begin{cases} Y(x_{je}) \approx -j_e \beta^2 G_0(x_{je}, x_{je}) + j_c \beta^2 G_0(x_{je}, x_{jc}), \\ Y(x_{jc}) \approx -j_e \beta^2 G_0(x_{jc}, x_{je}) + j_c \beta^2 G_0(x_{jc}, x_{jc}). \end{cases} \quad (174)$$

В формулах (174) диагональные по индексам e и c слагаемые описывают вклад в падения напряжения V_{be} и V_{cb} эмиттерного и коллекторного p - n -переходов соответственно. Недиagonальные слагаемые в (174) описывают влияние коллекторного перехода на падение напряжения на эмиттерном переходе и наоборот. В случае, когда членами $G_0(x_{je}, x_{jc})$ и $G_0(x_{jc}, x_{je})$ в (174) можно пренебречь, формулы (169) будут описывать падение напряжения на двух встречно включенных диодах.

С учетом (174) перепишем выражения (169) в виде

$$\begin{cases} V_{be} \approx j_e r_{eb} + \ln(1 + j_e \alpha_{ee} - j_c \alpha_{ec}), \\ V_{cb} \approx j_c r_{bc} - \ln(1 + j_e \alpha_{ce} - j_c \alpha_{cc}). \end{cases} \quad (175)$$

Для коэффициентов справедливы соотношения

$$\begin{cases} \alpha_{ee} = -\beta^2 G_0(x_{je}, x_{je}), \alpha_{ec} = -\beta^2 G_0(x_{je}, x_{jc}), \\ \alpha_{ce} = -\beta^2 G_0(x_{jc}, x_{je}), \alpha_{cc} = -\beta^2 G_0(x_{jc}, x_{jc}). \end{cases} \quad (176)$$

В соответствии с результатами предыдущего параграфа, диагональные элементы матрицы α_{ab} ($a = e, c; b = e, c$) имеют смысл обратных плотностей токов насыщения эмиттерного и коллекторного p - n -переходов

$$\begin{cases} \alpha_{ee} = j_{se}^{-1} = j_{se0}^{-1} [(1 - \exp\{-2\eta(x_{je})\})(1 - \exp\{-2\eta^*(x_{je})\}) / (1 - \exp\{-2\lambda_0^{int}\})], \\ \alpha_{cc} = j_{sc}^{-1} = j_{sc0}^{-1} [(1 - \exp\{-2\eta(x_{jc})\})(1 - \exp\{-2\eta^*(x_{jc})\}) / (1 - \exp\{-2\lambda_0^{int}\})], \end{cases} \quad (177)$$

где

$$j_{se0} = 2\beta^{-2} \lambda_0(x_{je}) g_0(x_{je}), \quad j_{sc0} = 2\beta^{-2} \lambda_0(x_{jc}) g_0(x_{jc}).$$

Для величин

$$J_{se0} = -J^* j_{se0} \text{ и } J_{sc0} = -J^* j_{sc0}$$

справедливы приближенные формулы, аналогичные (158), (159) и (160), позволяющие связать их с длинами и коэффициентами диффузии.

Недиagonальные элементы матрицы α_{ab} имеют вид

$$\begin{aligned} \alpha_{ec} = \alpha_{ce} = \\ = [(j_{se0} j_{sc0})^{1/2} (1 - \exp\{-2\lambda_0^{int}\})]^{-1} \exp\left\{-\int_{x_{je}}^{x_{jc}} \lambda_0(\xi) d\xi\right\} (1 - \exp\{-2\eta(x_{je})\}) (1 - \exp\{-2\eta^*(x_{jc})\}). \end{aligned} \quad (178)$$

Как и в случае биполярного диода, полученные формулы для плотностей токов насыщения эмиттерного и коллекторного p-n-переходов учитывают влияние неравновесных пограничных эффектов, то есть возможных пограничных слоев функции Y . Согласно (178) значение коэффициента α_{ce} , описывающего взаимное влияние эмиттерного и коллекторного p-n-переходов, существенно зависит не только от их расстояния до границ прибора, но и от расстояния между переходами. Фактор

$$\alpha = \exp\left\{-\int_{x_{je}}^{x_{jc}} \lambda_0(\xi) d\xi\right\} \quad (179)$$

является основным параметром, определяющим возможность использования рассматриваемого прибора в качестве транзистора. Из (179) легко получить простейшую оценку

$$\alpha \approx \exp\{-L_b/L_{neqb}\}, \quad (180)$$

где

$$L_b = x_c - x_e$$

- расстояние между эмиттерным и коллекторным p-n-переходами или, что то же самое, длина базы, а L_{neqb} – длина неравновесности электронно-дырочного газа в некоторой средней точке в базе. В случае достаточно длинной базы, когда выполнено условие

$$L_b \gg L_{neqb}, \quad (181)$$

коэффициент α_{ce} стремится к нулю и из формул (175) получаем выражения

$$V_{be} \approx j_e r_{eb} + \ln(1 + j_e \alpha_{ee}), \quad V_{cb} \approx j_c r_{bc} - \ln(1 - j_c \alpha_{cc}),$$

описывающие ВАХ двух диодов. Один из них соответствует эмиттерному p-n-переходу, а второй коллекторному. Условие (166), связывающее плотности токов j_c и j_e означает просто выполнение закона Киркгоффа в точке x_b , соответствующей базовому контакту.

В случае, когда длина базы прибора того же порядка, что и длина релаксации электронов и дырок в ней к состоянию химического равновесия, то есть когда справедлива оценка

$$L_b \sim L_{neqb}, \quad (182)$$

прибор может быть применен для усиления электрических сигналов. Действительно, используя формулы (166) и (175) получим для напряжения на коллекторном контакте

$$V_{cb} \approx j_c r_{bc} - \ln(1 + j_b \alpha_{ce} - j_c(\alpha_{cc} - \alpha_{ce})). \quad (183)$$

Пренебрегая омическим падением напряжения на участке $[x_b, x_c]$, найдем

$$j_c \approx (\alpha_{cc} - \alpha_{ce})^{-1}(1 - \exp\{-V_{cb}\}) + j_b \alpha_{ce}(\alpha_{cc} - \alpha_{ce})^{-1}. \quad (184)$$

В нормальном режиме работы биполярного транзистора коллекторный p-n-переход обратно смещен. Для n-p-n-транзистора это означает, что в этом режиме

$$V_{cb} > 0.$$

При условии выхода на режим насыщения коллекторного p-n-перехода, когда

$$V_{cb} \gg 1 \quad (V_{cb} \gg k_B T/e \text{ в размерных единицах})$$

из (184) получим

$$j_c \approx (1 + j_b \alpha_{ce})/(\alpha_{cc} - \alpha_{ce}), \quad (185)$$

откуда для статического коэффициента усиления (передачи) прибора в схеме с общим эмиттером β_0 [27] найдем приближенное выражение

$$\beta_0 = \partial j_c / \partial j_b \approx \alpha_{ce}/(\alpha_{cc} - \alpha_{ce}). \quad (186)$$

Если пренебречь влиянием пограничных слоев функции Y на характеристики транзистора, то из выражения (186) с учетом результатов (177), (178) и (180) получим

$$\beta_0 \approx (j_{sc0}/j_{se0})^{1/2} \exp\{-L_b/L_{neqb}\} / (1 - (j_{sc0}/j_{se0})^{1/2} \exp\{-L_b/L_{neqb}\}). \quad (187)$$

Из этой формулы следует, что рассматриваемый прибор позволяет получать значительное усиление электрических сигналов в случаях, когда длина его базы сравнима с длиной релаксации электронов и дырок к химическому равновесию.

Справедлива оценка, уточняющая результат (182)

$$L_b \sim (L_{neqb}/2) \ln(j_{sc0}/j_{se0}). \quad (188)$$

Используя формулы (186), (187) и (166) нетрудно найти приближенное выражение для статического коэффициента усиления (передачи) в схеме с общей базой

$$\alpha_0 = \partial j_c / \partial j_e \approx \alpha_{ce}/\alpha_{cc} \approx (j_{sc0}/j_{se0})^{1/2} \exp\{-L_b/L_{neqb}\}.$$

Таким образом, неравновесность электронно-дырочного газа, порождаемая протеканием тока через р-п-переходы транзистора, является физической основой применения этого прибора для усиления электрических сигналов. Поскольку, как показано на Рисунках 14 и 15, в нормальном режиме работы транзистора эмиттерный р-п-переход находится в прямосмещенном состоянии, а коллекторный – в обратносмещенном, малые изменения плотности тока базы (см. Рисунок 16) вызывают значительные изменения тока коллектора (см. Рисунок 17).

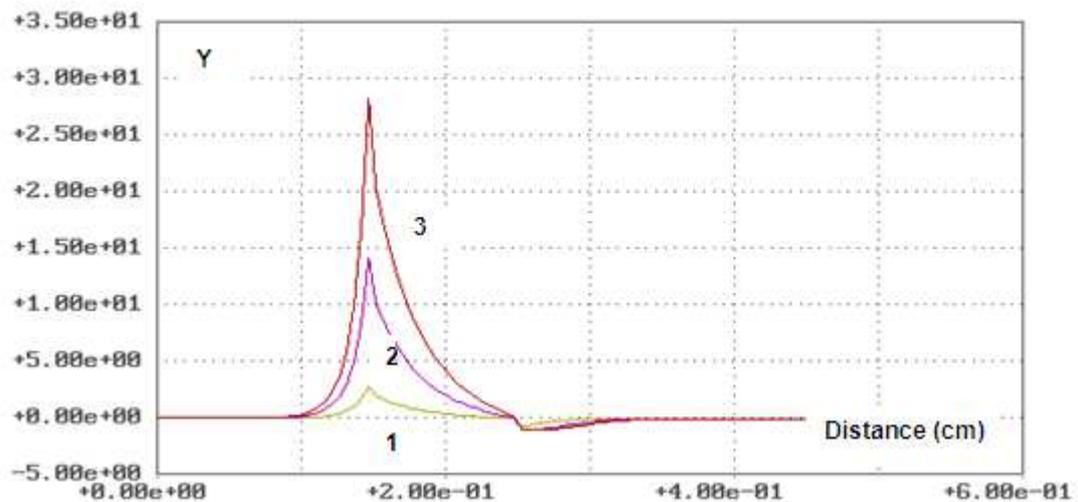


Рисунок 14. Характерные распределения функции Y в биполярном п-р-п-транзисторе при нулевом токе базы в случае, когда эмиттерный р-п-переход смещен в прямом направлении, а коллекторный – в обратном ($1 - j_c = 1 \cdot 10^{-13} \text{ A/cm}^2$; $2 - j_c = 5 \cdot 10^{-13} \text{ A/cm}^2$; $3 - j_c = 1 \cdot 10^{-12} \text{ A/cm}^2$).

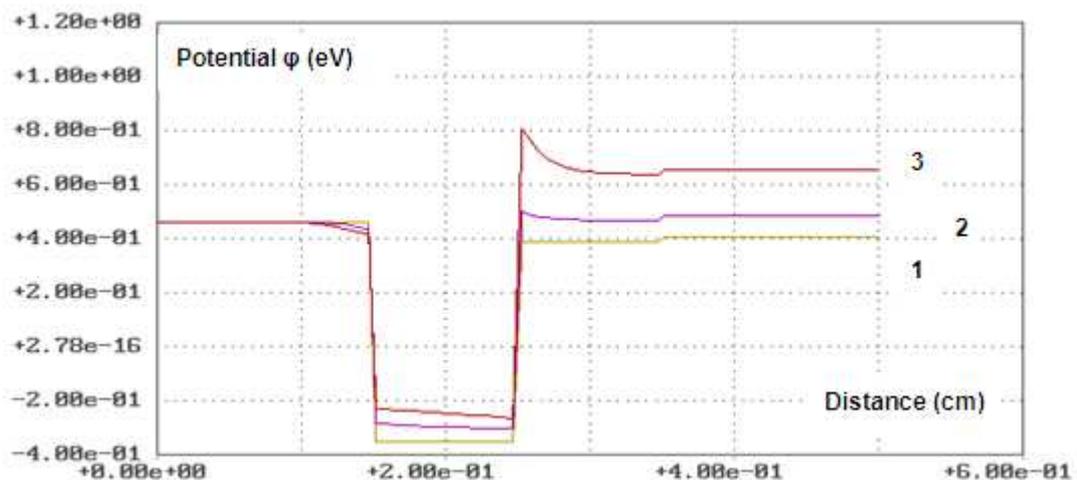


Рисунок 15. Характерное распределение электростатического потенциала в биполярном п-р-п-транзисторе при схеме включения с общим эмиттером при $j_b = 0 \text{ A/cm}^2$ и: $1 - j_c = 0 \text{ A/cm}^2$; $2 - j_c = 2 \cdot 10^{-13} \text{ A/cm}^2$; $3 - j_c = 1 \cdot 10^{-12} \text{ A/cm}^2$.

Этот эффект обусловлен изменением количества неравновесных носителей заряда, инжектируемых в область базы из эмиттера и достигающих коллекторного р-п-перехода. Модулируя поток неравновесных носителей заряда электрическим сигналом, приложенным между базой и эмиттером, можно добиться значительного усиления этого сигнала на коллекторном контакте.

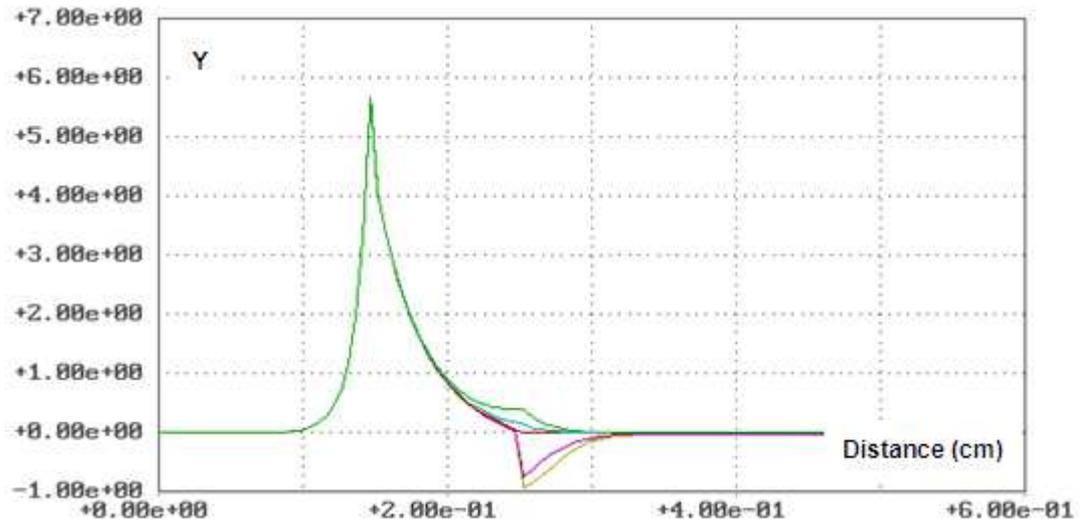


Рисунок 16. Изменение функции Y в биполярном п-р-п-транзисторе при фиксированном токе эмиттера и изменении тока базы. Кривые снизу вверх: $j_b = 0 \text{ A/cm}^2$; $j_b = -1 \cdot 10^{-14} \text{ A/cm}^2$; $j_b = -5 \cdot 10^{-14} \text{ A/cm}^2$; $j_b = -2 \cdot 10^{-13} \text{ A/cm}^2$; $j_b = -5 \cdot 10^{-13} \text{ A/cm}^2$.

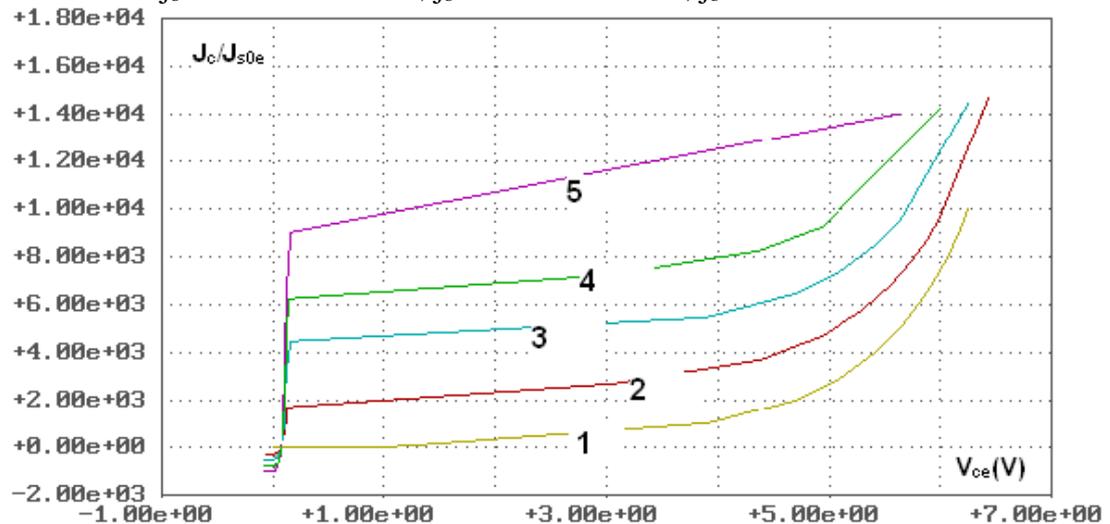


Рисунок 17. Семейство выходных ВАХ биполярного п-р-п-транзистора при схеме включения с общим эмиттером: 1 - $j_b = 0$; 2 - $j_b = -2,5 \cdot 10^2 J_{s0e}$; 3 - $j_b = -5 \cdot 10^2 J_{s0e}$; 4 - $j_b = -7,5 \cdot 10^2 J_{s0e}$; 5 - $j_b = -1 \cdot 10^3 J_{s0e}$. J_{s0e} – плотность тока насыщения р-п-перехода эмиттер-база.

3.3. Полупроводниковый тиристор

Полупроводниковый тиристор - еще один прибор, функционирование которого основано на неравновесных эффектах. Из всех возможных разновидностей этого прибора [27,62-64] остановимся на простейшем диодном тиристоре. Схематично его двумерное сечение показано на Рисунке 18.

Диодный тиристор представляет собой полупроводниковую структуру с тремя p-n-переходами и двумя омическими контактами. Как показано на Рисунке 18 в одномерном сечении прибора омическим контактам соответствуют точки $x = x_e$ и $x = x_c$, а p-n-переходам – точки $x = x_{j1}$, $x = x_{j2}$ и $x = x_{j3}$.

На практике [27,62,63] часто используют тиристоры с тремя или четырьмя контактами. Это контакты к областям между p-n-переходами. Их обычно называют затворами. С их помощью можно управлять характеристиками прибора, изменяя плотности токов в различных областях полупроводниковой структуры.

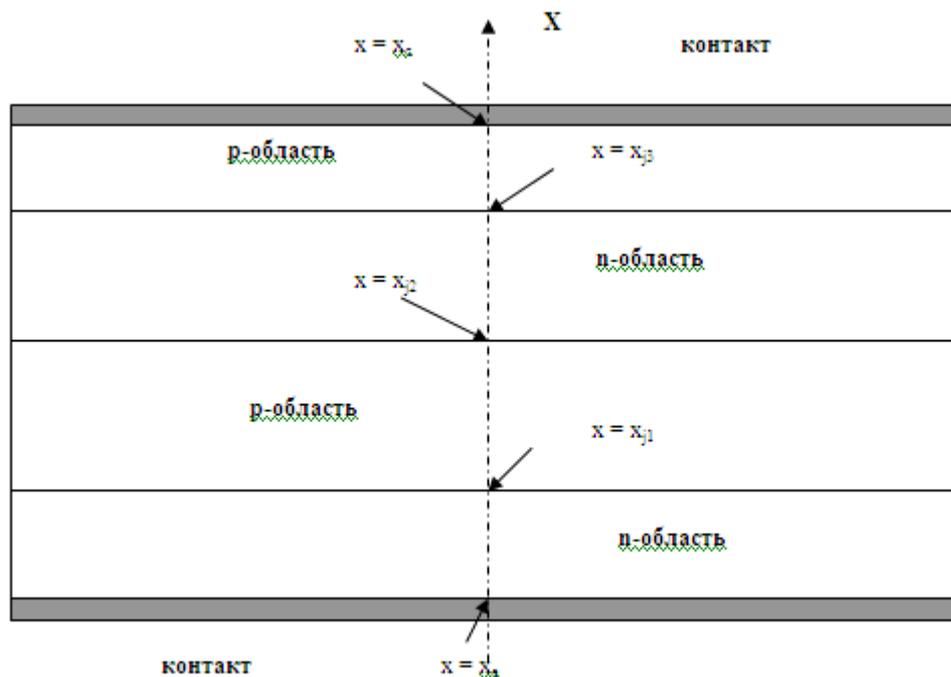


Рисунок 18. Двумерное сечение простейшего варианта полупроводниковой структуры, соответствующей диодному тиристору.

Для проведения аналитических оценок положим приближенно, что

$$Q_-/Q_+ \approx \begin{cases} 1 & \text{при } x_e \leq x < x_{j1}, x_{j2} < x < x_{j3} \text{ (в n-областях)} \\ -1 & \text{при } x_{j1} < x < x_{j2}; x_{j3} < x \leq x_c \text{ (в p-областях)}. \end{cases} \quad (189).$$

В точках x_e и x_c справедливы граничные условия, аналогичные (149)

$$Y(x_e) = Y(x_c) = 0, F(x_e) = 2V_e, F(x_c) = 2V_c.$$

Величина

$$V = V_c - V_e$$

является напряжением, приложенным к тиристор. Поскольку мы рассматриваем одномерную стационарную задачу, то в соответствии с уравнениями (44) плотность тока через прибор постоянна

$$j = \text{const.}$$

Так же как и для диода, применяя формулу (53), с учетом приближения (189) получим

$$V \approx jr + \ln(1 + Y(x_{j1})) - \ln(1 + Y(x_{j2})) + \ln(1 + Y(x_{j3})), \quad (190)$$

где удельное сопротивление r определено выражением (151)

$$r = \int_{x_e}^{x_c} dx/Q_+.$$

Функция $Y(x)$ вычисляется по формуле (171) в которой для источника $f_0(x)$ справедливо приближенное выражение

$$f_0(x) = j(\beta^2/2)(d(Q_-/Q_+)/dx) \approx -j\beta^2\{\delta(x - x_{j1}) - \delta(x - x_{j2}) + \delta(x - x_{j3})\}.$$

В результате имеем для $Y(x)$

$$Y(x) \approx -j\beta^2\{G_0(x, x_{j1}) - G_0(x, x_{j2}) + G_0(x, x_{j3})\}. \quad (191)$$

Функция Грина $G(x, x_{ja})$ определена формулой (155), в которой x_j необходимо заменить на x_{ja} , где $a = 1, 2, 3$. Из этого соотношения и формулы (155) получим выражения для функции Y в точках p-n-переходов

$$\begin{cases} Y(x_{j1}) = j(\alpha_{11} + \alpha_{12} + \alpha_{13}), \\ Y(x_{j2}) = j(\alpha_{21} + \alpha_{22} + \alpha_{23}), \\ Y(x_{j3}) = j(\alpha_{31} + \alpha_{32} + \alpha_{33}). \end{cases} \quad (192)$$

Для коэффициентов α_{ab} в (192) из (191) и (155) найдем

$$\mathfrak{a}_{ab} = \mathfrak{a}_{ba} = \mathfrak{a}_{0ab} = \begin{cases} \exp\left\{-\int_{X_{ja}}^{X_{jb}} \lambda_0(\xi) d\xi\right\} (1 - \exp\{-2\eta(x_{ja})\}) (1 - \exp\{-2\eta^*(x_{jb})\}), & x_e \leq X_{ja} \leq X_{jb} \leq x_c, \\ \exp\left\{-\int_{X_{jb}}^{X_{ja}} \lambda_0(\xi) d\xi\right\} (1 - \exp\{-2\eta^*(x_{ja})\}) (1 - \exp\{-2\eta(x_{jb})\}), & x_e \leq X_{jb} \leq X_{ja} \leq x_c, \end{cases}$$

где

$$j_{sa0} = 2\beta^{-2} \lambda_0(x_{ja}) g_0(x_{ja}), \quad \mathfrak{a}_{0ab} = \mathfrak{a}_{oba} = [(j_{sa0} j_{sb0})^{1/2} (1 - \exp\{-2\lambda_0^{int}\})]^{-1}, \quad a, b = 1, 2, 3.$$

В результате, получим

$$\mathfrak{a}_{aa} = j_{sa}^{-1} = j_{sa0}^{-1} [(1 - \exp\{-2\eta(x_{ja})\}) (1 - \exp\{-2\eta^*(x_{ja})\})] / (1 - \exp\{-2\lambda_0^{int}\}), \quad a = 1, 2, 3,$$

$$\mathfrak{a}_{12} = \mathfrak{a}_{21} =$$

$$= [(j_{s10} j_{s20})^{1/2} (1 - \exp\{-2\lambda_0^{int}\})]^{-1} \exp\left\{-\int_{X_{j1}}^{X_{j2}} \lambda_0(\xi) d\xi\right\} (1 - \exp\{-2\eta(x_{j1})\}) (1 - \exp\{-2\eta^*(x_{j2})\}),$$

$$\mathfrak{a}_{13} = \mathfrak{a}_{31} =$$

$$= [(j_{s10} j_{s30})^{1/2} (1 - \exp\{-2\lambda_0^{int}\})]^{-1} \exp\left\{-\int_{X_{j1}}^{X_{j3}} \lambda_0(\xi) d\xi\right\} (1 - \exp\{-2\eta(x_{j1})\}) (1 - \exp\{-2\eta^*(x_{j3})\}),$$

$$\mathfrak{a}_{23} = \mathfrak{a}_{32} =$$

$$= [(j_{s20} j_{s30})^{1/2} (1 - \exp\{-2\lambda_0^{int}\})]^{-1} \exp\left\{-\int_{X_{j2}}^{X_{j3}} \lambda_0(\xi) d\xi\right\} (1 - \exp\{-2\eta(x_{j2})\}) (1 - \exp\{-2\eta^*(x_{j3})\}). \quad (193)$$

Напомним, что эти формулы приближенно справедливы только в области значений плотности тока j , при которых выполнены неравенства (66).

На Рисунке 19 показана вольт - амперная характеристика тиристора. Будем называть прямой ветвью ВАХ ту ее часть, которая соответствует режиму работы, когда ближайшие к контактным областям р-п-переходы являются прямо смещенными. В рассматриваемом случае п-р-п-р- структуры (см. Рисунок 18) прямая ветвь соответствует положительным значениям плотности тока ($j > 0$). Понятно, что в случае р-п-р-п- структуры прямая ветвь будет соответствовать отрицательным плотностям тока j . Обратной ветвью ВАХ будем называть ту ее часть, которая соответствует режиму работы, когда ближайшие к контактным областям р-п-переходы являются обратно смещенными.

На прямой ветви ВАХ тиристора выделяют следующие области (см. Рисунок 19): $0 \leq j < j_s$ – область запирания; $j_s < j < j_h$ – область отрицательного дифференциального сопротивления; $j_h < j$ – область отпираания. На обратной ветви ВАХ прибора выделяют область запираания обратной ветви и область пробоя.

Важнейшей особенностью тиристора, определяющей его применение в электронных устройствах и схемах [27,62-64], является S-образность ВАХ, то есть наличие на ВАХ участка с отрицательным дифференциальным сопротивлением. Нетрудно понять физические причины появления этого участка. Действительно согласно (190), суммарное падение напряжения на прямой ветви ВАХ определяется омическим членом, падением напряжения на прямо смещенных p-n-переходах x_{j1} и x_{j3} и падением напряжения на обратно смещенном p-n-переходе x_{j2} . При плотностях тока порядка токов насыщения p-n-переходов вплоть до значения плотности тока включения j_s вклад обратно смещенного p-n-перехода x_{j2} в V является превалирующим. При этом, как показано на Рисунке 20, p-n-переход x_{j2} остается обедненным носителями заряда.

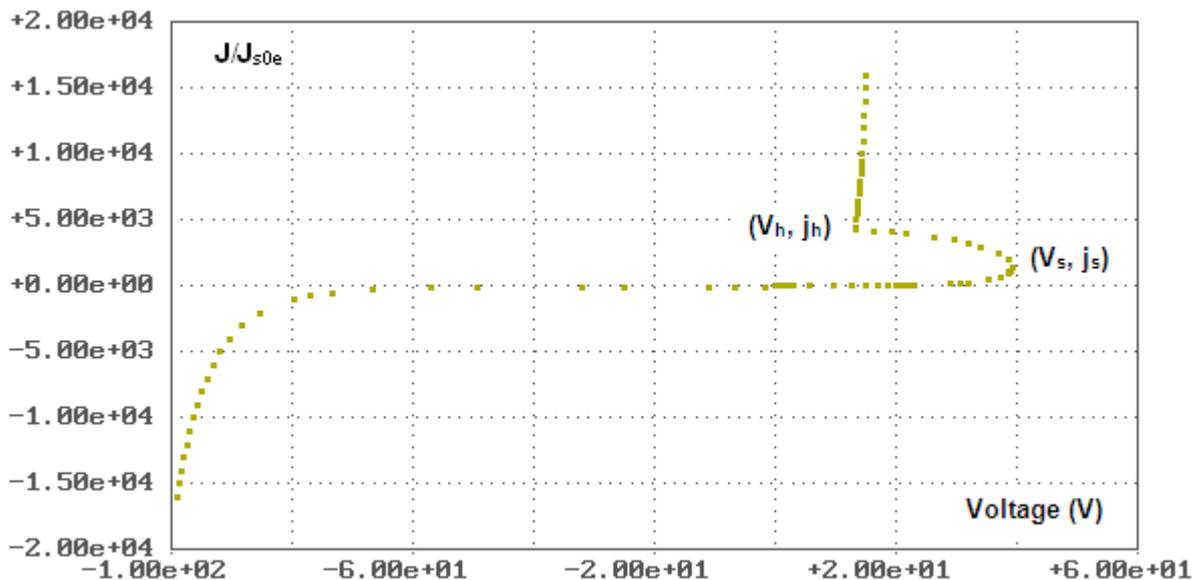


Рисунок 19. ВАХ тиристора, на которой $V_s = 39,3$ В и $j_s = 1,5 \cdot 10^3 J_{s0e}$ – напряжение и плотность тока включения; $V_h = 13,4$ В и $j_h = 4,3 \cdot 10^3 J_{s0e}$ – остаточное напряжение и плотность тока выключения. J_{s0e} – плотность тока насыщения p-n-перехода x_{j1} .

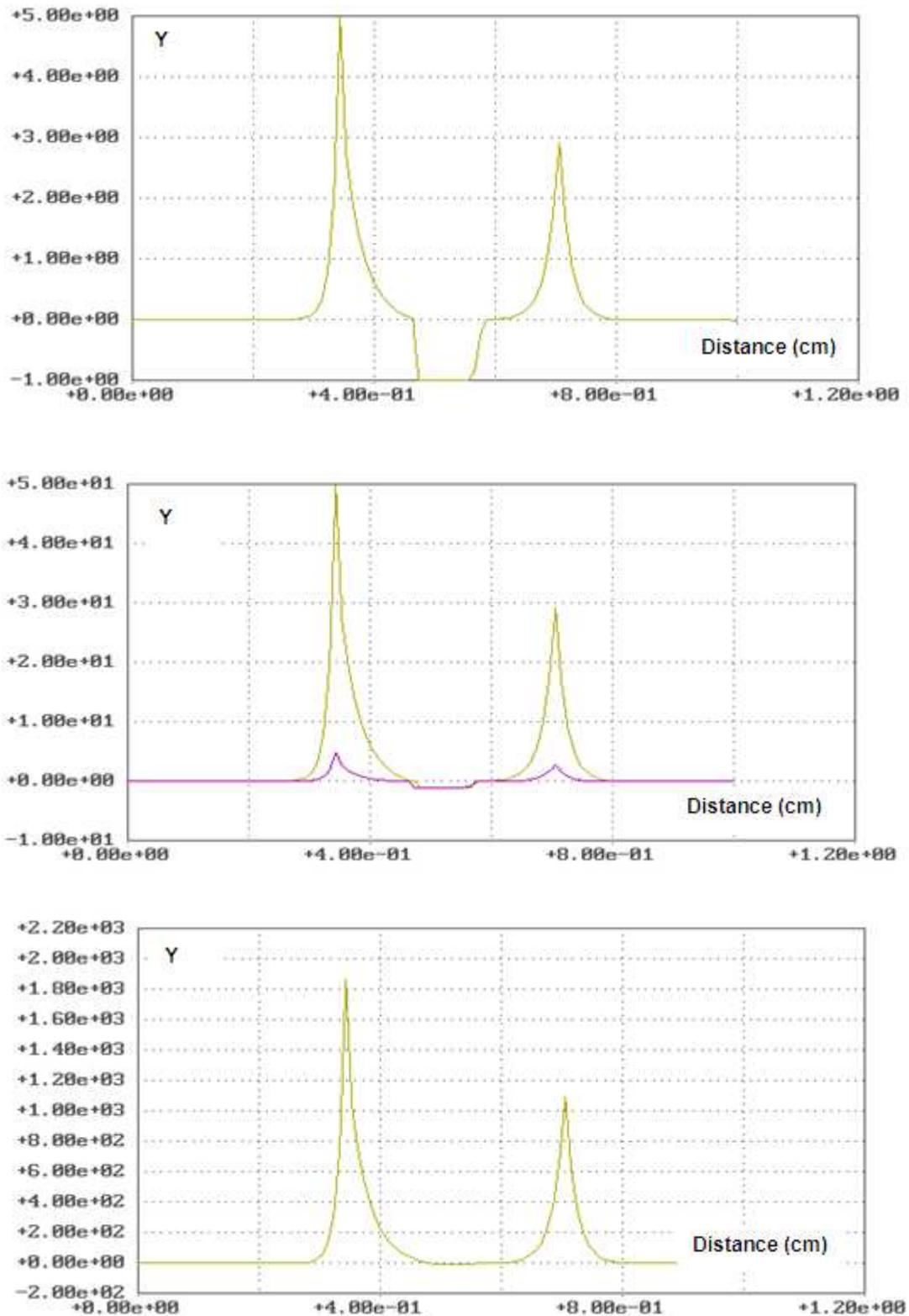


Рисунок 20. Поведение функции Y в тиристоре на участке запирания ($0 < j < j_s$). Верхний рисунок: $V = 2,7$ В, $j = 4 \cdot 10^{-13}$ А/см²; средний рисунок: $V = 15,2$ В, $j = 4 \cdot 10^{-12}$ А/см²; нижний рисунок соответствует точке включения: $V_s = 39,3$ В, $j_s = 1,5 \cdot 10^{-10}$ А/см².

Как мы уже отмечали, уровень обеднения ограничен значениями Y вблизи -1. Параллельно, с ростом тока происходит обогащение носителями заряда прямо

смещенных р-п-переходов. Если расстояния между р-п-переходами не очень велики, а точнее меньше или порядка диффузионной длины для неосновных носителей заряда, то варьируя уровни легирования областей прибора, можно добиться ситуации, когда носители заряда, инжектированные в область обеднения из прямо смещенных р-п-переходов, начнут обогащать ее еще при токах, меньших, чем ток пробоя р-п-перехода, расположенного в точке x_{j2} . В этом случае при $j > j_s$ величина $Y(x_{j2})$ начинает увеличиваться и падение напряжение на тиристоре уменьшается с ростом тока. Это соответствует участку ВАХ прибора с отрицательным дифференциальным сопротивлением. Поведение функции Y в тиристоре на участке отрицательного дифференциального сопротивления показано на Рисунке 21.

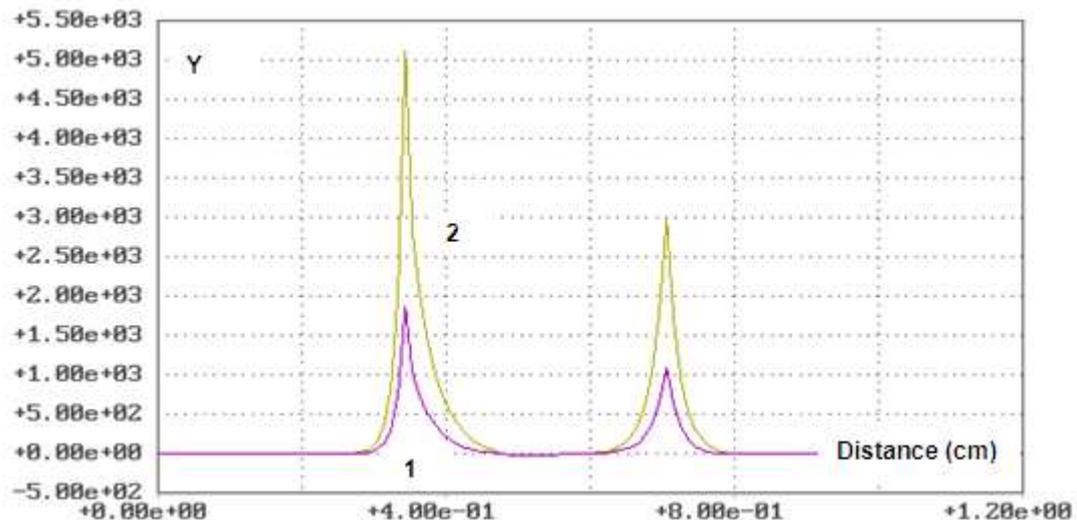


Рисунок 21. Поведение функции Y в тиристоре на участке отрицательного дифференциального сопротивления ($j_s < j < j_h$): 1 - $V_s = 39,3$ В, $j_s = 1,5 \cdot 10^{-10}$ А/см²; 2 - $V_h = 13,4$ В, $j_h = 4,3 \cdot 10^{-10}$ А/см².

Далее, рост плотности тока приводит к тому, что при $j > j_h$ величина $Y(x_{j2})$ практически перестает изменяться (см. Рисунок 22). На этом участке ВАХ рост величин $Y(x_{j1})$ и $Y(x_{j3})$ с ростом j обеспечивает, согласно (190), положительное дифференциальное сопротивление прибора.

Поведение ВАХ тиристора на обратной ветви менее интересно. Переходы, соответствующие точкам x_{j1} и x_{j3} , оказываются обратно смещенными. Носителей заряда, инжектируемых в области обеднения из прямо смещенного р-п-перехода в точке x_{j2} , оказывается недостаточно для обогащения этих областей. Поэтому

обратная ветвь ВАХ прибора аналогична соответствующему участку ВАХ обычного диода.

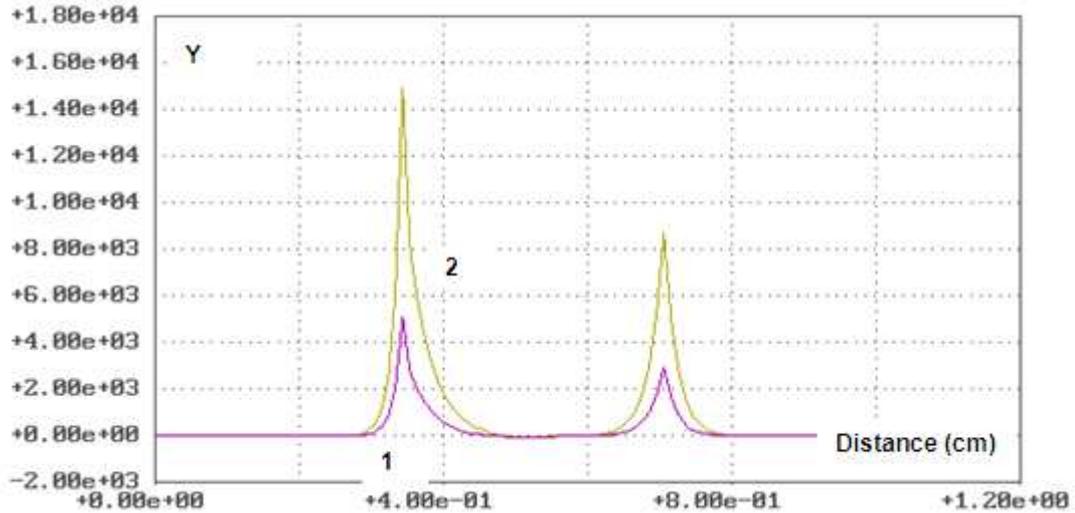


Рисунок 22. Поведение функции Y в тиристоре на участке отпирания ($j_h < j$): 1 - $V_h = 13,4$ В, $j_h = 4,3 \cdot 10^{-10}$ А/см²; 2 - $V = 14,9$ В, $j = 1,3 \cdot 10^{-09}$ А/см².

Формулы (193) кроме прочего описывают вклад пограничных эффектов в ВАХ тиристора. Если этим вкладом пренебречь и использовать грубые оценки

$$\exp\left\{-\int_{x_{j1}}^{x_{j2}} \lambda_0(\xi) d\xi\right\} \sim \exp\{-L_{21}/L_{\text{neq}21}\}, \quad \exp\left\{-\int_{x_{j2}}^{x_{j3}} \lambda_0(\xi) d\xi\right\} \sim \exp\{-L_{32}/L_{\text{neq}32}\},$$

где $L_{21} = x_{j2} - x_{j1}$, $L_{32} = x_{j3} - x_{j2}$, $L_{\text{neq}21}$, $L_{\text{neq}32}$ – средние длины неравновесности (релаксации) в областях между соответствующими p-n-переходами, с учетом (192) для $Y(x_{j2})$ при $j > 0$ получим

$$Y(x_{j2}) \sim -1 + j[(j_{s10}j_{s20})^{-1/2} \exp\{-L_{21}/L_{\text{neq}21}\} + (j_{s20}j_{s30})^{-1/2} \exp\{-L_{32}/L_{\text{neq}32}\}].$$

Это соотношение иллюстрирует приведенные выше рассуждения, поясняющие причины возникновения S-образной ВАХ у тиристора. В частности, из него следует требование на плотность тока насыщения среднего p-n-перехода

$$j_{s20} > j_{s10}j_{s30}/[(j_{s30})^{1/2} \exp\{-L_{21}/L_{\text{neq}21}\} + (j_{s10})^{1/2} \exp\{-L_{32}/L_{\text{neq}32}\}]^2. \quad (194)$$

Если неравенство (194) нарушается, то лавинный пробой среднего p-n-перехода происходит раньше, чем его обогащение неравновесными носителями заряда.

Если к областям полупроводника, расположенным в тиристоре между p-n-переходами, присоединить омические контакты, то характеристиками прибора можно управлять, изменяя плотность тока на каждом из p-n-переходов в

отдельности. Такие структуры описаны, например, в книгах [27,62-64]. При этом существенно, что и принцип действия тиристора и принципы управления этим прибором основаны на использовании неравновесных эффектов в полупроводниковой структуре.

3.4. Выводы

В первых трех главах диссертации проанализировано влияние неравновесных эффектов на функционирование полупроводниковых приборов, используемых, в частности, в качестве элементов электронных микросхем. Под неравновесными понимаются эффекты, обусловленные отклонением электронного и дырочного газов от состояния локального химического равновесия, характеризуемого равенством их химических потенциалов

$$F_n = F_p.$$

Температура электронов и дырок предполагается при этом одинаковой и равной температуре кристаллической решетки (окружающей среды). Класс явлений, обусловленных температурной неравновесностью, здесь не рассматривается.

Для анализа работы полупроводниковых приборов в условиях отклонения от химического равновесия уравнения дрейфово-диффузионной модели переноса заряда были переформулированы в новых переменных. Это позволило явным образом выделить уравнение для функции

$$Y = (np - n_i^2)/n_i^2 = \exp\{(F_n - F_p)/k_B T\} - 1,$$

являющейся количественной мерой указанной неравновесности. Причиной возникновения неравновесности является протекание тока в направлениях, перпендикулярных p-n-переходам.

Новый базис переменных позволил провести аналитическое исследование уравнений переноса заряда в полупроводниках и построить ряд приближенных решений для электростатического потенциала ϕ и для функции Y . Эти решения представляют существенный интерес, поскольку позволяют провести качественный анализ работы полупроводниковых приборов. Кроме того, они

могут быть использованы как начальное приближение для численного решения задач о переносе заряда в полупроводниковых структурах.

Показано, что в обычных условиях применения функционирование полупроводниковых диода, биполярного транзистора и тиристора определяются неравновесными свойствами носителей заряда. При этом, для понимания основных качественных особенностей работы этих приборов, достаточно исследовать решения уравнений переноса заряда в приближении квазиэлектронейтральности.

При разработке квантовых приборов, имеющих электрические характеристики, аналогичные характеристикам традиционных полупроводниковых элементов, логично попытаться найти и использовать в качестве основы функционирования физические эффекты, аналогичные неравновесности электронного и дырочного газов. Следует ожидать, что такие явления будут происходить в окрестностях областей, в которых резко изменяются свойства наноструктур, то есть вблизи контактов, гетерограниц, границ между нанообъектами с различными размерами и т.п.

ГЛАВА 4. Модель переноса заряда в мезоскопических структурах

Построение моделей, позволяющих описывать перенос заряда в структурах нанометровых размеров, является одной из актуальных задач современной математической физики [65-73]. На путях ее решения возникают трудности, обусловленные тем, что в областях с размерами порядка $1 \text{ нм} = 10^{-9} \text{ м} = 10^{-7} \text{ см}$ одинаково значимыми оказываются как квантовые, так и классические явления. Поэтому такие области часто называют мезоскопическими, промежуточными между областями, описываемыми либо чисто классическими, либо чисто квантовыми законами.

Как известно, классическая физика не описывает квантовых явлений. Квантовая теория считается более общей, включающей в себя классическую. Однако она содержит в себе классическое описание только как асимптотическое [60]. Общепринятой последовательной модели, описывающей с одинаковой достоверностью квантовые и классические явления, в настоящее время не существует.

Функционирование большинства электронных приборов, использующих наноструктуры в качестве активных областей, основано на квантовых эффектах [1-3,12-16,68,71,74-82]. По этой причине их часто называют квантовыми приборами. При расчете характеристик таких приборов квантовые явления, очевидно, необходимо учитывать.

Постановка граничных условий в виде приложенных к контактам напряжений или измеримых макроскопических токов (потоков), втекающих и вытекающих из контактов, подразумевает классическое описание. Конструктивно, квантовые приборы практически всегда содержат кроме нанометровых еще и области вполне макроскопических размеров. Причем границы между мезоскопическими и макроскопическими областями могут быть определены, как правило, лишь приблизительно. Строгое определение границ

мезоскопических областей и граничных условий на них возможно только в рамках последовательного описания явлений в таких областях.

Попытки построить чисто квантовое описание физических явлений в областях нанометровых размеров оказывается нереализуемыми на практике в связи со слишком большим количеством описываемых объектов (как правило, не менее десяти) и их возможных состояний [68-71]. При этом количество этих объектов кажется недостаточным для перехода к стандартному статистическому описанию.

4.1. Уравнения многофазной модели переноса заряда

Для анализа функционирования и расчета характеристик квантовых приборов используется модель, подробно описанная в книге [71]. Будем называть ее многофазной моделью переноса заряда.

Изначально, она была построена как феноменологическая, позволяющая рассчитывать характеристики резонансно-туннельных диодов (РТД) адекватно эксперименту, а также оптимизировать конструктивные параметры этих квантовых приборов [83-92]. Эта же модель была применена для анализа работы и расчета характеристик приборов на основе квантовых проводов [77-82,93-98], а также для объяснения некоторых необычных свойств наноструктур [99-102].

4.1.1. Интерпретируемые состояния носителей заряда

Положим, что одночастичное уравнение Шредингера [26,60] с соответствующими каждой конкретной задаче граничными и начальными условиями описывает интерпретируемые состояния носителей заряда для этой задачи

$$i\hbar\partial\Psi_0(t,\mathbf{r};\lambda)/\partial t + (\hbar^2/2)(\nabla,(m^{*-1}\nabla\Psi_0(t,\mathbf{r};\lambda)) - U(t,\mathbf{r})\Psi_0(t,\mathbf{r};\lambda) = 0, \quad (195)$$

В уравнении (195) $\Psi_0(t,\mathbf{r};\lambda)$ - волновая функция носителя заряда в состоянии, характеризуемом обобщенным индексом λ , который нумерует интерпретируемые

состояния носителей заряда и представляет собой набор параметров (квантовых чисел), соответствующих описываемому состоянию; m^* - эффективная масса носителя заряда; $U(\mathbf{t}, \mathbf{r})$ - детерминированный (не случайный) потенциал, с которым взаимодействуют носители заряда. Оператор U эрмитов

$$U(\mathbf{t}, \mathbf{r}) = (U(\mathbf{t}, \mathbf{r}))^\dagger.$$

Под интерпретируемыми будем понимать те состояния из пространства допустимых состояний решаемой задачи для уравнения (195), которым можно дать ясную физическую интерпретацию. Часто далее будем называть интерпретируемые состояния λ -состояниями.

Для каждого λ -состояния определим плотность вероятности $\rho(\mathbf{t}, \mathbf{r}; \lambda)$ и пространственную плотность потока вероятности $\mathbf{j}(\mathbf{t}, \mathbf{r}; \lambda)$ обнаружить частицу в λ -состоянии в момент времени t в точке \mathbf{r}

$$\begin{cases} \rho(\mathbf{t}, \mathbf{r}; \lambda) = \Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda) \Psi_0(\mathbf{t}, \mathbf{r}; \lambda), \\ \mathbf{j}(\mathbf{t}, \mathbf{r}; \lambda) = -(\i\hbar/2m^*)(\Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda) \nabla \Psi_0(\mathbf{t}, \mathbf{r}; \lambda) - (\nabla \Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda)) \Psi_0(\mathbf{t}, \mathbf{r}; \lambda)). \end{cases} \quad (196)$$

Как следует из уравнения (195), при каждом значении λ справедлив закон сохранения потока вероятности

$$\partial \rho(\mathbf{t}, \mathbf{r}; \lambda) / \partial t + (\nabla, \mathbf{j}(\mathbf{t}, \mathbf{r}; \lambda)) = 0. \quad (197)$$

4.1.2. Уравнение для детерминированного потенциала

Детерминированным будем называть взаимодействие, которое полностью определено: то есть либо известны аналитические выражения для соответствующих потенциалов или сил, либо сформулированы уравнения, из которых эти величины могут быть найдены.

Для моделирования переноса заряда в квантовых приборах, не использующих магнитные и спиновые эффекты, потенциал U можно представить в виде

$$U(\mathbf{t}, \mathbf{r}) = U^{\text{ext}}(\mathbf{r}) - e\phi(\mathbf{t}, \mathbf{r}), \quad (198)$$

где U^{ext} – внешний потенциал, задаваемый, например, возможным разрывом зон на гетерограницах, а φ – самосогласованный электростатический потенциал, который является решением уравнения Пуассона, справедливым при частотах, удовлетворяющих условию (2)

$$(\nabla, \varepsilon \nabla \varphi(t, \mathbf{r})) = e(n(t, \mathbf{r}) - N_{\text{int}}(\mathbf{r})). \quad (199)$$

Величина $n(t, \mathbf{r})$ является концентрацией подвижных носителей заряда, а $N_{\text{int}}(\mathbf{r})$ – концентрация неподвижных (встроенных) носителей заряда, которую мы будем полагать не зависящей от времени. Величину n считаем положительной для носителей отрицательного знака заряда (электронов), а N_{int} полагает положительной для встроенного положительного заряда.

Уравнения (198) и (199) определяют потенциал U при известных U^{ext} и N_{int} только если известно, каким образом можно вычислить концентрацию n .

4.1.3. Интерпретируемые состояния и фазы газа носителей заряда

Положим, что кроме детерминированного потенциала U носители заряда взаимодействуют еще и с некоторым случайным потенциалом

$$U^{\text{st}}(t, \mathbf{r}) = V(t, \mathbf{r}) \xi^{\text{st}}(t, \mathbf{r}). \quad (200)$$

К случайному будем относить взаимодействие, характеристики которого известны лишь частично, а по отношению к неизвестным характеристикам предполагается возможным применить какую-либо статистическую гипотезу. Оператор U^{st} будем считать эрмитовым. В выражении (200) $\xi^{\text{st}}(t, \mathbf{r})$ – случайная функция координат и/или времени, а $V(t, \mathbf{r})$ – детерминированный эрмитов оператор, включающий в себя вклады различных физических процессов, про которые мы не имеем информации, достаточной для детерминированного описания [103,104]. К таким процессам в твердом теле можно отнести рассеяние носителей заряда: на фононах, на примесных атомах, на неоднородностях границы, друг на друге и т.п.

Случайное взаимодействие является причиной случайных индуцированных переходов носителей заряда между интерпретируемыми состояниями. Будем считать, что эти переходы приводят к образованию в каждом λ -состоянии некоторой самостоятельной компоненты газа носителей заряда, характеризуемой своими химическим потенциалом $F(t, \mathbf{r}; \lambda)$ и температурой $T(t, \mathbf{r}; \lambda)$. То есть предполагается, что в присутствии случайного взаимодействия газ носителей заряда представляет собой многокомпонентную смесь.

Определим условия локального термодинамического равновесия между компонентами, соответствующими состояниям λ и λ'

$$F(t, \mathbf{r}; \lambda) = F(t, \mathbf{r}; \lambda'), \quad (201.1)$$

$$T(t, \mathbf{r}; \lambda) = T(t, \mathbf{r}; \lambda'), \quad (201.2)$$

которое достигается при равенстве их химических потенциалов и температур. Соотношения (201.1) определяют локальное химическое равновесие, а соотношения (201.2) – локальное температурное равновесие. Эти условия идентичны условиям равновесия термодинамических фаз [105,106]. Поэтому по аналогии с термодинамическими фазами компоненту, соответствующую λ -состоянию, будем называть λ -фазой газа носителей заряда.

В этой работе температуру будем считать одинаковой для всех λ -фаз, равной температуре окружающей среды T и не зависящей от времени и пространственных координат. Это требование означает глобальное температурное равновесие λ -фаз носителей заряда и их температурное равновесие с окружающей средой.

В каждой λ -фазе определим средние по случайному процессу плотности числа носителей заряда $n(t, \mathbf{r}; \lambda)$ и их пространственного потока $\mathbf{n}(t, \mathbf{r}; \lambda)$

$$n(t, \mathbf{r}; \lambda) = \rho(t, \mathbf{r}; \lambda) f(s(t, \mathbf{r}; \lambda)), \quad \mathbf{n}(t, \mathbf{r}; \lambda) = \mathbf{j}(t, \mathbf{r}; \lambda) f(s(t, \mathbf{r}; \lambda)). \quad (202)$$

Здесь $f(s)$ - функция распределения Ферми-Дирака

$$f(s) = 1/(\exp\{s\} + 1), \quad (203)$$

для аргумента, которой справедливо соотношение, аналогичное соотношению для аргумента функции распределения Ютнера, используемой в релятивистской кинетической теории [107]

$$s(\mathbf{t}, \mathbf{r}; \lambda) = \beta \{E(\mathbf{t}, \mathbf{r}; \lambda) - F(\mathbf{t}, \mathbf{r}; \lambda) - m^*(\mathbf{u}(\mathbf{t}, \mathbf{r}; \lambda), \mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda))\}. \quad (204)$$

В выражении (204)

$$\beta = (k_B T)^{-1} \quad (205)$$

- обратная температура,

$$\begin{cases} E(\mathbf{t}, \mathbf{r}; \lambda) = (i\hbar/2\rho(\mathbf{t}, \mathbf{r}; \lambda))(\Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda)\partial\Psi_0(\mathbf{t}, \mathbf{r}; \lambda)/\partial t - (\partial\Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda)/\partial t)\Psi_0(\mathbf{t}, \mathbf{r}; \lambda)), \\ \mathbf{u}(\mathbf{t}, \mathbf{r}; \lambda) = \mathbf{j}(\mathbf{t}, \mathbf{r}; \lambda)/\rho(\mathbf{t}, \mathbf{r}; \lambda) \end{cases} \quad (206)$$

- микроскопические энергия и скорость носителей заряда в λ -состоянии, $\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)$ – гидродинамическая скорость носителей заряда в λ -фазе. Отметим, что по определению скорость $\mathbf{u}(\mathbf{t}, \mathbf{r}; \lambda)$ является скоростью потока вероятности.

Суммированием выражений (202) по всем интерпретируемым состояниям (λ -фазам) получим макроскопические плотности числа частиц $n(\mathbf{t}, \mathbf{r})$ и пространственного потока $\mathbf{n}(\mathbf{t}, \mathbf{r})$

$$n(\mathbf{t}, \mathbf{r}) = \sum_{\lambda} n(\mathbf{t}, \mathbf{r}; \lambda), \quad \mathbf{n}(\mathbf{t}, \mathbf{r}) = \sum_{\lambda} \mathbf{n}(\mathbf{t}, \mathbf{r}; \lambda), \quad (207)$$

которые связаны с плотностями заряда и тока простыми формулам

$$Q(\mathbf{t}, \mathbf{r}) = -en(\mathbf{t}, \mathbf{r}), \quad \mathbf{J}(\mathbf{t}, \mathbf{r}) = -e\mathbf{n}(\mathbf{t}, \mathbf{r}). \quad (208)$$

4.1.4. Уравнение для гидродинамической скорости носителей заряда

Для вектора гидродинамической скорости $\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)$ постулируем уравнение, аналогичное используемому в теории плазмы полупроводников [36]

$$(\partial/\partial t + (\mathbf{u}(\mathbf{t}, \mathbf{r}; \lambda), \nabla))\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda) + \mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)/\tau(\lambda) = -\nabla F(\mathbf{t}, \mathbf{r}; \lambda)/m^*. \quad (209)$$

Оно связывает гидродинамическую скорость с градиентом химического потенциала, то есть с «термодинамической силой» [40,106].

Это параболическое уравнение, необратимое во времени, в котором положительный параметр $\tau(\lambda)$ определяет при $t > 0$ время релаксации решения начальной задачи Коши к устойчивому аттрактору

$$\mathbf{v}_0(\mathbf{t}, \mathbf{r}; \lambda) = -\tau(\lambda)\nabla F(\mathbf{t}, \mathbf{r}; \lambda)/m^*. \quad (210)$$

По своему физическому смыслу $\tau(\lambda)$ является временем релаксации импульса [26,66,103], которым в классической теории переноса заряда

определяется подвижность. Для корректного вычисления времени $\tau(\lambda)$ необходимо иметь математическую модель, описывающую взаимодействие носителей заряда со случайным потенциалом.

Если λ -состояния стационарны, а волновые функции $\Psi_0(\mathbf{r};\lambda)$ нормированы на одну частицу в некотором объеме и образуют полный набор, то для определения $\tau(\lambda)$ можно воспользоваться хорошо известным соотношением нестационарной теории возмущений [26,60,103]

$$(\tau(\lambda))^{-1} = (2\pi/\hbar) \sum_{\lambda'} \int d\mathbf{r} \Psi_0^*(\mathbf{r},\lambda) V(\mathbf{r}) \Psi_0(\mathbf{r},\lambda')^2 \delta(E(\lambda') - E(\lambda)). \quad (211)$$

4.1.5. Уравнения баланса числа носителей заряда

Для того чтобы многофазная модель переноса заряда была замкнутой, необходимо уметь вычислять химические потенциалы $F(t,\mathbf{r};\lambda)$. Случайное взаимодействие U^{st} приводит к переходам носителей заряда между λ -состояниями. То есть в его присутствии для плотностей числа частиц и плотностей потоков, определенных выражениями (202), вместо уравнений сохранения потока вероятности (197) должно выполняться уравнение баланса

$$\partial n(t,\mathbf{r};\lambda)/\partial t + (\nabla, \mathbf{n}(t,\mathbf{r};\lambda)) = \sum_{\lambda' \neq \lambda} R_{\lambda\lambda'}. \quad (212)$$

Величина $R_{\lambda\lambda'}$ - это интенсивность переходов частиц между λ и λ' - фазами. Поскольку случайное взаимодействие не порождает и не уничтожает носители заряда, для каждой двух λ и λ' - фаз должно выполняться условие

$$R_{\lambda\lambda'} = -R_{\lambda'\lambda}, \quad (213)$$

означающее, что изменение числа частиц в λ -фазе за счет их перехода/прихода в/из λ' -фазу(ы) равно соответствующему изменению числа частиц в λ' -фазе за счет их прихода/перехода из/в λ -фазу(ы). То есть величины $R_{\lambda\lambda'}$ антисимметричны по индексам λ и λ' . Из (213), очевидно, следует

$$\sum_{\lambda} \sum_{\lambda'} R_{\lambda\lambda'} = 0. \quad (214)$$

С учетом определения (207) из (212) и (214) получается закон сохранения суммарного потока

$$\partial n(\mathbf{t}, \mathbf{r}) / \partial t + (\nabla, \mathbf{n}(\mathbf{t}, \mathbf{r})) = 0. \quad (215)$$

Согласно законам термодинамики [35,104,106], система уравнений (212) должна, кроме прочего, при $t > 0$ описывать релаксацию λ -фаз к состоянию локального химического равновесия, определенного соотношениями (201.1). Когда оно выполнено для всех λ и λ' , все химические потенциалы равны между собой и равны некоторому равновесному значению $F(\mathbf{t}, \mathbf{r})$.

Если ограничиться простейшей линейной зависимостью интенсивности переходов между фазами от разности их химических потенциалов [106], то с учетом соотношений (202) – (204) и (212) – (214), требованию релаксации к состоянию локального химического равновесия, определяемому формулами (201.1), можно удовлетворить, полагая

$$R_{\lambda\lambda'} \sim -\gamma(F(\mathbf{t}, \mathbf{r}; \lambda) - F(\mathbf{t}, \mathbf{r}; \lambda')),$$

при условии

$$\gamma > 0.$$

Будем считать, что количество частиц, переходящих из λ -фазы в λ' -фазу в каждый момент времени в каждой точке пропорционально их количеству $n(\mathbf{t}, \mathbf{r}; \lambda)$ в λ -фазе в этот момент времени в этой точке. Аналогично, количество частиц переходящих из λ' -фазы в λ -фазу будем считать пропорциональным $n(\mathbf{t}, \mathbf{r}; \lambda')$. Тогда, суммируя сделанные предположения, можно представить $R_{\lambda\lambda'}$ в вид

$$R_{\lambda\lambda'} = -\hbar^{-1}(F(\mathbf{t}, \mathbf{r}; \lambda) - F(\mathbf{t}, \mathbf{r}; \lambda'))(n(\mathbf{t}, \mathbf{r}; \lambda)\Gamma_{\lambda\lambda'} + n(\mathbf{t}, \mathbf{r}; \lambda')\Gamma_{\lambda'\lambda}), \quad (216)$$

где безразмерные коэффициенты $\Gamma_{\lambda\lambda'}$ положительны

$$\Gamma_{\lambda\lambda'} > 0 \text{ при всех } \lambda, \lambda'. \quad (217)$$

Величины $\Gamma_{\lambda\lambda'}$ и $\Gamma_{\lambda'\lambda}$ определяют интенсивности переходов из λ -состояния в λ' -состояние и из λ' -состояния в λ -состояние в результате взаимодействия носителей заряда со случайным потенциалом U^{st} . Вычислить эти величины можно только при наличии математической модели, описывающей это взаимодействие.

Из уравнений (212), используя выражения (216) и (202), получим уравнения баланса числа носителей заряда в виде

$$\partial n(\mathbf{t}, \mathbf{r}; \lambda) / \partial t + (\nabla, \mathbf{n}(\mathbf{t}, \mathbf{r}; \lambda)) = -\hbar^{-1} \sum_{\lambda'} (F(\mathbf{t}, \mathbf{r}; \lambda) - F(\mathbf{t}, \mathbf{r}; \lambda')) (n(\mathbf{t}, \mathbf{r}; \lambda) \Gamma_{\lambda\lambda'} + n(\mathbf{t}, \mathbf{r}; \lambda') \Gamma_{\lambda'\lambda}). \quad (218)$$

Поскольку посредством формул (202), (203), (204) и (209) плотности $n(\mathbf{t}, \mathbf{r}; \lambda)$ и $\mathbf{n}(\mathbf{t}, \mathbf{r}; \lambda)$ связаны с химическим потенциалом $F(\mathbf{t}, \mathbf{r}; \lambda)$ и его градиентом $\nabla F(\mathbf{t}, \mathbf{r}; \lambda)$, уравнения (218) представляют собой систему уравнений для химических потенциалов λ -фаз газа носителей заряда.

4.1.6. Стационарные уравнения многофазной модели переноса заряда

Поскольку ниже мы ограничимся исследованием только стационарных явлений, обусловленных переносом заряда в мезоскопических структурах, то выпишем отдельно уравнения модели в этом случае. Детерминированный потенциал U не зависит от времени, а волновые функции для интерпретируемых λ -состояний представимы в виде

$$\Psi_0(\mathbf{t}, \mathbf{r}; \lambda) = \Psi_0(\mathbf{r}, \lambda) \exp\{-(i/\hbar)E(\lambda)t\}, \quad (219)$$

где энергии $E(\lambda)$ являются числами, не зависят от времени и координат. Кроме этого, будем полагать, что гидродинамическая скорость носителей заряда определена выражением (210). В результате, получим систему уравнений

$$\begin{cases} (\hbar^2/2)(\nabla, (m^*)^{-1} \nabla \Psi_0(\mathbf{r}, \lambda)) + (E(\lambda) - U(\mathbf{r}))\Psi_0(\mathbf{r}, \lambda) = 0, \\ (\nabla, \varepsilon \nabla \varphi(\mathbf{r})) = e(n(\mathbf{r}) - N_{\text{int}}(\mathbf{r})), \\ (\nabla, \mathbf{n}(\mathbf{r}, \lambda)) = -\hbar^{-1} \sum_{\lambda'} (F(\mathbf{r}, \lambda) - F(\mathbf{r}, \lambda'))(n(\mathbf{r}, \lambda) \Gamma_{\lambda \lambda'} + n(\mathbf{r}, \lambda') \Gamma_{\lambda' \lambda}), \end{cases} \quad (220)$$

где

$$\begin{cases} U(\mathbf{r}) = U^{\text{ext}}(\mathbf{r}) - e\varphi(\mathbf{r}), \\ n(\mathbf{r}) = \sum_{\lambda} n(\mathbf{r}, \lambda), \\ n(\mathbf{r}, \lambda) = \rho(\mathbf{r}, \lambda) f(s(\mathbf{r}, \lambda)), \quad \mathbf{n}(\mathbf{r}, \lambda) = \mathbf{j}(\mathbf{r}, \lambda) f(s(\mathbf{r}, \lambda)), \\ \rho(\mathbf{r}, \lambda) = \Psi_0^*(\mathbf{r}, \lambda) \Psi_0(\mathbf{r}, \lambda), \quad \mathbf{j}(\mathbf{r}, \lambda) = -(i\hbar/2m^*) (\Psi_0^*(\mathbf{r}, \lambda) \nabla \Psi_0(\mathbf{r}, \lambda) - (\nabla \Psi_0^*(\mathbf{r}, \lambda)) \Psi_0(\mathbf{r}, \lambda)), \\ f(s) = 1/(\exp\{s\} + 1), \\ s(\mathbf{r}, \lambda) = \beta \{E(\lambda) - F(\mathbf{r}, \lambda) + \tau(\lambda)(\mathbf{u}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda))\}, \\ \mathbf{u}(\mathbf{r}, \lambda) = \mathbf{j}(\mathbf{r}, \lambda) / \rho(\mathbf{r}, \lambda). \end{cases} \quad (221)$$

Для макроскопических плотностей потока, тока и заряда справедливы выражения (207) и (208)

$$n(\mathbf{r}) = \sum_{\lambda} n(\mathbf{r}, \lambda), \mathbf{J}(\mathbf{r}) = -en(\mathbf{r}), Q(\mathbf{r}) = -en(\mathbf{r}).$$

4.1.7. Длина релаксации к локальному химическому равновесию

Уравнения баланса числа частиц в системе (220) описывают пространственное распределение химических потенциалов λ -фаз. Подстановка выражений для плотности потока и функции распределения из (221) в последние уравнения системы (220) приводит к уравнениям баланса в виде

$$\beta \rho(\mathbf{r}, \lambda) f'(s(\mathbf{r}, \lambda)) (\mathbf{u}(\mathbf{r}, \lambda), \nabla(\tau(\lambda)(\mathbf{u}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda))) - \nabla F(\mathbf{r}, \lambda)) = \\ -\hbar^{-1} \sum_{\lambda'} (\mathbf{F}(\mathbf{r}, \lambda) - \mathbf{F}(\mathbf{r}, \lambda')) (n(\mathbf{r}, \lambda) \Gamma_{\lambda\lambda'} + n(\mathbf{r}, \lambda') \Gamma_{\lambda'\lambda}), \quad (222)$$

где

$$f'(s) = \partial f(s) / \partial s < 0.$$

Из уравнений (222) следует, что роль диффузионной длины для химических потенциалов играет величина

$$L_{\text{rel}} = (\beta u^2(\mathbf{r}, \lambda) \tau(\lambda) \hbar)^{1/2} = L_0 (m^* u^2(\mathbf{r}, \lambda) / k_B T)^{1/2}, \quad (223)$$

которую мы будем называть длиной релаксации к локальному химическому равновесию или просто длиной релаксации химических потенциалов. Уравнения (222) могут иметь для $F(\mathbf{r}, \lambda)$ как экспоненциально растущие, так и экспоненциально затухающие по координатам решения. Однако если существует некоторая пространственная граница, на которой в качестве граничного условия принято локальное химическое равновесие, определенное соотношениями (215), то уравнение (222) описывает пространственную релаксацию (затухание) отклонений от этого равновесия на длинах порядка L_{rel} .

Физическая целесообразность постановки указанных граничных условий следует из уравнения (218), которое в пространственно однородной ситуации однозначно описывает релаксацию системы носителей заряда к состоянию локального химического равновесия за характерный промежуток времени

$$\tau_0 = \hbar / k_B T. \quad (224)$$

Длина L_0 , которой определяется значение L_{rel} , вычисляется по формуле

$$L_0 = (\hbar\tau(\lambda)/m^*)^{1/2}. \quad (225)$$

Если в качестве $\tau(\lambda)$ брать значения усредненных времен релаксации импульса для объемных материалов, а в качестве m^* – соответствующие эффективные массы, то для электронов получаются значения, приведенные в Таблице 1.

Для сравнения в таблице приведены длины размерного квантования

$$L_{dq} = (3\pi^2\hbar^2/2m^*k_B T)^{1/2} = (3\pi^2/2)^{1/2}\lambda_T \quad (226)$$

в этих же материалах. По определению L_{dq} - это пороговая величина, при которой расстояние между первыми двумя уровнями энергии частиц становится равным энергии тепловых флуктуаций $k_B T$. При размерах меньших L_{dq} расстояние между уровнями энергии электронов становится большим $k_B T$. Мезоскопическими следует считать структуры, размеры которых сопоставимы с L_{dq} .

Таблица 1.

Характерная длина релаксации химических потенциалов и длина размерного квантования для электронов при $T = 300^0K$

Материал	Длина релаксации химических потенциалов электронов L_0 (нм)	Длина размерного квантования для электронов L_{dq} (нм)
Si	10	15
GaAs	23	25
InSb	71	58
Металлы	~ 2	~ 6

Величина

$$\lambda_T = (\hbar^2/m^*k_B T)^{1/2}$$

- это длина температурной когерентности электронов [73,75,103]. Она совпадает с L_0 , если в качестве $\tau(\lambda)$ в выражение (225) подставить τ_0 . Длиной температурной когерентности определяется характерное расстояние, на котором случайное взаимодействие разрушает фазовую когерентность электронных волн. Согласно (226)

$$L_{dq} > \lambda_T.$$

То есть эффекты, обусловленные размерным квантованием, начинают проявляться при существенно больших размерах структур, чем явления, связанные с интерференцией электронных волн.

Напротив, из сравнения величин L_0 и L_{dq} в различных материалах следует, что длина релаксации химических потенциалов близка к длине размерного квантования, и учет физических эффектов, обусловленных релаксацией, может быть принципиально важным для описания переноса заряда в мезоскопических структурах. Квантовые и неравновесные эффекты в твердотельных электронных приборах начинают проявляться при одних и тех же характерных размерах.

При анализе явлений переноса, характеризующихся макроскопическими масштабами длин, много большими L_0 , различия между λ -фазами газа носителей заряда исчезают и можно считать, что все они находятся в состоянии локального химического равновесия и характеризуются единым химическим потенциалом $F(\mathbf{r})$.

4.1.8. Закон Ома и формула Цу - Есаки

Рассмотрим выражение для плотности потока в пределе малых $\tau(\lambda)$. Будем считать, что справедливо неравенство

$$|F(\mathbf{r}, \lambda)| \gg |\tau(\lambda)(\mathbf{u}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda))|, \quad (227)$$

позволяющее представить функцию распределения в виде суммы:

$$f(s_0 + s_1) \approx f(s_0) + s_1 f'(s_0),$$

где

$$s_0(\mathbf{r}, \lambda) = \beta(E(\lambda) - F(\mathbf{r}, \lambda)), \quad s_1(\mathbf{r}, \lambda) = \beta\tau(\lambda)(\mathbf{u}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda)).$$

Откуда получим:

$$f(s(\mathbf{r}, \lambda)) \approx f(s_0(\mathbf{r}, \lambda)) + \beta\tau(\lambda)(\mathbf{u}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda))f'(s_0(\mathbf{r}, \lambda)). \quad (228)$$

Физически неравенство (227) означает, что характерная длина изменения химического потенциала

$$L_F = |F|/|\nabla F|$$

много больше, чем длина свободного пробега частицы

$$L_{sc} = |\tau u|,$$

то есть (227) можно переписать в виде

$$L_F(\mathbf{r}, \lambda) \gg L_{sc}(\mathbf{r}, \lambda).$$

В этом случае функция распределения приближенно представима в виде (228), который внешне напоминает стандартный результат классической теории переноса заряда [26]. Однако важным отличием, приводящим к серьезным физическим следствиям, является зависимость химического потенциала от λ . Для суммарного потока из (207) и (228) получим

$$n(\mathbf{r}) \approx \sum_{\lambda} \mathbf{j}(\mathbf{r}, \lambda) f(s_0(\mathbf{r}, \lambda)) + \beta \sum_{\lambda} \tau(\lambda) \mathbf{j}(\mathbf{r}, \lambda) (\mathbf{u}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda)) f'(s_0(\mathbf{r}, \lambda)). \quad (229)$$

Функция распределения $f(s_0)$ и ее производная $f'(s_0)$ зависят от энергии λ -состояния, химического потенциала и температуры. Для большинства прикладных задач справедливо утверждение, что для всякого λ найдется такое λ_1 , что

$$\mathbf{j}(\mathbf{r}, \lambda) = -\mathbf{j}(\mathbf{r}, \lambda_1), \text{ при } E(\lambda) = E(\lambda_1). \quad (230)$$

Это условие выполняется, например, для интерпретируемых λ -состояний в виде плоских волн, соответствующих одинаковым по абсолютному значению, но противоположным по направлению импульсам. При выполнении условия (230) суммирование в выражении для $n(\mathbf{r}, \lambda)$ по λ будет приводить к возникновению следующих комбинаций

$$\begin{aligned} \mathbf{j}(\mathbf{r}, \lambda) f(s(\mathbf{r}, \lambda)) + \mathbf{j}(\mathbf{r}, \lambda_1) f(s(\mathbf{r}, \lambda_1)) &= \mathbf{j}(\mathbf{r}, \lambda) (f(s(\mathbf{r}, \lambda)) - f(s(\mathbf{r}, \lambda_1))) \approx \\ &\approx \mathbf{j}(\mathbf{r}, \lambda) \{ (f(s_0(\mathbf{r}, \lambda)) - f(s_0(\mathbf{r}, \lambda_1))) + \\ &+ \tau(\lambda) \beta (\mathbf{j}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda)) f'(s_0(\mathbf{r}, \lambda)) / \rho(\mathbf{r}, \lambda) + \\ &+ \tau(\lambda_1) \beta (\mathbf{j}(\mathbf{r}, \lambda), \nabla F(\mathbf{r}, \lambda_1)) f'(s_0(\mathbf{r}, \lambda_1)) / \rho(\mathbf{r}, \lambda_1) \}. \end{aligned} \quad (231)$$

Первое слагаемое в выражении (231) линейно по микроскопическому потоку $\mathbf{j}(\mathbf{r}, \lambda)$, а второе – квадратично.

В пределе

$$\tau(\lambda) = 0$$

с учетом соотношения (231) для плотности потока $\mathbf{n}(\mathbf{r}, \lambda)$ получим

$$\mathbf{n}(\mathbf{r}, \lambda)|_{\tau(\lambda)=0} \approx \sum_{\lambda} \mathbf{j}(\mathbf{r}, \lambda) \{f(s_0(\mathbf{r}, E(\lambda), F_1(\mathbf{r}, \lambda))) - f(s_0(\mathbf{r}, E(\lambda), F_r(\mathbf{r}, \lambda)))\}. \quad (232)$$

Здесь мы заменили суммирование по λ_1 эквивалентным суммированием по λ , и обозначили

$$F_1(\mathbf{r}, \lambda) = F(\mathbf{r}, \lambda), F_r(\mathbf{r}, \lambda) = F(\mathbf{r}, \lambda_1)$$

Функции распределения, стоящие в фигурной скобке в формуле (232) отличаются только значением химического потенциала. При

$$F_1(\mathbf{r}, \lambda) = F_r(\mathbf{r}, \lambda)$$

плотность потока обращается в нуль. Таким образом, выражение (232) описывает поток носителей заряда, порождаемый локальной разностью химических потенциалов между их фазами.

Если рассматривать мезоскопическую структуру, имеющую два контакта, и полагать, что есть две фазы носителей заряда, различающихся по признаку попадания носителей заряда в структуру из одного или другого контакта, то из выражения (232) получается известная формула Цу – Есаки [14,65], которую обычно используют для вычисления туннельного тока [12-16,67]. При туннелировании плотность микроскопического потока \mathbf{j} пропорциональна коэффициенту прохождения частиц через потенциальный рельеф.

В случае, когда различием химических потенциалов λ -фаз можно пренебречь, например, для структур макроскопических размеров, из (229) для плотности потока с учетом формулы (231) получается выражение

$$\mathbf{n}(\mathbf{r}) \approx -\mu(\mathbf{r})\mathbf{n}(\mathbf{r})\nabla F(\mathbf{r})/e, \quad (233)$$

которое представляет собой закон Ома в дифференциальной форме. Для изотропных структур подвижность μ вычисляется по формуле

$$\mu(\mathbf{r}) = -2e\beta \sum_{\lambda} \tau(\lambda) \rho(\mathbf{r}, \lambda) u^2(\mathbf{r}, \lambda) f'(s_0(\mathbf{r}, \lambda)) / \mathbf{n}(\mathbf{r}), \quad (234)$$

где

$$u^2(\mathbf{r}, \lambda) = (\mathbf{u}(\mathbf{r}, \lambda), \mathbf{u}(\mathbf{r}, \lambda)).$$

Стандартное выражение для μ (см. [26]) получается из (234), если в качестве λ -состояний выбраны плоские волны. Такой выбор соответствует случаю однородной структуры, когда потенциал U постоянен или изменяется настолько плавно, что применимо квазиклассическое приближение [59,60].

Таким образом, соотношение многофазной модели для плотности потока (тока) носителей заряда позволяет учитывать как квантовые, так и классические явления. Его можно рассматривать и как обобщение формулы Цу – Есаки [65,14] для туннельного тока на случай наличия омического сопротивления, и как обобщение обычного закона Ома в дифференциальной форме [26] на случай не квазиклассических потенциальных рельефов.

4.2. Об обосновании многофазной модели переноса заряда

4.2.1. Квантовая динамика заряженных частиц, взаимодействующих со случайным потенциалом

Уравнению Шредингера (195), которое используется в многофазной модели для описания интерпретируемых состояний носителей заряда, соответствует плотность лагранжиана

$$\begin{cases} L_0(t, \mathbf{r}) = L_\Psi(t, \mathbf{r}) + L^{\text{int}}(t, \mathbf{r}), \\ L_\Psi(t, \mathbf{r}) = (i\hbar/2)(\Psi^*(t, \mathbf{r})(\partial\Psi(t, \mathbf{r})/\partial t) - (\partial\Psi^*(t, \mathbf{r})/\partial t)\Psi(t, \mathbf{r})) - (\hbar^2/2m^*)(\nabla\Psi^*(t, \mathbf{r}), \nabla\Psi(t, \mathbf{r})) \\ L^{\text{int}}(t, \mathbf{r}) = -\Psi^*(t, \mathbf{r})U(t, \mathbf{r})\Psi(t, \mathbf{r}). \end{cases} \quad (235)$$

Для каждой конкретной задачи уравнение (195) с соответствующими граничными или (и) начальными условиями имеет некоторый набор решений $\{\Psi_0(t, \mathbf{r}; \lambda)\}$, которым можно дать ту или иную физическую интерпретацию. Обобщенный индекс λ представляет собой набор параметров (квантовых чисел), характеризующих интерпретируемые состояния, и нумерует их. Количество интерпретируемых или, как мы их называем, λ -состояний, обозначим через K . Поскольку все реальные физические системы пространственно ограничены, K является целым числом. При рассмотрении «очень больших» систем, которые

Существенно отметить, что в отличие от уравнения (195), которое абсолютно идентично при всех λ , то есть оно одно описывает весь набор из K интерпретируемых состояний, уравнения (238) при различных λ представляют собой, вообще говоря, различные уравнения.

При предположении

$$|U^{st}(t, \mathbf{r})| \ll |U(t, \mathbf{r})|, \quad (239)$$

из (238) следует приближенно уравнение (195), описывающее интерпретируемые λ -состояния. Кроме того, если ввести суммарную волновую функцию

$$\Psi(t, \mathbf{r}) = \sum_{\lambda} \Psi(t, \mathbf{r}; \lambda), \quad (240)$$

то для нее получим уравнение Шредингера

$$i\hbar \partial \Psi(t, \mathbf{r}) / \partial t + (\hbar^2 / 2m^*) (\nabla \Psi(t, \mathbf{r})) - (U(t, \mathbf{r}) + U^{st}(t, \mathbf{r})) \Psi(t, \mathbf{r}) = 0. \quad (241)$$

Это уравнение также приближенно переходит в уравнение (195) при выполнении условия (239).

Обычная процедура введения случайного взаимодействия в динамические уравнения состоит в добавлении к детерминированной силе или потенциалу, фигурирующим в этих уравнениях, соответствующих случайных силы или потенциала [103,104,109]. В случае квантовой динамики, описываемой уравнением Шредингера, такой подход привел бы к уравнению (241). Мы поступили иначе, предположив, что случайное взаимодействие приводит к «расщеплению» одного уравнения на систему из K , вообще говоря, различных уравнений.

Для каждого из полей $\Psi(t, \mathbf{r}; \lambda)$ определим плотность вероятности и плотность потока вероятности

$$\begin{cases} \sigma(t, \mathbf{r}; \lambda) = \Psi^*(t, \mathbf{r}; \lambda) \Psi(t, \mathbf{r}; \lambda), \\ \mathbf{i}(t, \mathbf{r}; \lambda) = -(i\hbar / 2m^*) \{ \Psi^*(t, \mathbf{r}; \lambda) (\nabla \Psi(t, \mathbf{r}; \lambda)) - (\nabla \Psi^*(t, \mathbf{r}; \lambda)) \Psi(t, \mathbf{r}; \lambda) \}. \end{cases} \quad (242)$$

Из уравнения (238) получим уравнение баланса

$$\begin{aligned} \partial \sigma(t, \mathbf{r}; \lambda) / \partial t + (\nabla \cdot \mathbf{i}(t, \mathbf{r}; \lambda)) = \\ = -(i/\hbar(K-1)) \sum_{\lambda'} (\Psi^*(t, \mathbf{r}; \lambda) U^{st}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda') - \Psi^*(t, \mathbf{r}; \lambda') U^{st}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda)). \end{aligned} \quad (243)$$

Определим суммарные плотности вероятности и потоков вероятности

$$\sigma(\mathbf{t}, \mathbf{r}) = \sum_{\lambda} \sigma(\mathbf{t}, \mathbf{r}; \lambda), \quad \mathbf{i}(\mathbf{t}, \mathbf{r}) = \sum_{\lambda} \mathbf{i}(\mathbf{t}, \mathbf{r}; \lambda). \quad (244)$$

Суммируя (243) по λ , получим закон сохранения суммарного потока вероятности обнаружить какое-либо из полей $\Psi(\mathbf{t}, \mathbf{r}; \lambda)$ в точке \mathbf{r} в момент времени t

$$\partial \sigma(\mathbf{t}, \mathbf{r}) / \partial t + (\nabla, \mathbf{i}(\mathbf{t}, \mathbf{r})) = 0. \quad (245)$$

Выражение (243) являются исходным для получения уравнений баланса числа частиц многофазной модели переноса заряда (212) и (218). Правая часть формулы (243), очевидно, удовлетворяет условиям (213) и (214). Следует, однако, иметь в виду, что в уравнении (243) фигурируют случайные плотности потока вероятности и вероятности, случайные волновые функции и случайный потенциал, а в уравнениях (212) и (218) все величины детерминированы.

4.2.2. Термодинамика заряженных частиц, взаимодействующих со случайным потенциалом

Будем полагать, что измеряемые, то есть макроскопические величины плотности заряда и тока получаются из соответствующих случайных плотностей вероятности и потока вероятности после их усреднения по случайному процессу, суммирования по всем интерпретируемым состояниям и умножения на единичный заряд. В качестве последнего примем для определенности заряд электрона $-e$. В результате получим соотношения для вычисления макроскопических плотностей заряда и тока

$$\begin{cases} Q(\mathbf{t}, \mathbf{r}) = -en(\mathbf{t}, \mathbf{r}), \quad \mathbf{J}(\mathbf{t}, \mathbf{r}) = -e\mathbf{n}(\mathbf{t}, \mathbf{r}), \\ n(\mathbf{t}, \mathbf{r}) = \sum_{\lambda} n(\mathbf{t}, \mathbf{r}; \lambda), \quad \mathbf{n}(\mathbf{t}, \mathbf{r}) = \sum_{\lambda} \mathbf{n}(\mathbf{t}, \mathbf{r}; \lambda), \\ n(\mathbf{t}, \mathbf{r}; \lambda) = \langle \sigma(\mathbf{t}, \mathbf{r}; \lambda) \rangle, \quad \mathbf{n}(\mathbf{t}, \mathbf{r}; \lambda) = \langle \mathbf{i}(\mathbf{t}, \mathbf{r}; \lambda) \rangle. \end{cases} \quad (246)$$

Здесь символ $\langle \dots \rangle$ означает усреднение по случайному процессу величины, взятой в «скобки».

Будем считать, что в среднем случайное взаимодействие не изменяет самих интерпретируемых состояний. В том числе, оно не изменяет и характеристик интерпретируемых состояний (квантовых чисел), входящих в обобщенный индекс λ . Случайное взаимодействие порождает переходы между этими состояниями.

Переходы приводят к образованию в каждом из интерпретируемых состояний компоненты (λ -фазы) газа носителей заряда, подчиняющейся законам термодинамики. Постараемся формализовать перечисленные свойства системы носителей заряда в присутствии случайного взаимодействия.

Для описания каждой из λ -фаз будем использовать термодинамические переменные:

- $F(\mathbf{t}, \mathbf{r}; \lambda)$ – химический потенциал;
- $T(\mathbf{t}, \mathbf{r}; \lambda)$ – температуру;

и, дополнительно:

- $\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)$ – гидродинамическую скорость (скорость λ -фазы газа носителей заряда как целого).

Скорость $\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)$ связана с химическим потенциалом и температурой λ -фазы уравнением [36,40,106]

$$(\partial/\partial t + (\mathbf{u}(\mathbf{t}, \mathbf{r}; \lambda), \nabla))\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda) + \mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)/\tau(\mathbf{t}, \mathbf{r}; \lambda) = -(\nabla F(\mathbf{t}, \mathbf{r}; \lambda) + s(\mathbf{t}, \mathbf{r}; \lambda)k_B \nabla T(\mathbf{t}, \mathbf{r}; \lambda))/m^*, \quad (247)$$

в котором градиенты последних играют роль гидродинамических сил. Функцию $s(\mathbf{t}, \mathbf{r}; \lambda)$ определим позже.

Кроме того, мы будем оперировать микроскопическими характеристиками λ -состояний, которые вычисляются с использованием волновых функций $\Psi_0(\mathbf{t}, \mathbf{r}; \lambda)$, являющихся решениями уравнения (238) при $U^{st} = 0$ или, что то же самое, уравнения (195).

Потребуем, чтобы случайное взаимодействие в среднем не изменяло микроскопическую энергию

$$E(\mathbf{t}, \mathbf{r}; \lambda) = (i\hbar/2) \langle (\Psi^*(\mathbf{t}, \mathbf{r}; \lambda) \partial \Psi(\mathbf{t}, \mathbf{r}; \lambda) / \partial t - (\partial \Psi^*(\mathbf{t}, \mathbf{r}; \lambda) / \partial t) \Psi_0(\mathbf{t}, \mathbf{r}; \lambda)) \rangle / \langle \sigma(\mathbf{t}, \mathbf{r}; \lambda) \rangle = \\ = (i\hbar/2) (\Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda) \partial \Psi_0(\mathbf{t}, \mathbf{r}; \lambda) / \partial t - (\partial \Psi_0^*(\mathbf{t}, \mathbf{r}; \lambda) / \partial t) \Psi_0(\mathbf{t}, \mathbf{r}; \lambda)) / \rho(\mathbf{t}, \mathbf{r}; \lambda), \quad (248)$$

и скорость

$$\mathbf{u}(\mathbf{t}, \mathbf{r}; \lambda) = \langle \mathbf{i}(\mathbf{t}, \mathbf{r}; \lambda) \rangle / \langle \sigma(\mathbf{t}, \mathbf{r}; \lambda) \rangle = \mathbf{j}(\mathbf{t}, \mathbf{r}; \lambda) / \rho(\mathbf{t}, \mathbf{r}; \lambda) \quad (249)$$

носителей заряда в интерпретируемых состояниях.

Положим, что рассматриваемые нами носители заряда представляют собой частицы (квазичастицы), то есть неделимые на части объекты, имеющие заряд $-e$ и эффективную массу m^* . Будем считать, что уравнение (195) описывает

состояния именно этих частиц. Неделимость частиц проявляется в том, что в каждом из интерпретируемых состояний может находиться только их целое количество k . При отсутствии частиц в λ -состоянии их количество в нем полагается равным нулю ($k=0$). Возможное появление в формализме описания нецелого количества числа частиц связано с тем, что нахождение частиц в фиксированных λ -состояниях и фиксированных точках пространства-времени носит вероятностный характер.

Определим вероятность нахождения k частиц в момент времени t в точке \mathbf{r} в λ -состоянии как

$$w_k = \exp\{\beta(t, \mathbf{r}; \lambda)\Omega(t, \mathbf{r}; \lambda) - ks(t, \mathbf{r}; \lambda)\}, \quad (250)$$

где

$$\beta(t, \mathbf{r}; \lambda) = (k_B T(t, \mathbf{r}; \lambda))^{-1},$$

а $\Omega(t, \mathbf{r}; \lambda)$ – потенциал Гиббса λ -фазы газа носителей заряда. Заметим, что согласно (250) функция $s(t, \mathbf{r}; \lambda)$ является аддитивной характеристикой частиц.

Поскольку какое-то количество частиц в момент времени t в точке \mathbf{r} в λ -состоянии достоверно находится, то суммирование выражения (250) по всем возможным значениям k должно дать единицу. В результате для потенциала Гиббса получим

$$\Omega(t, \mathbf{r}; \lambda) = -k_B T(t, \mathbf{r}; \lambda) \ln \sum_k (\exp\{-s(t, \mathbf{r}; \lambda)\})^k. \quad (251)$$

Этот потенциал должен удовлетворять термодинамическому тождеству [105]

$$d\Omega(t, \mathbf{r}; \lambda) = -S(t, \mathbf{r}; \lambda)dT(t, \mathbf{r}; \lambda) - N(t, \mathbf{r}; \lambda)dF(t, \mathbf{r}; \lambda), \quad (252)$$

связывающему между собой потенциал Гиббса Ω , энтропию S , температуру T , среднее число частиц N и химический потенциал F . Существенно, что среднее число частиц $N(t, \mathbf{r}; \lambda)$ может отличаться от их допустимого целого количества k в λ -состоянии в момент времени t в точке \mathbf{r} , и быть нецелым. На основании тождества (252) получим

$$N(t, \mathbf{r}; \lambda) = -d\Omega(t, \mathbf{r}; \lambda)/dF(t, \mathbf{r}; \lambda). \quad (253)$$

В квантовой теории используется [105,107] два типа статистики частиц: Ферми-Дирака и Бозе-Эйнштейна. В первом случае количество частиц, называемых фермионами, в каждом состоянии может быть либо 0, либо 1. Во

втором случае количество частиц, называемых бозонами, в каждом состоянии может быть любым целым числом: 0,1,2,3 ... Производя суммирование в (251), получим

$$\Omega(t, \mathbf{r}; \lambda) = \begin{cases} -\beta(t, \mathbf{r}; \lambda)^{-1} \ln(1 + \exp\{-s(t, \mathbf{r}; \lambda)\}) - \text{для фермионов,} \\ \beta(t, \mathbf{r}; \lambda)^{-1} \ln(1 - \exp\{-s(t, \mathbf{r}; \lambda)\}) - \text{для бозонов.} \end{cases} \quad (254)$$

При этом существенно, что выражение для бозонов получено в предположении $s(t, \mathbf{r}; \lambda) > 0$. С учетом соотношений (253) и (254) найдем

$$N(t, \mathbf{r}; \lambda) = \begin{cases} (\exp\{s(t, \mathbf{r}; \lambda)\} + 1)^{-1} - \text{для фермионов,} \\ (\exp\{s(t, \mathbf{r}; \lambda)\} - 1)^{-1} - \text{для бозонов.} \end{cases} \quad (255)$$

Для того, чтобы применить полученный результат к моделированию переноса заряда необходимо связать величину $N(t, \mathbf{r}; \lambda)$ со средней плотностью вероятности обнаружить частицу в λ -состоянии в точке x . По определению будем считать, что эта связь дается соотношением

$$N(t, \mathbf{r}; \lambda) = \langle \sigma(t, \mathbf{r}; \lambda) \rangle / \rho(t, \mathbf{r}; \lambda). \quad (256)$$

То есть мы полагаем, что среднее по случайному процессу число носителей заряда в λ -состоянии в точке x равно средней плотности вероятности обнаружить носитель заряда в этом состоянии в этой точке, выраженной в единицах этой плотности вероятности в отсутствие случайного взаимодействия. Такое определение имеет смысл только, если выполнено условие (236).

На основании формул (246) и (256) имеем

$$n(t, \mathbf{r}; \lambda) = \rho(t, \mathbf{r}; \lambda) N(t, \mathbf{r}; \lambda). \quad (257)$$

Это выражение позволяет трактовать $N(t, \mathbf{r}; \lambda)$ как функцию распределения

$$f(s(t, \mathbf{r}; \lambda)) = N(t, \mathbf{r}; \lambda). \quad (258)$$

В результате, используя (249), (256) и (258), получим

$$n(t, \mathbf{r}; \lambda) = \rho(t, \mathbf{r}; \lambda) f(s(t, \mathbf{r}; \lambda)), \quad \mathbf{n}(t, \mathbf{r}; \lambda) = \mathbf{j}(t, \mathbf{r}; \lambda) f(s(t, \mathbf{r}; \lambda)), \quad (259)$$

где функция $f(s(t, \mathbf{r}; \lambda))$ определена соотношениями (255). Поскольку, как правило, переносчиками заряда в мезоскопических структурах являются электроны, то есть

фермионы, то далее для них будет использоваться соответствующее выражение из формулы (255).

Аддитивную по отношению к числу частиц величину $s(t, \mathbf{r}; \lambda)$ определим выражением

$$s(t, \mathbf{r}; \lambda) = [E(t, \mathbf{r}; \lambda) - F(t, \mathbf{r}; \lambda) - m^*(\mathbf{u}(t, \mathbf{r}; \lambda), \mathbf{v}(t, \mathbf{r}; \lambda))] / k_B T(t, \mathbf{r}; \lambda). \quad (260)$$

За исключением последнего слагаемого в квадратной скобке (260) – это практически стандартное выражение, используемое в статистической физике [26,35,109] в качестве аргумента функции распределения частиц по энергиям. В него входит собственно энергия λ -состояния и химический потенциал, которым измеряется изменение энергии λ -фазы при удалении из нее или добавлении в нее одной частицы. «Нестандартное» слагаемое

$$-m^*(\mathbf{u}(t, \mathbf{r}; \lambda), \mathbf{v}(t, \mathbf{r}; \lambda))$$

отражает сдвиг энергии λ -состояния за счет движения λ -фазы как целого.

4.2.3. Уравнения баланса числа заряженных частиц, взаимодействующих со случайным потенциалом

Уравнения баланса числа носителей заряда в присутствии случайного взаимодействия получаются после усреднения уравнений баланса для плотности потоков вероятности (243) с учетом формул (246)

$$\partial n(t, \mathbf{r}; \lambda) / \partial t + (\nabla, \mathbf{n}(t, \mathbf{r}; \lambda)) =$$

$$-(i/\hbar(K-1)) \sum_{\lambda'} (\langle \Psi^*(t, \mathbf{r}; \lambda) U^{st}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda') \rangle - \langle \Psi^*(t, \mathbf{r}; \lambda') U^{st}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda) \rangle). \quad (261)$$

Для того чтобы они стали уравнениями для химических потенциалов, необходимо определить зависимость правой части (261) от этих потенциалов и средних по случайному процессу плотностей числа частиц.

Установление этой зависимости является, фактически, наложением дополнительных требований на случайное взаимодействие. Основываясь на рассуждениях, приведенных в параграфе 4.1.5, представим правую часть (261) в виде

$$-(i/\hbar(K-1)) (\langle \Psi^*(t, \mathbf{r}; \lambda) U^{st}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda') \rangle - \langle \Psi^*(t, \mathbf{r}; \lambda') U^{st}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda) \rangle) =$$

$$\begin{aligned}
&= 2\text{Im}(\langle \Psi^*(t, \mathbf{r}; \lambda) U^{\text{st}}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda') \rangle) / (\hbar(K - 1)) = \\
&= -\hbar^{-1} (F(t, \mathbf{r}; \lambda) - F(t, \mathbf{r}; \lambda')) (n(t, \mathbf{r}; \lambda) \Gamma_{\lambda\lambda'} + n(t, \mathbf{r}; \lambda') \Gamma_{\lambda'\lambda}). \quad (262)
\end{aligned}$$

С учетом определений для плотностей $n(t, \mathbf{r}; \lambda)$ и $n(t, \mathbf{r}; \lambda')$ соотношение (262) обеспечивает релаксацию решений уравнений (261) к состоянию локального химического равновесия, если коэффициенты $\Gamma_{\lambda\lambda'}$ положительны. По своему смыслу они представляют собой амплитуды прямых переходов непосредственно между различными состояниями λ и λ' , обусловленные случайным взаимодействием и нормированные на плотности вероятности и количество возможных переходов. Сами амплитуды

$$W_{\lambda\lambda'} = W_{\lambda'\lambda} = |\int dtd^3x \langle \Psi^*(t, \mathbf{r}; \lambda) U^{\text{st}}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda') \rangle|, \quad (263)$$

очевидно, симметричны по индексам λ и λ' . Несимметрия возникает при нормировке $W_{\lambda\lambda'}$ на плотность вероятности обнаружить частицу в момент времени t в точке \mathbf{r} в состоянии λ или λ'

$$\Gamma_{\lambda\lambda'} = T_{\lambda\lambda'} / k_B T, \quad \Gamma_{\lambda'\lambda} = T_{\lambda'\lambda} / k_B T,$$

$$T_{\lambda\lambda'} = W_{\lambda\lambda'} / ((K - 1) \int dtd^3x \rho(t, \mathbf{r}; \lambda)), \quad T_{\lambda'\lambda} = W_{\lambda'\lambda} / ((K - 1) \int dtd^3x \rho(t, \mathbf{r}; \lambda')), \quad (264)$$

а T – температура окружающей среды.

Соотношения (263) и (264) позволяют дать определение для средней температуры λ -фазы носителей заряда $T(\lambda)$. Положим, что

$$T(\lambda) = \int dtd^3x T(t, \mathbf{r}; \lambda) \rho(t, \mathbf{r}; \lambda) / \int dtd^3x \rho(t, \mathbf{r}; \lambda) \quad (265)$$

и

$$\begin{aligned}
T(\lambda) &= T \sum_{\lambda' \neq \lambda} \Gamma_{\lambda\lambda'} = \\
&= (k_B)^{-1} \sum_{\lambda' \neq \lambda} W_{\lambda\lambda'} / ((K - 1) \int dtd^3x \rho(t, \mathbf{r}; \lambda)) = \\
&= (k_B)^{-1} \sum_{\lambda' \neq \lambda} |\int dtd^3x \langle \Psi^*(t, \mathbf{r}; \lambda) U^{\text{st}}(t, \mathbf{r}) \Psi(t, \mathbf{r}; \lambda') \rangle| / ((K - 1) \int dtd^3x \Psi^*_0(t, \mathbf{r}; \lambda) \Psi_0(t, \mathbf{r}; \lambda)). \quad (266)
\end{aligned}$$

Таким образом, средняя температура λ -фазы пропорциональна средней энергии обусловленных случайным взаимодействием переходов из этой фазы во все другие.

Возможный вариант построения многофазной модели переноса заряда при различных температурах λ -фаз описан в книге [71]. Здесь мы будем полагать, что фазы газа носителей заряда находятся в локальном температурном равновесии

между собой и с окружающей средой, то есть все $T(x, \lambda)$ равны между собой и равны температуре окружающей среды

$$T(\mathbf{t}, \mathbf{r}; \lambda) = T(\mathbf{t}, \mathbf{r}; \lambda') = T(\mathbf{t}, \mathbf{r}), \text{ при всех } \lambda \text{ и } \lambda'. \quad (267)$$

Температуру окружающей среды будем считать постоянной и не зависящей от времени и координат

$$\partial T(\mathbf{t}, \mathbf{r}) / \partial t = 0, \nabla T(\mathbf{t}, \mathbf{r}) = 0. \quad (268)$$

Требования (267) и (268) в совокупности представляют собой условия глобального температурного равновесия. В этом случае справедливо соотношение, вытекающее из (267), (268) и определений (265), (266)

$$\sum_{\lambda' \neq \lambda} \Gamma_{\lambda\lambda'} = 1. \quad (269)$$

С учетом сделанных предположений и данных определений, уравнения баланса числа носителей заряда примут вид

$$\partial n(\mathbf{t}, \mathbf{r}; \lambda) / \partial t + (\nabla, \mathbf{n}(\mathbf{t}, \mathbf{r}; \lambda)) = \hbar^{-1} \sum_{\lambda'} (F(\mathbf{t}, \mathbf{r}; \lambda) - F(\mathbf{t}, \mathbf{r}; \lambda')) (n(\mathbf{t}, \mathbf{r}; \lambda) \Gamma_{\lambda\lambda'} + n(\mathbf{t}, \mathbf{r}; \lambda') \Gamma_{\lambda'\lambda}),$$

тождественный уравнениям (218). Из них, очевидно, следует закон сохранения суммарного потока (215)

$$\partial n(\mathbf{t}, \mathbf{r}) / \partial t + (\nabla, \mathbf{n}(\mathbf{t}, \mathbf{r})) = 0.$$

4.2.4. Время релаксации τ

В уравнении (247) фигурирует время $\tau(\mathbf{t}, \mathbf{r}; \lambda)$, определяющее при условии

$$\tau(\mathbf{t}, \mathbf{r}; \lambda) > 0$$

скорость релаксации гидродинамической скорости $\mathbf{v}(\mathbf{t}, \mathbf{r}; \lambda)$ при $t > 0$ к решению

$$\mathbf{v}_0(\mathbf{t}, \mathbf{r}; \lambda) = -\tau(\mathbf{t}, \mathbf{r}; \lambda) (\nabla F(\mathbf{t}, \mathbf{r}; \lambda) + s(\mathbf{t}, \mathbf{r}; \lambda) k_B \nabla T(\mathbf{t}, \mathbf{r}; \lambda)) / m^*. \quad (270)$$

Как было показано ранее, этим временем определяется диффузионная длина для химического потенциала $F(\mathbf{t}, \mathbf{r}; \lambda)$ или, что то же самое, длина релаксации λ -фазы к состоянию локального химического равновесия. Это же время фигурирует в выражении для подвижности носителей заряда (234), которую можно ввести при условии локального химического равновесия электронных фаз [26].

По своему физическому смыслу это время $\tau(\mathbf{t}, \mathbf{r}; \lambda)$ очень близко к величине

$$\tau_0(\mathbf{t}, \mathbf{r}; \lambda) = \hbar / (k_B T(\mathbf{t}, \mathbf{r}; \lambda)), \quad (271)$$

которой, согласно (218) и (265), определяется время релаксации химического потенциала λ -фазы. В случае, когда

$$\tau_0(\mathbf{t}, \mathbf{r}; \lambda) \rightarrow \infty$$

уравнения (218) имеют единственное решение для $F(\mathbf{t}, \mathbf{r}; \lambda)$, соответствующее состоянию локального химического равновесия между фазами (201.1). В случае, когда

$$\tau(\mathbf{t}, \mathbf{r}; \lambda) \rightarrow \infty$$

уравнения (247) при сделанных предположениях относительно температурного равновесия имеет единственное решение

$$\nabla F(\mathbf{t}, \mathbf{r}; \lambda) = 0,$$

которое означает химическое равновесие между пространственно разделенными точками λ -фазы. Таким образом, $\tau_0(\mathbf{t}, \mathbf{r}; \lambda)$ является характерным временем релаксации к химическому равновесию в пространстве состояний носителей заряда, а время $\tau(\mathbf{t}, \mathbf{r}; \lambda)$ играет аналогичную роль в обычном пространстве.

Температуру каждого из λ -состояний, а следовательно и время τ_0 , мы связали через соотношения (265) и (266) с амплитудой переходов между λ -состояниями. Логично таким же образом определить время $\tau(\mathbf{t}, \mathbf{r}; \lambda)$. При этом, согласно сказанному выше, в определении должны фигурировать диагональные элементы матрицы $W_{\lambda\lambda}$. Положим, что для $\tau(\mathbf{t}, \mathbf{r}; \lambda)$ справедливо интегральное уравнение, аналогичное уравнению (265)

$$\tau^{-1}(\lambda) = \int d\mathbf{t} d^3\mathbf{x} \tau^{-1}(\mathbf{t}, \mathbf{r}; \lambda) \rho(\mathbf{t}, \mathbf{r}; \lambda) / \int d\mathbf{t} d^3\mathbf{x} \rho(\mathbf{t}, \mathbf{r}; \lambda), \quad (272)$$

где $\tau(\lambda)$ – усредненное по пространству и времени время релаксации импульса. Именно оно фигурирует в уравнениях (208) и (209), которые мы используем при моделировании характеристик квантовых приборов. Это время будем вычислять на основе соотношения

$$\tau^{-1}(\lambda) = W_{\lambda\lambda} (\hbar \int d\mathbf{t} d^3\mathbf{x} \rho(\mathbf{t}, \mathbf{r}; \lambda))^{-1}, \quad (273)$$

где, как следует из определения (263),

$$W_{\lambda\lambda} = \left| \int d\mathbf{t} d^3\mathbf{x} \langle \Psi^*(\mathbf{t}, \mathbf{r}; \lambda) U^{\text{st}}(\mathbf{t}, \mathbf{r}) \Psi(\mathbf{t}, \mathbf{r}; \lambda) \rangle \right|.$$

Согласно определению (273) время релаксации $\tau(\lambda)$ положительно. Решения уравнения (272) также можно искать в классе неотрицательных значений $\tau(\mathbf{t}, \mathbf{r}; \lambda)$. Тогда это гарантирует устойчивость решений уравнения (247).

Если воспользоваться теорией возмущений, то величину $\tau^{-1}(\lambda)$ в первом приближении можно связать с мнимой частью «амплитуды рассеяния вперед» [68,69,108] под действием случайного взаимодействия. То есть с амплитудой рассеяния из λ -состояния в λ -состояние по всем возможным пространственно-временным траекториям. В случае, когда волновые функции $\Psi_0(\mathbf{t}, \mathbf{r}; \lambda)$ нормированы на одну частицу в объеме и представляют собой полный набор, из (273) приближенно получается формула (210) для времени релаксации импульса.

4.2.5. Необратимость уравнений многофазной модели

Отталкиваясь от обратимых во времени уравнений динамики (238), мы построили необратимые во времени уравнения баланса числа частиц (218). Представляет интерес вопрос: результатом каких допущений и преобразований в модели является необратимость?

Определим в явном виде понятие обратимости во времени для уравнений (238). Если произвести замену переменных

$$\xi = -t, \quad (274)$$

и обозначить

$$\mathbf{x} = (\mathbf{t}, \mathbf{r}), \mathbf{x}^* = (\xi, \mathbf{r}), \quad (275)$$

то в силу эрмитовости операторов U и U^{st} из (44) для функции

$$\chi(\mathbf{x}^*, \lambda) = \chi(\xi, \mathbf{r}; \lambda) = \Psi^*(-\mathbf{t}, \mathbf{r}; \lambda) \quad (276)$$

получаются уравнения, идентичные по форме уравнениям (238)

$$\begin{aligned} i\hbar \partial \chi(\mathbf{x}^*, \lambda) / \partial \xi + (\hbar^2/2)(\nabla, m^{*-1} \nabla \chi(\mathbf{x}^*, \lambda)) - U(\mathbf{x}^*) \chi(\mathbf{x}^*, \lambda) = \\ = (1/(K-1)) \sum_{\lambda' \neq \lambda} U^{\text{st}}(\mathbf{x}^*) \chi(\mathbf{x}^*, \lambda'), \end{aligned} \quad (277)$$

с потенциалами

$$U(\mathbf{x}^*) = U(\xi, \mathbf{r}) = U(-\mathbf{t}, \mathbf{r}) \text{ и } U^{\text{st}}(\mathbf{x}^*) = U^{\text{st}}(\xi, \mathbf{r}) = U^{\text{st}}(-\mathbf{t}, \mathbf{r}).$$

Если считать, что

$$U(x^*) = U(x), U^{st}(x^*) = U^{st}(x), \quad (278)$$

то уравнения (277) и (278) для функций $\chi(\xi, \mathbf{r}; \lambda)$ и $\Psi(t, \mathbf{r}; \lambda)$ оказываются тождественными. Требование (278) означает идентичное поведение потенциалов, описывающих взаимодействие полей $\chi(\xi, \mathbf{r}; \lambda)$ и $\Psi(t, \mathbf{r}; \lambda)$ с окружающей средой, при изменении знака u времени t .

Процедуру, при помощи которой из уравнений для функции $\Psi(t, \mathbf{r}; \lambda)$ получены уравнения для функции $\chi(\xi, \mathbf{r}; \lambda)$, будем называть СТ – преобразованием [108]. Эта операция, для произвольной функции означает изменение знака (обращение) времени и комплексное сопряжение

$$A^{CT}(t, \mathbf{r}) = A^*(-t, \mathbf{r}). \quad (279)$$

Величины и соотношения не меняющиеся при СТ – преобразовании будем называть СТ – инвариантными. Понятие СТ-инвариантности будем считать тождественным понятию обратимости во времени.

Уравнения (238) при выполнении условий (278) являются СТ – инвариантными, то есть обратимыми во времени. Аналогичное утверждение справедливо и в отношении всех следствий из этих уравнений, в частности, в отношении уравнений баланса (243). Действительно, легко видеть, что

$$\begin{aligned} \sigma^{CT}(x, \lambda) &= \sigma(x^*, \lambda), \mathbf{i}^{CT}(x, \lambda) = -\mathbf{i}(x^*, \lambda), \\ \{i/\hbar(K-1)\sum_{\lambda'}(\Psi^*(x, \lambda)U^{st}(x)\Psi(x, \lambda') - \Psi^*(x, \lambda')U^{st}(x)\Psi(x, \lambda))\}^{CT} &= \\ &= -i/\hbar(K-1)\sum_{\lambda'}(\chi^*(x^*, \lambda)U^{st}(x^*)\chi(x^*, \lambda') - \chi^*(x^*, \lambda')U^{st}(x^*)\chi(x^*, \lambda)), \end{aligned}$$

Откуда для СТ-сопряженных величин следуют уравнения

$$\begin{aligned} \partial\sigma^{CT}(x^*, \lambda)/\partial\xi + (\nabla, \mathbf{i}^{CT}(x^*, \lambda)) &= \\ &= -(i/\hbar(K-1))\sum_{\lambda'}(\chi^*(x^*, \lambda)U^{st}(x^*)\chi(x^*, \lambda') - \chi^*(x^*, \lambda')U^{st}(x^*)\chi(x^*, \lambda)), \end{aligned} \quad (280)$$

тождественные уравнениям (243). Это означает обратимость во времени уравнений баланса (243).

Произведем СТ – преобразование уравнений (218). Для входящих в них величин имеем

$$\mathbf{n}^{CT}(x, \lambda) = \rho(x^*, \lambda)f(s(x^*, \lambda)), \mathbf{n}^{CT}(x, \lambda) = -\mathbf{j}(x^*, \lambda)f(s(x^*, \lambda)), \Gamma_{\lambda\lambda'}^{CT} = \Gamma_{\lambda\lambda'}.$$

В результате, используя уравнение непрерывности (197), получим

$$\begin{aligned} \rho(x^*, \lambda) \partial f(s(x^*, \lambda)) / \partial \xi + (\mathbf{j}(x^*, \lambda), \nabla f(s(x^*, \lambda))) = \\ = \hbar^{-1} \sum_{\lambda'} (F(x^*, \lambda) - F(x^*, \lambda')) (\rho(x^*, \lambda) f(s(x^*, \lambda)) \Gamma_{\lambda \lambda'} + \rho(x^*, \lambda') f(s(x^*, \lambda')) \Gamma_{\lambda' \lambda}), \end{aligned} \quad (281)$$

Это уравнение не совпадает с аналогичным по форме записи уравнением для $f(s(x, \lambda))$, получающимся из (218)

$$\begin{aligned} \rho(x, \lambda) \partial f(s(x, \lambda)) / \partial t + (\mathbf{j}(x, \lambda), \nabla f(s(x, \lambda))) = \\ = -\hbar^{-1} \sum_{\lambda'} (F(x, \lambda) - F(x, \lambda')) (\rho(x, \lambda) f(s(x, \lambda)) \Gamma_{\lambda \lambda'} + \rho(x, \lambda') f(s(x, \lambda')) \Gamma_{\lambda' \lambda}). \end{aligned} \quad (282)$$

Различие уравнений (281) и (282) состоит в разных знаках их правых частей. Оно означает различное поведение химических потенциалов как функций t и ξ . Уравнения (282) описывают при $t > 0$ релаксацию химических потенциалов полей с различными λ к состоянию локального химического равновесия (215). Уравнения (281) при $\xi > 0$ описывают экспоненциальное удаление с ростом ξ от этого равновесия.

Продемонстрируем сказанное на простейшем примере. Рассмотрим пространственно однородный случай, когда химические потенциалы не зависят от координат. Положим, что существует некоторое выделенное λ -состояние, для которого при $t = \xi = 0$ справедливо начальное условие

$$F(0, \lambda) = F_0, \quad (283)$$

а для всех остальных $\lambda' \neq \lambda$ химические потенциалы равны нулю и от времени не зависят. Химическое равновесие в этом случае будет достигаться при равенстве нулю $F(\xi, \lambda)$ и $F(t, \lambda)$.

Из (281) и (282) получим уравнения

$$\partial F(\xi, \lambda) / \partial \xi = F(\xi, \lambda) q^*, \quad (284)$$

$$\partial F(t, \lambda) / \partial t = -F(t, \lambda) q, \quad (285)$$

где

$$q^* = -(1/\tau_0) (\rho(x^*, \lambda) f'(s(x^*, \lambda)))^{-1} \sum_{\lambda' \neq \lambda} (\rho(x^*, \lambda) f(s(x^*, \lambda)) \Gamma_{\lambda \lambda'} + \rho(x^*, \lambda') f(s(x^*, \lambda')) \Gamma_{\lambda' \lambda}) > 0,$$

$$q = -(1/\tau_0) (\rho(x, \lambda) f'(s(x, \lambda)))^{-1} \sum_{\lambda' \neq \lambda} (\rho(x, \lambda) f(s(x, \lambda)) \Gamma_{\lambda \lambda'} + \rho(x, \lambda') f(s(x, \lambda')) \Gamma_{\lambda' \lambda}) > 0,$$

$$f'(s) = \partial f(s) / \partial s < 0.$$

Если дополнительно считать, что

$$F_0 \ll k_B T,$$

а λ -состояния стационарны, то зависимостью q и q^* от t и ξ можно пренебречь. В результате получим приближенные решения начальных задач (284), (283) и (285), (283)

$$F(\xi, \lambda) \approx F_0 \exp\{q^* \xi\}, \quad (286)$$

$$F(t, \lambda) \approx F_0 \exp\{-qt\}. \quad (287)$$

Решение (287) при $t > 0$ описывает релаксацию функции $F(t, \lambda)$ к нулю и выравнивание химических потенциалов всех фаз. Решение (286) при $\xi > 0$ описывает экспоненциально быстрое удаление функции $F(\xi, \lambda)$ от начального значения, и не приводит к выравниванию химических потенциалов различных фаз.

Таким образом, в отличие от уравнений (243), СТ-преобразование уравнений (218) не привело к уравнениям, тождественным исходным. Правые части уравнений (218) ведут себя при СТ – преобразовании иначе, чем правые части уравнений (243). А именно, при СТ – преобразовании не происходит смены знака правых частей уравнений (218).

Сказанное означает, что уравнения (218) не являются прямым следствием уравнений (243). При получении уравнений (218) сделаны предположения, эквивалентные постулированию новых уравнений. Это постулирование является результатом согласованного определения средних по случайному процессу от плотностей вероятности и потока вероятности (выражения (258)) и от правых частей уравнений (243) (выражения (262) – (264)). При этом произошел переход от описания в терминах плотности вероятности и плотности потока вероятности к описанию в терминах плотности числа частиц и плотности потока частиц.

Кроме уравнений (218), необратимыми во времени являются уравнения (247) для гидродинамической скорости $\mathbf{v}(t, \mathbf{r}; \lambda)$. Эти уравнения изначально постулируются и не имеют аналога в виде некоторого следствия уравнений динамики. Различия в поведении $\mathbf{v}(t, \mathbf{r}; \lambda)$ при положительных и отрицательных значениях времени приводят согласно определениям (257) – (260) к различию в поведении функции распределения $f(s(t, \mathbf{r}; \lambda))$ при $t > 0$ и $t < 0$. То есть усреднение плотности и потока вероятности по случайному процессу выступает в

многофазной модели в качестве процедуры, изменяющей характер временной зависимости усредняемых величин.

Выбранный вид функции распределения вместе с уравнениями (247) для гидродинамической скорости являются причиной возникновения электрического (омического) сопротивления (см. формулы (233) и (234)), а также приводят к появлению в уравнениях (218) характерной длины релаксации химических потенциалов.

Таким образом, в рассматриваемой модели присутствие случайного взаимодействия само по себе не приводит к необратимости во времени. Она возникает как следствие выбранной процедуры усреднения по случайному процессу исходных обратимых соотношений. Следствием этой необратимости являются:

- наличие состояния химического равновесия;
- наличие электрического (омического) сопротивления.

4.3. Граничные условия для стационарных уравнений переноса заряда в мезоскопических структурах

Для того, чтобы моделировать перенос заряда в квантовых приборах, кроме уравнений необходимо сформулировать еще и граничные условия для входящих в эти уравнения величин. Эти условия могут быть очень разнообразными, существенно зависящими от постановки решаемой задачи. Здесь мы сформулируем только общие условия, используемые обычно при моделировании простейших приборов.

Важным достоинством многофазной модели переноса заряда является возможность сформулировать большую часть граничных условий в тех же терминах, что и в дрейфово-диффузионной модели (см. параграф 1.4). При этом, место химических потенциалов электронов и дырок в мезоскопических структурах играют химические потенциалы λ -фаз электронного газа.

На Рисунке 23 показана абстрактная область некоторого квантового прибора с границами нескольких типов Γ_0 , Γ_1, Γ_2 и Γ_3 . В большинстве случаев геометрические и технологические особенности конструкции не определяют эти границы однозначно, и их выбор является одним из важных этапов постановки задачи моделирования. При этом должны быть удовлетворены некоторые физические требования. В частности, для постановки граничных условий для уравнения Шредингера необходимо знать его физически интерпретируемые решения вне области Ω . Как правило, полагают, что это решения при постоянном потенциале (плоские волны, их суперпозиция, экспоненциально затухающие волновые функции и т.п.). Поэтому важным является требование постоянства либо, как минимум, квазиклассического поведения потенциала U всюду вне области Ω , которое часто называют «условием плоских зон» [26].

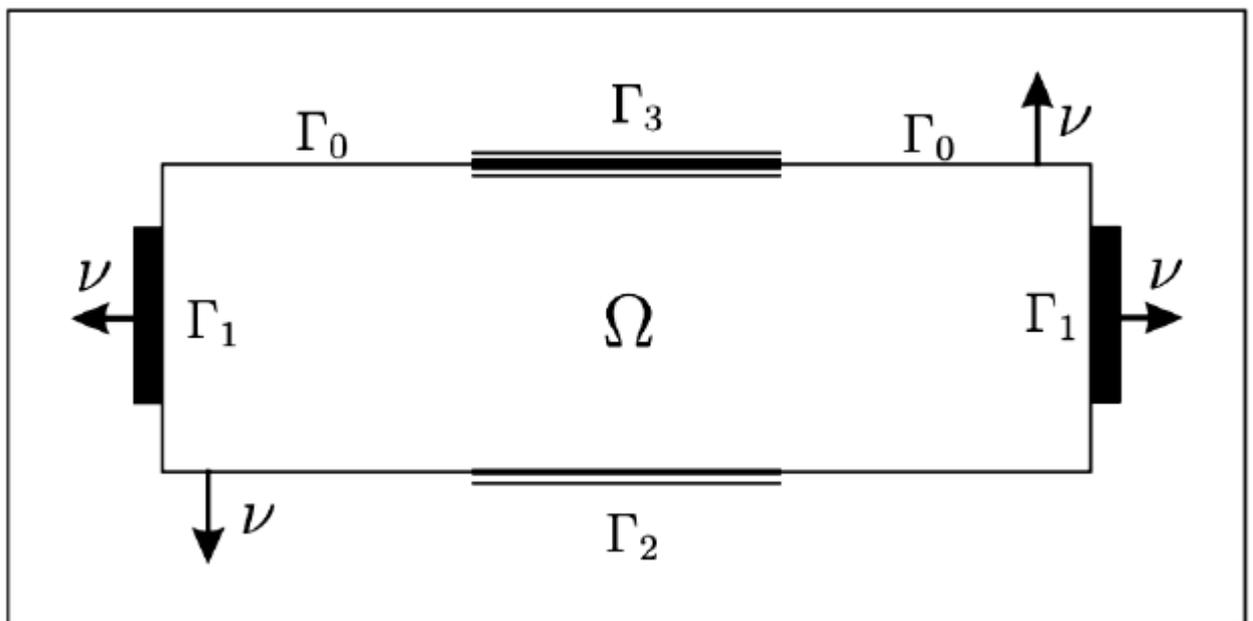


Рисунок 23. Некоторые типы границ квантовых приборов (область Ω): Γ_0 – границы, переносом заряда через которые можно пренебречь; Γ_1 – контакты, на которых задано напряжение; Γ_2 – полевые контакты (затворы); Γ_3 – токовые контакты, ν - вектор внешней нормали к границе.

Формально граничные условия для волновых функций $\Psi(\lambda)$ на всех типах границ достаточно просты. Они являются следствием очевидного требования непрерывности потоков плотности вероятности $\mathbf{j}(\lambda)$ на границе и выражаются в

непрерывности функции $\Psi(\lambda)$ и ее нормальной производной на Γ ($\Gamma = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$)

$$\Psi(\Gamma_{-0}, \lambda) = \Psi(\Gamma_{+0}, \lambda), \quad \partial\Psi(\Gamma_{-0}, \lambda)/\partial\mathbf{v} = \partial\Psi(\Gamma_{+0}, \lambda)/\partial\mathbf{v}. \quad (288)$$

Обозначения Γ_{+0} и Γ_{-0} приняты для малых внешней и внутренней окрестностей границы Γ соответственно.

Посредством Γ_0 мы будем обозначать границы, переносом заряда через которые можно пренебречь, и полагать на Γ_0

$$(\mathbf{v}, \nabla F(\lambda)) = 0, \quad (289.1)$$

$$(\mathbf{v}, \nabla \varphi) = 0. \quad (289.2)$$

Условие (289.2) является в общем случае приближенным. Оно корректно только если предполагать, что диэлектрическая проницаемость ε в области Ω много больше, чем во внешней по отношению к Ω области, либо если прилегающая к границе часть области Ω квазиэлектронейтральна. Будем считать это предположение выполненным.

Границы типа Γ_1 соответствуют контактам, на которых известны приложенные напряжения V_i (i – номер контакта). На них справедливы условия $F_i(\lambda) = -eV_i$. (290)

Напряжение V_i не зависит от λ , что согласно (201.1) означает химическое равновесие электронов в контактах. Последнее требование является определяющим при выборе областей постановки граничных условий (290). Для корректного описания физики процессов переноса выбор границ типа Γ_1 необходимо осуществлять на расстояниях от активных областей (таких как гетеропереходы, переходы металл-полупроводник и пр.), много превышающих характерную длину релаксации неравновесных электронных фаз L_0 .

Для электростатического потенциала φ на границах Γ_1 логично требовать [27] выполнения условия электронейтральности

$$n - N_{\text{int}} = 0. \quad (291)$$

При этом необходимо, чтобы граница Γ_1 была расположена на расстоянии, много большем, чем длина Дебая L_D , от всех неоднородностей потенциала U^{ext} и областей, где имеются градиенты концентрации N_{int} .

В квантовом случае условие (291) является неявным условием для φ в отличие от аналогичного условия в квазиклассической теории полупроводников [26,27], проанализированного в первой части, когда φ входит в n явным образом через функцию распределения. В квантовой теории n вычисляется суммированием по λ соответствующих парциальных вкладов $n(\lambda)$, ни в одном из которых φ явно не присутствует. Потенциал φ входит в уравнение Шредингера и, таким образом, определяет вид волновых функций $\Psi(\lambda)$, которыми в свою очередь определяется $n(\lambda)$.

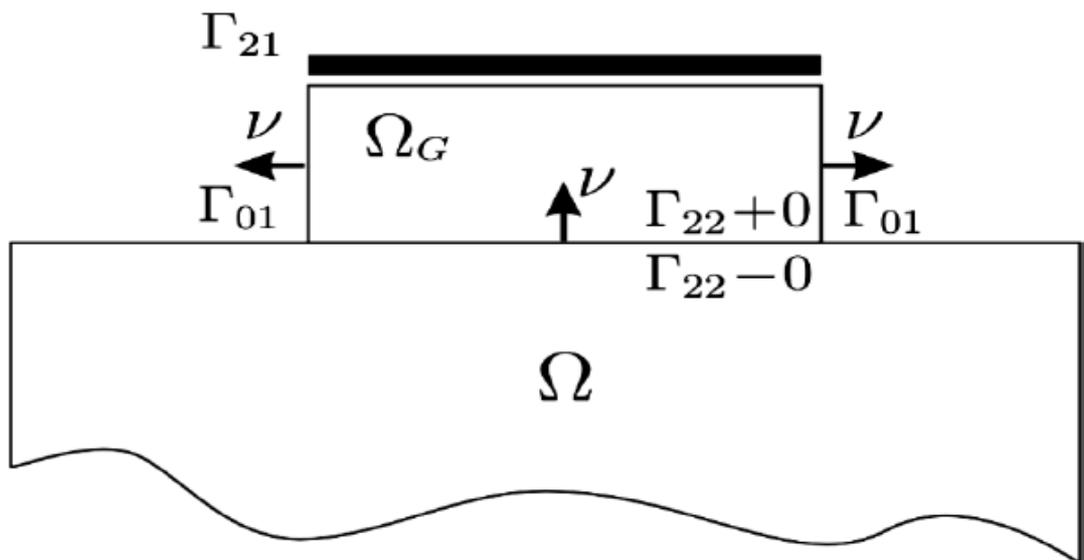


Рисунок 24. Полевой затвор Ω_G , граничащий с областью переноса заряда.

Достаточно распространены в электронных приборах контакты типа Γ_2 , так называемые полевые затворы. Реально они представляют собой некоторый слой непроводящего материала (см. Рисунок 24, область Ω_G), с одной стороны находящегося в контакте с исследуемой областью Ω (граница Γ_{22}), а с другой стороны, имеющего обычный омический контакт (граница Γ_{21} на Рисунке 24). На Γ_{21} задается напряжение затвора V_G , постоянное на всей границе и связанное с электростатическим потенциалом φ некоторым линейным законом

$$\varphi(\Gamma_{21}) = V_G + \text{const}, \quad (292)$$

где const определяется обычно работой выхода электрона, электронным сродством и пр. (см. [26, 27]). Далее, в области Ω_G решается уравнение Лапласа

$$\Delta\varphi = 0 \quad (293)$$

с условием (292) на Γ_{21} и условиями

$$\partial\varphi/\partial\nu(\Gamma_{02}) = 0, \quad (294)$$

$$\varepsilon\partial\varphi/\partial\nu(\Gamma_{22-0}) - \varepsilon_1\partial\varphi/\partial\nu(\Gamma_{22+0}) = Q, \quad (295)$$

где ε_1 - диэлектрическая проницаемость области затвора, Q – возможно существующий на границе Γ_{22} поверхностный заряд.

Никаких условий на химический потенциал F на границе Γ_2 (Γ_{22}) не ставится. В часто встречающемся случае, когда ширина области Ω_G , то есть размер в направлении, перпендикулярном области Ω , много меньше длины, то есть размера вдоль области Ω , можно приближенно считать, что поле распределено под затвором линейно в направлении, перпендикулярном Ω , и, исходя из этого, строить различные приближенные описания.

Наконец, в некоторых случаях встречаются границы типа Γ_3 , на которых задана нормальная составляющая градиента химического потенциала

$$\partial F/\partial\nu(\Gamma_3) = S(\Gamma_3), \quad (296)$$

где $S(\Gamma_3)$ – некоторая известная функция. Такие граничные условия можно, например, ставить на условных границах между резонансно - туннельными структурами и подложками [88], задавая тем самым распределение тока на границе квантового прибора и классической области.

В заключение этого параграфа перечислим сформулированные здесь требования, позволяющие физически корректно выбрать границы квантового прибора:

- постоянство либо, как минимум, квазиклассическое поведение потенциала U вне области квантового прибора;
- контактные области, в которых задаются приложенные напряжения, должны быть расположены на расстояниях, много больших длины

релаксации неравновесных электронных фаз L_0 от активных областей прибора;

- контактные области должны быть расположены на расстояниях, много больших, чем длина Дебая L_D от всех неоднородностей потенциала U^{ext} и областей, где имеются градиенты концентрации N_{int} .

ГЛАВА 5. Моделирование статических характеристик резонансно-туннельных диодов

Приборы, созданные на основе резонансно-туннельных структур (РТС), такие как резонансно-туннельные диоды (РТД), уже перестали быть экзотическими и находят постоянно расширяющуюся сферу применения в устройствах СВЧ и КВЧ диапазонов [1-3,110-113]. Технология производства резонансно-туннельных приборов является сегодня практически промышленной.

Генераторы на основе РДТ могут быть источниками электромагнитного излучения в широком диапазоне частот: от нескольких ГГц до нескольких ТГц [15,16,112,113]. Исследования этого прибора, начатые около тридцати лет назад, остаются актуальными и сегодня в связи с разворачивающимися в мире практическими работами по освоению терагерцового диапазона частот [4].

Наличие у РТД участка ВАХ с отрицательной дифференциальной проводимостью позволяет использовать этот прибор в вычислительной технике [1,110]. На его основе могут быть построены элементы памяти и вычислительные ячейки, обеспечивающие терагерцовые частоты переключения.

Однако далеко не все проблемы, с которыми столкнулись разработчики первых РТД, решены до конца. Многие вопросы остаются до сих пор без удовлетворительных ответов. Касаются они не только собственно РТС и РТД, но являются отражением более общих проблем, связанных с пониманием основных принципов функционирования квантовых приборов.

К таковым относится проблема адекватного описания статических характеристик приборов. Состоит она в том, что, следуя только простейшим представлениям квантовой механики и кинетической теории, невозможно дать сколь-нибудь удовлетворительное описание вольт - амперных характеристик приборов. Результаты расчетов и экспериментальные данные различаются в разы, а часто и на порядки (см. Рисунок 25).

Анализ большого массива данных показывает, что подобная ситуация является типичной для всех общеизвестных моделей, с нею сталкивались постоянно и сталкиваются до сих пор все исследовательские группы во всех центрах. Обстоятельный анализ проблем моделирования резонансно-туннельных структур дан в статье [114] и цитируемой в ней литературе.

Рассмотренная выше модель позволяет решить проблему адекватного эксперименту теоретического описания статических характеристик РТД. С попыток моделирования ВАХ РТД и начались последовательные работы по построению многофазной модели переноса заряда. Она была разработана еще в период создания первых советских РТД (конец восьмидесятых - начало девяностых годов двадцатого столетия) [83,86,115]. Работы были инициированы Ю.С. Тиходеевым, руководившим коллективом разработчиков. Активное участие в экспериментальных и теоретических исследованиях принимали О.Т. Гаврилов, И.И. Квяткевич, А.А. Чернявский, Ю.А. Иванов, С.А. Дианов, С.В. Крючков.

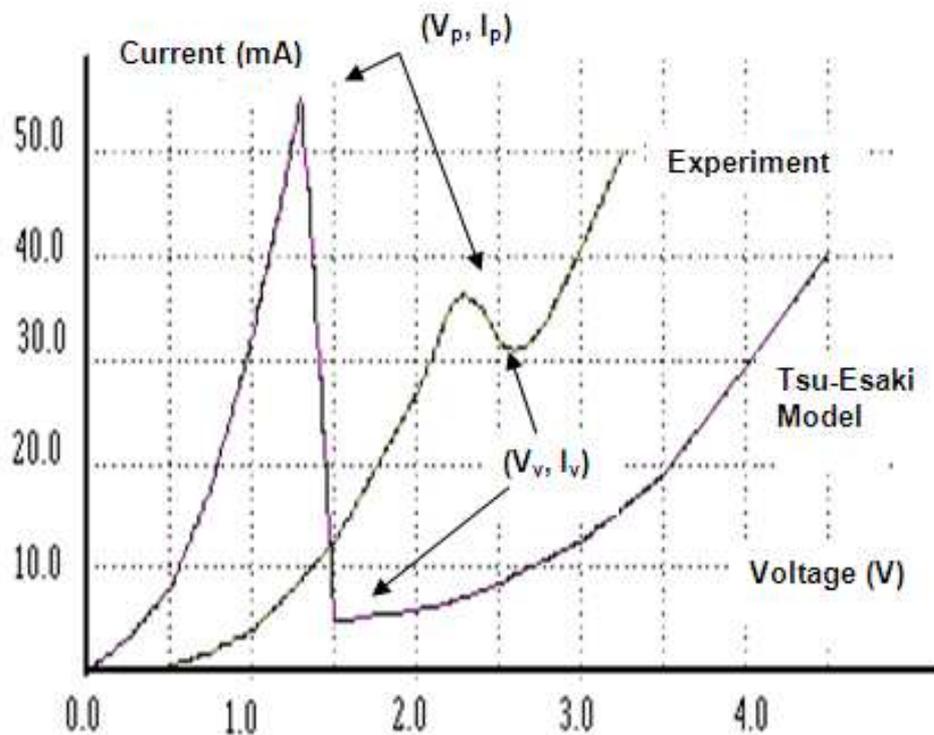


Рисунок 25. Типичные экспериментальная и расчетная ВАХ для AlGaAs РТД на полу-изолирующей подложке.

Проблема адекватного описания характеристик практически всех квантовых приборов имеет, как минимум, две существенных составляющих:

- описание переноса заряда в мезоскопической структуре при ненулевой температуре;
- описание взаимодействия классических и квантовых областей приборов, функционирующих на квантовых принципах.

Ниже будут предложены возможные пути решения обеих этих проблем.

5.1. Перенос заряда в РТС

5.1.1. Резонансное туннелирование электронов

Резонансно-туннельная структура представляет собой активную часть электронного прибора, названного резонансно-туннельным диодом, и является аналогом хорошо известного оптического интерферометра Фабри - Перо [116]. Подробное описание принципов функционирования РТС и РТД можно найти в работах [1,12-16,74,117].

В основе действия прибора лежит эффект резонансного туннелирования электронов. Суть его состоит в том, что для некоторых потенциальных рельефов возможна ситуация, когда коэффициент прохождения электронов через эти рельефы близок к нулю практически при всех энергиях, меньших некоторого значения, за исключением выделенных резонансных энергий, при которых этот коэффициент близок к единице. В идеальном случае строго симметричной РТС коэффициент прохождения электронов через структуру при резонансных энергиях равен единице.

Простейшей РТС является потенциальный рельеф, состоящий из двух одинаковых потенциальных барьеров и потенциальной ямы между ними (см. Рисунок 26). В случае прямоугольных барьеров путем элементарных аналитических расчетов легко показать, что в зависимости от ширины и высоты барьеров, а также от ширины ямы в такой РТС может быть один или несколько

резонансных уровней энергии. На Рисунке 27 показано поведение коэффициента прохождения электронов через потенциал, изображенный на Рисунке 27, где ширина потенциальной ямы составила 50 А, а высота барьеров – 1 эВ. В такой структуре наблюдается один резонансный уровень энергии $E_r \approx 0,54$ эВ. При ширине ямы в 70 А наблюдается уже два резонансных уровня: $E_{r1} \approx 0,32$ эВ и $E_{r2} \approx 0,69$ эВ (см. Рисунок 28). На Рисунке 28 область построения выбрана таким образом, чтобы продемонстрировать, что кроме резонансного туннелирования наблюдается еще надбарьерное отражение. Этот эффект проявляется в осцилляциях коэффициента прохождения при энергиях больших высоты барьеров (1 эВ).

Резонансное туннелирование характерно для любых периодических потенциальных рельефов. При увеличении числа барьеров и ям можно получить уже не только уровни, но и минизоны [11], а при бесконечном числе подобных ячеек – зонную структуру типа одномерного кристалла.

Практическая реализация простейшей двухбарьерной РТС возможна, например, в многослойных полупроводниковых гетероструктурах, выращенных при помощи технологии молекулярно-лучевой эпитаксии (МЛЭ). На Рисунке 29 схематично показана такая GaAs/AlGaAs РТС. Находящийся сверху слой n+GaAs представляет собой верхнюю контактную область, далее следует эмиттерный разделительный слой, часто называемый эмиттерным спейсером, предназначенный для предотвращения попадания примеси из контактной области в следующую за ним область первого барьера из AlGaAs (в соединениях типа $Al_xGa_{1-x}As$, регулируя содержание алюминия, можно изменять высоту потенциального барьера по отношению к GaAs), далее область квантовой ямы (GaAs), за нею второй потенциальный барьер (AlGaAs), коллекторный спейсер и нижняя контактная область, граничащая с подложкой.

РТС, имеющая два контакта, к которым может быть приложена разность потенциалов, представляет собой квантовый прибор, названный резонансно-туннельным диодом (РТД) (см. Рисунок 29).

Как правило, один из контактов подсоединяется к РТС через n^+ область. Технологически бывает удобно размещать n^+ область на полуизолирующей подложке (см. Рисунок 30). В этом случае оба контакта РТД находятся на одной поверхности полупроводниковой пластины. Возможно подсоединение контакта к n^+ области снизу без полуизолирующей подложки (РТД на n^+ подложке, см. Рисунок 30). Влиянию параметров РТС и подложки на характеристики РТД посвящено достаточно много работ, например, [73,84,85,88,90-92,117,118].

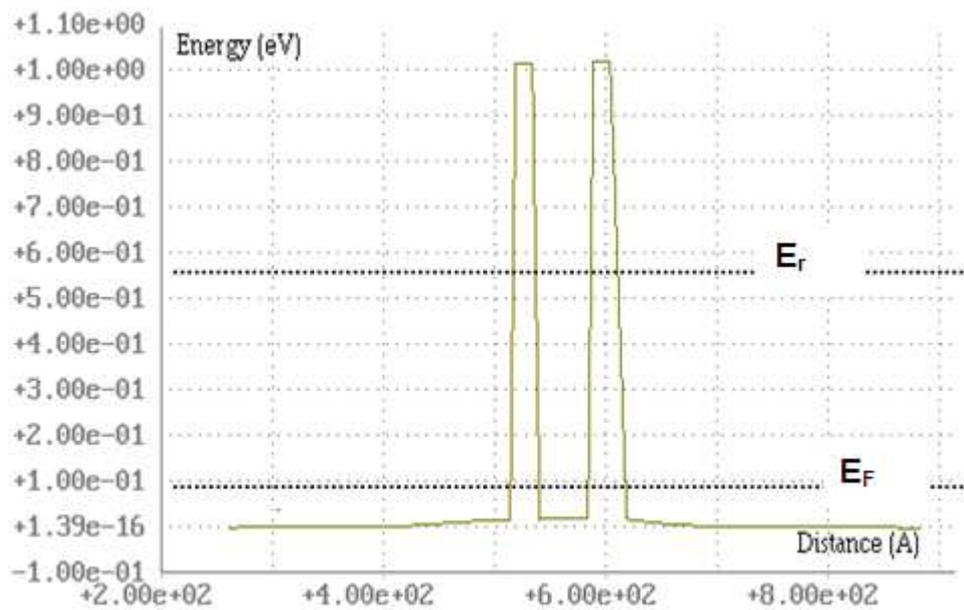


Рисунок 26. Потенциальный рельеф $U = U^{\text{ext}}$ - еф для AlGaAs РТС с параметрами слоев: n^+ GaAs/500Å/GaAs/15Å/AlGaAs/20Å/GaAs/50Å/AlGaAs/20Å/GaAs/15Å/ n^+ GaAs/500Å при нулевом смещении.

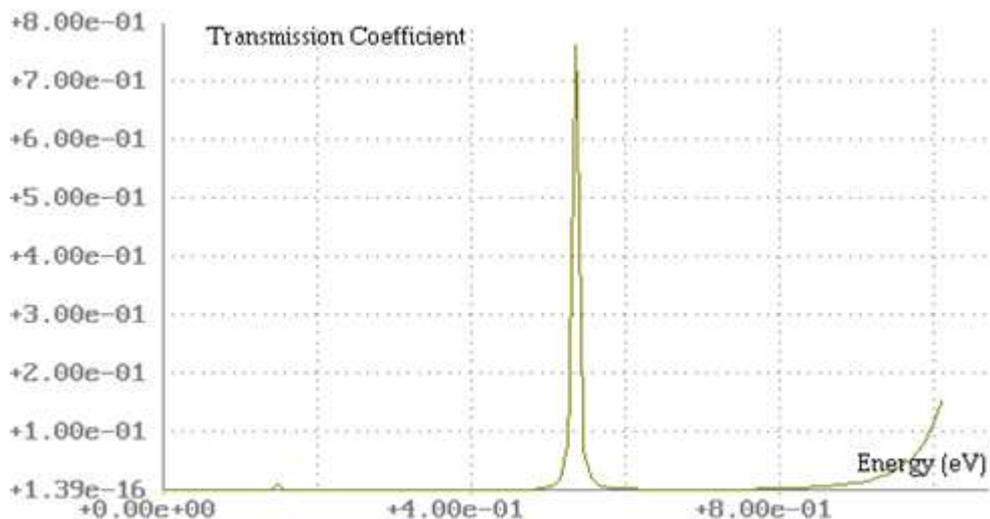


Рисунок 27. Коэффициент прохождения электронов через потенциальный рельеф РТС, представленной на Рисунке 26.

Не нарушая общности, будем считать нулевым напряжение, приложенное к верхнему контакту РТД, а через V будем обозначать напряжение на контакте к подложке. Кроме того, будем полагать, что система координат выбрана таким образом, что ось Oz направлена перпендикулярно гетерослоям в сторону подложки (см. Рисунок 30).

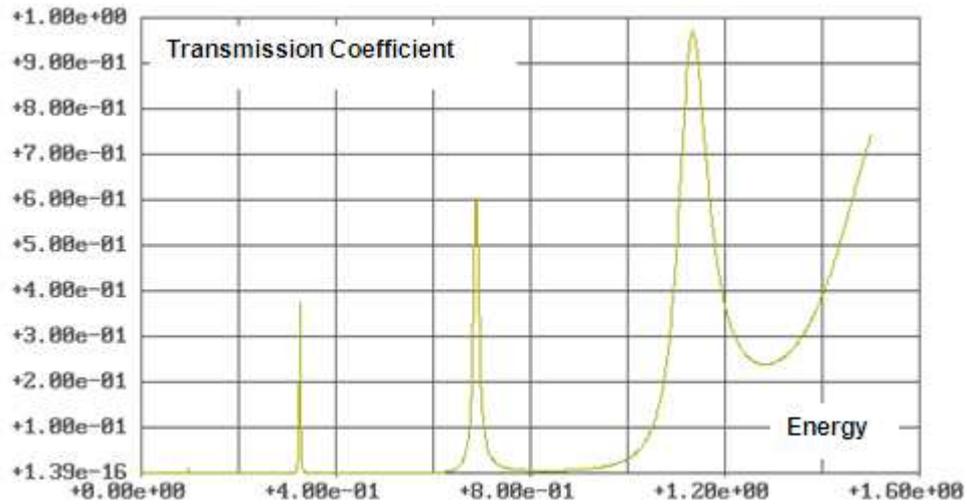


Рисунок 28. Коэффициент прохождения электронов через РТС, аналогичную изображенной на Рисунке 26, с шириной барьеров в 20 А и ямы – 70 А.

На Рисунке 27 показана типичная зависимость коэффициента прохождения электронов (T) через РТС от энергии (расчет проводился для потенциала, показанного на Рисунке 26). Видно, что при $E_T = 0,54$ эВ наблюдается резонансное поведение этого коэффициента, который от практически нулевых значений при других энергиях, меньших высоты барьеров, возрастает в узкой энергетической области вблизи E_T до близкого к единице значения.

Описанная ситуация означает, что электроны практически не переносятся через РТС при энергиях меньших высоты барьеров и отличных от E_T , поскольку микроскопический поток j_z в состояниях с этими энергиями равен нулю, а значит равен нулю и парциальный вклад $n_z(E)$ этих состояний в полный поток n_z .

Перенос заряда в РТС осуществляется надбарьерными электронами и электронами с энергиями очень близкими к E_T (в идеальном случае - только с энергией равной E_T). Поскольку барьеры выбираются достаточно высокими (порядка 1 эВ), а энергия Ферми (E_F) много ниже высоты барьеров, то в соответствии со статистикой, над барьерами практически нет частиц, и

надбарьерным переносом заряда можно пренебречь. Если параметры барьеров и ямы согласованы с уровнем легирования контактных областей, то состояние с энергией E_T достаточно заполнено и практически весь поток электронов распространяется через это резонансное состояние (резонансный уровень).

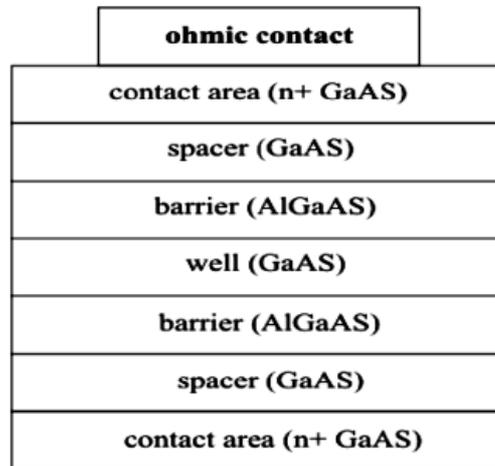


Рисунок 29. Многослойная GaAs/AlGaAs структура, реализующая РТС.

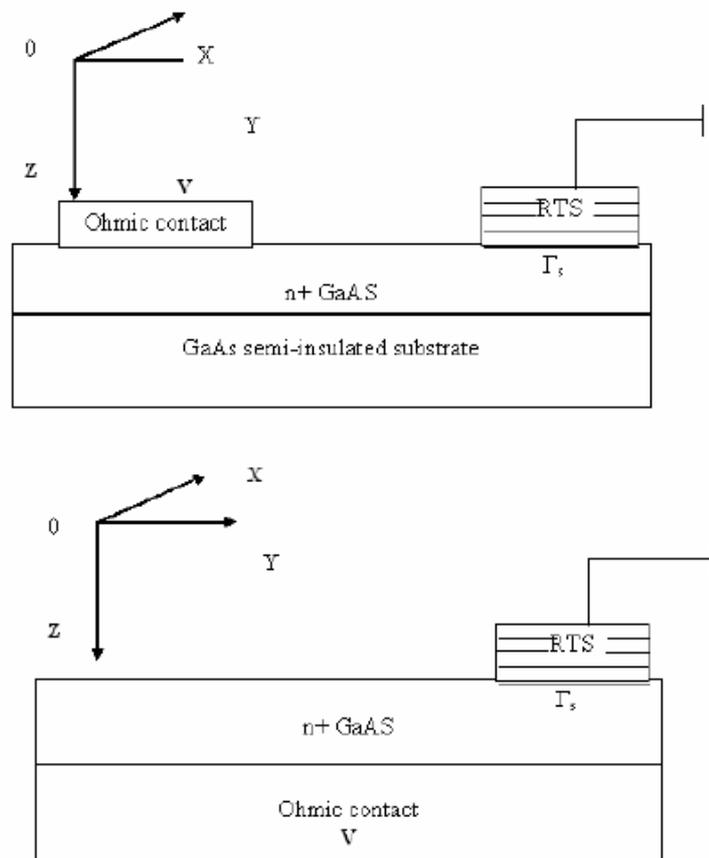


Рисунок 30. Два варианта РТД:

- РТС на полуизолирующей подложке;
- РТС на n+ - подложке.

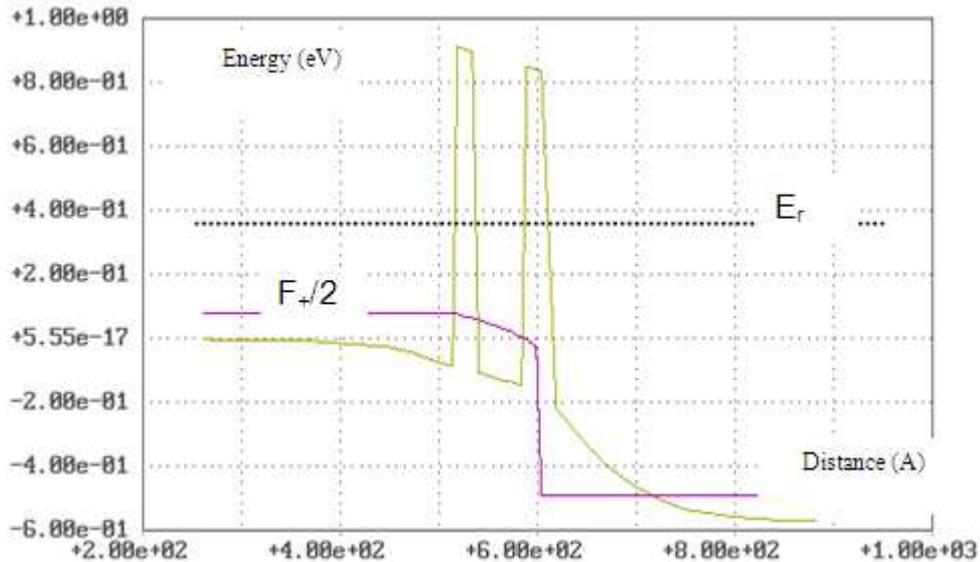


Рисунок 31. Потенциальный рельеф $U = U^{\text{ext}} - e\varphi$ и полусумма химических потенциалов $F_{+}/2$ «левых» и «правых» электронов для РТС с параметрами слоев: n+GaAs/500Å/GaAs/15Å/AlGaAs/20Å/GaAs/50Å/AlGaAs/20Å/GaAs/15Å/n+GaAs/500Å при смещении 0,6 В на границе РТС с подложкой.

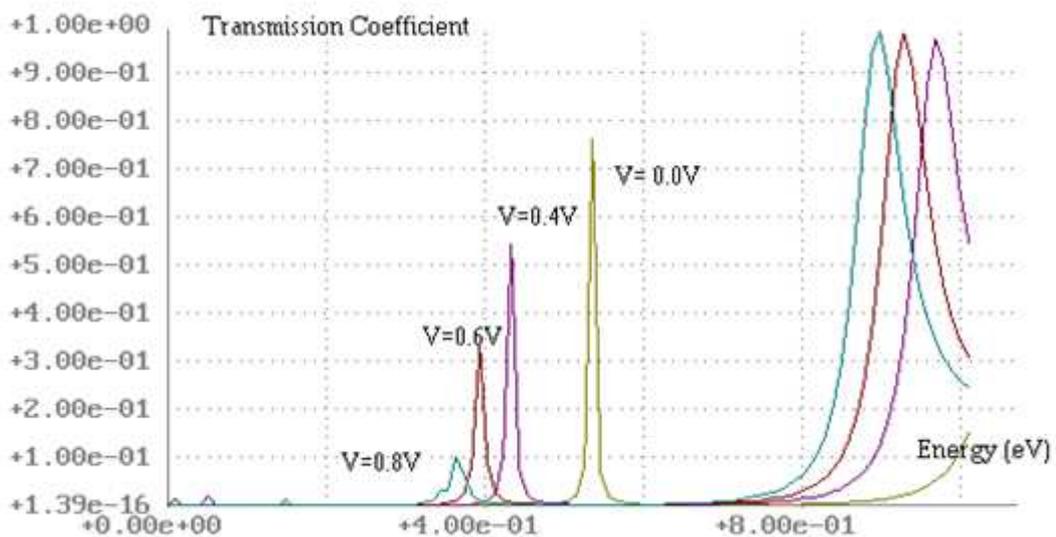


Рисунок 32. Коэффициент прохождения электронов через РТС (Рисунок 26) при различных напряжениях.

При изменении напряжения V потенциальный рельеф РТС искажается (см. Рисунок 31), и соответственно меняется зависимость от энергии коэффициента прохождения электронов (см. Рисунок 32). При некотором значении $V = V_p$ («напряжение пика» на Рисунок 25) резонанс становится настолько незначительным, что ток через структуру начинает уменьшаться с ростом напряжения. Дальнейший рост тока начинается при $V = V_v$ («напряжение долины» на Рисунок 25) и уже не связан с резонансным механизмом переноса.

5.1.2. Одномерная стационарная двухфазная модель переноса заряда

5.1.2.1. Интерпретируемые состояния

Горизонтальные размеры РТС (плоскость xOy на Рисунке 30), как правило, достаточно велики и составляют величины порядка микрометров. Эти размеры практически всегда много больше характерной длины размерного квантования L_{dq} (см. формулу (226)). При $L < L_{dq}$ расстояние между уровнями энергии квазичастиц в кристалле становится больше, чем $k_B T$, и их энергетический спектр уже нельзя считать квазинепрерывным [119,120]. Для электронов в GaAs при комнатной температуре (300^0K) L_{dq} приблизительно 25 нм (250 А). Для достижения подобных горизонтальных размеров необходимо использование нанолитографии [6-9,121,122]. Кроме того, в плоскости xOy потенциальный рельеф стандартной РТС практически однороден, а точнее не содержит неоднородностей с размерами меньшими характерных длин волн квазичастиц.

Эти два обстоятельства позволяют представить волновую функцию интерпретируемых состояний электронов $\Psi(\mathbf{r},\lambda)$ в виде

$$\Psi(\mathbf{r},\lambda) = \Psi(z,\lambda_z)\exp\{ip_x x/\hbar + ip_y y/\hbar\}, \quad (297)$$

где p_x и p_y – компоненты вектора квазиимпульса в соответствующих направлениях. Обобщенный индекс λ расшифровывается теперь следующим образом

$$\lambda = \{\sigma, p_x, p_y, \lambda_z\}.$$

Параметр σ может принимать значения $(1,-1)$, которые соответствуют двум возможным спиновым состояниям квазичастиц с идентичными другими квантовыми числами и тождественными с точностью до постоянного фазового множителя нерелятивистскими волновыми функциями. Этой тождественностью обусловлено то, что суммирование по σ квадратичных по волновым функциям выражений сводится к простому умножению на 2. Ниже будет определен точный смысл λ_z для интерпретируемых состояний.

Подстановка (297) в уравнение Шредингера из системы (220) дает

$$-(\hbar^2/2)d((1/m^*(z))d\Psi(z,\lambda_z)/dz)/dz + (U(z) + E_{\perp}(z) - E)\Psi(z,\lambda_z)=0. \quad (298)$$

Здесь учтено, что эффективная масса в гетероструктуре может быть различной в различных слоях и, следовательно, зависит от z , поэтому и

$$E_{\perp}(z) = (p_x^2 + p_y^2)/2m^*(z), \quad (299)$$

также зависит от этой координаты. Обычно при моделировании (см., например, [67,73,114]) это обстоятельство не берется в расчет, энергия E представляется в виде

$$E = E_{\perp} + E_z,$$

а E_{\perp} считается постоянной. Мы будем строить модель с учетом зависимости (299).

Обозначим через z_0 и z_V точки, расположенные в верхней и нижней n +GaAs - областях, примыкающих к спейсерам. Причем, z_0 будем выбирать непосредственно на границе с областью металлизации, а z_V – на некоторой условной границе РТС - подложка (граница Γ_s). Эта граница должна быть расположена достаточно далеко от потенциальных барьеров так, чтобы электроны на ней были равновесными, объемный заряд отсутствовал, а потенциал U был постоянным. Этого можно достичь, выбрав Γ_s на расстоянии, много большем, чем L_0 и L_D (длины релаксации неравновесных электронов и длина Дебая в легированном материале).

Для определенности будем далее считать, что напряжение, приложенное к омическому контакту с подложкой, неотрицательно и положим, что

$$E = E_{\perp}(z_0) + E_z. \quad (300)$$

В случае отрицательных напряжений на подложке для проведения численных расчетов удобно в выражении (300) выбрать $E_{\perp}(z_V)$ вместо $E_{\perp}(z_0)$, хотя принципиально этот выбор ничего не меняет.

С учетом соотношений (299) и (300) из уравнения (297) получим

$$-(\hbar^2/2)d((1/m^*(z))d\Psi(z,\lambda_z)/dz)/dz + (U^{\text{eff}}(z,p_{\perp}^2) - E_z)\Psi(z,\lambda_z)=0, \quad (301)$$

где

$$U^{\text{eff}}(z, p_{\perp}^2) = U(z) + \Delta E_{\perp}(z), \quad (302)$$

$$p_{\perp}^2 = p_x^2 + p_y^2,$$

$$\Delta E_{\perp}(z) = E_{\perp}(z) - E_{\perp}(z_0). \quad (303)$$

Очевидно, что в случае, когда эффективную массу можно считать не зависящей от z , $\Delta E_{\perp}(z)$ обращается в нуль, а $U^{\text{eff}}(z, p_{\perp}^2) = U(z)$.

В областях $z < z_0$ (левее z_0) и $z > z_V$ (правее z_V), которые будем называть левой и правой контактными областями к РТС, потенциал $U^{\text{eff}}(z, p_{\perp}^2)$ постоянен, а нетривиальное решение уравнения (301) $\Psi(z, \lambda_z)$ при $E_z > U^{\text{eff}}(z_0, p_{\perp}^2)$ и $E_z > U^{\text{eff}}(z_V, p_{\perp}^2)$ представляет собой некоторую суперпозицию плоских волн с квазиимпульсами $p_z(z_0, p_{\perp}^2)$ и $p_z(z_V, p_{\perp}^2)$

$$p_z(z_0, p_{\perp}^2) = (2m^*(z_0)(E_z - U^{\text{eff}}(z_0, p_{\perp}^2)))^{1/2},$$

$$p_z(z_V, p_{\perp}^2) = (2m^*(z_V)(E_z - U^{\text{eff}}(z_V, p_{\perp}^2)))^{1/2}.$$

В областях $z < z_0$ при $E_z < U^{\text{eff}}(z_0, p_{\perp}^2)$ и $z > z_V$ при $E_z < U^{\text{eff}}(z_V, p_{\perp}^2)$ нетривиальные решения уравнения (301) представляют собой суперпозиции затухающих и возрастающих с ростом z экспонент с декрементами

$$q_z(z_0, p_{\perp}^2) = (2m^*(z_0)(U^{\text{eff}}(z_0, p_{\perp}^2) - E_z))^{1/2},$$

$$q_z(z_V, p_{\perp}^2) = (2m^*(z_V)(U^{\text{eff}}(z_V, p_{\perp}^2) - E_z))^{1/2}.$$

Определим набор интерпретируемых состояний для решаемой задачи. Вид волновых функций электронов при различных E_z в области $z_0 < z < z_V$ определяется уравнением (301) и граничными условиями в точках z_0 и z_V . Интерпретировать решения внутри РТС можно только связав их через граничные условия с теми решениями в левой и правой контактных областях, физическая интерпретация которых абсолютно прозрачна. В предельном случае постоянного потенциального рельефа мы обязаны получить в контактных областях и во всей структуре состояния, соответствующие свободному электронному газу, то есть плоские волны единичной амплитуды, распространяющиеся вправо и влево вдоль оси Oz , для каждого значения $E_z > U^{\text{eff}}$ и нулевые (тривиальные) решения для $E_z < U^{\text{eff}}$.

Исходя из сказанного, интерпретируемые состояния в рассматриваемой структуре можно разбить на две группы:

1) решения уравнения (301), относящиеся к непрерывному спектру и порождаемые плоскими волнами единичной амплитуды, распространяющимися при $E_z > U^{\text{eff}}$ из контактных областей в сторону РТС;

2) решения уравнения (301), относящиеся к дискретному спектру и возникающие только в том случае, если внутри области $z_0 < z < z_V$ имеются участки, где $U^{\text{eff}}(z, p_{\perp}^2) < \min\{U^{\text{eff}}(z_0, p_{\perp}^2), U^{\text{eff}}(z_V, p_{\perp}^2)\}$.

Будем называть состояния непрерывного спектра «левыми» и «правыми» в зависимости от того, из какой контактной области левой или правой распространяется исходная волна, а соответствующие волновые функции обозначать как $\Psi_l(E_z, p_{\perp}^2, z)$ и $\Psi_r(E_z, p_{\perp}^2, z)$. Волновые функции состояний дискретного спектра обозначим через $\Psi_d(E_z, p_{\perp}^2, z)$.

Рассмотрим сначала «левые» и «правые» состояния. Если $E_z > U^{\text{eff}}(z_0, p_{\perp}^2)$, то при $z < z_0$ волновые функции таких состояний будут иметь вид

$$\Psi_l(E_z, p_{\perp}^2, z) = \exp\{ip_z(z_0, p_{\perp}^2)z/\hbar\} + A_l(E_z, p_{\perp}^2)\exp\{-ip_z(z_0, p_{\perp}^2)z/\hbar\}.$$

Соответственно, при $z > z_V$ для этих же состояний выбираем либо прошедшую волну

$$\Psi_l(E_z, p_{\perp}^2, z) = B_l(E_z, p_{\perp}^2) \exp\{i p_z(z_V, p_{\perp}^2)z/\hbar\},$$

если $E_z > U^{\text{eff}}(z_V, p_{\perp}^2)$, либо затухающее с возрастанием z решение

$$\Psi_l(E_z, p_{\perp}^2, z) = B_l(E_z, p_{\perp}^2) \exp\{-q_z(z_V, p_{\perp}^2)z/\hbar\},$$

если $E_z < U^{\text{eff}}(z_V, p_{\perp}^2)$.

Аналогично, для «правых» состояний, если $E_z > U^{\text{eff}}(z_V, p_{\perp}^2)$, то при $z > z_V$ волновые функции будут иметь вид

$$\Psi_r(E_z, p_{\perp}^2, z) = \exp\{-i p_z(z_V, p_{\perp}^2)z/\hbar\} + A_r(E_z, p_{\perp}^2)\exp\{i p_z(z_V, p_{\perp}^2)z/\hbar\}.$$

При $z < z_0$ для этих же состояний выбираем либо прошедшую волну

$$\Psi_r(E_z, p_{\perp}^2, z) = B_r(E_z, p_{\perp}^2) \exp\{-ip_z(z_0, p_{\perp}^2)z/\hbar\},$$

если $E_z > U^{\text{eff}}(z_0, p_{\perp}^2)$, либо затухающее с убыванием z решение

$$\Psi_r(E_z, p_{\perp}^2, z) = B_r(E_z, p_{\perp}^2)\exp\{q_z(z_0, p_{\perp}^2)z/\hbar\},$$

если $E_z < U^{\text{eff}}(z_0, p_{\perp}^2)$.

В соответствии с граничными условиями (288) в точках z_0 и z_V для сколь угодно малого положительного ε имеем

$$\begin{aligned}\Psi_{l,r}(E_z, p_{\perp}^2, z_0 + \varepsilon) &= \Psi_{l,r}(E_z, p_{\perp}^2, z_0 - \varepsilon), \quad \Psi_{l,r}(E_z, p_{\perp}^2, z_V + \varepsilon) = \Psi_{l,r}(E_z, p_{\perp}^2, z_V - \varepsilon), \\ d\Psi_{l,r}(E_z, p_{\perp}^2, z_0 + \varepsilon)/dz &= d\Psi_{l,r}(E_z, p_{\perp}^2, z_0 - \varepsilon)/dz, \\ d\Psi_{l,r}(E_z, p_{\perp}^2, z_V + \varepsilon)/dz &= d\Psi_{l,r}(E_z, p_{\perp}^2, z_V - \varepsilon)/dz,\end{aligned}$$

откуда с учетом выписанных выше решений в контактных областях получаем в точке z_0

$$\begin{cases} ip_z(z_0, p_{\perp}^2) \Psi_l(E_z, p_{\perp}^2, z_0) + \hbar d\Psi_l(E_z, p_{\perp}^2, z_0)/dz = 2ip_z(z_0, p_{\perp}^2) \exp\{ip_z(z_0, p_{\perp}^2)z_0/\hbar\}, \\ \hbar d\Psi_r(E_z, p_{\perp}^2, z_0)/dz + ip_z(z_0, p_{\perp}^2) \Psi_r(E_z, p_{\perp}^2, z_0) = 0, \end{cases} \quad (304.1)$$

при $E_z > U^{\text{eff}}(z_0, p_{\perp}^2)$,

$$\begin{cases} \Psi_l(E_z, p_{\perp}^2, z_0) = 0, \\ \hbar d\Psi_r(E_z, p_{\perp}^2, z_0)/dz - q_z(z_0, p_{\perp}^2) \Psi_r(E_z, p_{\perp}^2, z_0) = 0, \end{cases} \quad (304.2)$$

при $E_z < U^{\text{eff}}(z_0, p_{\perp}^2)$ и $E_z > U^{\text{eff}}(z_V, p_{\perp}^2)$;

в точке z_V

$$\begin{cases} ip_z(z_V, p_{\perp}^2) \Psi_r(E_z, p_{\perp}^2, z_V) - \hbar d\Psi_r(E_z, p_{\perp}^2, z_V)/dz = 2ip_z(z_V, p_{\perp}^2) \exp\{-ip_z(z_V, p_{\perp}^2)z_V/\hbar\}, \\ \hbar d\Psi_l(E_z, p_{\perp}^2, z_V)/dz - ip_z(z_V, p_{\perp}^2) \Psi_l(E_z, p_{\perp}^2, z_V) = 0, \end{cases} \quad (305.1)$$

при $E_z > U^{\text{eff}}(z_V, p_{\perp}^2)$;

$$\begin{cases} \Psi_r(E_z, p_{\perp}^2, z_V) = 0, \\ \hbar d\Psi_l(E_z, p_{\perp}^2, z_V)/dz + q_z(z_V, p_{\perp}^2) \Psi_l(E_z, p_{\perp}^2, z_V) = 0, \end{cases} \quad (305.2)$$

при $E_z < U^{\text{eff}}(z_V, p_{\perp}^2)$ и $E_z > U^{\text{eff}}(z_0, p_{\perp}^2)$.

Уравнение (301) и граничные условия (304), (305) позволяют полностью определить «левые» и «правые» интерпретируемые состояния квазичастиц (электронов) для одномерной задачи. Нетрудно видеть, что выбранная нормировка волновых функций соответствует ситуации, когда при однородном потенциальном рельефе решение задачи (301), (304) и (305) дает плоские волны единичной амплитуды (коэффициенты $A_{l,r}$ и $B_{l,r}$ равны нулю в этом случае), что приводит к стандартным волновым функциям квазичастиц в объемном материале.

Для нахождения состояний дискретного спектра необходимо решить соответствующую задачу для уравнения (301). Для каждого значения E_z , удовлетворяющего неравенству

$$\min\{U^{\text{eff}}(z, p_{\perp}^2)\} \leq E_z \leq \min\{U^{\text{eff}}(z_0, p_{\perp}^2), U^{\text{eff}}(z_V, p_{\perp}^2)\},$$

будем решать уравнение (301) с граничными условиями

$$\begin{cases} \hbar d\Psi_d(E_z, p_{\perp}^2, z_V)/dz + q_z(z_V, p_{\perp}^2)\Psi_d(E_z, p_{\perp}^2, z_V) = 0, \\ \hbar d\Psi_d(E_z, p_{\perp}^2, z_0)/dz - q_z(z_0, p_{\perp}^2)\Psi_d(E_z, p_{\perp}^2, z_0) = 0, \end{cases} \quad (306)$$

соответствующими «сшиванию» в точках $z = z_{0,V}$ решения $\Psi_d(E_z, p_{\perp}^2, z)$ внутри области $z \in [z_0, z_V]$ с решениями уравнения (301), экспоненциально затухающими вне этой области, пропорциональными $\exp\{\pm q_z(z_{0,V}, p_{\perp}^2)z\}$. Под «сшиванием» подразумевается требование непрерывности волновой функции и ее первой производной, обеспечивающее сохранение потока плотности вероятности.

Нетривиальные решения уравнения (301) $\Psi_d(E_z, p_{\perp}^2, z)$ будут удовлетворять граничным условиям (306) только при некоторых значениях энергии E_z

$$E_z = E_z(k),$$

которые определяют энергетический спектр дискретных состояний. Индекс k принимает положительные целые значения от 1 до K и нумерует связанные состояния в порядке возрастания их энергии. Значение $k = 1$ соответствует состоянию с минимальной возможной энергией. Волновые функции задачи (301), (306) необходимо нормировать на «одну частицу»

$$\int_{-\infty}^{+\infty} |\Psi_d(E_z(k), p_{\perp}^2, z)|^2 dz = 1.$$

С учетом описанной процедуры выбора интерпретируемых состояний обобщенный индекс λ_z может быть расшифрован как

$$\lambda_z = \{E_z, p_{\perp}^2, \text{dir}\}, \text{dir} = (l, r, d)$$

и соответственно

$$\lambda = \{\sigma, p_x, p_y, E_z, \text{dir}\}.$$

Удобно также записать

$$\lambda = \{\sigma, p_x, p_y, E_z, \text{dir}\} = \{\lambda_l(\sigma, p_x, p_y, E_z), \lambda_r(\sigma, p_x, p_y, E_z), \lambda_d(\sigma, p_x, p_y, E_z)\}.$$

Таким образом, мы полностью определили интерпретируемые состояния электронов для решаемой задачи как квазичастиц, характеризуемых проекцией σ спина на некоторое направление, значениями двух компонент вектора квазиимпульса p_x и p_y , энергией E_z , направлением движения «падающей» волны вдоль оси Oz в контактных областях

$$\text{dir} = l \text{ или } r,$$

либо принадлежностью к дискретному спектру

$$\text{dir} = d.$$

Квадраты модулей коэффициентов $A_{l,r}$

$$R_{l,r}(E_z, p_{\perp}^2) = |A_{l,r}(E_z, p_{\perp}^2)|^2,$$

имеют физический смысл коэффициентов отражения «левых» и «правых» электронов от потенциального рельефа, создаваемого РТС. По определению коэффициент отражения равен отношению отраженного потока к падающему потоку. В случае, когда $E_z > U^{\text{eff}}(z_0, p_{\perp}^2)$ и $E_z > U^{\text{eff}}(z_V, p_{\perp}^2)$ справедливо равенство [123]

$$R_l(E_z, p_{\perp}^2) = R_r(E_z, p_{\perp}^2) = R(E_z, p_{\perp}^2),$$

а для потоков справедливы выражения

$$\begin{cases} j_l = p_z(z_0, p_{\perp}^2)(1 - R(E_z, p_{\perp}^2))/m^*(z_0) = p_z(z_V, p_{\perp}^2)|B_l(E_z, p_{\perp}^2)|^2/m^*(z_V), \\ j_r = -p_z(z_V, p_{\perp}^2)(1 - R(E_z, p_{\perp}^2))/m^*(z_V) = -p_z(z_0, p_{\perp}^2)|B_r(E_z, p_{\perp}^2)|^2/m^*(z_0). \end{cases} \quad (307)$$

Коэффициент прохождения $T(E_z, p_{\perp}^2)$ электронов через потенциальный рельеф определяется как отношение прошедшего потока к падающему потоку, откуда следуют равенства

$$T_l(E_z, p_{\perp}^2) = T_r(E_z, p_{\perp}^2) = T(E_z, p_{\perp}^2) = 1 - R(E_z, p_{\perp}^2),$$

а также соотношения для $B_{l,r}(E_z, p_{\perp}^2)$

$$|B_l(E_z, p_{\perp}^2)|^2 = p_z(z_0, p_{\perp}^2) m^*(z_V) T(E_z, p_{\perp}^2) / (p_z(z_V, p_{\perp}^2) m^*(z_0)),$$

$$|B_r(E_z, p_{\perp}^2)|^2 = p_z(z_V, p_{\perp}^2) m^*(z_0) T(E_z, p_{\perp}^2) / (p_z(z_0, p_{\perp}^2) m^*(z_V)).$$

Следует иметь в виду, что если для некоторого E_z справедливо хотя бы одно из неравенств

$$E_z < U^{\text{eff}}(z_0, p_{\perp}^2) \text{ или } E_z < U^{\text{eff}}(z_v, p_{\perp}^2),$$

то оба потока j_l и j_r равны нулю. Очевидно, что равны нулю и пространственные потоки для состояний дискретного спектра (d-состояний).

В заключение этого пункта отметим, что нигде при формулировке уравнений, граничных условий и выборе интерпретируемых состояний никак не использовались специфические особенности РТС. Поэтому все, что здесь изложено, имеет прямое отношение не только к РТС, но и к любой квантовой структуре, которую можно с определенной точностью описать в рамках одномерного приближения.

5.1.2.2. Уравнения для химических потенциалов

Сделаем еще несколько предположений. Во-первых, будем считать, что для всех частиц, соответствующих «левым» электронным состояниям, химический потенциал одинаков и равен F_l , а для всех частиц, соответствующих «правым» электронным состояниям, он также одинаков и равен F_r , и, наконец, для всех частиц, соответствующих состояниям дискретного спектра, он одинаков и равен F_d . Таким образом, «левые», «правые» и d-электроны представляют собой три разные фазы электронного газа.

Сказанное означает, что мы пренебрегли возможной неравновесностью между электронами, составляющими «левую», «правую» и d-фазы, а учитываем только возможную неравновесность между ансамблями «левых», «правых» и d-электронов. В этом случае из уравнений (218) получим

$$\left\{ \begin{array}{l} dI_l(z)/dz = -\hbar^{-1}(F_l(z) - F_r(z))\sum_{\lambda_l}\sum_{\lambda_r}\{n(z,\lambda_l)\Gamma_{\lambda_l,\lambda_r} + n(z,\lambda_r)\Gamma_{\lambda_r,\lambda_l}\} - \\ \quad -\hbar^{-1}(F_l(z) - F_d(z))\sum_{\lambda_l}\sum_{\lambda_d}\{n(z,\lambda_l)\Gamma_{\lambda_l,\lambda_d} + n(z,\lambda_d)\Gamma_{\lambda_d,\lambda_l}\}, \\ dI_r(z)/dz = -\hbar^{-1}(F_r(z) - F_l(z))\sum_{\lambda_r}\sum_{\lambda_l}\{n(z,\lambda_r)\Gamma_{\lambda_r,\lambda_l} + n(z,\lambda_l)\Gamma_{\lambda_l,\lambda_r}\} - \\ \quad -\hbar^{-1}(F_r(z) - F_d(z))\sum_{\lambda_r}\sum_{\lambda_d}\{n(z,\lambda_r)\Gamma_{\lambda_r,\lambda_d} + n(z,\lambda_d)\Gamma_{\lambda_d,\lambda_r}\}, \\ 0 = -\hbar^{-1}(F_d(z) - F_l(z))\sum_{\lambda_d}\sum_{\lambda_l}\{n(z,\lambda_d)\Gamma_{\lambda_d,\lambda_l} + n(z,\lambda_l)\Gamma_{\lambda_l,\lambda_d}\} - \\ \quad -\hbar^{-1}(F_d(z) - F_r(z))\sum_{\lambda_d}\sum_{\lambda_r}\{n(z,\lambda_d)\Gamma_{\lambda_d,\lambda_r} + n(z,\lambda_r)\Gamma_{\lambda_r,\lambda_d}\}, \end{array} \right. \quad (308)$$

где мы учли, что плотность потока d-электронов равна нулю. В уравнениях (308) и далее при одномерном описании будем обозначать пространственную плотность потока электронов буквой I. В данном случае

$$I = n_z,$$

$$I_{l,r}(z) = \sum_{\lambda_{l,r}} j_{l,r}(\lambda_{l,r}) f(s(z, \lambda_{l,r})).$$

Для остальных величин сохранены принятые ранее обозначения

$$\left\{ \begin{array}{l} n_{l,r,d}(z) = \sum_{\lambda_{l,r,d}} n_{l,r,d}(z, \lambda_{l,r,d}), \quad n_{l,r,d}(z, \lambda_{l,r,d}) = \rho_{l,r,d}(\lambda_{l,r,d}) f(s(z, \lambda_{l,r,d})), \\ \lambda_{l,r,d} = \{\sigma, p_x, p_y, E_z, \text{dir}\}, \quad \text{dir} = l, r, d, \\ j_l(E_z, p_{\perp}^2) = p_z(z_0, p_{\perp}^2) T(E_z, p_{\perp}^2) / m^*(z_0), \\ j_r(E_z, p_{\perp}^2) = -p_z(z_V, p_{\perp}^2) T(E_z, p_{\perp}^2) / m^*(z_V), \\ \rho_{l,r,d}(E_z, p_{\perp}^2, z) = \Psi_{l,r,d}^*(E_z, p_{\perp}^2, z) \Psi_{l,r,d}(E_z, p_{\perp}^2, z), \\ f(s) = (\exp(s) + 1)^{-1}, \\ s_{l,r}(E_z, p_{\perp}^2, z) = \beta \{E_z + E_{\perp}(p_{\perp}^2, z) - F_{l,r}(z) + \tau v_{l,r}(E_z, p_{\perp}^2, z) dF_{l,r}/dz\}, \\ v_{l,r}(E_z, p_{\perp}^2, z) = j_r(E_z, p_{\perp}^2) / \rho_{l,r}(E_z, p_{\perp}^2, z), \\ s_d(E_z, p_{\perp}^2, z) = \beta \{E_z(k) + E_{\perp}(p_{\perp}^2, z) - F_d(z)\}. \end{array} \right. \quad (309)$$

Последнее уравнение в системе (308) является алгебраическим и легко разрешается относительно F_d

$$F_d(z) = [F_l(z) \sum_{\lambda_d} \sum_{\lambda_l} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l} + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d}\} + \\ + F_r(z) \sum_{\lambda_d} \sum_{\lambda_r} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r} + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d}\}] / [\sum_{\lambda_d} \sum_{\lambda_l} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l} + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d}\} + \\ + \sum_{\lambda_d} \sum_{\lambda_r} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r} + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d}\}]. \quad (310)$$

Подстановка (310) в оставшиеся два уравнения системы (308) приводит к появлению слагаемых, пропорциональных $F_l - F_r$, вместо слагаемых, пропорциональных $F_{l,r} - F_d$

$$\left\{ \begin{array}{l} dI_l(z)/dz = -\hbar^{-1} (F_l(z) - F_r(z)) \sum_{\lambda_l} \sum_{\lambda_r} \{n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_r} + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_l}\} - \\ - \hbar^{-1} (F_l(z) - F_r(z)) [\sum_{\lambda_l} \sum_{\lambda_d} \{n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d} + n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l}\} \sum_{\lambda_d} \sum_{\lambda_r} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r} \\ + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d}\}] / [\sum_{\lambda_d} \sum_{\lambda_l} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l} + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d}\} + \\ + \sum_{\lambda_d} \sum_{\lambda_r} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r} + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d}\}], \\ dI_r(z)/dz = -\hbar^{-1} (F_r(z) - F_l(z)) \sum_{\lambda_r} \sum_{\lambda_l} \{n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_l} + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_r}\} - \\ - \hbar^{-1} (F_r(z) - F_l(z)) [\sum_{\lambda_r} \sum_{\lambda_d} \{n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d} + n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r}\} \sum_{\lambda_d} \sum_{\lambda_l} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l} \\ + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d}\}] / [\sum_{\lambda_d} \sum_{\lambda_l} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l} + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d}\} + \\ + \sum_{\lambda_d} \sum_{\lambda_r} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r} + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d}\}]. \end{array} \right. \quad (311)$$

Отметим, что множители при F_1 и F_r в первом, и втором уравнениях системы (311) отрицательны. С учетом формул для $i_{1,r}$ в (309) это обстоятельство обеспечивает не осциллирующий характер решений уравнений (311) относительно химических потенциалов.

С целью дальнейшего упрощения предположим дополнительно, что вероятности переходов между d-состояниями и состояниями «левых» и «правых» электронов много меньше, чем вероятности переходов между «левыми» и «правыми» электронными состояниями. Это предположение позволяет пренебречь последними слагаемыми в правых частях уравнений (311).

Если положить, что

$$\Gamma_{\lambda_d, \lambda_l} = \Gamma_{\lambda_d, \lambda_r} = \Gamma_{\lambda_l, \lambda_d} = \Gamma_{\lambda_r, \lambda_d} = 0,$$

и совсем не учитывать такие переходы, то последнее из уравнений (308) (или уравнение (310)) будет удовлетворяться тождественно при любых значениях F_d . Это вполне понятно, поскольку именно переходы между различными электронными фазами делают их химические потенциалы взаимосвязанными. Однако физически величина F_d не может быть произвольной, поскольку ею определяется концентрация электронов в связанных состояниях. Поэтому мы будем пренебрегать величинами $\Gamma_{\lambda_d, \lambda_l}$, $\Gamma_{\lambda_d, \lambda_r}$, $\Gamma_{\lambda_l, \lambda_d}$ и $\Gamma_{\lambda_r, \lambda_d}$ в уравнениях для $F_{1,r}$, но будем считать их все же отличными от нуля.

Формулу (310) затруднительно применить для практических расчетов F_d и концентраций электронов, поскольку входящие в нее коэффициенты Γ вычислить сложно. Однако в большинстве задач нет причин, по которым нарушалась бы симметрия между «левыми» и «правыми» состояниями по отношению к переходам электронов в состояния дискретного спектра и обратно. Поэтому положим, что справедливо «детальное равновесие»

$$\sum_{\lambda_d} \sum_{\lambda_l} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_l} + n(z, \lambda_l) \Gamma_{\lambda_l, \lambda_d}\} = \sum_{\lambda_d} \sum_{\lambda_r} \{n(z, \lambda_d) \Gamma_{\lambda_d, \lambda_r} + n(z, \lambda_r) \Gamma_{\lambda_r, \lambda_d}\}.$$

Тогда из (310) получим

$$F_d = (F_1 + F_r)/2 = F_+. \quad (312)$$

Сделанные допущения правомерны, как первое приближение, поскольку разность химических потенциалов «левых» и «правых» электронов, обусловленная ненулевым приложенным напряжением, является основным процессом, определяющим особенности переноса заряда в одномерной квантовой структуре. Приложенное напряжение будет порождать неравновесность и между электронами, движущимися в одном направлении. Однако в рамках предлагаемой простейшей двухфазной модели мы пренебрегаем этими процессами.

Неравновесность электронов непрерывного и дискретного спектров будем учитывать упрощенно, основываясь на формуле (312). Название «двухфазная модель» обусловлено тем, что третья фаза, соответствующая электронам дискретного спектра, при сделанных предположениях непосредственно в переносе заряда не участвует. Однако плотность заряда, создаваемая электронами в связанных состояниях будет учитываться при решении уравнения Пуассона, то есть будет давать вклад в самосогласованный электростатический потенциал.

Пренебрежение возможными переходами электронов из связанных d -состояний в состояния непрерывного спектра обусловлено стремлением сосредоточиться на простейшей модели, позволяющей описывать неравновесные квантовые процессы. Однако нет никаких принципиальных трудностей для того, чтобы включить в рассмотрение и третью d -фазу.

С учетом сделанных предположений и в соответствии с формулой (269) справедливы равенства

$$\sum_{\lambda_r} \Gamma_{\lambda_l, \lambda_r} = \sum_{\lambda_l} \Gamma_{\lambda_r, \lambda_l} = 1.$$

В результате вместо уравнений (209) для средних потоков «левых» и «правых» электронов получим

$$\begin{cases} dI_l(z)/dz = -\hbar^{-1}(F_l(z) - F_r(z))(n_l(z) + n_r(z)), \\ dI_r(z)/dz = -\hbar^{-1}(F_r(z) - F_l(z))(n_l(z) + n_r(z)). \end{cases} \quad (313)$$

Это и есть окончательные уравнения переноса заряда двухфазной модели.

Для моделирования характеристик РТС в точках z_0 и z_V необходимо поставить граничные условия на F_l и F_r . Точка z_0 является типичной границей

типа Γ_1 , на которой согласно (96) и принятым выше соглашениям справедливо условие

$$F_l(z_0) = F_r(z_0) = 0. \quad (314.1)$$

Точка z_V принадлежит границе Γ_s , которая удалена на расстояние, много большее чем L_0 , от активной области РТС и, поэтому, также является для РТС границей типа Γ_1

$$F_l(z_V) = F_r(z_V) = -eV. \quad (314.2)$$

5.1.2.3. Выражения для плотностей числа частиц и потоков

Суммирование по $\lambda_{l,r}$ в формулах (309) для средних значений плотностей числа частиц и их потоков сводится к суммированию по спиновым состояниям, приводящему к появлению множителя 2, интегрированию по всем возможным значениям p_x и p_y , а также интегрированию по всем возможным значениям импульса p_z :

$$\begin{aligned} \sum_{\lambda_{l,r}} \dots &= \sum_{\sigma} (2\pi\hbar)^{-3} \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int dp_z (\dots) = 2(2\pi\hbar)^{-3} \int_0^{\infty} p_{\perp} p_{\perp} \int_0^{2\pi} d\phi \int dp_z (\dots) = \\ &= 2\pi(2\pi\hbar)^{-3} \int_0^{\infty} dp_{\perp}^2 \int dp_z (\dots) = 2\pi (m^*(z_{\min}))^{1/2} (2\pi\hbar)^{-3} \int_0^{\infty} dp_{\perp}^2 \int_{U^{\min}(p_{\perp}^2)}^{U^{\max}(p_{\perp}^2)} dE_z (\dots) / (2(E_z - U^{\min}(p_{\perp}^2)))^{1/2}. \end{aligned}$$

Величинами $U^{\min}(p_{\perp}^2)$ и $U^{\max}(p_{\perp}^2)$ определяются нижний и верхний пределы интегрирования по энергии E_z . Очевидно, что эти величины будут разными для электронов, относящихся к различным фазам. Учитывая определения, сделанные при выборе интерпретируемых состояний, получим

$$\begin{aligned} U^{\min}_l(p_{\perp}^2) &= U^{\text{eff}}(z_0, p_{\perp}^2), \quad U^{\max}_l(p_{\perp}^2) = \infty, \quad z_{\min,l} = z_0, \\ U^{\min}_r(p_{\perp}^2) &= U^{\text{eff}}(z_V, p_{\perp}^2), \quad U^{\max}_r(p_{\perp}^2) = \infty, \quad z_{\min,r} = z_V. \end{aligned} \quad (315)$$

Для состояний дискретного спектра имеем аналогичное выражение, учитывающее условие нормировки волновых функций

$$\sum_{\lambda_d} \dots = \sum_{\sigma} (2\pi\hbar)^{-2} \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \sum_{k=1}^K (\dots) = 2\pi(2\pi\hbar)^{-2} \int_0^{\infty} dp_{\perp}^2 \sum_{k=1}^K (\dots).$$

В результате для средних значений локальных плотностей числа частиц и потоков получим выражения

$$n_{l,r}(z) = 2\pi(m^*(z_{\min,l,r}))^{1/2}(2\pi\hbar)^{-3} \times \int_0^\infty dp_\perp^2 \int_{U^{\min}_{l,r}(p_\perp^2)}^\infty dE_z \Psi_{l,r}^*(E_z, p_\perp^2, z) \Psi_{l,r}(E_z, p_\perp^2, z) f(s_{l,r}(E_z, p_\perp^2, z)) / (2(E_z - U^{\min}_{l,r}(p_\perp^2)))^{1/2}, \quad (316)$$

$$n_d(z) = 2\pi(2\pi\hbar)^{-2} \times \int_0^\infty dp_\perp^2 \sum_{k=1}^K \Psi_d^*(E_z(k), p_\perp^2, z) \Psi_d(E_z(k), p_\perp^2, z) f(s_d(E_z, p_\perp^2, z)), \quad (317)$$

$$I_l = 2\pi(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_{U^{\text{eff}}(z_0, p_\perp^2)}^\infty dE_z T(E_z, p_\perp^2) f(s_l(E_z, p_\perp^2, z)), \quad (318)$$

$$I_r = -2\pi(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_{U^{\text{eff}}(z_v, p_\perp^2)}^\infty dE_z T(E_z, p_\perp^2) f(s_r(E_z, p_\perp^2, z)). \quad (319)$$

Следует иметь в виду, что поскольку коэффициент прохождения $T(E_z, p_\perp^2)$ равен нулю для всех состояний с энергией меньше, чем

$$U^{\text{cur}}(p_\perp^2) = \max\{U^{\text{eff}}(z_0, p_\perp^2), U^{\text{eff}}(z_v, p_\perp^2)\}, \quad (320)$$

то в выражениях (318) и (319) нижний предел интегрирования по E_z можно заменить на $U^{\text{cur}}(p_\perp^2)$. Важно, что выражения для потоков, в отличие от хорошо известной формулы Тсу-Есаки, являются локальными. Для полного потока получим

$$I = I_l + I_r = 2\pi(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_{U^{\text{cur}}(p_\perp^2)}^\infty dE_z T(E_z, p_\perp^2) \{f(s_l(E_z, p_\perp^2, z)) - f(s_r(E_z, p_\perp^2, z))\}. \quad (321)$$

Когда зависимостью E_\perp от z (см. соотношение (299)) можно пренебречь, коэффициент прохождения T перестает зависеть от p_\perp^2 , и интегралы по p_\perp^2 в выражениях (316) – (319) и (321) можно снять аналитически. В этом случае функция распределения примет привычный для одномерной модели логарифмический вид [14,67]. В частности, для суммарного потока из (321) получим

$$I = (m^* k_B T / 2\pi^2 \hbar^3) \int dE_z T(E_z) \ln \left\{ \frac{1 + \exp\{-s_l(E_z, z)\}}{1 + \exp\{-s_r(E_z, z)\}} \right\} U^{\text{cur}}(p_{\perp}^2).$$

5.1.2.4. Уравнение Пуассона в РТС

Потенциальный рельеф для электронов в уравнении Шредингера (301) определяется в соответствии с выражением (302) как

$$U^{\text{eff}}(z, p_{\perp}^2) = U(z) + \Delta E_{\perp}(z).$$

В свою очередь, потенциал $U(z)$ представляет собой сумму внешнего потенциала U^{ext} , обусловленного разрывом зон на гетерограницах слоев, формирующих РТС, и самосогласованного электростатического потенциала ϕ

$$U(z) = U^{\text{ext}}(z) - e\phi(z).$$

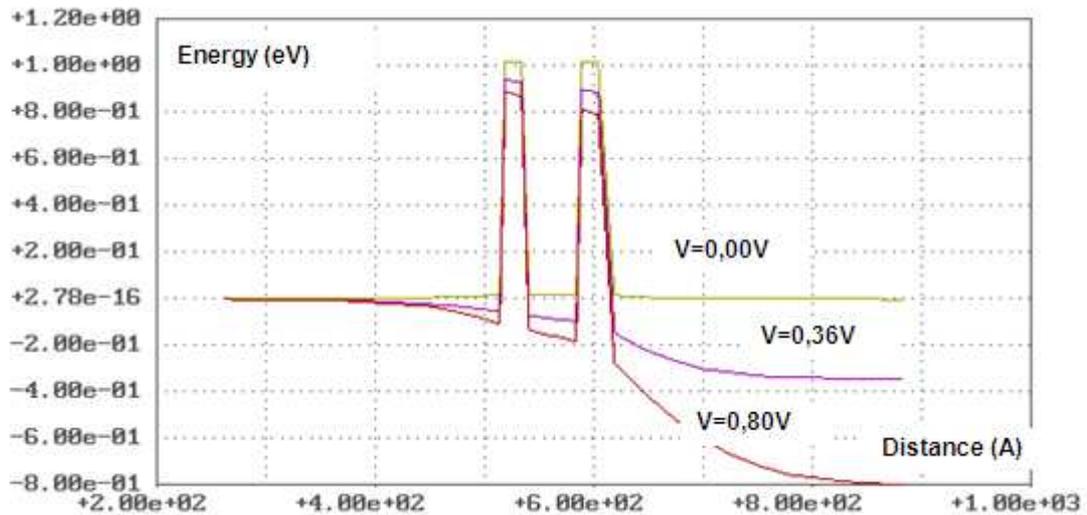


Рисунок 33. Потенциальные рельефы РТС при различных приложенных напряжениях ($V = 0,00 \text{ В}; 0,36 \text{ В}; 0,80 \text{ В}$).

В одномерном случае для ϕ из уравнения Пуассона (5) получим

$$d(\epsilon d\phi/dz)/dz = e(n - N_{\text{int}}). \quad (322)$$

В этом уравнении n – локальная концентрация электронов:

$$n = n_l + n_r + n_d. \quad (323)$$

Слагаемые в выражении (323) определены соотношениями (316) и (317), пределы интегрирования в которых зависят от значений $U^{\text{eff}}(z, p_{\perp}^2)$ в точках z_0 , z_V и $z_{\text{min},d}$.

Таким образом, концентрация n неявно зависит от ϕ как через пределы интегрирования в формулах (316) и (317), так и через волновые функции $\Psi_1(E_z, p_\perp^2, z)$, $\Psi_r(E_z, p_\perp^2, z)$ и $\Psi_d(E_z, p_\perp^2, z)$, квадраты модулей которых фигурируют в выражениях (316) и (317).

Поскольку точки z_0 и z_V расположены на расстояниях, много больших длины Дебая L_D , от активной области РТС в качестве граничных условий для ϕ мы выбираем условия электронейтральности (97) в этих точках

$$\begin{cases} n(z_0) - N_{\text{int}}(z_0) = 0, \\ n(z_V) - N_{\text{int}}(z_V) = 0. \end{cases} \quad (324)$$

В отличие от классической дрейфово-диффузионной модели, где электростатический потенциал ϕ входит в функции распределения носителей заряда явным образом, неявная зависимость концентрации n от ϕ , характерная для квантовой теории, усложняет процедуру определение потенциала в граничных точках. Тем не менее, краевая задача (322), (324) является полностью определенной и однозначно разрешимой при заданных внешних параметрах.

Подчеркнем, что электростатический потенциал ϕ вносит решающий вклад в общий потенциальный рельеф для электронов в РТС при ненулевых смещениях. Это наглядно видно при сравнении потенциалов U с одинаковыми U^{ext} при различных приложенных напряжениях, которые отличаются только вкладом слагаемого $-e\phi$ (см. Рисунок 33). Именно электростатическим потенциалом описываются все искажения потенциального рельефа, обусловленные встроенными зарядами и приложенными напряжениями.

5.2. Вольт - амперная характеристика РТС

Для моделирования переноса заряда в РТС необходимо решить систему уравнений (301), (313) и (322) с граничными условиями (304.1), (304.2), (305.1), (305.2), (314.1), (314.2) и (324). Это достаточно сложная задача. Уравнения переноса и Пуассона нелинейны, а уравнение Шредингера необходимо

многократно решать при каждом значении энергии E_z и поперечного импульса p_{\perp} в итерационном процессе самосогласования по потенциалу ϕ . Мы не будем останавливаться на использованных методах численного решения, поскольку они в целом стандартны [29,30-33,83].

Далее результаты расчетов будут сравниваться с результатами измерений. В этой связи необходимо напомнить, что само по себе понятие ВАХ РТС является условным, поскольку эта характеристика не поддается непосредственному измерению. Измерить удастся только ВАХ РТД, которая существенно зависит от геометрии и электрофизических свойств подложки, размеров и формы контактов и т.д. Как показано в ряде работ [15,84,85,88,89,91], наиболее близкой к ВАХ РТС оказывается ВАХ РТД на высоколегированной n^{+-} подложке с малыми размерами контакта к РТС (диаметр контакта менее 5 мкм). При анализе проблемы взаимодействия квантовых и классических областей РТД мы подробно исследуем этот вопрос.

Всюду далее мы будем рассматривать только РТС с одним резонансным уровнем энергии E_r . Это обусловлено стремлением проанализировать явления, определяющие перенос заряда в квантовых приборах, и понять их влияние на измеряемые характеристики. Усложнение ситуации многоуровневыми РТС только затруднит анализ. Понимание физики переноса заряда в простейших случаях облегчает анализ более сложных, но вторичных ситуаций. В этом параграфе мы не будем обсуждать явления типа гистерезиса ВАХ РТД и пр., поскольку они не имеют отношения непосредственно к РТС, а связаны либо с особенностями взаимодействия РТС - подложка, либо с дополнительными физическими явлениями, обусловленными специальными свойствами материала.

5.2.1. Соглашение о токах

Прежде, чем перейти к анализу характеристик РТС, сделаем одно короткое замечание. Электрический ток через i -тый контакт J_i определяется, как интеграл по площади контакта от нормальной к контакту составляющей вектора плотности

тока \mathbf{J} . Для того, чтобы сохранить правильные знаки произведений тока на напряжение на каждом из контактов (выделяемая или поглощаемая мощности), в качестве положительного принимается втекающий в контакт ток, то есть при интегрировании берется проекция \mathbf{J} на внутреннюю нормаль к контакту. Учитывая уже принятые определения (см. Рисунки 23 и 24), для тока J_i получим

$$J_i = - \int_{\Gamma_i} d\sigma(\mathbf{v}, \mathbf{J}), \quad (325)$$

где \mathbf{v} - вектор внешней нормали к i -тому контакту.

При сопоставлении результатов измерений с результатами расчетов по одномерным моделям обычно измеренный ток через контакт делят на его площадь, что с учетом выражений (325) и (246) приводит к правилу

$$J = eI, \quad (326)$$

которым мы и будем дальше руководствоваться.

5.2.2. Сравнение теории и эксперимента, формула Цу - Есаки

На Рисунке 12 показаны:

- экспериментальная ВАХ РТД (кривая 1) на n^+ -подложке, которую мы будем интерпретировать как ВАХ РТС с параметрами слоев $n^+GaAs/500A/GaAs/15A/AlGaAs/20A/GaAs/50A/AlGaAs/20A/GaAs/15A/n^+GaAs/500A$;
- результат расчета ВАХ этой РТС в рамках стандартной самосогласованной модели, использующей для вычисления плотности тока формулу Цу - Есаки (кривая 3);
- результаты расчета ВАХ той же РТС в рамках двухфазной модели (кривая 2).

Видно не только качественное, но и неплохое количественное соответствие кривых 1 и 2. Относительно кривой 3 можно заметить лишь, что представленная на ней ВАХ N-образна, также как и экспериментальная.

Такое «условно качественное» соответствие теории и эксперимента характерно практически для всех общеизвестных моделей переноса заряда в мезоскопических структурах [67,68,73,114] и связано с отсутствием в этих моделях последовательного учета стохастического рассеяния носителей заряда.

Модель, в рамках которой получена кривая 3, физически прозрачна и чаще других используется для анализа квантовых приборов [67,114]. Она во многом совпадает с двухфазной моделью. В частности, в ней также определяется потенциальный рельеф для электронов и самосогласованным образом решаются уравнения Шредингера и Пуассона. Как правило, не учитывается зависимость эффективного потенциала от поперечного импульса, но она не играет в случае РТС определяющей роли в соответствии теории и эксперимента. Отличия состоят в применении для вычисления тока формулы Цу - Есаки и произволе при вычислении плотности заряда (концентрации) электронов внутри РТС. Произвол этот связан с тем, что не используются никакие уравнения переноса, а значения концентраций частиц внутри РТС определяется на основе некоторых упрощенных гипотез, имеющих очень ограниченную область применимости.

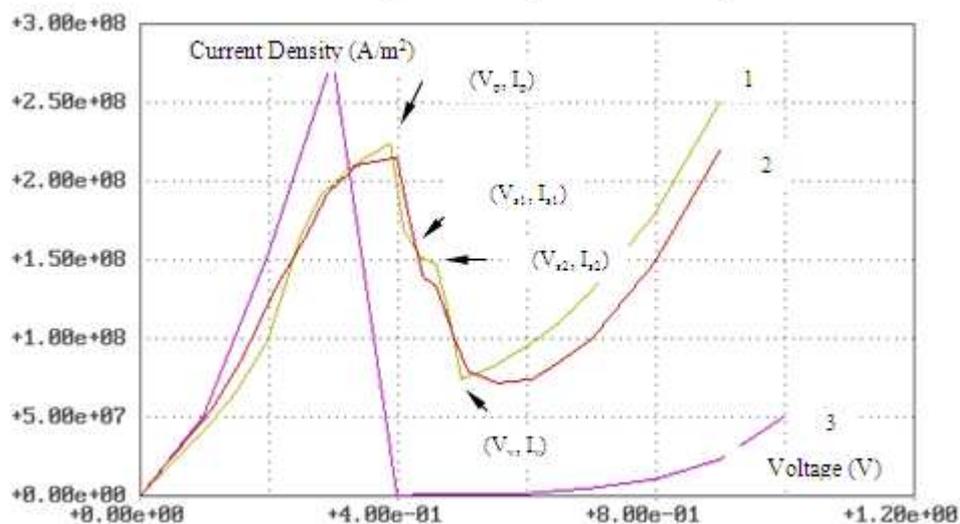


Рисунок 34. Экспериментальная и теоретические ВАХ РТД на основе GaAs/AlGaAs РТС с параметрами слоев:
 n+GaAs/500Å/GaAs/15Å/AlGaAs/20Å/GaAs/50Å/AlGaAs/20Å/GaAs/15Å/n+GaAs/500Å
 на хорошо проводящей n+-подложке и с диаметром контакта к РТС 5 μm.

Неоднократно упоминавшаяся формула Цу - Есаки для потока электронов [14,67,124] имеет вид

$$I_{TE} = (m^* k_B T / 2\pi^2 \hbar^3) \int_{\max\{U(z_0), U(z_V)\}}^{\infty} dE T(E) \ln\{(1 + \exp\{-\beta E\}) / (1 + \exp\{-\beta(E + eV)\})\}. \quad (327)$$

Она напоминает приведенное выше выражение для потока, следующее из (321) при пренебрежении зависимостью эффективного потенциала от поперечного импульса. Формально эти соотношения отличаются только аргументами экспонент. Однако именно в этих аргументах содержится принципиально важное различие: выражение (127) и следствия из него локальны, поскольку величины

$$s_{l,r}(E_z, z) = \beta \{ E_z - F_{l,r}(z) + \tau v_{l,r}(E_z) dF_{l,r}(z)/dz \}$$

зависят от химических потенциалов «левых» и «правых» электронов $F_{l,r}(z)$ и их пространственных производных, определенных во всех точках z , а выражение (327) не локально и является функцией только приложенного напряжения.

Сравнивая соотношения (321) и (327) нетрудно понять, каким образом формула Цу - Есаки может быть получена из выражения для потока (321). Действительно, если в (321) пренебречь «полевыми вкладами», пропорциональными времени релаксации импульса τ , то различие между формулами сведется только к зависимости I от $F_{l,r}$. Далее, можно предположить, что химический потенциал «левых» электронов равен нулю (то есть своему граничному значению на контакте к РТС) практически везде внутри РТС за исключением, быть может, малой области у границы РТС – подложка, а химический потенциал «правых» электронов равен $-eV$ (то есть своему значению на границе РТС – подложка) практически везде в РТС, за исключением, быть может, малой области вблизи контакта к РТС. В этом случае формула (321) будет совпадать с (327) всюду внутри РТС, например, в некоторой точке z_w , которую можно выбрать в середине квантовой ямы. Таким образом, формулу (327) можно считать предельным случаем выражения (321) при предположениях практического постоянства химических потенциалов электронов и длин изменения химических потенциалов, много больших длины свободного пробега электронов (см. неравенство (227)). Если полагать, что для характерной скорости электронов справедлива оценка

$$v \sim (k_B T / m^*)^{1/2},$$

то с учетом выражения (325) для L_0 эти предположения сведутся к неравенству

$$\hbar \tau / m^* \gg \tau^2 k_B T / m^*,$$

из которого получим условие применимости формулы Цу - Есаки:

$$\tau \ll \tau_0 = \hbar / k_B T.$$

Это условие «очень быстрой» релаксации электронов к состоянию термодинамического равновесия. При комнатной температуре $\tau_0 \approx 2,54 \cdot 10^{-14}$ с. Сомнительно, что приведенное условие часто реализуется на практике.

Хотя исходно формула Цу - Есаки получена просто как разность потоков электронов с «левого и правого берегов» некоторого потенциального рельефа при условии приложенного между «берегами» внешнего напряжения V , в ней физически заложены отмеченные нами ограничения. Она предполагает мгновенную стохастизацию носителей заряда и не позволяет понять, каким образом «квантовый ток» в РТС переходит в обычный дрейфово – диффузионный ток в контактах и подложке? Применение выражения (327) для предварительных оценок вполне корректно. Основная проблема «стандартной» модели не в формуле для расчета плотности тока, а в отсутствии последовательного алгоритма для нахождения плотностей числа частиц внутри прибора.

5.2.3. Дрейфово–диффузионный ток и нарушение химического равновесия

Для понимания особенностей переноса заряда в РТС часто недостаточно только результатов численного моделирования. Желательно иметь еще и аналитические выражения, позволяющие делать оценки и аппроксимации.

Перепишем уравнения переноса (313) в более удобном для анализа виде

$$\begin{cases} a_l(z) d^2 F_l(z) / dz^2 - b_l(z) dF_l(z) / dz = \hbar^{-1} (F_l(z) - F_r(z)) (n_l(z) + n_r(z)), \\ a_r(z) d^2 F_r(z) / dz^2 + b_r(z) dF_r(z) / dz = \hbar^{-1} (F_r(z) - F_l(z)) (n_l(z) + n_r(z)), \end{cases} \quad (328)$$

где для коэффициентов справедливы выражения

$$\begin{cases} a_{l,r}(z) = \pm 2\pi\beta(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_0^\infty dE_z T(E_z, p_\perp^2) (-f'(s_{l,r}(E_z, p_\perp^2, z))) \tau v_{l,r}(E_z, p_\perp^2, z), \\ b_{l,r} = c_{l,r} - d_{l,r}, \end{cases} \quad (329.1)$$

$$\begin{cases} c_{l,r}(z) = 2\pi\beta(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_0^\infty dE_z T(E_z, p_\perp^2) (-f'(s_{l,r}(E_z, p_\perp^2, z))), \\ d_{l,r}(z) = 2\pi\beta(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_0^\infty dE_z T(E_z, p_\perp^2) (-f'(s_{l,r}(E_z, p_\perp^2, z))) d(\tau v_{l,r}(E_z, p_\perp^2, z))/dz. \end{cases} \quad (329.2)$$

Следует отметить, что коэффициенты $a_{l,r}$ и $c_{l,r}$ по определению неотрицательны, а коэффициенты $d_{l,r}$ могут иметь любой знак, совпадающий со знаком пространственной производной от произведения $\tau v_{l,r}$.

Введем новые переменные, применение которых более удобно при аналитическом исследовании уравнений двухфазной модели

$$F_+ = F_l + F_r, F_- = F_l - F_r, \quad (330.1)$$

откуда следует:

$$F_l = (F_+ + F_-)/2, F_r = (F_+ - F_-)/2. \quad (330.2)$$

Величина $F_-(z)$ описывает отклонение системы «левых» и «правых» электронов от состояния локального химического равновесия, а величина $F_+(z)$ определяет распределение приложенного напряжения внутри структуры. Согласно (314.1) и (314.2) для F_- и F_+ справедливы граничные условия

$$\begin{cases} F_-(z_0) = F_-(z_V) = 0, \\ F_+(z_0) = 0, F_+(z_V) = -2eV. \end{cases} \quad (331)$$

Из уравнений (328) нетрудно получить уравнение, которое условно можно назвать уравнением для уровня неравновесности F_- .

$$\begin{aligned} a_l a_r d^2 F_- / dz^2 + (1/2)(a_l b_r - a_r b_l) dF_- / dz - (1/2)(a_l b_r + a_r b_l) dF_+ / dz = \\ = \hbar^{-1} F_-(a_l + a_r)(n_l + n_r). \end{aligned} \quad (332)$$

Условно потому, что в само уравнение (332) входит первая производная от F_+ и во все коэффициенты $a_{l,r}$ и $b_{l,r}$ также входят F_- , F_+ и их производные. Однако остальное описание можно построить таким образом, чтобы вторая производная

от F_- фигурировала бы только в уравнении (332), которое в этом случае следует считать уравнением для F_- .

Если просуммировать уравнения (313), то мы получим закон сохранения полного потока

$$d(I_l(z) + I_r(z))/dz = 0$$

или

$$I_l(z) + I_r(z) = I = \text{const.} \quad (333)$$

Соотношение (333) можно рассматривать как нелинейное алгебраическое уравнение, связывающее производные от F_- и F_+ . Как показано в первой главе этой работы, в классической теории уравнение, аналогичное (333), позволяет явно выразить градиент F_+ через поток и градиент F_- [38,47,54]. В рассматриваемом случае это можно сделать лишь приближенно, раскладывая соотношения для потоков в ряд по τ и оставляя только нулевой и первый члены разложения

$$I \approx I_q + I_{dd}, \quad (334.1)$$

где

$$\begin{cases} I_q(z) = 2\pi(2\pi\hbar)^{-3} \int_0^\infty dp_\perp^2 \int_{U^{\text{cur}}(p_\perp^2)}^\infty dE_z T(E_z, p_\perp^2) \{f(s_{0l}(E_z, p_\perp^2, z)) - f(s_{0r}(E_z, p_\perp^2, z))\}, \\ I_{dd} = - (1/2)(a_{l0} - a_{r0})dF_-/dz - (1/2)(a_{l0} + a_{r0})dF_+/dz, \\ s_{0l,r}(E_z, p_\perp^2, z) = \beta \{E_z + E_\perp(p_\perp^2, z) - F_{l,r}(z)\}, \\ a_{l0,r0} = a_{l,r}(s_{l,r} = s_{0l,r}). \end{cases} \quad (334.2)$$

Величины I_q и I_{dd} можно интерпретировать как квантовую и дрейфово-диффузионную составляющие потока. Такое разделение возможно только в линейном по τ приближении и невозможно в общем случае. Из (39.2) получим приближенное выражение для F_+

$$dF_+(z)/dz = -2I_{dd}/(a_{l0}+a_{r0}) - ((a_{l0} - a_{r0})/(a_{l0} + a_{r0}))dF_-(z)/dz. \quad (335)$$

Легко понять, что полученное выражение аналогично уравнению (46) классической теории. Из (335) следует, что падение напряжения в структуре может быть приближенно представлено как сумма чисто омического члена, определяемого первым слагаемым в правой части (335), и члена, связанного с

локальной химической неравновесностью «левых» и «правых» электронов, определяемого вторым слагаемым в правой части соотношения (335). Действительно, проинтегрировав (335) по z от z_0 до z_V и используя граничные условия (331) получим

$$V = \int_{z_0}^{z_V} dz I_{dd} / e(a_{l0} + a_{r0}) - \int_{z_0}^{z_V} dz ((a_{l0} - a_{r0}) / 2e(a_{l0} + a_{r0})) dF_{-}(z) / dz. \quad (336)$$

Это выражение является аналогом формулы (53) и представляет собой обобщение закона Ома на случай, когда в переносе заряда участвуют две фазы заряженных частиц, способных отклоняться от состояния локального химического равновесия. Соотношения (334) - (336) позволяют интерпретировать коэффициенты $a_{l,r}$ как проводимости для «левых» и «правых» электронов, деленные на заряд электрона. Далее мы будем называть их просто проводимостями.

В линейном по τ приближении из (332) и (335) получим

$$a_{l0} a_{r0} d^2 F_{-} / dz^2 + ((a_{l0}^2 b_{r0} - a_{r0}^2 b_{l0}) / (a_{l0} + a_{r0})) dF_{-} / dz = \\ = \hbar^{-1} F_{-}(a_{l0} + a_{r0})(n_l + n_r) - I_{dd} ((a_{l0} b_{r0} + a_{r0} b_{l0}) / (a_{l0} + a_{r0})), \quad (337)$$

где

$$b_{l0,r0} = b_{l,r}(s_{l,r} = s_{0l,r}).$$

Из уравнения (337) следует важный вывод: источником отклонения «левых» и «правых» электронов от состояния локального химического равновесия является дрейфово-диффузионный ток. В случае, когда $I_{dd} = 0$ аттрактором уравнения (337) является решение $F_{-} = 0$, то есть состояние локального химического равновесия между «левыми» и «правыми» электронами. Если же $I_{dd} \neq 0$, то аттрактором уравнения (337) является величина

$$F_{-}^{att} = \hbar I_{dd} ((a_{l0} b_{r0} + a_{r0} b_{l0}) / ((a_{l0} + a_{r0})^2 (n_l + n_r))). \quad (338)$$

Нетрудно показать, что с учетом граничных условий (331) уравнение (337) при $I_{dd} = 0$ имеет только тривиальное решение $F_{-} = 0$, то есть в отсутствии дрейфово – диффузионного тока «левые» и «правые» электроны будут всегда находиться в локальном химическом равновесии.

В заключение этого параграфа приведем рисунки, демонстрирующие поведение F_{\dots} и $F_{I,r}$ внутри РТС, характеристики которой изображены на Рисунке 34.

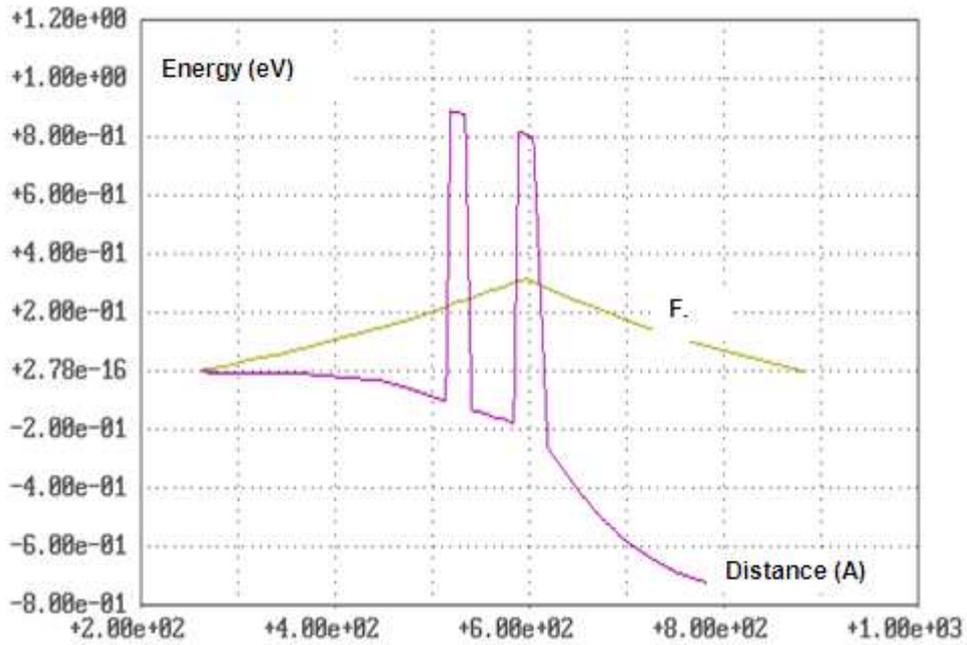


Рисунок 35. Распределение F внутри РТС при $V = 0,67$ В.

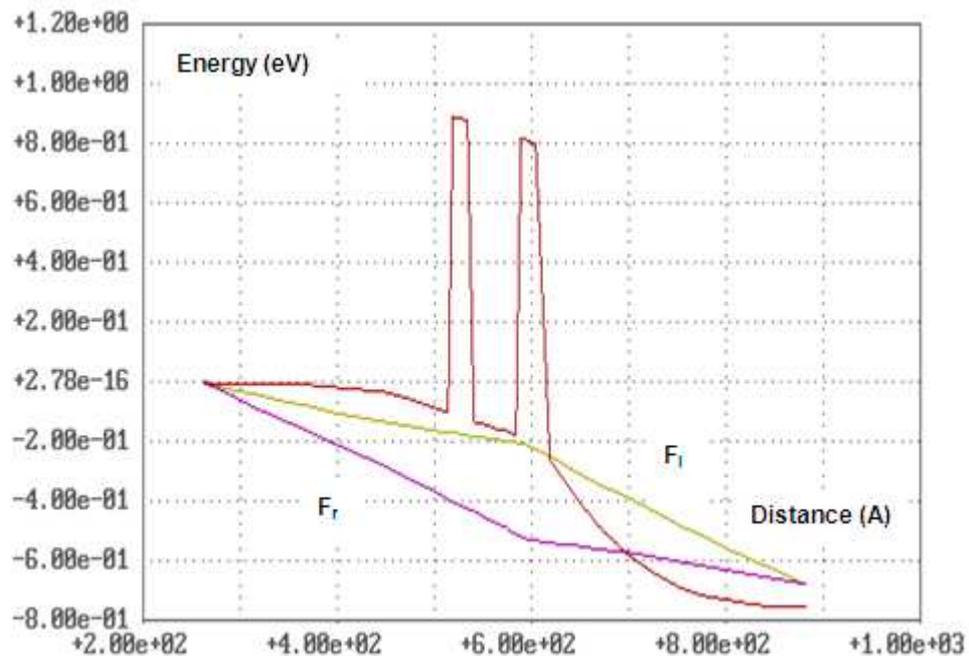


Рисунок 36. Распределение химических потенциалов F_1 и F_r внутри РТС при $V = 0,67$ В.

5.2.4. Характерные участки ВАХ РТС

Представленная на Рисунке 34 экспериментальная ВАХ РТС содержит практически все характерные участки, наблюдаемые для структуры с одним резонансным уровнем. Кривые 1 и 2 на Рисунке 34 можно разделить на пять характерных участков:

- первый участок возрастания тока: $V \in [0, V_p]$;
- первый участок спада тока: $V \in [V_p, V_{s1}]$;
- участок ступеньки: $V \in [V_{s1}, V_{s2}]$;
- второй участок спада тока: $V \in [V_{s2}, V_v]$;
- второй участок возрастания тока $V \in [V_v, \infty]$.

На участке ступеньки для различных РТС ток может либо падать, но с меньшей крутизной ВАХ, чем на участках спада, либо быть практически постоянным. Разный ход ВАХ в этой области связан только с конструктивными особенностями прибора, но не с физикой явления.

На первом участке возрастания тока перенос заряда происходит через резонансный уровень энергии, для которого коэффициент прохождения электронов через потенциальный рельеф близок к единице. Коэффициент прохождения электронов для других энергий, меньших высоты барьеров, практически равен нулю (см. Рисунки 27 и 32). Надбарьерным переносом заряда для рассматриваемой РТС можно полностью пренебречь, поскольку при комнатной температуре и высоте барьеров в 1 эВ над барьерами практически нет электронов.

Как видно на Рисунке 34 даже на этом начальном участке формула Цу - Есаки описывает ВАХ РТС довольно плохо. Для того чтобы понять, с чем это связано, рассмотрим предельный случай

$$\tau \rightarrow 0 \text{ и } |F_v| \ll k_B T.$$

Уравнения (328) и следующее из них уравнение (332) оказываются в этом случае сингулярно возмущенными и их можно исследовать, используя стандартную в

таких случаях технику [44-46]. Для коэффициентов в (328) и (332) справедливы соотношения

$$a_l \approx a_r = a \rightarrow 0, d_{l,r} \rightarrow 0, c_l \approx c_r = c,$$

учитывая которые получим вырожденные уравнения

$$\begin{cases} dF/dz = 0, \\ dF_+/dz = -(2/c\hbar)F_+(n_l + n_r). \end{cases} \quad (339)$$

Решения системы (339) имеют вид

$$\begin{aligned} F_- &= F_-^0 = \text{const}, \\ F_+ &= - (2/\hbar c)F_-^0 \int_{z_0}^z dz(n_l + n_r). \end{aligned}$$

Откуда, учитывая граничные условия (331) для F_+ , найдем

$$\begin{cases} F_-^0 = eV\hbar c / \int_{z_0}^{z_V} dz(n_l + n_r), \\ F_+ = - 2eV \int_{z_0}^z dz(n_l + n_r) / \int_{z_0}^{z_V} dz(n_l + n_r). \end{cases} \quad (340)$$

Решение (340) для F_-^0 не удовлетворяет граничным условиям (331). Его необходимо дополнить еще пограничными слоями. Из уравнения (332) получим в рассматриваемом пределе

$$d^2F/dz^2 - (c/a)dF_+/dz = \lambda^2 F_-, \quad (341)$$

где

$$\lambda^2 = (2/\hbar a)(n_l + n_r).$$

Из (341) также как из (328) в пределе $a = 0$ получается вырожденное уравнение, совпадающее со вторым уравнением в системе (339). Это означает, что формула (340) описывает вырожденное решение уравнения (341). В главном порядке теории сингулярных возмущений [44-46] получим приближенное решение (341), удовлетворяющее граничным условиям (331)

$$F_- = F_-^0 \{ 1 - (\exp(-\lambda(z-z_0)) + \exp(-\lambda(z_V-z)))/(1 + \exp(-\lambda(z_V-z_0))) \}. \quad (342)$$

Поскольку при $a \rightarrow 0$ ($\tau \rightarrow 0$) величину λ следует считать большой, то решение (342) совпадает с F^0 повсюду в РТС, кроме пограничных слоев вблизи контактов. Оценку размеров этих пограничных слоев мы сделаем чуть позже. Это означает, что для вычисления тока мы можем приближенно использовать выражение (321), подставив в него $F_{l,r}$, вычисленные из решения (340). В рассматриваемом приближении аргументы функций распределения в (321), соответствующих вкладам в поток «левых» и «правых» частиц, будут отличаться на величину F^0 , что напоминает формулу Цу - Есаки. Однако как видно из (340) F^0 существенно отличается от приложенного напряжения.

Сделаем простейшие оценки для F^0 .

$$\int_{z_0}^{z_V} dz (n_l + n_r) \sim (n_l(z_w) + n_r(z_w))L, \quad L = z_V - z_0,$$

где z_w – некоторая точка внутри РТС,

$$c \sim (n_l(z_w) + n_r(z_w))\beta(2E_r/m^*)^{1/2} = v_{tr}\beta(n_l(z_w) + n_r(z_w)),$$

$$v_{tr} = (2E_r/m^*)^{1/2},$$

где E_r – энергия резонансного уровня, а v_{tr} – соответствующая этому уровню скорость электрона, в том случае, когда его волновая функция является плоской волной. В результате для F^0 справедлива простая оценка

$$F^0 \sim (\tau_0/\tau_{tr})eV, \quad (343)$$

где $\tau_0 = \hbar/k_B T$, а

$$\tau_{tr} = L/v_{tr}$$

– характерное время пролета электрона с энергией равной резонансной через РТС.

При комнатной температуре $\tau_0 \approx 2,57 \cdot 10^{-14}$ с. При резонансной энергии порядка 0,54 эВ для электрона в GaAs $v_{tr} \sim 10^8$ см/с. Длина РТС порядка 10^{-6} см. Таким образом $\tau_0/\tau_{tr} \sim 1$. Эта оценка вместе с (343) объясняет качественное соответствие на первом участке возрастания тока хода ВАХ РТС, рассчитанной по формуле Цу - Есаки, с экспериментальной кривой. Количественного соответствия, очевидно, не наблюдается. Существенно, что при различных значениях энергии E_r , которая зависит от параметров гетерослоев, образующих

структуру, отличия токов, посчитанных в модели Цу - Есаки, от экспериментальных будут серьезно варьироваться.

Еще одна важная оценка касается величины λ или обратной к ней величины

$$L_{\text{rel}} = \lambda^{-1},$$

имеющей смысл длины релаксации к состоянию локального химического равновесия. Нетрудно получить, что

$$L_{\text{rel}} \sim L_0(E_r/k_B T)^{1/2}.$$

Для объемного GaAs при комнатной температуре L_0 около 240 А. Это означает, что L_{rel} порядка 1000 А, то есть больше или сравнима с длиной РТС. В результате, как хорошо видно на Рисунке 37, реально никаких пограничных слоев для F_r в РТС не наблюдается, а оценки в пределе $\tau \rightarrow 0$ позволяют лишь качественно понять поведение разности химических потенциалов.

Наличие участка с отрицательной дифференциальной проводимостью на ВАХ РТС физически связано с быстрым уменьшением коэффициента прохождения электронов через резонансный уровень энергии $T(E_r)$ при $V > V_p$. Этим легко объяснить первый участок спада тока. Однако возникновение ступеньки на ВАХ РТС никак не объясняется этими простыми рассуждениями.

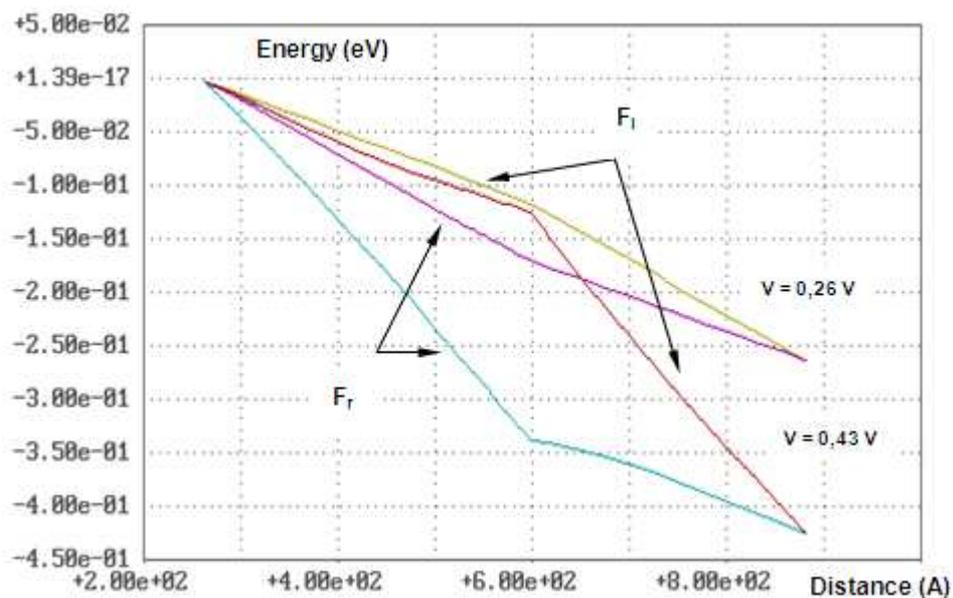


Рисунок 37. Характерные распределения химических потенциалов внутри РТС на первом участке возрастания тока и на падающем участке непосредственно перед образованием ступеньки.

На Рисунке 37 показано характерное поведение химических потенциалов «левых» и «правых» электронов на первом участке роста тока и на падающем участке непосредственно перед образованием ступеньки. Видно, что на падающем участке наблюдается резкое возрастание F_{\perp} внутри РТС по сравнению с первым участком роста. Если на первом участке возрастания тока максимальное значение F_{\perp} составляет около пятой части от V , то на падающем участке это уже примерно половина приложенного напряжения V . Данный эффект связан с тем, что из-за резкого уменьшения $T(E_r)$ на падающем участке ВАХ превалирующую роль начинает играть дрейфово-диффузионный механизм переноса электронов через РТС. На первом участке роста тока основной вклад в ток давал квантовый член I_q в формуле (334.1). Возрастающий на первом участке спада дрейфово-диффузионный ток приводит в соответствии с уравнением (337) к быстрому росту F_{\perp} внутри РТС практически до предельного значения.

Если переписать выражения для $s_{l,r}$ через $F_{\perp,+}$, то становится полностью ясным механизм образования ступеньки на ВАХ РТС. Действительно

$$\begin{cases} s_l = \beta \{ E_z + E_{\perp} - (1/2)(F_+ + F_{\perp}) + (1/2)\tau v_l d(F_+ + F_{\perp})/dz \}, \\ s_r = \beta \{ E_z + E_{\perp} - (1/2)(F_+ - F_{\perp}) + (1/2)\tau v_r d(F_+ - F_{\perp})/dz \}. \end{cases} \quad (344)$$

Величина F_{\perp} входит в s_l и s_r с разными знаками. В результате возрастания потенциала F_{\perp} , который растет с ростом напряжения, и остается при этом положительным, начинает возрастать первое слагаемое в формуле (321), обусловленное током «левых» электронов. Второе слагаемое в (321), определяющее ток «правых» электронов, уменьшается с ростом F_{\perp} . В результате, на падающем участке ВАХ РТС появляется ступенька. Ее возникновение никак не описывается и не объясняется в моделях, основанных на формуле Цу - Есаки.

Однако рост F_{\perp} ограничен величиной V , которая очень незначительно меняется на участке ВАХ РТС с отрицательной дифференциальной проводимостью. Дальнейшее уменьшение $T(E_r)$ вновь приводит к резкому уменьшению тока и образуется второй участок спада тока. Но падение тока происходит только до некоторого значения I_v , которое в эксперименте примерно в

два раза меньше, чем I_p . Модель Цу - Есаки дает для I_v значения в десятки и даже сотни раз меньше, чем I_p . Экспериментально этого не наблюдается.

Пройдя через точку I_v , ток вновь начинает быстро возрастать. Характер этого роста и наблюдаемые значения I_v никак не объясняются в простых моделях, а ход ВАХ РТС, начиная с участка ступеньки, описывается ими неудовлетворительно.

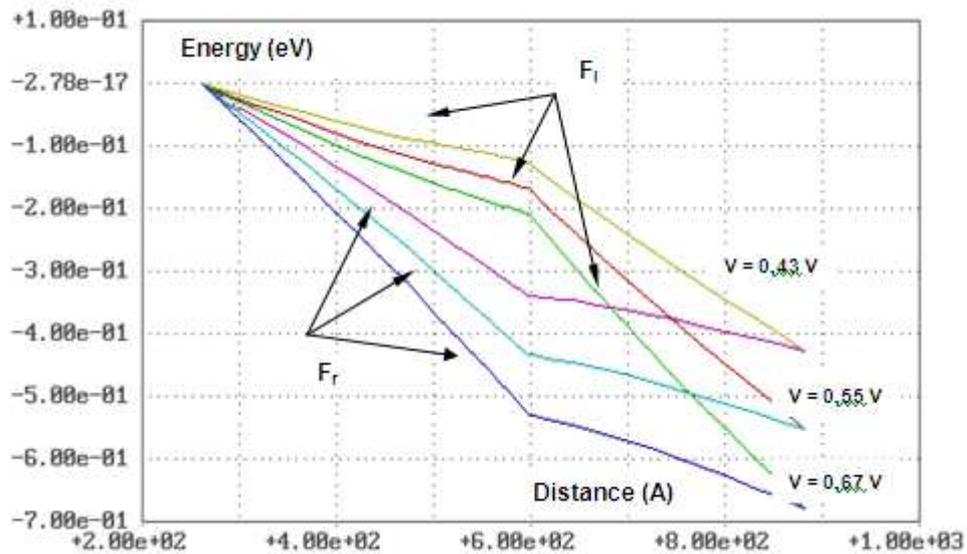


Рисунок 38. Характерные распределения химических потенциалов внутри РТС на участках спада и втором участке возрастания тока.

В рамках предлагаемой модели большие значения I_v и второй участок возрастания тока описываются очень простым механизмом, и это описание хорошо коррелирует с экспериментальными данными. При напряжениях больших V_p искажение потенциального рельефа для электронов в РТС становится очень значительным. Связанное с этим искажением уменьшение коэффициента прохождения электронов через резонансный уровень ведет к обеднению РТС носителями заряда. Как результат этих двух взаимосвязанных явлений резко возрастают электрическое поле и градиенты химических потенциалов. Как видно на Рисунках 36 - 38 потенциалы $F_{l,r}$ монотонно убывают внутри РТС. С учетом того, что величина v_l - положительна, а величина v_r - отрицательна, становится понятно, что слагаемые в $s_{l,r}$, пропорциональные градиентам F_l и F_r соответственно, ведут к возрастанию тока. Первое слагаемое в (321) возрастает с ростом абсолютной величины производной от F_l , а второе слагаемое уменьшается

с ростом модуля производной от F_r . Этот полевой эффект полностью описывает поведение ВАХ РТС на втором участке роста тока и объясняет наблюдаемые значения I_v .

Таким образом, удается довольно просто объяснить все основные особенности ВАХ РТС. Это связано с тем, что многофазная модель позволяет учесть квантовые и классические, баллистические и стохастические явления, определяющие перенос заряда в РТС. В частности, учитывается «последовательное туннелирование», обусловленное стохастическим рассеянием электронов [16,73]. Численное моделирование дает хорошее соответствие с экспериментом (см., например, Рисунок 34).

5.3. Вольт – амперная характеристика РТД

Выше показано, что для РТД, представляющего собой РТС на хорошо проводящей $n+$ подложке, можно добиться приемлемого соответствия теории и эксперимента. Важными конструктивно-топологическими условиями достижения такого соответствия являются малый размер верхнего контакта к РТС и примитивная топология прибора, который в двумерном сечении вдоль оси OZ должен представлять собой симметричный относительно центра контакта прямоугольник. Если эти условия не выполнены, то полученные в рамках одномерной модели ВАХ РТС могут существенно отличаться от измеряемых.

На Рисунке 39 представлены экспериментальная ВАХ РТД на полуизолирующей подложке и ВАХ соответствующей РТС, рассчитанная в рамках одномерной модели. Хорошо видно, что эти характеристики существенно отличаются друг от друга. Очень часто это отличие бывает еще более сильным, чем показано на Рисунке 39 [84,85,89-91].

Простейшее объяснение расхождений теории и эксперимента состоит в предположении, что подложку можно рассматривать просто как нагрузочное сопротивление, подключенное к РТС последовательно [3,15,113,25]. Такой подход позволяет понять смещение ВАХ РТД вправо по отношению к ВАХ РТС,

а также изменение углов наклона различных участков ВАХ. Однако остаются совершенно непонятными существенное уменьшение отношения тока пика к току долины и более плавный наклон участка ВАХ с отрицательной дифференциальной проводимостью на экспериментальной ВАХ по сравнению с теоретической.

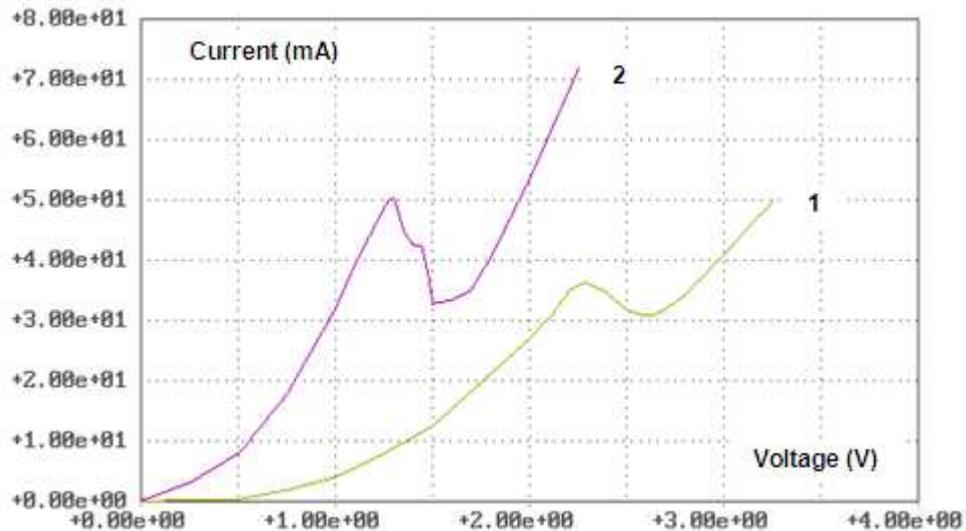


Рисунок 39. Экспериментальная ВАХ РТД на полуизолирующей подложке (кривая 1) и соответствующая ей ВАХ РТС, полученная в рамках одномерной модели (кривая 2).

Действительно, если подложка это просто нагрузочное сопротивление, то ВАХ РТД может быть легко пересчитана из ВАХ РТС на основе элементарных законов электротехники

$$V_{\text{rtd}}(I) = V_{\text{rts}}(I) + IR,$$

где V_{rtd} – напряжение на РТД при текущем значении тока I , $V_{\text{rts}}(I)$ – известная из расчета ВАХ РТС функция, R – сопротивление подложки. Учет нагрузочного сопротивления приводит к растяжению характеристики вдоль оси напряжений и к появлению дополнительно к N-образности еще и S-образности ВАХ. При этом токи пика и долины не изменяются, что демонстрируется на Рисунках 40.1 и 40.2. На Рисунке 40.2 представлена кривая, полученная при $R = 24$ Ом, которая максимально приближается к экспериментальной, насколько это возможно в модели нагрузочного сопротивления. Отсюда можно заключить, что такое описание ВАХ РТД нельзя признать удовлетворительным.

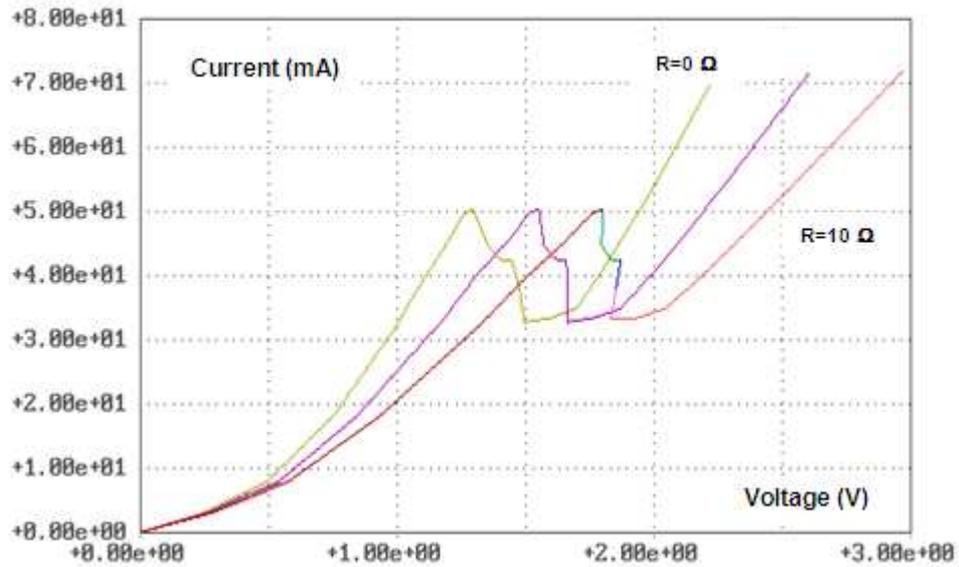


Рисунок 40.1. Вольт - амперные характеристики, полученные из кривой 2, представленной на Рисунке 39, при различных нагрузочных сопротивлениях.

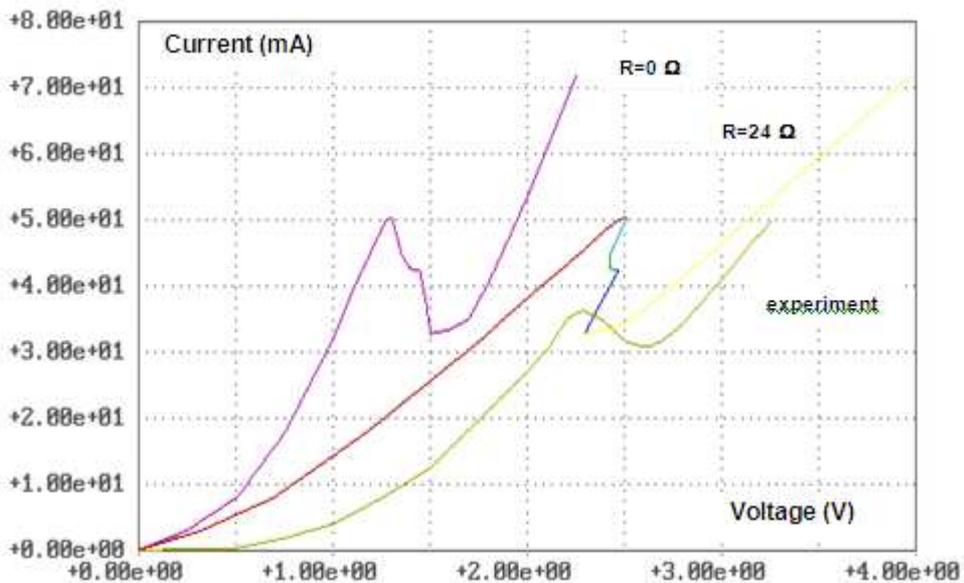


Рисунок 40.2. Вольт - амперная характеристика ($R = 24 \Omega$), полученная из кривой 2, представленной на Рисунке 39, в модели нагрузочного сопротивления, «максимально приближающаяся» к экспериментальной.

Как отмечалось, двухфазная модель позволяет в общем случае описывать и многомерные и нестационарные явления, обусловленные переносом заряда. Совершенно ясно, что распространение одномерной модели на область подложки приведет к результатам, идентичным простому учету подложки, как нагрузочного сопротивления, включенного последовательно с РТС. Отличия теории и эксперимента могут быть объяснены только при учете существенной

неоднородности переноса заряда в РТД с топологией, отличной от простейшей, описанной выше.

Однако в то время, когда проводились основные расчеты характеристик РТД, не имелось вычислительных и временных возможностей реализовать последовательные многомерные расчеты. Поэтому был реализован другой подход, предложенный О.Т. Гавриловым [84,85,88,89], который будем условно называть гибридным.

5.3.1. Двумерная гибридная модель РТД

Суть этой модели состоит в том, что перенос заряда в РТС рассчитывается в рамках одномерного подхода, а перенос заряда в подложке описывается в двумерном приближении на основе стандартных уравнений дрейфа – диффузии. При этом на условной границе Γ_3 , моделирующей контакт РТС с подложкой, ставятся граничные условия (296), позволяющие учесть распределение нормальной составляющей плотности тока вдоль этой границы. В результате в области подложки получаем следующую краевую задачу

$$\begin{cases} (\nabla, \rho \nabla F) = 0, \\ \mathbf{J} = \rho \nabla F / e, \\ (\mathbf{v}, \mathbf{J})_{\Gamma_0} = 0, F(\Gamma_1) = -eV_{\text{rtd}}, (\mathbf{v}, \mathbf{J})_{\Gamma_3} = J_{\text{rts}}(F(\Gamma_3)). \end{cases} \quad (345)$$

Здесь ρ - проводимость подложки; \mathbf{v} - внешняя нормаль к границе подложки; Γ_1 – граница, соответствующая омическому контакту к подложке; $J_{\text{rts}}(F(\Gamma_3))$ - плотность тока РТС, вычисленная на основе одномерной модели. Как уже отмечалось, выбор Γ_3 осуществляется таким образом, чтобы на этой границе электроны можно было бы считать химически равновесными.

Как показали расчеты, проведенные в двумерном приближении для прямоугольных подложек, решение краевой задачи (345) существенно улучшает соответствие между теорией и экспериментом. Это наглядно продемонстрировано на Рисунке 41, где представлены экспериментальная и расчетные ВАХ РТД на основе РТС, теоретическая ВАХ которой показана на Рисунке 39.

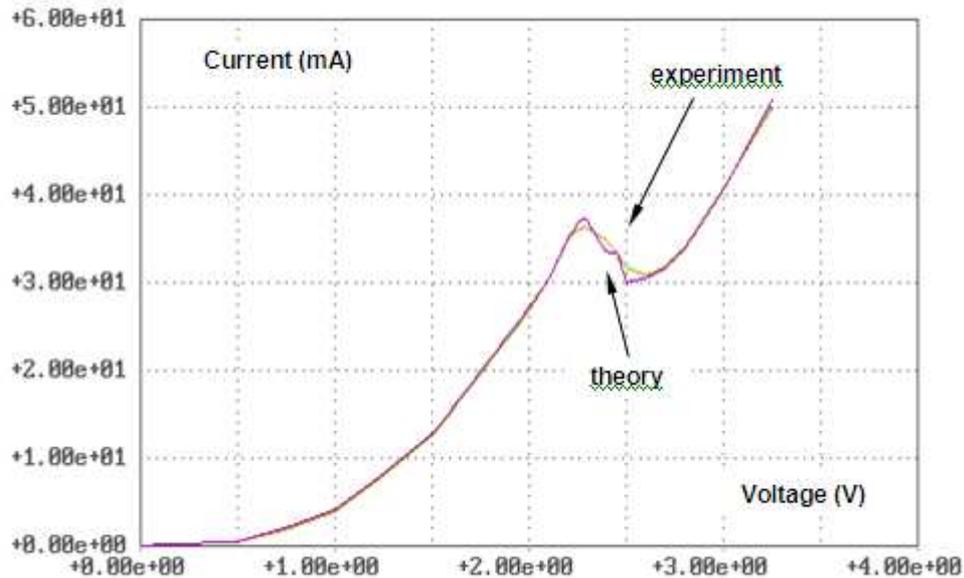


Рисунок 41. Экспериментальная и теоретическая ВАХ РТД на полуизолирующей подложке. В расчетах учтены эффекты взаимодействия РТС с подложкой.

Анализ результатов расчетов показывает, что взаимодействие РТС с подложкой отнюдь не сводится к тривиальному эффекту нагрузочного сопротивления. Необычный вид последнего из граничных условий задачи (345) приводит к двум интересным следствиям:

- распределение нормальной составляющей плотности тока на границе Γ_3 имеет нестандартный для решений уравнения Лапласа вид (см. Рисунок 42);
- при определенных соотношениях между проводимостью РТС на участке ВАХ с отрицательной дифференциальной проводимостью, проводимостью подложки и геометрическими размерами структуры задача (345) может иметь неединственное решение [85] и даже бесконечный спектр собственных решений [88].

Первая из отмеченных особенностей означает, что далеко не вся площадь «контакта» между РТС и подложкой «работает» в одинаковом режиме. Растекание тока в подложке приводит к неоднородному распределению химического потенциала F вдоль границы Γ_3 . В результате, при фиксированном напряжении V_{rtd} на омическом контакте Γ_1 какая-то часть границы Γ_3 может соответствовать участкам ВАХ РТС с отрицательной проводимостью, другие же части - участкам ВАХ с положительной проводимостью. Все это в совокупности

приводит к нетривиальной зависимости ВАХ РТД от проводимости подложки, ее топологии и площадей контактных областей. Этот вывод был подтвержден экспериментальными исследованиями [89].

Вторая особенность делает возможным использование подложки РТД для оригинальных приборных решений. В частности, она может выступать в качестве резонатора электромагнитных колебаний СВЧ и КВЧ диапазонов [88].

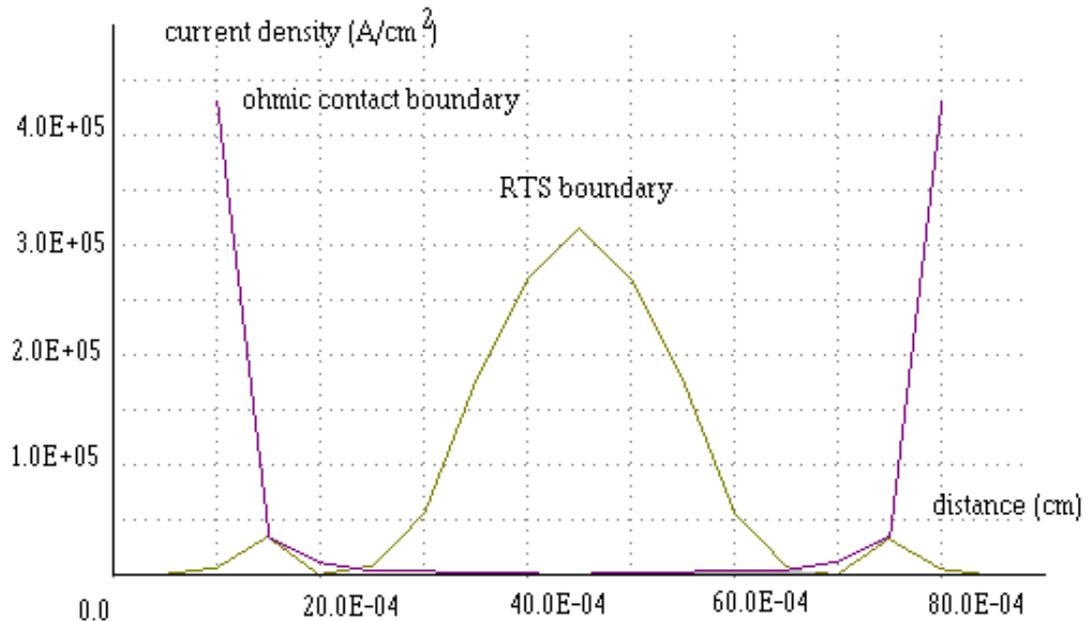


Рисунок 42. Распределение нормальной составляющей плотности тока вдоль границы между РТС и n^+ - подложкой и распределения нормальной составляющей плотности тока вдоль омического контакта такого же размера для структуры той же топологии и при том же суммарном токе через контакт.

5.3.2. Простейший анализ гибридной модели РТД

Если ограничиться двумерным приближением и считать область подложки односвязной, то задачу (345) можно решить аналитически. Действительно, при помощи конформного отображения любую двумерную односвязную область можно отобразить в полуплоскость [55] $\{u, v\}$, показанную на Рисунке 43. Краевая задача (345) превращается при этом в хорошо известную задачу Келдыша - Седова [55] для гармонической функции

$$q(\tau) = -j_u + ij_v.$$

Граничные условия для q выглядят следующим образом

$\text{Re}(q(u,0)) = 0$ при $u \in (a,b)$,

$$\text{Im}(q(u,0)) = \begin{cases} -J_{\text{rts}}, & u \in (\alpha, \beta), \\ 0, & u \in (-\infty, \alpha) \cup (\beta, a) \cup (b, \infty). \end{cases} \quad (346)$$

Формальное решение этой задачи при условии $q(\infty) = 0$ имеет вид

$$q(\tau) = (-1/\pi) \int_{\alpha}^{\beta} du J_{\text{rts}}(u) \left\{ \left[\frac{((a-\tau)(b-u))}{(a-u)(b-\tau)} \right]^{1/2} / (u-\tau) + ((\tau-a)(\tau-b))^{-1/2} \right\}. \quad (347)$$

Положим, что в некоторой точке $u_0 \in (\alpha, \beta)$

$$F(u_0) = -eV_0.$$

Для потенциала $F(u)$ при $u \in (\alpha, \beta)$ получим интегральное уравнение с сингулярным ядром

$$\begin{aligned} F(u, u_0) = & -eV_0 + (e/\pi\rho) \int_{\alpha}^{\beta} dt J_{\text{rts}}(F(t)) \{ \ln|(t-u)/(t-u_0)| + \\ & + 2[1 - ((b-t)/(a-t))^{1/2}] \ln[((b-u_0)^{1/2} + (a-u_0)^{1/2}) / ((b-u)^{1/2} + (a-u)^{1/2})] + \\ & + 2\ln[(((b-t)(a-u_0))^{1/2} + ((a-t)(b-u_0))^{1/2}) / (((b-t)(a-u))^{1/2} + ((a-t)(b-u))^{1/2})] \}. \end{aligned} \quad (348)$$

Если найти $F(u, u_0)$ на границе Γ_3 , то уже не составляет труда вычислить $V(V_0)$ и $J(V_0)$, исходя из соотношений для тока через контакты

$$J = \int_{\Gamma_3} (v, j) ds = \int_{\Gamma_1} (v, j) ds.$$

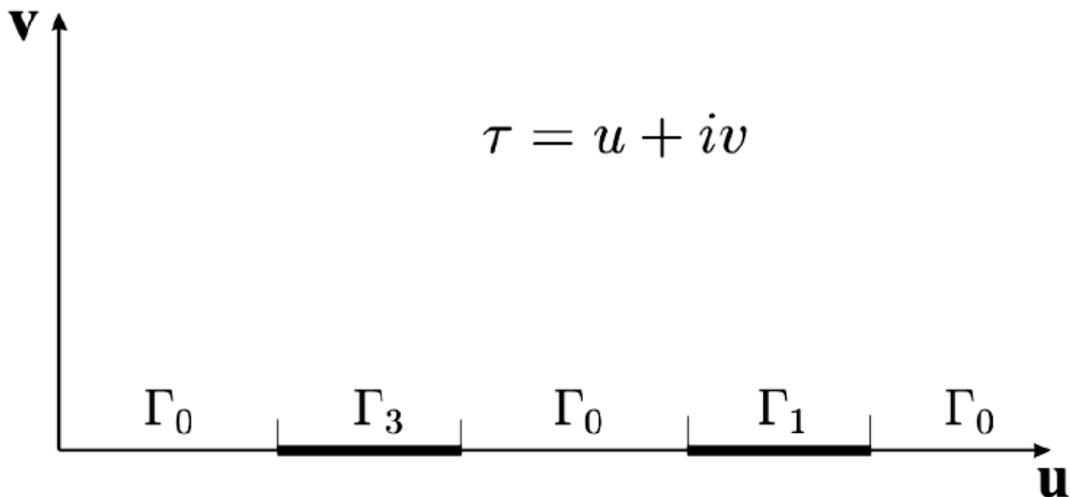


Рисунок 43. Область подложки после конформного отображения.

Задачу можно еще упростить, если принять предположение, часто выполняющееся на практике

$$(b - a) \gg (\beta - \alpha).$$

Тогда из (348) получим

$$\begin{cases} F(u, u_0) = -eV_0 + J_{rts}(F(u))G(u, u_0), \\ G(u, u_0) = (e/\pi\rho)\{(\beta - u)\ln(\beta - u) + (u - \alpha)\ln(u - \alpha) - (\beta - u_0)\ln(\beta - u_0) - (u_0 - \alpha)\ln(u_0 - \alpha)\}. \end{cases} \quad (349)$$

Функция $G(u, u_0)$ имеет одинаковые значения в точках $u = \alpha$ и $u = \beta$

$$G(\alpha, u_0) = G(\beta, u_0) = (e/\pi\rho)\{(\beta - \alpha)\ln(\beta - \alpha) - (\beta - u_0)\ln(\beta - u_0) - (u_0 - \alpha)\ln(u_0 - \alpha)\}$$

и минимум в точке $u_{\min} = (\alpha + \beta)/2$.

$$G_{\min}(u_{\min}, u_0) = (e/\pi\rho)\{(\beta - \alpha)\ln((\beta - \alpha)/2) - (\beta - u_0)\ln(\beta - u_0) - (u_0 - \alpha)\ln(u_0 - \alpha)\}.$$

Если положить $u_0 = u_{\min}$, что всегда можно сделать, то из (340) получим

$$\Delta = -(F(u_0) - F(\alpha))/e = -(F(u_0) - F(\beta))/e = (\ln 2/\pi\rho)J_{rts}(F(\alpha))(\beta - \alpha). \quad (350)$$

Величина Δ показывает насколько отличается эффективное приложенное к одномерной РТС напряжение в середине контакта РТС – подложка (граница Γ_3) от его значений на краях контакта. Собственно, этим различием определяется, каким участкам ВАХ РТС соответствуют середина и края контакта. Как видно из (350) в случае малых контактов и высоких проводимостей подложки величиной Δ можно пренебречь. Действительно, при

$$\Delta \ll V_v - V_p,$$

где, как обычно, V_v и V_p – напряжения долины и пика ВАХ РТС, химический потенциал вдоль границы Γ можно считать постоянным. Из этого неравенства и соотношения (350) получим условие

$$\rho/(\beta - \alpha) \gg (\pi/\ln 2)J_p/(V_v - V_p), \quad (351)$$

при выполнении которого статическая ВАХ РТД может быть описана в простейшей модели: ВАХ РТС и подложка в виде нагрузочного сопротивления. Через J_p обозначена плотность тока пика ВАХ РТС. Если неравенство (351) справедливо, то неоднородностью химического потенциала (эффективного приложенного к РТС напряжения) вдоль границы РТС - подложка можно пренебречь.

Характерные значения J_p составляют 10^5 А/см², а разность $V_v - V_p$ порядка 10^{-1} В. Поэтому для удельной проводимости получим оценку

$$\rho/(\beta - \alpha) \gg 10^6 \text{ Ом}^{-1} \cdot \text{см}^{-2}.$$

При характерных размерах контактных площадок в 10 мкм для GaAs это соответствует уровням легирования выше, чем 10^{19} см⁻³. Только при этих условиях корректно представление о подложке как о нагрузочном сопротивлении.

В тех же случаях, когда соотношение (351) не выполнено, для адекватного описания ВАХ РТД необходимо решать задачу (345). В большинстве случаев РТД изготавливают с достаточно большими контактами и на подложках с уровнем легирования меньшим, чем требует неравенство (351). Поэтому учет влияния переноса заряда в подложке на ВАХ РТД становится практически актуальной задачей.

Анализ выражения (348) показывает, что решение задачи при условии

$$\text{sign}(F) = \text{sign}(J),$$

соответствующем положительной выделяемой мощности, существует, но может быть неединственным. Действительно, поскольку J_{rts} немонотонная функция F , то пересечение кривых $F + eV_0$ и $J_{\text{rts}}(F)G$ может происходить как в одной, так и в трех точках, в зависимости от того, какому участку ВАХ РТС соответствует эффективное приложенное напряжение в анализируемой точке границы Γ_3 . Такая ситуация может наблюдаться при всех $u \in (\alpha, \beta)$, то есть на несчетном множестве, что означает возможность существования несчетного количества решений задач (345) или (348). Численное моделирование полностью подтверждает сделанные выводы.

В заключение этого параграфа отметим, что здесь в практической плоскости рассмотрена одна из важных проблем наноэлектроники, а именно проблема описания взаимодействия классических и квантовых областей приборов. Для всех квантовых приборов эта задача возникает в том или ином конкретном выражении, зависящем от конструктивных особенностей приборов и способа подключения их к другим приборам и устройствам, в том числе к измерительному оборудованию.

ГЛАВА 6. Приборы на основе квантовых проводов и их статические характеристики

Уже с середины 80-х годов прошлого века объекты, названные квантовыми проводами (КП), и возможные электронные приборы на их основе привлекают пристальное внимание исследователей [2,77,126-128]. Интерес к ним продиктован двумя обстоятельствами:

- возможностью создания сверхбыстродействующих электронных устройств с рабочими частотами более 100 ГГц;
- перспективой повышения уровня интеграции элементов в схемах до величин порядка 10^{10} на квадратный сантиметр и более.

Создание технологий, позволяющих осуществлять массовое производство подобных приборов и схем на их основе, будет означать новый качественный этап в развитии электроники. Предполагаемые малые габариты и вес, высокое быстродействие и низкое энергопотребление приборов и схем указывают на то, что они найдут применение в:

- системах связи и радиоуправления;
- вычислительной технике;
- бытовой и научно-исследовательской электронной технике и т.п.

Рабочие частоты в сотни ГГц могут быть достигнуты в приборах на основе РТС или других гетероструктур с нанометровыми размерами вертикальных слоев и вполне макроскопическими планарными размерами. Такие приборы успешно применяются для создания сверхвысокочастотных устройств и схем малой степени интеграции [2]. В генераторах на основе РТД уже давно получены частоты СВЧ - сигналов до 700 ГГц [15,125].

Приборы на основе квантовых проводов с планарной топологией теоретически являются перспективными как с точки зрения их частотных свойств, так и для применения в интегральных схемах. Однако практических результатов получено пока в этой области немного. Это связано:

- с технологическими проблемами, стоящими на пути создания планарных объектов нанометровых размеров;
- с отсутствием достаточной ясности в вопросе, какие же собственно приборы следует создавать?

В последние годы наметились реальные пути решения технологических проблем создания КП с удовлетворительными электрофизическими свойствами как на основе углеродных нанотрубок [2,129,130], так и на основе традиционных полупроводниковых материалов [2,6-8,121,122,127]. Это внушает надежду на то, что создание элементной базы наноэлектроники становится задачей уже не столь отдаленного будущего.

6.1. Квантовые провода

Квантовым проводом принято называть объект, энергетический спектр носителей заряда в котором квантован в двух направлениях (мы будем полагать, что это направления OY и OZ) и непрерывен (квазинепрерывен) в третьем (OX – в нашем случае). Именно в этом третьем направлении возможен перенос заряда в КП. По этой причине КП часто называют «одномерным квантовым объектом».

Будем придерживаться следующей терминологии:

- трехмерными или объемными будем называть объекты, энергетический спектр носителей заряда в которых непрерывен во всех трех пространственных направлениях;
- двумерными будем называть объекты, энергетический спектр носителей заряда в которых дискретен в одном из направлений и непрерывен в двух других;
- одномерными будем называть объекты, энергетический спектр носителей заряда в которых дискретен в двух направлениях и непрерывен в одном направлении;
- нульмерными будем называть объекты, энергетический спектр носителей заряда в которых дискретен во всех трех направлениях.

Напомним, что поскольку все рассматриваемые объекты имеют конечные геометрические размеры, под непрерывным энергетическим спектром понимается квазинепрерывный спектр с расстояниями между уровнями энергии, много меньшими, чем $k_B T$.

При ненулевых температурах энергетический спектр электронов можно считать квантованным, если расстояние между двумя нижними уровнями энергии больше, чем $k_B T$. Как уже отмечалось (см. формулу (226)), простейшие оценки показывают, что это требование будет выполнено, если для линейного размера структуры в направлении размерного квантования справедливо неравенство

$$L < L_{dq} = (3\pi^2 \hbar^2 / 2m^* k_B T)^{1/2} .$$

Численные значения L_{dq} приведены ранее. Ими определяются характерные размеры структур из различных материалов, которые имеет смысл называть мезоскопическими. Существенно, что длина размерного квантования и длина релаксации электронов к химическому равновесию оказываются для мезоскопических структур близкими величинами. Это означает, что при описании функционирования квантовых приборов учет неравновесных эффектов обязателен.

Существует довольно много возможных способов создания квантовых проводов. Они могут быть сформированы, например, при помощи прямого осаждения проводящего материала на изолирующую подложку при помощи литографа на основе сканирующего туннельного микроскопа (СТМ) или атомно-силового микроскопа (АСМ) (см., например, [131]). Другой способ создания КП состоит в использовании поверхностных структур в качестве масок для последующего травления и формирования КП в полупроводнике. В частности, подложкой может служить модулированно - легированная гетероструктура [17,117,132] как показано на Рисунке 44. Такой вариант КП привлекателен тем, что как и в НЕМТ – транзисторе [132], легирующая примесь оказывается пространственно отделенной от электронов в проводящем канале, а это существенно уменьшает их рассеяние и, соответственно, повышает подвижность.

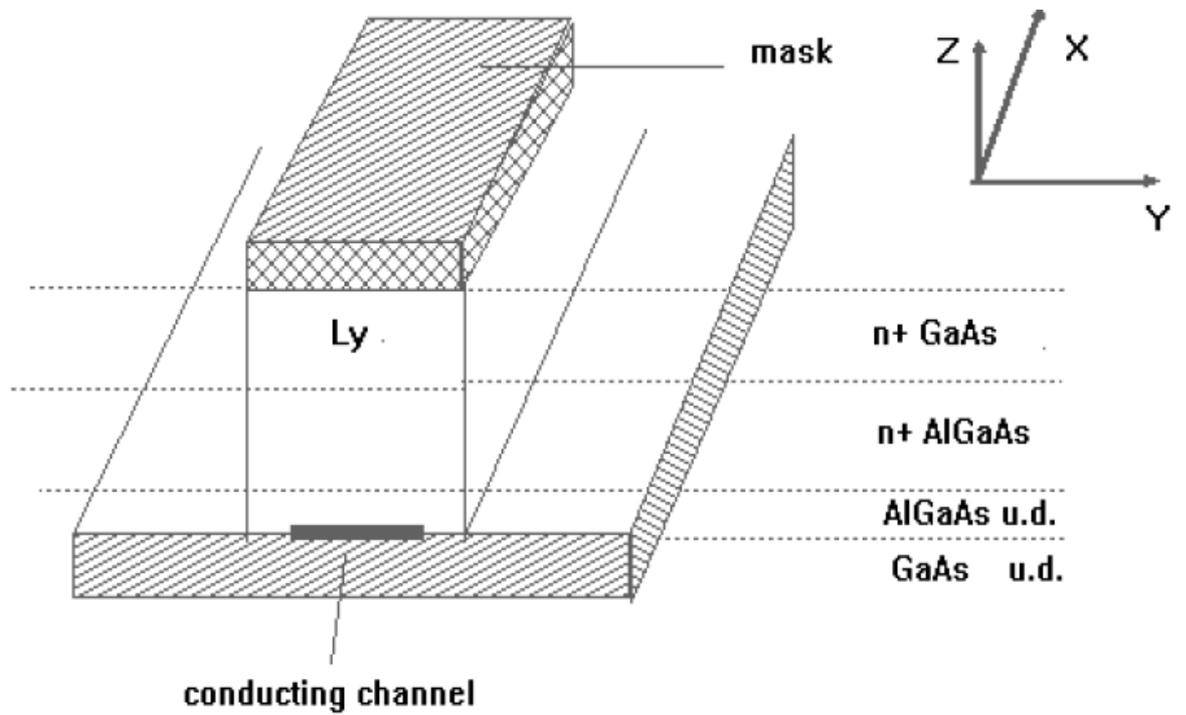


Рисунок 44. Квантовый провод на основе модулированно - легированной GaAs/AlGaAs гетероструктуры.

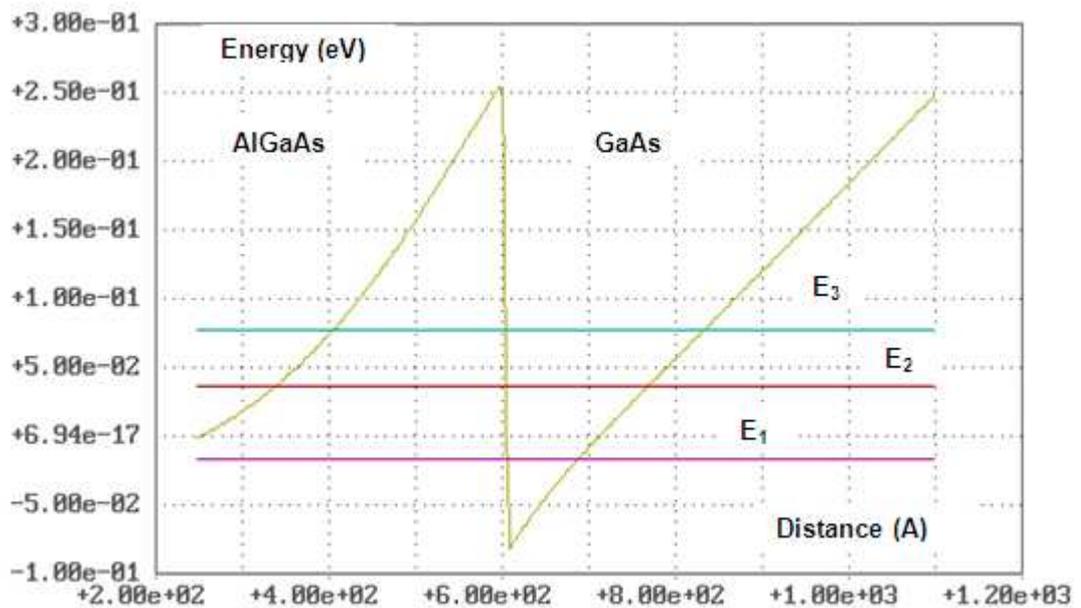


Рисунок 45. Типичный потенциальный рельеф модулированно - легированной AlGaAs/GaAs гетероструктуры в направлении OZ (см. Рисунок 44) и первые три уровня энергии в потенциальной яме на гетерогранице.

На Рисунке 45 изображены характерный потенциальный рельеф в направлении OZ и расположение трех нижних уровней энергии электронов для КП, изображенного на Рисунке 44. Следует отметить, что только первый уровень энергии является действительно дискретным, остальные – квазидискретны [60].

На рис. 46.1-46.3 показаны волновые функции электронов на трех нижних уровнях энергии.

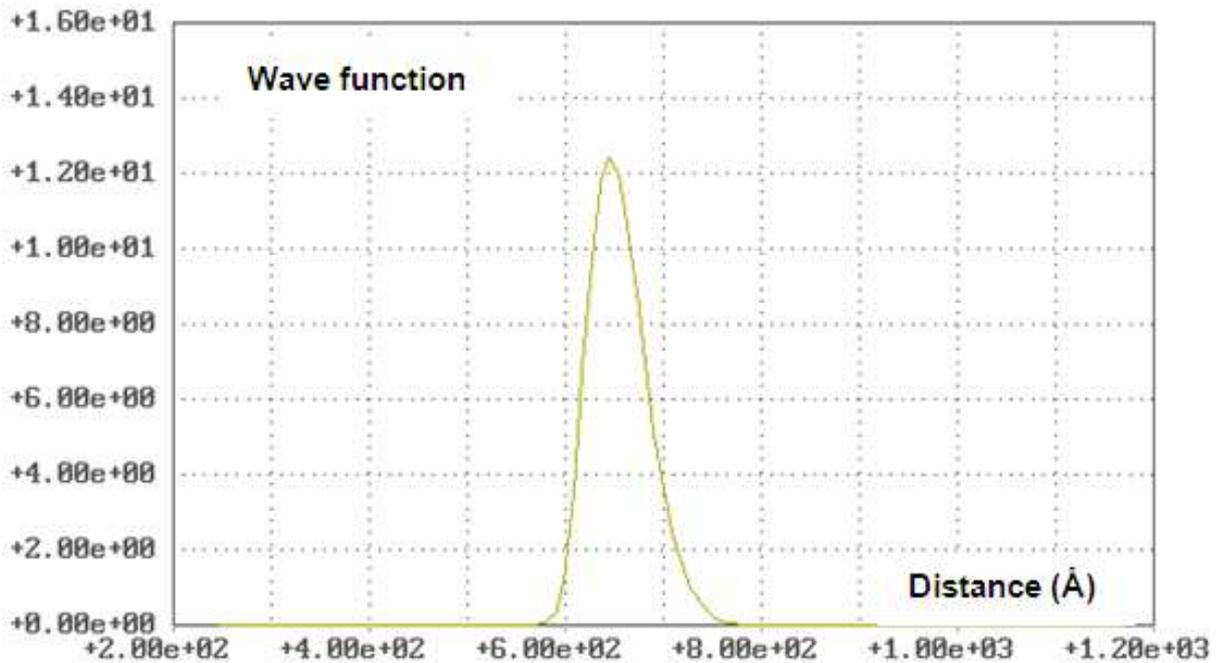


Рисунок 46.1. Волновая функция электрона на первом уровне энергии в потенциальном рельефе, изображенном на Рисунке 45.

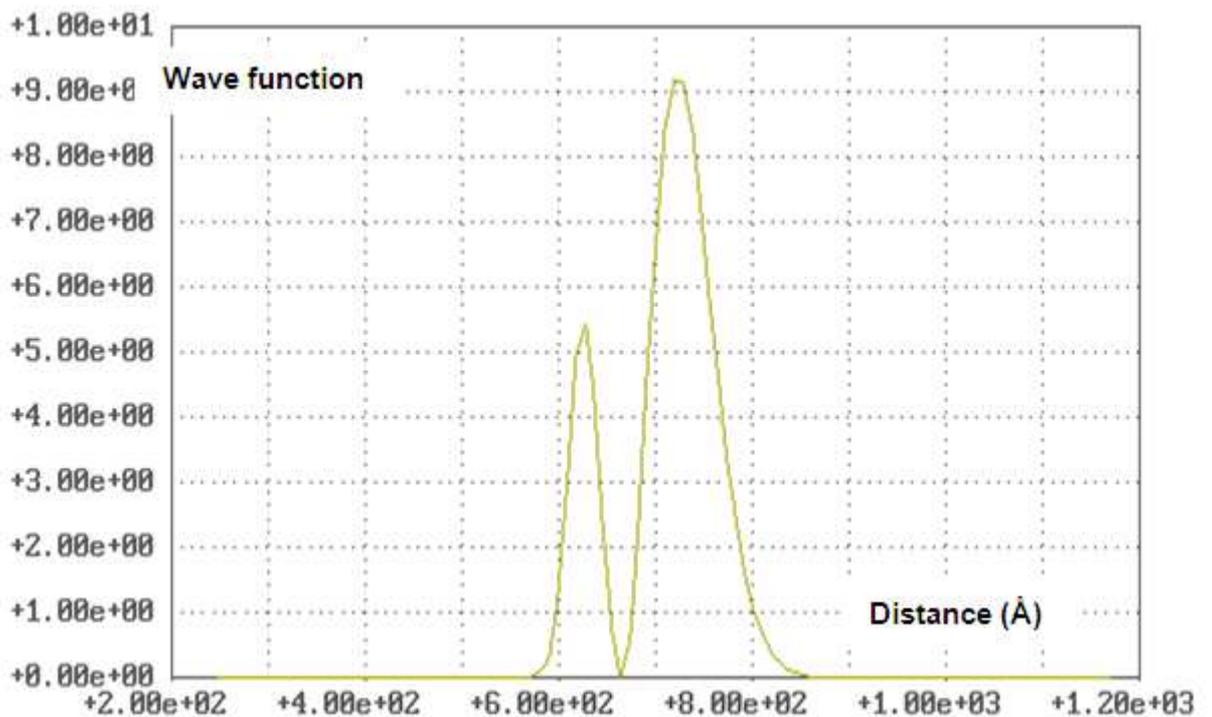


Рисунок 46.2. Волновая функция электрона на втором уровне энергии в потенциальном рельефе, изображенном на Рисунке 45.

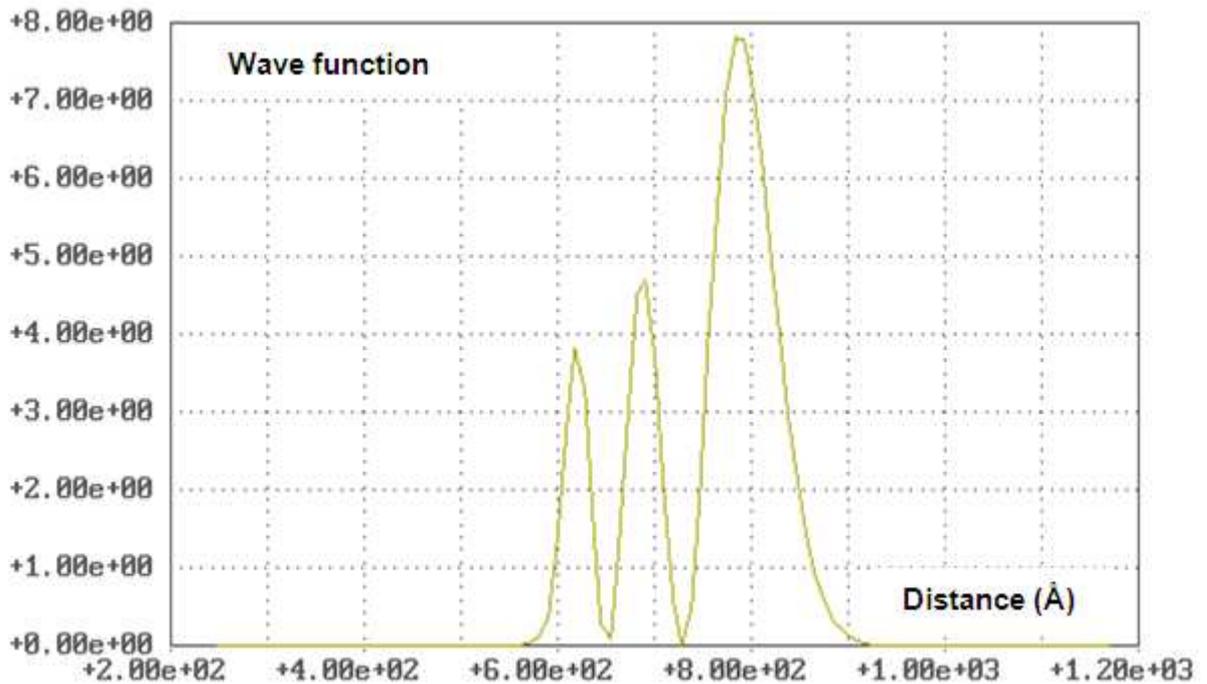


Рисунок 46.3. Волновая функция электрона на третьем уровне энергии в потенциальном рельефе, изображенном на Рисунке 45.

В качестве КП, как мы уже отмечали, может выступать углеродная нанотрубка – объект с очень интересными физическими и химическими свойствами. Исследованию нанотрубок, различным способам их создания и некоторым применениям в электронике посвящена достаточно обширная литература [2,129,130]. В частности, на основе КП из углеродной нанотрубки созданы полевые квантовые транзисторы с хорошими электрическими характеристиками [2,128].

В дальнейшем при моделировании переноса заряда в КП и приборах на их основе будем считать, что потенциальный рельеф для электронов в КП в OZ и OY направлениях представляет собой некоторые потенциальные ямы, параметры которых вычисляются или задаются в зависимости от специфики конкретной задачи. Потенциальный рельеф в направлении переноса заряда OX мы будем вычислять самосогласованным образом.

6.2. Модель переноса заряда в квантовом проводе

Развитый выше аппарат позволяет, в принципе, построить полную трехмерную модель переноса заряда в КП и приборах на их основе. Однако применение такой достаточно сложной модели будет в полной мере оправдано только в том случае, когда целью моделирования является оптимизация конструкции и параметров технологии изготовления уже созданного прибора.

Когда речь идет о моделировании, как о средстве поиска новых приборных конструкций и выявления основных эффектов, лежащих в основе их функционирования, то вполне достаточным является упрощенный подход, близкий к тому, что применялся выше для описания переноса заряда в РТС.

Поскольку граничные условия, связывающие химические потенциалы электронов с измеряемыми величинами - приложенными напряжениями, можно сформулировать только на достаточно большом расстоянии от областей неоднородности потенциального рельефа, то мы обязаны включать в моделируемую область не только квантовый провод, но и контакты. Последние, как правило, представляют собой области объемного (трехмерного) материала. Поэтому далее под КП будет пониматься структура, включающая в себя контактные области и одномерный проводящий канал (ПК), расположенный между этими областями. Проводящий канал будем считать однородным.

Заметим, что, как и в случае РТД, квантовый прибор декомпозирован на собственно активную область, обладающую квантовыми свойствами – проводящий канал, и классические контактные области, без которых невозможно измерение характеристик прибора.

6.2.1. Интерпретируемые состояния электронов

Предположим, что КП сосредоточен внутри прямоугольного параллелепипеда, ребра которого параллельны осям координат OX , OY и OZ и имеют длины L_x , L_y и L_z соответственно (см. Рисунок 47). Вершины

параллелепипеда расположены в точках с координатами (x_e, y_e, z_e) , $(x_e, y_e + L_y, z_e)$, $(x_e, y_e, z_e + L_z)$, $(x_e, y_e + L_y, z_e + L_z)$, (x_c, y_e, z_e) , $(x_c, y_e + L_y, z_e)$, $(x_c, y_e, z_e + L_z)$, $(x_c, y_e + L_y, z_e + L_z)$.

Боковые грани КП, то есть сечения КП плоскостями $y = y_e$, $y = y_e + L_y$, $z = z_e$ и $z = z_e + L_z$, будем считать границами типа Γ_0 , а торцы, образованные пересечениями КП с плоскостями $x = x_e$ и $x = x_c$ – границами типа Γ_1 (контактами).

Проводящий канал внутри квантового провода ограничен плоскостями $x = x_{e0}$ и $x = x_{c0}$. Для координат x_{e0} и x_{c0} справедливы неравенства

$$x_e < x_{e0} < x_{c0} < x_c$$

Отрезки $[x_e, x_{e0}]$ и $[x_{c0}, x_c]$ представляют собой контактные области.

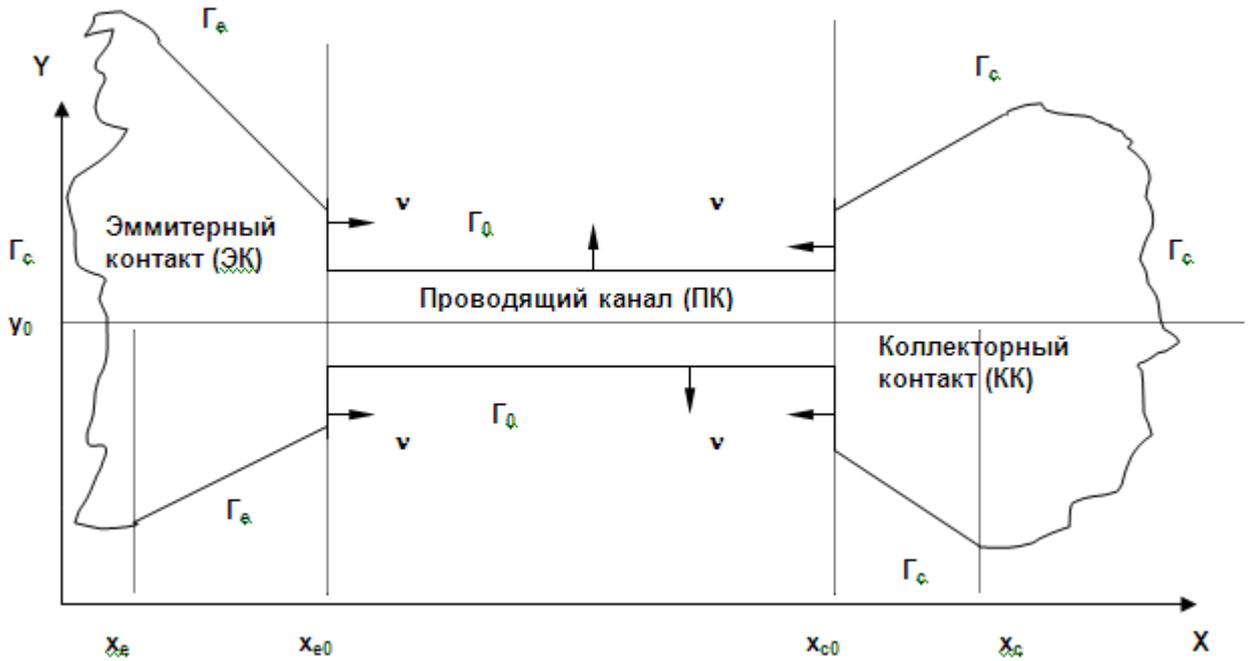


Рисунок 47. Иллюстрация к описанию квантового провода.

Предположим, что потенциал U^{ext} в уравнении Шредингера (26) представим в виде

$$U^{ext}(x, y, z) = U^{ext}(x) + U^{ext}(x, y) + U^{ext}(x, z).$$

Будем считать, что внутри каждого из отрезков $[x_e, x_{e0}]$, $[x_{e0}, x_{c0}]$, $[x_{c0}, x_c]$ потенциалы $U^{ext}(x, y)$ и $U^{ext}(x, z)$ от переменной x не зависят, то есть являются кусочнопостоянными функциями x . Тогда в ПК и в контактных областях справедливо представление потенциала $U^{ext}(x, y, z)$

$$U^{ext}(x, y, z) = U^{ext}(x) + U^{ext}(y) + U^{ext}(z) \tag{352}$$

где функциями $U^{\text{ext}}(y)$ и $U^{\text{ext}}(z)$ задаются рельефы квантовых ям в $0Y$ и $0Z$ направлениях. Не будем заранее ограничивать возможный вид потенциалов $U^{\text{ext}}(y)$ и $U^{\text{ext}}(z)$. Потребуем только, чтобы они были непрерывными функциями координат постоянными вне квантового провода, то есть в областях $y \in (-\infty, y_e) \cup (y_e + L_y, \infty)$ и $z \in (-\infty, z_e) \cup (z_e + L_z, \infty)$ на каждом из рассматриваемых отрезков. Потенциал $U^{\text{ext}}(x)$ будем считать непрерывным всюду в области КП.

Подставив (352) в уравнение Шредингера из системы (220), получим уравнение, справедливое на каждом из участков КП

$$\{(\hbar^2/2)(\nabla, (m^{*-1})\nabla) + (E(\lambda) - U^{\text{ext}}(x) - U^{\text{ext}}(y) - U^{\text{ext}}(z) + e\phi(x, y, z))\}\Psi(x, y, z, \lambda) = 0. \quad (353)$$

Волновую функцию $\Psi(x, y, z, \lambda)$ будем искать в виде

$$\Psi(x, y, z, \lambda) = \Psi(x, y, z, E_x)\Psi(y, E_y)\Psi(z, E_z), \quad (354)$$

где функции $\Psi(y, E_y)$ и $\Psi(z, E_z)$ удовлетворяют уравнениям

$$(\hbar^2/2)d((1/m^*(y))d\Psi(y, E_y)/dy)/dy + (E_y - U^{\text{ext}}(y))\Psi(y, E_y) = 0, \quad (355)$$

$$(\hbar^2/2)d((1/m^*(z))d\Psi(z, E_z)/dz)/dz + (E_z - U^{\text{ext}}(z))\Psi(z, E_z) = 0. \quad (356)$$

Очевидно, что на каждом участке КП, где $U^{\text{ext}}(y)$ и $U^{\text{ext}}(z)$ не зависят от x , решениями (355) и (356) будут, вообще говоря, разные волновые функции $\Psi(y, E_y)$ и $\Psi(z, E_z)$. Однако в рассматриваемом приближении мы не будем решать задачу их сшивания на границах отрезков, а будем использовать выражение (354) для всего КП, подразумевая, что на каждом из отрезков в нем стоит своя волновая функция, соответствующая энергиям $E_{y,z}$. В качестве начала отсчета для E_y и E_z выберем минимальные в КП значения потенциалов $U^{\text{ext}}(y)$ и $U^{\text{ext}}(z)$, соответственно.

Сделанные приближения являются достаточно грубыми. В частности, мы допускаем разрывность волновых функций на границах ПК и контактных областей. Физически эти допущения оправдываются тем, что в контактных областях энергетический спектр носителей заряда является квазинепрерывным, и всегда найдутся волновые функции, соответствующие энергиям состояний дискретного спектра в ПК. Более корректное рассмотрение требует решения трехмерной задачи для уравнения Шредингера.

Чтобы решить уравнения (355) и (356) и найти допустимые значения энергий E_y и E_z , необходимо сформулировать граничные условия для волновых функций $\Psi(y, E_y)$ и $\Psi(z, E_z)$. Физические допущения, на которых основываются граничные условия для уравнений (355), (356) и (353), ничем не отличаются от тех, что были использованы в задаче об РТС. Точно также состояния непрерывного спектра должны переходить в плоские волны в пределе плоских потенциальных рельефов, а состояния дискретного спектра должны быть нормированы на «одну частицу».

Очевидно, что для описания КП наиболее существенными являются решения уравнений (355) и (356), относящиеся к дискретному спектру. Согласно данному выше определению КП, в одномерном проводящем канале состояния непрерывного спектра отделены от нижнего энергетического уровня более, чем на $k_B T$, поэтому их вклад в концентрацию носителей заряда и потоки в равновесных ситуациях оказывается существенно меньшим вклада состояний дискретного спектра. Тем не менее, мы будем учитывать все возможные электронные состояния, описываемые уравнениями (355) и (356) и граничными условиями, приведенными ниже, поскольку в различных режимах функционирования КП и приборов на их основе вклад различных электронных состояний в электрические характеристики может меняться.

Волновые функции состояний дискретного спектра будем обозначать как $\Psi_d(\dots)$. Состояния непрерывного спектра будем называть «левыми» и «правыми» в зависимости от того, откуда вдоль соответствующей оси координат распространяется падающая волна: из $-\infty$ или из $+\infty$. Волновые функции «левых» и «правых» состояний обозначим как $\Psi_l(\dots)$ и $\Psi_r(\dots)$.

При формулировке граничных условий с целью сокращения количества формул удобно ввести некоторую обобщенную координату ξ , принимающую при необходимости значения x , y или z . Уравнения (355) и (356) запишутся тогда как одно уравнение

$$(\hbar^2/2)d((1/m^*(\xi))d\Psi(\xi, E_\xi)/d\xi)/d\xi + (E_\xi - U^{\text{ext}}(\xi))\Psi(\xi, E_\xi) = 0. \quad (357)$$

при $\xi = y$ и $\xi = z$ соответственно.

Определим импульсы

$$p_{\xi}(\xi_e) = (2m^*(\xi_e)(E_{\xi} - U^{\text{ext}}(\xi_e)))^{1/2}, E_{\xi} > U^{\text{ext}}(\xi_e),$$

$$p_{\xi}(\xi_c) = (2m^*(\xi_c)(E_{\xi} - U^{\text{ext}}(\xi_c)))^{1/2}, E_{\xi} > U^{\text{ext}}(\xi_c).$$

и декременты затухания

$$q_{\xi}(\xi_e) = (2m^*(\xi_e)(U^{\text{ext}}(\xi_e) - E_{\xi}))^{1/2}, E_{\xi} < U^{\text{ext}}(\xi_e),$$

$$q_{\xi}(\xi_c) = (2m^*(\xi_c)(U^{\text{ext}}(\xi_c) - E_{\xi}))^{1/2}, E_{\xi} < U^{\text{ext}}(\xi_c).$$

Необходимым условием существования дискретного спектра для уравнения (357) является неравенство

$$\min\{U^{\text{ext}}(\xi)\}_{\xi \in [\xi_e, \xi_c]} < \min\{U^{\text{ext}}(\xi_e), U^{\text{ext}}(\xi_c)\}.$$

Будем считать его выполненным. Для каждого значения энергии E_{ξ} , удовлетворяющего условию

$$\min\{U^{\text{ext}}(\xi)\}_{\xi \in [\xi_e, \xi_c]} \leq E_{\xi} \leq \min\{U^{\text{ext}}(\xi_e), U^{\text{ext}}(\xi_c)\},$$

будем решать уравнение (357) с граничными условиями

$$\begin{cases} \hbar d\Psi_d(\xi_c, E_{\xi})/d\xi + q_{\xi}(\xi_c)\Psi_d(\xi_c, E_{\xi}) = 0, \\ \hbar d\Psi_d(\xi_e, E_{\xi})/d\xi - q_{\xi}(\xi_e)\Psi_d(\xi_e, E_{\xi}) = 0, \end{cases} \quad (358)$$

соответствующими «сшиванию» в точках $\xi = \xi_{c,e}$ решения $\Psi_d(\xi, E_{\xi})$ внутри области $\xi \in [\xi_e, \xi_c]$ с решениями уравнения (161), экспоненциально затухающими вне этой области, пропорциональными $\exp\{\pm q_{\xi}(\xi_{e,c}) \xi\}$. Под «сшиванием», как обычно, подразумевается требование непрерывности волновой функции и ее первой производной, обеспечивающее сохранение потока плотности вероятности (см. условия (288)).

Нетривиальные решения краевой задачи (357), (358) дадут волновые функции $\Psi_d(\xi, E_{\xi})$, а энергии

$$E_{\xi} = E_{\xi}(k_{\xi}),$$

при которых такие решения существуют, дадут энергетический спектр дискретных состояний. Индекс k_{ξ} принимает положительные целые значения от 1

до K_ξ и нумерует связанные состояния в порядке возрастания их энергии, значение $k_\xi = 1$ соответствует состоянию с минимальной возможной энергией.

Решения уравнения (357), удовлетворяющие граничному условию (358), будем нормировать согласно соотношению

$$\int_{-\infty}^{\infty} |\Psi_d(\xi, E_\xi)|^2 d\xi = 1. \quad (359)$$

Очевидно, что заменив ξ на y или z , мы получим вместо (358) и (359) граничные условия и условия нормировки для состояний дискретного спектра уравнений (355) и (356).

Все решения уравнения (357) с энергиями

$$E_\xi > \min\{U^{\text{ext}}(\xi_e), U^{\text{ext}}(\xi_c)\}$$

относятся к непрерывному спектру. Если $E_\xi > U^{\text{ext}}(\xi_e)$ то при $\xi < \xi_e$ волновые функции будут иметь вид

$$\Psi_l(\xi, E_\xi) = \exp\{ip_\xi(\xi_e)\xi/\hbar\} + A_{l\xi}(E_\xi)\exp\{-ip_\xi(\xi_e)\xi/\hbar\}.$$

Соответственно, при $\xi > \xi_c$ для этих же состояний выбираем либо прошедшую волну

$$\Psi_l(\xi, E_\xi) = B_{l\xi}(E_\xi)\exp\{ip_\xi(\xi_c)\xi/\hbar\},$$

если $E_\xi > U^{\text{ext}}(\xi_c)$, либо затухающее с возрастанием ξ решение

$$\Psi_l(\xi, E_\xi) = B_{l\xi}(E_\xi)\exp\{-q_\xi(\xi_c)\xi/\hbar\},$$

если $E_\xi < U^{\text{ext}}(\xi_c)$.

Аналогично, для «правых» состояний, если $E_\xi > U^{\text{ext}}(\xi_c)$, то при $\xi > \xi_c$ волновые функции будут иметь вид

$$\Psi_r(\xi, E_\xi) = \exp\{-ip_\xi(\xi_c)\xi/\hbar\} + A_{r\xi}(E_\xi)\exp\{ip_\xi(\xi_c)\xi/\hbar\}.$$

При $\xi < \xi_e$ для этих же состояний выбираем либо прошедшую волну

$$\Psi_r(\xi, E_\xi) = B_{r\xi}(E_\xi)\exp\{-ip_\xi(\xi_e)\xi/\hbar\},$$

если $E_\xi > U^{\text{ext}}(\xi_e)$, либо затухающее с убыванием ξ решение

$$\Psi_r(\xi, E_\xi) = B_{r\xi}(E_\xi)\exp\{q_\xi(\xi_e)\xi/\hbar\},$$

если $E_\xi < U^{\text{ext}}(\xi_e)$.

Подставляя эти выражения в граничные условия (288) и исключая из получившихся уравнений коэффициенты А и В, найдем в точке ξ_e

$$\begin{cases} ip_{\xi}(\xi_e)\Psi_l(\xi_e, E_{\xi}) + \hbar d\Psi_l(\xi_e, E_{\xi})/d\xi = 2ip_{\xi}(\xi_e)\exp\{ip_{\xi}(\xi_e)\xi_e/\hbar\}, \\ \hbar d\Psi_r(\xi_e, E_{\xi})/d\xi + ip_{\xi}(\xi_e)\Psi_r(\xi_e, E_{\xi}) = 0 \end{cases} \quad (360.1)$$

при $E_{\xi} > U^{\text{ext}}(\xi_e)$,

$$\begin{cases} \Psi_l(\xi_e, E_{\xi}) = 0, \\ \hbar d\Psi_r(\xi_e, E_{\xi})/d\xi - q_{\xi}(\xi_e)\Psi_r(\xi_e, E_{\xi}) = 0 \end{cases} \quad (360.2)$$

при $E_{\xi} < U^{\text{ext}}(\xi_e)$ и $E_{\xi} > U^{\text{ext}}(\xi_c)$;

и в точке ξ_c

$$\begin{cases} ip_{\xi}(\xi_c)\Psi_r(\xi_c, E_{\xi}) - \hbar d\Psi_r(\xi_c, E_{\xi})/d\xi = 2ip_{\xi}(\xi_c)\exp\{-ip_{\xi}(\xi_c)\xi_c/\hbar\}, \\ \hbar d\Psi_l(\xi_c, E_{\xi})/d\xi - ip_{\xi}(\xi_c)\Psi_l(\xi_c, E_{\xi}) = 0 \end{cases} \quad (361.1)$$

при $E_{\xi} > U^{\text{ext}}(\xi_c)$

$$\begin{cases} \Psi_r(\xi_c, E_{\xi}) = 0, \\ \hbar d\Psi_l(\xi_c, E_{\xi})/d\xi + q_{\xi}(\xi_c)\Psi_l(\xi_c, E_{\xi}) = 0 \end{cases} \quad (361.2)$$

при $E_{\xi} < U^{\text{ext}}(\xi_c)$ и $E_{\xi} > U^{\text{ext}}(\xi_e)$.

Граничные условия (360) и (361) позволяют найти волновые функции состояний непрерывного спектра для уравнения (357). Замена в (360) и (361) переменной ξ на y и z дает требуемые граничные условия для уравнений (355) и (356).

Подставив выражение (354) в уравнение (353), с учетом (355) и (356) получим уравнение для $\Psi(x, y, z, E_x)$

$$\begin{cases} (\hbar^2/2)\partial((1/m^*)\partial\Psi(x, y, z, E_x)/\partial x)/\partial x + \\ (\hbar^2/2)\partial((1/m^*)\partial\Psi(x, y, z, E_x)/\partial y)/\partial y + (\hbar^2/2)\partial((1/m^*)\partial\Psi(x, y, z, E_x)/\partial z)/\partial z + \\ (\hbar^2/m^*)(\partial\Psi(x, y, z, E_x)/\partial y)(\partial \ln \Psi(y, E_y)/\partial y) + (\hbar^2/m^*)(\partial\Psi(x, y, z, E_x)/\partial z)(\partial \ln \Psi(z, E_z)/\partial z) + \\ + (E_x - U^{\text{eff}}(x) + e\phi(x, y, z))\Psi(x, y, z, E_x) = 0, \end{cases} \quad (362)$$

где мы определили

$$E_x = E(\lambda) - E_{\perp}, \quad U^{\text{eff}}(x) = U^{\text{ext}}(x) + \Delta E_{\perp} \quad (363)$$

и

$$E_{\perp} = E_y + E_z - \Delta E_{\perp},$$

$$\Delta E_{\perp} = E_y(1) + E_z(1).$$

Поскольку квантование энергетического спектра в ПК приводит, фактически, к сдвигу дна зоны проводимости материала канала на величину энергии первого уровня, целесообразно включить этот сдвиг во встроенный потенциал. Волновая функция $\Psi(x,y,z,E_x)$ полностью определяет зависимость решений уравнения (353) от координаты x , а для пакетов $\Psi(y,E_y)$ и $\Psi(z,E_z)$ она является огибающей.

Как обычно, определим для них импульсы

$$p_x(x_e,y,z) = (2m^*(x_e)(E_x - U^{\text{eff}}(x_e,y,z)))^{1/2}, E_x > U^{\text{eff}}(x_e,y,z),$$

$$p_x(x_c,y,z) = (2m^*(x_c)(E_x - U^{\text{eff}}(x_c,y,z)))^{1/2}, E_x > U^{\text{eff}}(x_c,y,z)$$

и декременты затухания

$$q_x(x_e,y,z) = (2m^*(x_e)(U^{\text{eff}}(x_e,y,z) - E_x))^{1/2}, E_x < U^{\text{eff}}(x_e,y,z),$$

$$q_x(x_c,y,z) = (2m^*(x_c)(U^{\text{eff}}(x_c,y,z) - E_x))^{1/2}, E_x < U^{\text{eff}}(x_c,y,z),$$

где

$$U^{\text{eff}}(x,y,z) = U^{\text{eff}}(x) - e\phi(x,y,z).$$

Далее мы будем часто опускать аргументы y и z у функций p_x и q_x , подразумевая, что такая зависимость имеет место.

Как и для уравнения (357) необходимым условием существования дискретного спектра для уравнения (362) является неравенство

$$\min\{U^{\text{eff}}(x,y,z)\}_{x \in [x_e, x_c]} < \min\{U^{\text{eff}}(x_e,y,z), U^{\text{eff}}(x_c,y,z)\}.$$

Если это неравенство выполнено, то для каждого значения энергии E_x , удовлетворяющего условию

$$\min\{U^{\text{eff}}(x,y,z)\}_{x \in [x_e, x_c]} \leq E_x \leq \min\{U^{\text{eff}}(x_e,y,z), U^{\text{eff}}(x_c,y,z)\},$$

будем искать решения уравнения (362), удовлетворяющие граничным условиям

$$\Psi_d(x, y_{e,c}, z_{e,c}, E_x) = 0,$$

$$\begin{cases} \hbar d\Psi_d(x_c, y, z, E_x)/dx + q_x(x_c, y, z)\Psi_d(x_c, y, z, E_x) = 0, \\ \hbar d\Psi_d(x_e, y, z, E_x)/dx - q_x(x_e, y, z)\Psi_d(x_e, y, z, E_x) = 0, \end{cases}$$

(364)

соответствующим «сшиванию» на гранях $x=x_{e,c}$ решения внутри области $x \in [x_e, x_c]$ с экспоненциально затухающими по x решениями вне этой области, а на гранях $y=y_{e,c}$ и $z=z_{e,c}$ - с тривиальным решением.

Нетривиальные решения краевой задачи (362), (364) как обычно, дадут волновые функции $\Psi_d(x,y,z,E_x)$, а энергии

$$E_x = E_x(k_x),$$

при которых такие решения существуют, дадут энергетический спектр дискретных состояний. Индекс k_x принимает положительные целые значения от 1 до K_x и нумерует связанные состояния в порядке возрастания их энергии, значение $k_x = 1$ соответствует состоянию с минимальной возможной энергией.

Для нормировки волновых функций дискретного спектра выберем условие

$$\int_{-\infty}^{\infty} \int \int |\Psi_d(x,y,z,E_x)|^2 dx dy dz = 1. \quad (365)$$

Все решения уравнения (362) с энергиями

$$E_x > \min\{U^{\text{eff}}(x_e, y, z), U^{\text{eff}}(x_c, y, z)\}$$

относятся к непрерывному спектру. Будем требовать, чтобы эти решения не создавали потоков через боковые грани КП

$$\partial\Psi(x,y,z,E_x)/\partial y|_{y=y_e, y_c} = 0, \quad \partial\Psi(x,y,z,E_x)/\partial z|_{z=z_e, z_c} = 0. \quad (366)$$

Если $E_x > U^{\text{eff}}(x_e, y, z)$ то при $x < x_e$ волновые функции будут иметь вид

$$\Psi_l(x,y,z,E_x) = \exp\{ip_x(x_e)x/\hbar\} + A_{lx}(y,z,E_x)\exp\{-ip_x(x_e)x/\hbar\}.$$

Соответственно, при $x > x_c$ для этих же состояний выбираем либо прошедшую волну

$$\Psi_l(x,y,z,E_x) = B_{lx}(y,z,E_x)\exp\{ip_x(x_c)x/\hbar\},$$

если $E_x > U^{\text{eff}}(x_c, y, z)$, либо затухающее с возрастанием x решение

$$\Psi_l(x,y,z,E_x) = B_{lx}(y,z,E_x)\exp\{-q_x(x_c)x/\hbar\},$$

если $E_x < U^{\text{eff}}(x_c, y, z)$.

Аналогично для «правых» состояний, если $E_x > U^{\text{eff}}(x_c, y, z)$, то при $x > x_c$ волновые функции будут иметь вид

$$\Psi_r(x,y,z,E_x) = \exp\{-ip_x(x_c)x/\hbar\} + A_{rx}(y,z,E_x)\exp\{ip_x(x_c)x/\hbar\}.$$

При $x < x_e$ для этих же состояний выбираем либо прошедшую волну

$$\Psi_r(x,y,z,E_x) = B_{rx}(y,z,E_x)\exp\{-ip_x(x_e)x/\hbar\},$$

если $E_x > U^{\text{eff}}(x_e, y, z)$, либо затухающее с убыванием x решение

$$\Psi_r(x,y,z,E_x) = B_{rx}(y,z,E_x)\exp\{q_x(x_e)x/\hbar\},$$

если $E_x < U^{\text{eff}}(x_e,y,z)$.

Подставляя эти выражения в граничные условия (288) и исключая из получившихся уравнений коэффициенты А и В, получим в точках x_e и x_c граничные условия, полностью совпадающие с условиями (360) и (361), в которых только необходимо произвести замены

$$p_\xi(\xi_{e,c}) \rightarrow p_x(x_{e,c},y,z), q_\xi(\xi_{e,c}) \rightarrow q_x(x_{e,c},y,z), E_\xi \rightarrow E_x, U^{\text{ext}}(\xi_{e,c}) \rightarrow U^{\text{eff}}(x_{e,c},y,z), \xi \rightarrow x$$

Решение уравнений (362) с граничными условиями (364), (366) и (360), (361) позволяет найти волновые функции $\Psi_d(x,y,z,E_x)$, $\Psi_l(x,y,z,E_x)$ и $\Psi_r(x,y,z,E_x)$. Однако сформулированная краевая задача является трехмерной и достаточно сложной для численного решения. Ее можно упростить и свести к одномерной, если учесть, что вся зависимость $\Psi(x,y,z,E_x)$ от координат y и z определяется зависимостью от них электростатического потенциала $\phi(x,y,z)$. Этот потенциал является решением уравнения Пуассона из системы (220), которое применительно к рассматриваемой задаче мы подробно обсудим немного позже. Зависимость от координат правой части уравнения Пуассона определяется суммарной концентрацией носителей заряда n и концентрацией положительных зарядов N_{int} . Для упрощения задачи мы будем полагать, что N_{int} является функцией только координаты x . Концентрация n вычисляется путем суммирования средних плотностей числа частиц в каждом интерпретируемом состоянии, то есть с учетом их функций распределения. При суммировании быстрые осцилляции, характерные для волновых функций $\Psi(y,E_y)$ и $\Psi(z,E_z)$ в различных состояниях, частично взаимно компенсируются, и зависимость правой части уравнения для $\phi(x,y,z)$ от y и z оказывается, как правило, очень слабой.

Эти соображения позволяют в первом (грубом) приближении пренебречь зависимостью волновых функций $\Psi(x,y,z,E_x)$ от y и z и считать, что

$$\Psi(x,y,z,E_x) \approx \Psi(x,E_x), \tag{367}$$

а функцию $\Psi(x,E_x)$ искать как решение уравнения (362), в котором вместо потенциала $\phi(x,y,z)$ используется его усредненное по y и z значение

$$\varphi(x) = (1/L_y L_z) \int_{y_e z_e}^{y_c z_c} \varphi(x, y, z) dy dz . \quad (368)$$

В результате для $\Psi(x, E_x)$ получим уравнение

$$(\hbar^2/2) \partial((1/m^*) \partial \Psi(x, E_x) / \partial x) / \partial x + (E_x - U^{\text{eff}}(x) + e\varphi(x)) \Psi(x, E_x) = 0, \quad (369)$$

которое отличается от уравнений (355), (356) и (357) только наличием самосогласованного электростатического потенциала $\varphi(x)$. Для постановки краевой задачи уравнение (369) следует дополнить граничными условиями типа (358) и (360), (361), а волновые функции дискретного спектра нормировать согласно условию (359).

Теперь мы можем полностью определить используемые далее интерпретируемые состояния квазичастиц (электронов) в квантовом проводе. Эти состояния описываются волновой функцией

$$\Psi(x, y, z, \lambda) = \Psi_{\lambda_x}(x, y, z, E_x) \Psi_{\lambda_y}(y, E_y) \Psi_{\lambda_z}(z, E_z) \quad (370)$$

при точном решении уравнения (362) или волновой функцией

$$\Psi(x, y, z, \lambda) = \Psi_{\lambda_x}(x, E_x) \Psi_{\lambda_y}(y, E_y) \Psi_{\lambda_z}(z, E_z), \quad (371)$$

если используются приближенные волновые функции $\Psi(x, E_x)$, получаемые при решении уравнения (369). Индексы λ_ξ ($\xi = x, y$ или z) могут принимать значения d , l или r , то есть классифицируют волновые функции по типу граничных условий. Обобщенный индекс λ содержит также проекцию спина σ носителей заряда на некоторое фиксированное направление. Структура λ для интерпретируемых состояний электронов в КП такова:

$$\lambda = \{ \sigma, E_x, E_y, E_z, \lambda_x, \lambda_y, \lambda_z \}$$

С учетом определения полной энергии, которое следует из соотношений (363)

$$E(\lambda) = E_x + E_\perp,$$

определений импульсов и декрементов затухания, физический смысл состояний, описываемых волновой функцией (370) или (371) вполне ясен.

Ниже мы будем чаще использовать определение (371), поскольку при численном моделировании решалась приближенная задача. Однако все общие соотношения будут справедливы и при использовании волновой функции (370).

6.2.2. Уравнения для химических потенциалов

Точно так же, как в случае РТС, разделим интерпретируемые состояния электронов на три типа: «левые», «правые» и состояния дискретного спектра.

Введем обозначения

$$\begin{cases} \Psi_l(\mathbf{r}, \lambda_l) = \Psi_l(x, y, z, E_x) \Psi_{\lambda_y}(y, E_y) \Psi_{\lambda_z}(z, E_z), \\ \Psi_r(\mathbf{r}, \lambda_r) = \Psi_r(x, y, z, E_x) \Psi_{\lambda_y}(y, E_y) \Psi_{\lambda_z}(z, E_z), \\ \Psi_d(\mathbf{r}, \lambda_d) = \Psi_d(x, y, z, E_x) \Psi_{\lambda_y}(y, E_y) \Psi_{\lambda_z}(z, E_z). \end{cases} \quad (372)$$

Обобщенные индексы $\lambda_{l,r,d}$ имеют следующую структуру

$$\begin{aligned} \lambda_l &= \{\sigma, E_x, E_y, E_z, l, \lambda_y, \lambda_z\}, \\ \lambda_r &= \{\sigma, E_x, E_y, E_z, r, \lambda_y, \lambda_z\}, \\ \lambda_d &= \{\sigma, E_x, E_y, E_z, d, \lambda_y, \lambda_z\}, \end{aligned}$$

а индексы $\lambda_{y,z}$ пробегает все допустимые значения l, r и d . Таким образом, представление (372) классифицирует интерпретируемые состояния (370) по типу граничных условий в Ox направлении, то есть в направлении переноса заряда в КП.

Будем считать, что для всех объектов, соответствующих «левым» состояниям, химический потенциал одинаков и равен F_l , для всех объектов, соответствующих «правым» электронным состояниям, он также одинаков и равен F_r , и, наконец, для всех объектов, соответствующих состояниям дискретного спектра, он одинаков и равен F_d . Как и при моделировании РТС мы представили «левые», «правые» и d -электроны в виде трех разных фаз электронного газа.

Плотности потоков и концентрации носителей заряда вычисляются из стандартных выражений

$$\begin{cases} n_{l,r,d}(\mathbf{r}) = \sum_{\lambda_{l,r,d}} n(\mathbf{r}, \lambda_{l,r,d}), \quad \mathbf{n}_{l,r,d}(\mathbf{r}) = \sum_{\lambda_{l,r,d}} \mathbf{n}_{l,r,d}(\mathbf{r}, \lambda_{l,r,d}), \quad n(\mathbf{r}, \lambda_{l,r,d}) = \rho(\mathbf{r}, \lambda_{l,r,d}) f(s(\mathbf{r}, \lambda_{l,r,d})), \\ \rho(\mathbf{r}, \lambda_{l,r,d}) = \Psi_{l,r,d}^*(\mathbf{r}, \lambda_{l,r,d}) \Psi_{l,r,d}(\mathbf{r}, \lambda_{l,r,d}), \quad \mathbf{n}_{l,r,d}(\mathbf{r}, \lambda_{l,r,d}) = \mathbf{j}_{l,r,d}(\mathbf{r}, \lambda_{l,r,d}) f(s(\mathbf{r}, \lambda_{l,r,d})), \\ \mathbf{j}_{l,r}(\mathbf{r}, \lambda_{l,r,d}) = -i\hbar/2m^* (\Psi_{l,r}^*(\mathbf{r}, \lambda_{l,r}) \nabla \Psi_{l,r}(\mathbf{r}, \lambda_{l,r}) - (\nabla \Psi_{l,r}^*(\mathbf{r}, \lambda_{l,r})) \Psi_{l,r}(\mathbf{r}, \lambda_{l,r})), \quad \mathbf{j}_d(\mathbf{r}, \lambda_d) \equiv 0, \\ f(s) = (\exp(s) + 1)^{-1}, \quad s(\mathbf{r}, \lambda_{l,r,d}) = \beta \{ E_x + E_{\perp} - F_{l,r,d}(\mathbf{r}) + \tau(\mathbf{u}(\mathbf{r}, \lambda_{l,r,d}), \nabla F_{l,r,d}(\mathbf{r})) \}, \\ \mathbf{u}(\mathbf{r}, \lambda_{l,r}) = \mathbf{j}(\mathbf{r}, \lambda_{l,r}) / \rho(\mathbf{r}, \lambda_{l,r}), \quad \mathbf{u}(\mathbf{r}, \lambda_d) \equiv 0. \end{cases} \quad (373)$$

Поскольку нас интересуют статические характеристики приборов, то из уравнений (220) получим стационарные уравнения

$$\left\{ \begin{array}{l} (\nabla, \mathbf{n}_l(\mathbf{r})) = -\hbar^{-1} (F_l(\mathbf{r}) - F_r(\mathbf{r})) \sum_{\lambda_l} \sum_{\lambda_r} \{ n(\mathbf{r}, \lambda_l) \Gamma_{\lambda_l, \lambda_r} + n(\mathbf{r}, \lambda_r) \Gamma_{\lambda_r, \lambda_l} \} - \\ \quad -\hbar^{-1} (F_l(\mathbf{r}) - F_d(\mathbf{r})) \sum_{\lambda_l} \sum_{\lambda_d} \{ n(\mathbf{r}, \lambda_l) \Gamma_{\lambda_l, \lambda_d} + n(\mathbf{r}, \lambda_d) \Gamma_{\lambda_d, \lambda_l} \}, \\ (\nabla, \mathbf{n}_r(\mathbf{r})) = -\hbar^{-1} (F_r(\mathbf{r}) - F_l(\mathbf{r})) \sum_{\lambda_r} \sum_{\lambda_l} \{ n(\mathbf{r}, \lambda_r) \Gamma_{\lambda_r, \lambda_l} + n(\mathbf{r}, \lambda_l) \Gamma_{\lambda_l, \lambda_r} \} - \\ \quad -\hbar^{-1} (F_r(\mathbf{r}) - F_d(\mathbf{r})) \sum_{\lambda_r} \sum_{\lambda_d} \{ n(\mathbf{r}, \lambda_r) \Gamma_{\lambda_r, \lambda_d} + n(\mathbf{r}, \lambda_d) \Gamma_{\lambda_d, \lambda_r} \}, \\ 0 = -\hbar^{-1} (F_d(\mathbf{r}) - F_l(\mathbf{r})) \sum_{\lambda_d} \sum_{\lambda_l} \{ n(\mathbf{r}, \lambda_d) \Gamma_{\lambda_d, \lambda_l} + n(\mathbf{r}, \lambda_l) \Gamma_{\lambda_l, \lambda_d} \} - \\ \quad -\hbar^{-1} (F_d(\mathbf{r}) - F_r(\mathbf{r})) \sum_{\lambda_d} \sum_{\lambda_r} \{ n(\mathbf{r}, \lambda_d) \Gamma_{\lambda_d, \lambda_r} + n(\mathbf{r}, \lambda_r) \Gamma_{\lambda_r, \lambda_d} \}. \end{array} \right. \quad (374)$$

Рассуждая также как и при построении модели переноса заряда в РТС, мы приходим к формуле (312) для F_d и более простым уравнениям для потоков «левых» и «правых» электронов

$$\left\{ \begin{array}{l} (\nabla, \mathbf{n}_l(\mathbf{r})) = -\hbar^{-1} (F_l(\mathbf{r}) - F_r(\mathbf{r})) (n_l(\mathbf{r}) + n_r(\mathbf{r})), \\ (\nabla, \mathbf{n}_r(\mathbf{r})) = -\hbar^{-1} (F_r(\mathbf{r}) - F_l(\mathbf{r})) (n_l(\mathbf{r}) + n_r(\mathbf{r})), \end{array} \right. \quad (375)$$

которые не учитывают возможные неравновесные переходы между частицами дискретного и непрерывного спектров, но достаточны для первоначального описания переноса заряда в квантовых проводках. Далее, при необходимости, модель будет модифицироваться за счет отказа от чрезмерных в рассматриваемой физической ситуации упрощений.

Граничные условия для уравнений (312) выберем в виде

$$\left\{ \begin{array}{l} F_l(x_e, y, z) = F_r(x_e, y, z) = -eV_e, \\ F_l(x_c, y, z) = F_r(x_c, y, z) = -eV_c, \\ \partial F_{l,r}(x, y, z) / \partial y|_{y=y_e, y_c} = 0, \\ \partial F_{l,r}(x, y, z) / \partial z|_{z=z_e, z_c} = 0. \end{array} \right. \quad (376)$$

Напомним, что торцевые грани КП мы полагаем контактными областями, то есть границами типа Γ_1 , на которых заданы приложенные напряжения V_e и V_c соответственно. Боковые грани КП полагаем границами типа Γ_0 , через которые ток не протекает.

Следует заметить, что двух последних условий (376) достаточно, чтобы обеспечить нераспространение тока через боковые грани КП. В соответствии с

определениями (373) эти условия, гарантирующие нулевые градиенты химических потенциалов на соответствующих границах, делают на них нулевым последнее слагаемое в выражении для $s(\mathbf{r}, \lambda_{l,r,d})$, что означает отсутствие дрейфово-диффузионной составляющей тока через эти границы. Квантовая составляющая тока через боковые грани КП может, в принципе, существовать только в состояниях, у которых $\Psi_{\lambda_y}(y, E_y)$ и $\Psi_{\lambda_z}(z, E_z)$ относятся к непрерывному спектру. Для состояний дискретного спектра отсутствие тока гарантирует нулевой микроскопический поток. Однако в соответствии с выбранным нами разбиением интерпретируемых состояний по фазам электронного газа, микроскопические потоки, порождаемые «левыми» и «правыми» поперечными волновыми функциями, компенсируются в каждой из фаз электронного газа по отдельности, поскольку каждая фаза характеризуется локально только одним химическим потенциалом. В рамках обсуждаемой модели различаться могут только химические потенциалы самих фаз, но не частиц, их составляющих. Подчеркнем, что эти утверждения обусловлены только сделанными упрощениями. Они не верны в общем случае многофазной модели, когда каждому интерпретируемому электронному состоянию сопоставляется отдельная фаза электронного газа.

Уравнения для химических потенциалов (375) являются трехмерными. Для построения упрощенной одномерной модели, которую мы использовали в дальнейших расчетах, придется пренебречь зависимостью химических потенциалов F_l и F_r от координат y и z . Такое предположение является корректным, если поперечные размеры квантового провода меньше, чем длина релаксации неравновесных электронов в нем, то есть если

$$L_{z,y} < L_0. \quad (377)$$

По определению КП, его поперечные размеры меньше длины размерного квантования L_{dq} . Поскольку значения L_{dq} и L_0 для большинства материалов различаются несущественно, условие (377) не накладывает серьезных ограничений на область применимости одномерной модели.

Считаем далее, что химические потенциалы $F_{l,r}$ зависят только от координаты x . Обозначим через $I_{l,r}(x)$ x -компоненты плотностей потоков, а через $N_{l,r,d}(x)$ концентрации, усредненные по сечению КП

$$I_{l,r}(x) = 1/(L_y L_z) \int_{y_e Z_e}^{y_c Z_c} \int n_{x;l,r}(x,y,z) dy dz, \quad (378)$$

$$N_{l,r,d}(x) = 1/(L_y L_z) \int_{y_e Z_e}^{y_c Z_c} \int n_{l,r,d}(x,y,z) dy dz. \quad (379)$$

Интегрируя уравнения (375) по y и z в тех же пределах, что и в формулах (378) и (379), получим одномерные уравнения переноса заряда в КП

$$\begin{cases} dI_l(x)/dx = -\hbar^{-1}(F_l(x) - F_r(x))(N_l(x) + N_r(x)), \\ dI_r(x)/dx = -\hbar^{-1}(F_r(x) - F_l(x))(N_l(x) + N_r(x)). \end{cases} \quad (380)$$

При выводе уравнений (186) мы учли, что

$$\int_{y_e Z_e}^{y_c Z_c} \int \partial I_{y;l,r} / \partial y (x,y,z) dy dz = \int_{Z_e}^{Z_c} \{ I_{y;l,r}(x,y_c,z) - I_{y;l,r}(x,y_e,z) \} dz = 0,$$

$$\int_{y_e Z_e}^{y_c Z_c} \int \partial I_{z;l,r} / \partial z (x,y,z) dy dz = \int_{y_e}^{y_c} \{ I_{z;l,r}(x,y,z_c) - I_{z;l,r}(x,y,z_e) \} dy = 0.$$

Справедливость этих соотношений следует из граничных условий (376) и пояснений к ним.

В одномерном приближении вторые два условия (376) выполнены автоматически. Граничные условия для уравнений (380) следуют из соотношений (376) и имеют вид

$$F_l(x_e) = F_r(x_e) = -eV_e, \quad F_l(x_c) = F_r(x_c) = -eV_c. \quad (381)$$

6.2.3. Выражения для плотностей числа частиц и потоков

Учитывая сделанные упрощения, из выражений (373) получим соотношения для концентраций и потоков

$$\left\{ \begin{array}{l} N_{1,r,d}(x) = \sum_{\lambda_{1,r,d}} \rho(x, \lambda_{1,r,d}) f(s(x, \lambda_{1,r,d})), \quad I_{1,r}(x) = \sum_{\lambda_{1,r,d}} j_{1,r}(x, \lambda_{1,r,d}) f(s(x, \lambda_{1,r,d})), \\ \rho(x, \lambda_{1,r,d}) = 1/(L_y L_z) \int_{y_e Z_e}^{y_c Z_c} dy dz \Psi_{1,r,d}^*(\mathbf{r}, \lambda_{1,r,d}) \Psi_{1,r,d}(\mathbf{r}, \lambda_{1,r,d}), \\ j_{1,r}(x, \lambda_{1,r,d}) = -(\i\hbar/2L_y L_z) \int_{y_e Z_e}^{y_c Z_c} dy dz (1/m^*) (\Psi_{1,r}^*(\mathbf{r}, \lambda_{1,r}) \partial \Psi_{1,r}(\mathbf{r}, \lambda_{1,r}) / \partial x - (\partial \Psi_{1,r}^*(\mathbf{r}, \lambda_{1,r}) / \partial x) \Psi_{1,r}(\mathbf{r}, \lambda_{1,r})), \\ f(s) = (\exp(s) + 1)^{-1}, \quad s(x, \lambda_{1,r,d}) = \beta \{ E_x + E_{\perp} - F_{1,r,d}(x) + \tau u(x, \lambda_{1,r,d}) dF_{1,r,d}(x)/dx \}, \\ u(x, \lambda_{1,r}) = j(x, \lambda_{1,r}) / \rho(x, \lambda_{1,r}), \quad u(x, \lambda_d) \equiv 0. \end{array} \right. \quad (382)$$

В выражении для $s(x, \lambda_{1,r,d})$ мы заменили точное значение x -компоненты микроскопической скорости $u_x(\mathbf{r}, \lambda_{1,r,d})$ на ее приближенное значение $u(x, \lambda_{1,r,d})$, вычисляемое на основе усредненных по сечению ПК величин.

Суммирование по обобщенным индексам $\lambda_{1,r,d}$ в первых двух формулах (382) сводится к суммированию по спиновым состояниям, приводящему к появлению множителя 2, суммированию по всем возможным состояниям дискретного спектра (суммирование по k_x , k_y и k_z) и суммированию по всем возможным состояниям непрерывного спектра, то есть к интегрированию по импульсам p_y , p_z и p_x .

Суммирование по k_x , k_y и k_z – это именно простое суммирование, поскольку состояния дискретного спектра нормированы таким образом (см. выражения (359) и (366)), что модули соответствующих волновых функций имеют размерность обратной длины.

Суммирование по состояниям непрерывного спектра сводится к интегрированию по импульсу p_{ξ} ($\xi = x, y, z$) от нуля до бесконечности с нормировочным множителем $(2\pi\hbar)^{-1}$, обеспечивающим нужную размерность для

соответствующих квадратичных комбинаций волновых функций. Интеграл по импульсам можно представить в виде интеграла по энергиям

$$(2\pi\hbar)^{-1} \int_0^{\infty} dp_{\xi}(\dots) = (2\pi\hbar)^{-1} (m^*)^{1/2} \int_{U^{\text{ext}}(\xi_{\min})}^{\infty} dE_{\xi}(\dots) / (2(E_{\xi} - U^{\text{ext}}(\xi_{\min})))^{1/2}, \quad (383)$$

где для x-направления в качестве $U^{\text{ext}}(\xi)$ выбирается потенциал $U^{\text{ext}}(x) = e\phi(x)$, а точка ξ_{\min} – это та из двух точек ξ_e и ξ_c , в которой потенциал $U^{\text{ext}}(\xi)$ меньше.

Если понимать под суммированием по энергиям суммирование по состояниям дискретного спектра и интегрирование по состояниям непрерывного спектра в смысле выражения (383), то суммирование по $\lambda_{l,r,d}$ в формулах (382) можно представить в виде

$$\sum_{\lambda_{l,r,d}} = 2 \sum_{E_x} \sum_{E_y} \sum_{E_z} \sum_{\lambda_y=l,r,d} \sum_{\lambda_z=l,r,d} (\dots)_{l,r,d}. \quad (384)$$

Это выражение уже вполне пригодно для практического использования.

6.2.4. Уравнение для самосогласованного электростатического потенциала

Согласно (220) уравнение Пуассона, описывающее поведение электростатического потенциала ϕ в КП, имеет вид

$$(\nabla, \varepsilon \nabla \phi(\mathbf{r})) = e(n(\mathbf{r}) - N_{\text{int}}(\mathbf{r})), \quad (385)$$

где

$$n(\mathbf{r}) = n_l(\mathbf{r}) + n_r(\mathbf{r}) + n_d(\mathbf{r}). \quad (386)$$

Будем полагать, что на торцевых границах КП потенциал ϕ можно найти из условий электронейтральности

$$(n(\mathbf{r}) - N_{\text{int}}(\mathbf{r}))|_{\mathbf{r}=(x_e, y, z), (x_c, y, z)} = 0. \quad (387)$$

Напомним, что такие граничные условия корректны, если длины контактных областей

$$L_e = x_{e0} - x_e \text{ и } L_c = x_c - x_{c0}$$

много больше длины Дебая в контактной области.

На боковых гранях будем считать равной нулю нормальную производную электростатического потенциала

$$\left\{ \begin{array}{l} \partial\varphi(x,y,z)/\partial y|_{y=y_e,y_c} = 0, \\ \partial\varphi(x,y,z)/\partial z|_{z=z_e,z_c} = 0. \end{array} \right. \quad (388)$$

Эти граничные условия являются приближенными и в задачах о переносе заряда ими можно пользоваться, только если:

- диэлектрическая проницаемость в проводящем канале много больше, чем во внешней по отношению к квантовому проводу области;
- размеры областей, в которых нарушается условие локальной электронейтральности, много меньше характерных размеров структуры.

Будем считать, что эти предположения справедливы. В более общем случае, когда необходимо анализировать поверхностные явления, такие как поверхностные плазмоны [133,134], следует использовать более общие граничные условия (25).

Проинтегрируем уравнение (385) по внутреннему сечению квантового провода. С учетом определений (368) и (379), а также граничных условий (388) получим

$$d(\epsilon d\varphi(x)/dx)/dx = e(N(x) - N_{int}(x)), \quad (389)$$

где

$$N(x) = N_l(x) + N_r(x) + N_d(x). \quad (390)$$

Граничные условия (387) примут вид

$$(N(x) - N_{int}(x))|_{x=x_e,x_c} = 0. \quad (391)$$

Соотношения (389) - (391) полностью определяют граничную задачу для потенциала φ в одномерном приближении. Уравнения и граничные условия (385)-(388) позволяют рассматривать и трехмерные задачи.

Таким образом, мы сформулировали самосогласованную замкнутую одномерную модель, позволяющую приближенно описать перенос заряда в квантовом проводе.

6.3. Статические характеристики квантового провода

Построенная модель позволяет рассчитывать статические характеристики самых разнообразных квантовых проводов, в частности, проводов на основе полупроводниковых материалов, карбоновых нанотрубок, металлов, органических соединений. При этом могут быть учтены особенности каждого вида провода и способов его формирования. Эти особенности отражаются на потенциальных рельефах $U^{\text{ext}}(y)$ и $U^{\text{ext}}(z)$, временах релаксации τ , концентрациях N_{int} и т.д. В работах [78,80,93-95] опубликованы результаты вычислений, учитывающих специфику различных квантовых проводов. Обобщая опыт моделирования и сопоставления его результатов с доступными измерениями [128,131,135-138], можно утверждать, что характеристики квантовых проводов с однородным проводящим каналом во многом определяются некоторыми универсальными факторами. Ниже мы постараемся выделить эти факторы и проанализировать их влияние на электрические характеристики приборов. Такой подход позволяет использовать моделирование в качестве инструмента для поиска оптимальных характеристик, конструкций, материалов и технологий изготовления квантового прибора.

6.3.1. Результаты моделирования

Для максимального исключения влияния второстепенных деталей будем предполагать, что потенциальный рельеф ПК в OY и OZ направлениях представляет собой просто прямоугольные потенциальные ямы шириной

$$L_y = L_z = L$$

и глубиной 1 эВ. Для определенности выберем $m^* = 0,068 m_0$, то есть равной эффективной массе электронов в GaAs. При этом длина размерного квантования при комнатной температуре $L_{\text{dq}} = 25,3$ нм. Величиной L_{dq} , а значит и m^* , во многом определяются требования к технологии формирования квантового

провода. Положим, что длина провода L_x более чем в десять раз превышает L_{dq} . В расчетах, результаты которых приведены ниже, предполагалось, что $L_x = 320$ нм, длина ПК $L_{x0} = x_{c0} - x_{e0} = 220$ нм, а контактные области имеют длину по 50 нм каждая. Концентрация электронов проводимости в контактах выбиралась равной $1,20 \cdot 10^{25} \text{ м}^{-3}$, при этом длина релаксации химических потенциалов L_{rel} оказывается порядка 1,5 нм, то есть много меньше длины контактной области. Время релаксации импульса τ принималось равным $3,21 \cdot 10^{-13}$ с, а температура - 300°K .

Полагалось, что проводящий канал является локально электронейтральным. Физически это означает, что в случае бесконечно длинного ПК в отсутствии внешнего смещения в каждой его точке заряд электронов проводимости компенсируется зарядом положительных ионов кристаллической решетки. Исходная концентрация электронов проводимости в объемном материале, из которого изготовлен ПК, предполагалась равной $5 \cdot 10^{23} \text{ м}^{-3}$. Потенциальный рельеф и положение энергии Ферми для электронов при нулевом смещении

$$V_{ce} = V_c - V_e$$

в приборе, в котором в качестве ПК использован объемный материал, показан на Рисунке 48.

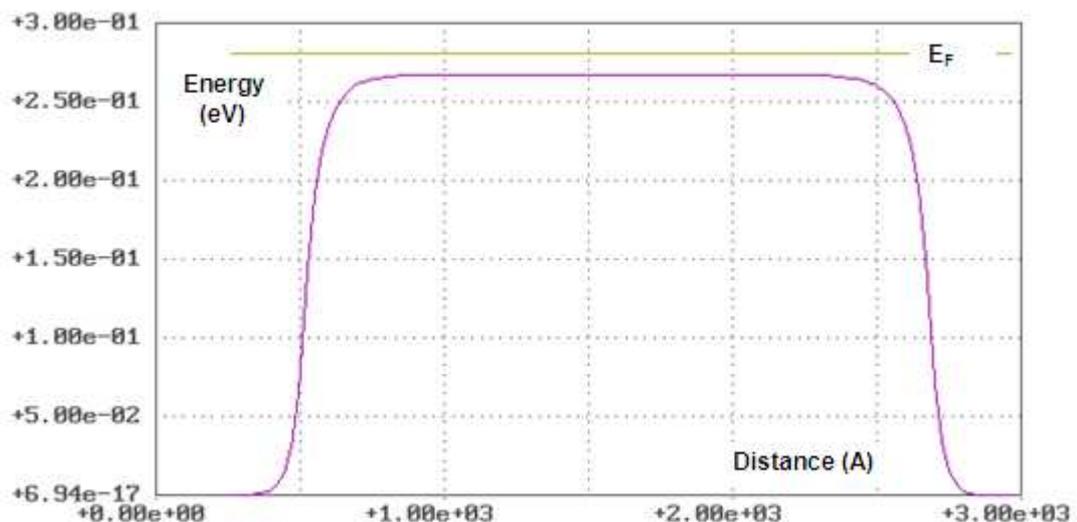


Рисунок 48. Потенциальный рельеф и энергия Ферми структуры в направлении Ox в случае, когда в качестве проводящего канала используется объемный материал ($V_{ce} = 0$).

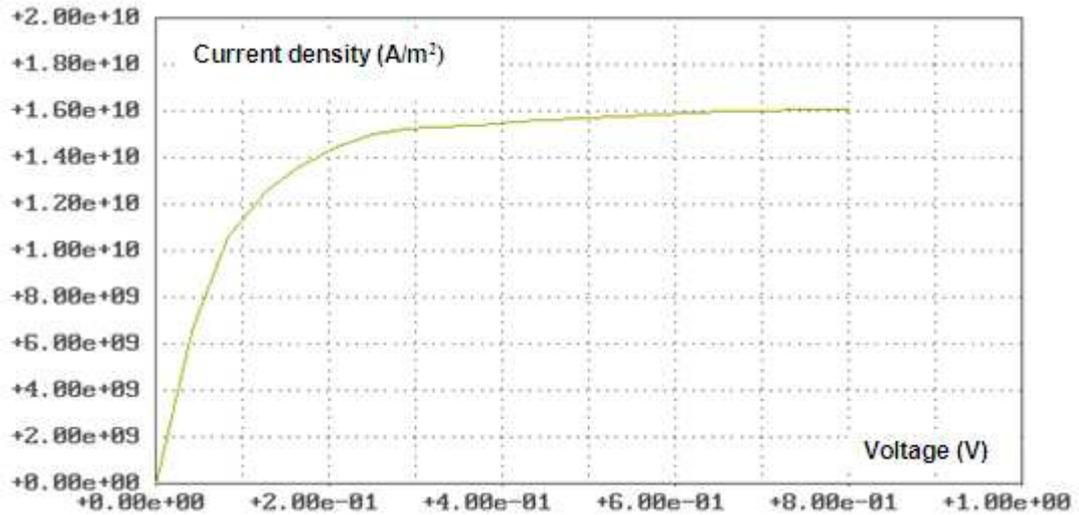


Рисунок 49. ВАХ резистора, потенциальный рельеф и энергия Ферми структуры которого показаны на Рисунке 48.

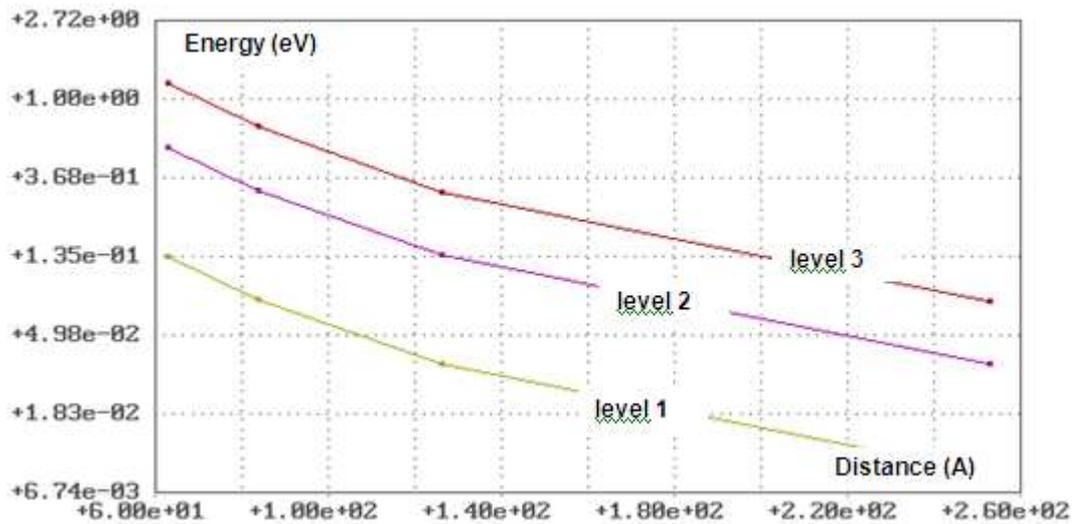


Рисунок 50. Первые три уровня размерного квантования как функция ширины прямоугольной потенциальной ямы. Точки слева направо соответствуют ширине $L = L_{dq}/4, L_{dq}/3, L_{dq}/2, L_{dq}$.

На Рисунке 49 показана рассчитанная ВАХ резистора с проводящим каналом из описанного объемного материала. Участок насыщения связан с эффектом насыщения дрейфовой скорости носителей заряда [17,27], который учитывался при моделировании.

При переходе от объемного материала к одномерному проводящему каналу начинают проявляться эффекты, обусловленные размерным квантованием. На Рисунке 50 демонстрируется зависимость положения трех первых уровней энергии электронов в одномерном проводящем канале относительно дна зоны проводимости от ширины канала L . Хорошо видно, что уровни разделены

энергетическими щелями, превышающими энергию тепловых флуктуаций $k_B T$ при комнатной температуре, равную приблизительно 0,0256 эВ.

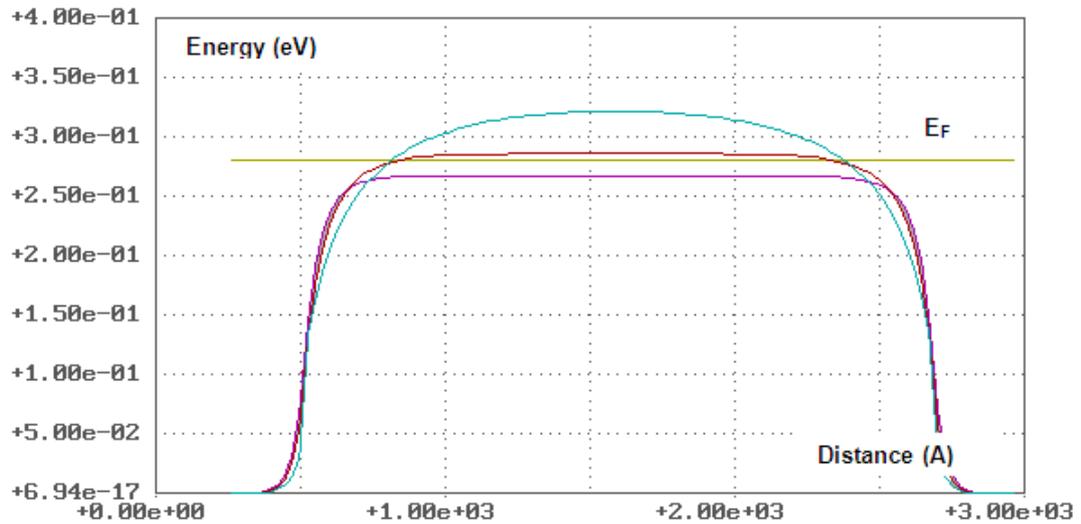


Рисунок 51. Эффективные потенциальные рельефы для электронов в направлении 0X при $V_{ce} = 0$ и различных L : 1 – объемный материал; 2 – $L = L_{dq}$; 3 – $L = L_{dq}/2$.

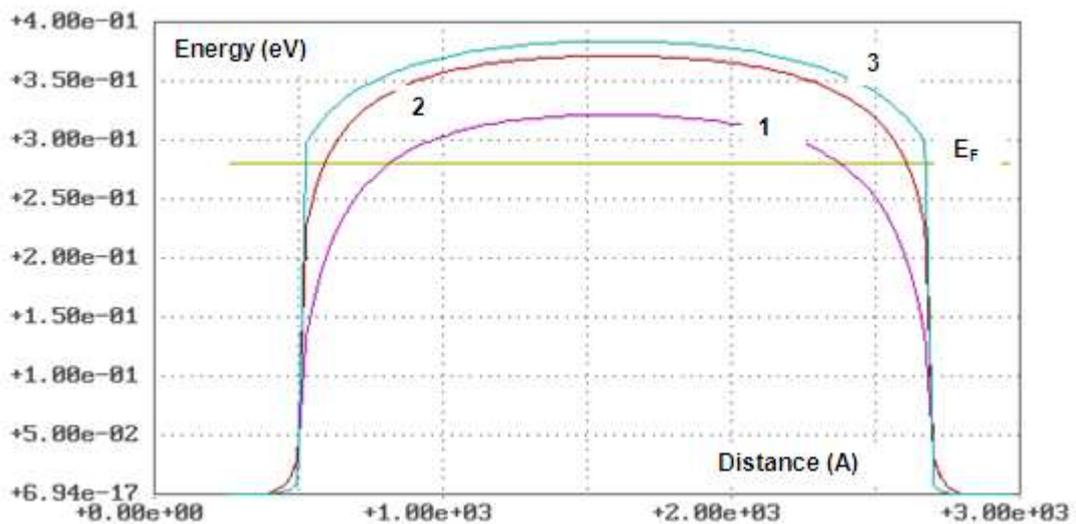


Рисунок 52. Эффективные потенциальные рельефы для электронов в направлении 0X при $V_{ce} = 0$ и различных L : 1 – $L = L_{dq}/2$; 2 – $L = L_{dq}/3$; 3 – $L = L_{dq}/4$.

Изменения в энергетическом спектре электронов проводимости приводят к соответствующим изменениям их эффективного потенциального рельефа и концентрации, что показано на Рисунках 51-53. Все это сказывается на выходных характеристиках прибора.

Рисунки 54 и 55 демонстрируют зависимость плотности тока в квантовых проводках с проводящими каналами различного поперечного размера от напряжения (удельные ВАХ). Изменения, происходящие с характеристиками КП

при уменьшении L , вполне предсказуемы. Важными на наш взгляд являются две особенности:

- появление N-образности на некоторых ВАХ;
- плотность тока на начальном линейном участке ВАХ КП значительно выше, чем должна быть у обычного резистора с той же концентрацией электронов проводимости и электрофизическими характеристиками объемного материала.

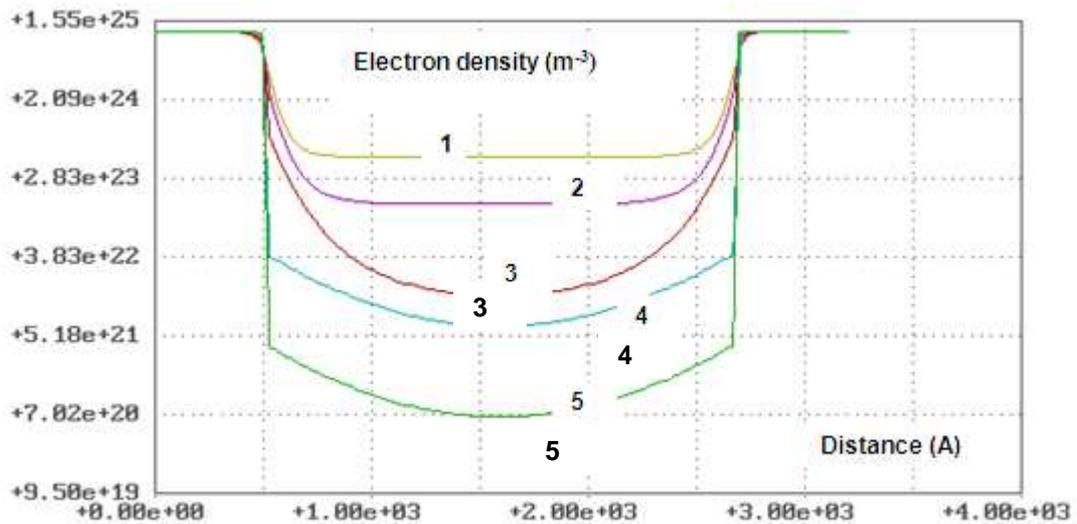


Рисунок 53. Распределение концентрации электронов в направлении Ox при $V_{ce} = 0$ и различных L : 1 – объемный материал; 2 – $L = L_{dq}$; 3 – $L = L_{dq}/2$; 4 – $L = L_{dq}/3$; 5 – $L = L_{dq}/4$.

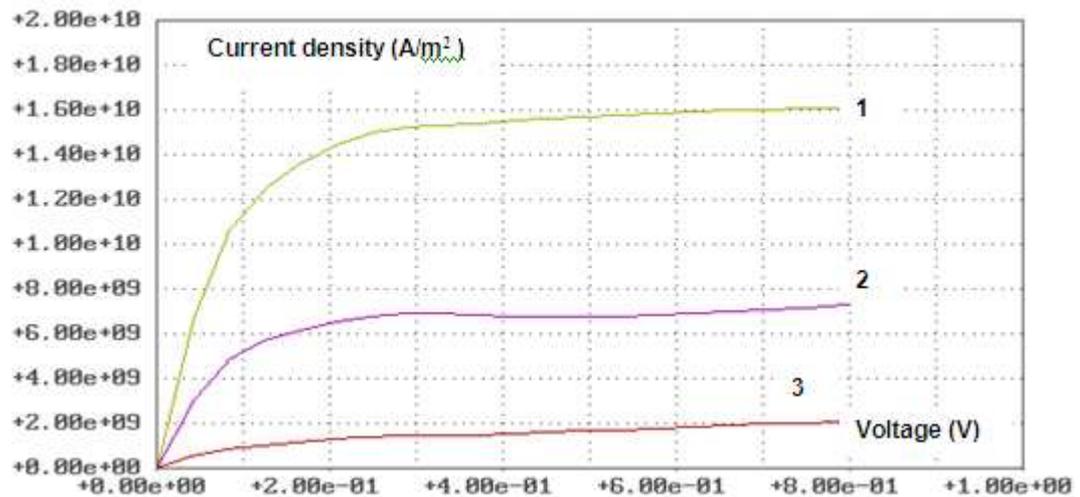


Рисунок 54. Удельные ВАХ КП при различных размерах L : 1 – объемный материал; 2 – $L = L_{dq}$; 3 – $L = L_{dq}/2$.

Первая особенность может быть объяснена по-разному. На наш взгляд она связана с одновременным проявлением эффектов размерного квантования и

насыщения дрейфовой скорости электронов. Размерное квантование приводит к возникновению изломов на ВАХ прибора, а наложение на эти изломы полевой зависимости времени релаксации импульса может проявляться в виде N – образности ВАХ. Это рассуждение легко проверить путем простых вычислений по поиску экстремумов дробно-нелинейных функций. На Рисунке 56 показаны ВАХ КП с $L = L_{dq}$, иллюстрирующие приведенное объяснение.

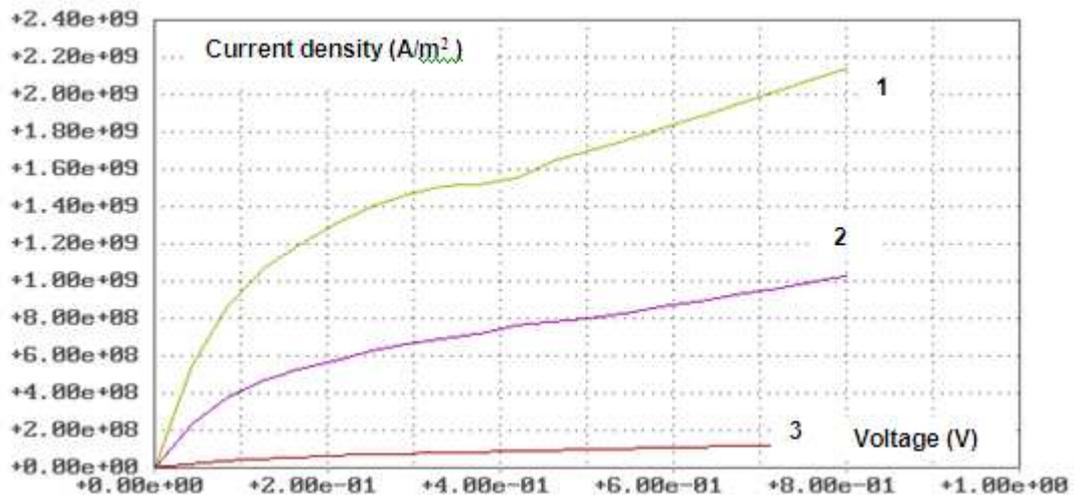


Рисунок 55. Удельные ВАХ КП при различных размерах L : 1 – $L = L_{dq}/2$; 2 – $L = L_{dq}/3$; 3 – $L = L_{dq}/4$.

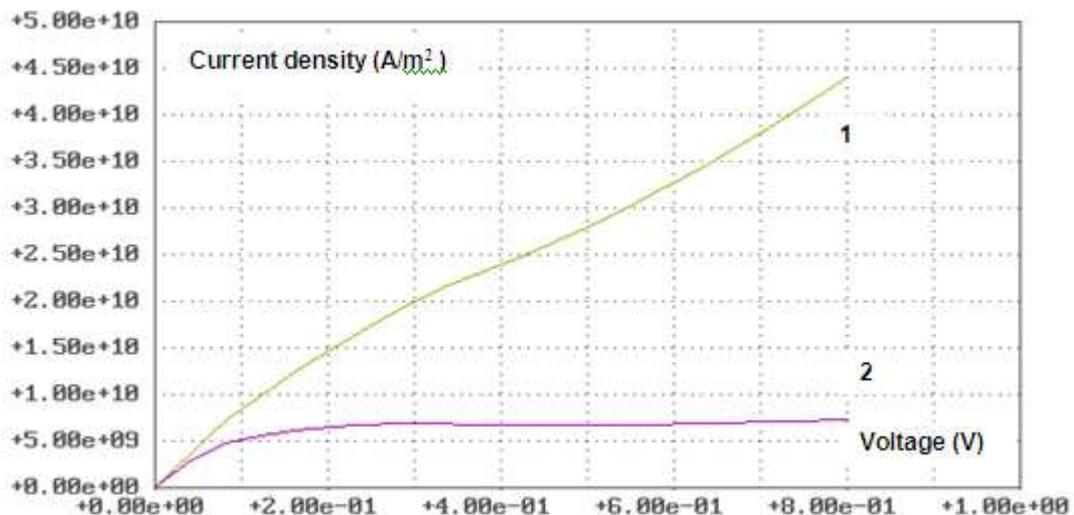


Рисунок 56. Удельные ВАХ КП с $L = L_{dq}$: 1 – без учета эффекта насыщения дрейфовой скорости электронов; 2 – с учетом этого эффекта.

Вторая особенность связана с механизмами проводимости в квантовом проводе. Кроме собственных электронов, существенную роль в процессах переноса заряда в проводе играют электроны, инжектируемые в проводящий

канал из контактных областей. Именно они определяют «избыточную» проводимость прибора. Существенная зависимость ВАХ квантового провода от концентрации электронов в контактах иллюстрируется кривыми на Рисунке 57. Процесс инжекции (или, иначе, испарительной эмиссии) носителей заряда из контактов в проводящий канал играет принципиально важную роль в формировании электрических характеристик квантовых проводов и приборов на их основе [94,95,139,140-143]. Этому явлению будет посвящен отдельный параграф.

На Рисунке 58 показаны расчетные ВАХ КП с различными поперечными размерами. В отличие от плотности тока (см. Рисунки 54, 55), которая при уменьшении L от L_{dq} до $L_{dq}/4$, уменьшается в 55 раз, то есть менее, чем на два порядка, ток уменьшается почти в 1000 раз. Это не удивительно, поскольку площадь проводящего канала уменьшается в 16 раз.

При дальнейшем уменьшении поперечных размеров проводящего канала начинает наблюдаться более сильная зависимость характеристик квантового провода от поперечного размера (L). Это связано с тем, что энергетическая щель между максимальным значением эффективного потенциала и энергией Ферми при $L < L_{dq}/4$ начинает существенно (в четыре раза) превосходить энергию тепловых флуктуаций $k_B T$ ($k_B T = 0,026$ эВ при $T = 300$ К). При этом качественно меняется форма эффективного потенциала для электронов в направлении переноса заряда, и резко уменьшается концентрация электронов проводимости в ПК (см. Рисунки 59 и 60).

Указанные изменения приводят к резкому уменьшению плотности тока в КП с размерами ПК меньшими, чем $L_{dq}/4$. На Рисунке 61 демонстрируются удельные ВАХ КП, у которых L различаются очень незначительно (менее чем на 10%), а плотности токов изменяются при этом на пять порядков. Показательной является зависимость плотности тока J в КП, отнесенной к плотности тока в приборе с ПК из объемного материала J_{vol} , от поперечного размера ПК, выраженного в единицах L_{dq} (см. Рисунок 62). Отчетливо видно, что резкое

уменьшение плотности тока начинается с размера L , равного примерно одной трети длины размерного квантования L_{dq} .

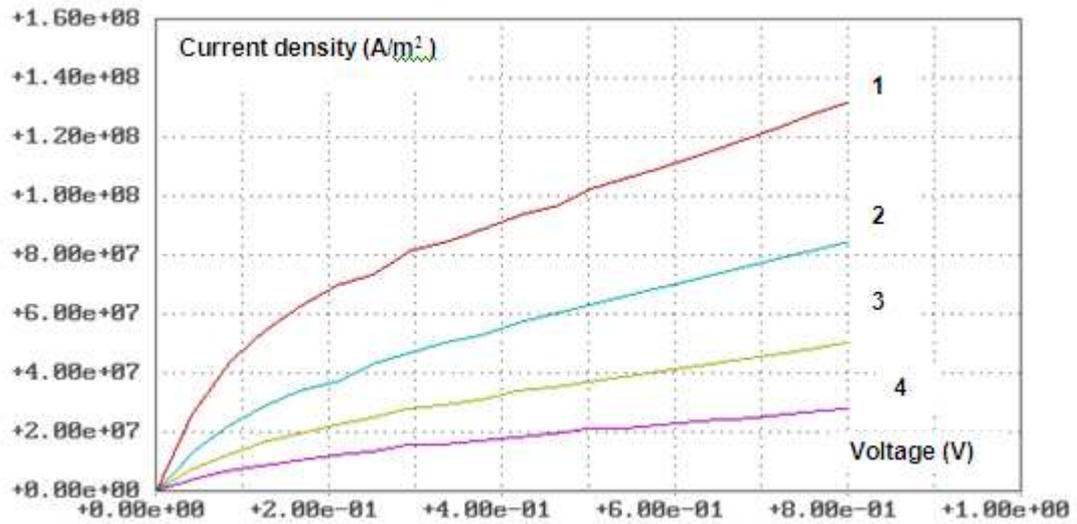


Рисунок 57. Удельные ВАХ КП с $L = L_{dq}/4$ при различных концентрациях электронов проводимости в контактных областях: 1 – $1,20 \cdot 10^{25} \text{ м}^{-3}$; 2 – $1,10 \cdot 10^{25} \text{ м}^{-3}$; 3 – $1,00 \cdot 10^{25} \text{ м}^{-3}$; 4 – $9 \cdot 10^{24} \text{ м}^{-3}$.

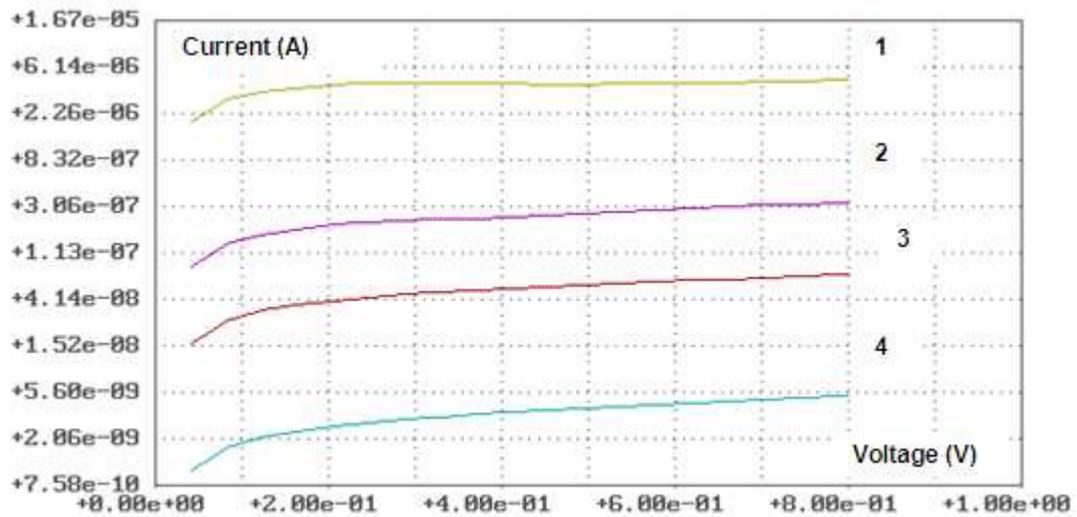


Рисунок 58. ВАХ КП с различными поперечными размерами: 1 – $L = L_{dq}$; 2 – $L = L_{dq}/2$; 3 – $L = L_{dq}/3$; 4 – $L = L_{dq}/4$.

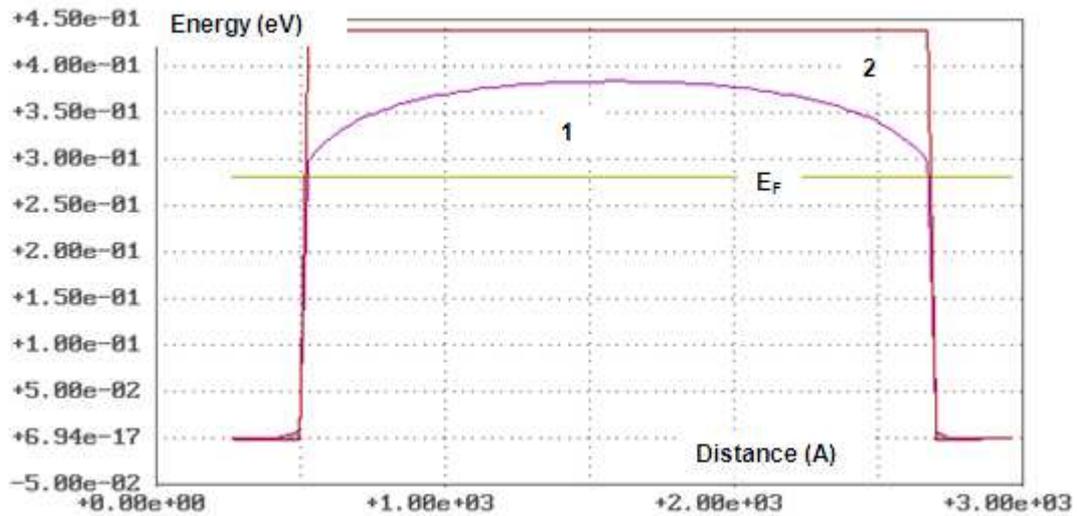


Рисунок 59. Эффективные потенциальные рельефы для электронов в направлении OX при $V_{ce} = 0$ и различных L : 1 – $L = L_{dq}/4$; 2 – $L = L_{dq}/5$

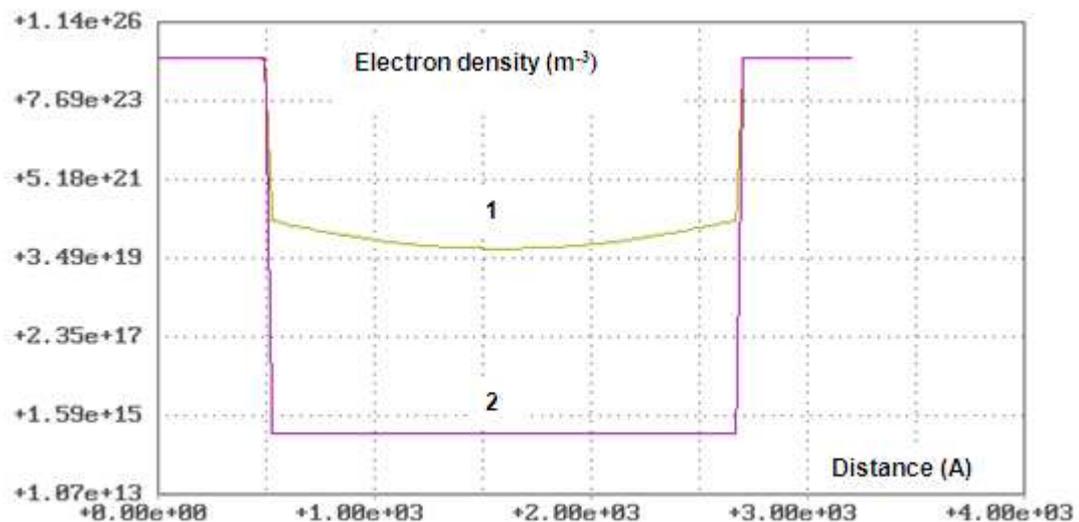


Рисунок 60. Распределение концентрации электронов в направлении OX при $V_{ce} = 0$ и различных L : 1 – $L = L_{dq}/4$; 2 – $L = L_{dq}/5$.

На кривой, изображенной на Рисунке 62, можно выделить две точки излома. Одна находится в окрестности уже упоминавшегося размера $L_{dq}/3$, а вторая – вблизи поперечного размера $L_{dq}/4$. На Рисунках 51, 52 и 59 показаны эффективные потенциальные рельефы для электронов в направлении переноса заряда при различных L . Отчетливо видно, что, начиная с поперечных размеров ПК, близких к $L_{dq}/3$, происходит качественное изменение формы потенциального рельефа. Оно связано с тем, что величина энергетической щели между дном зоны проводимости и первым уровнем размерного квантования становится сравнимой с энергией Ферми. Уменьшается влияние на потенциал заряда примеси в проводе, и

рельеф приобретает все более прямоугольную форму. Соответственно, уменьшается и концентрация электронов проводимости в канале. При $L < L_{dq}/4$, как мы уже отмечали, энергетическая щель между максимальным значением эффективного потенциала и энергией Ферми становится больше, чем $4k_B T$, и концентрация электронов в ПК убывает с уменьшением L практически экспоненциально.

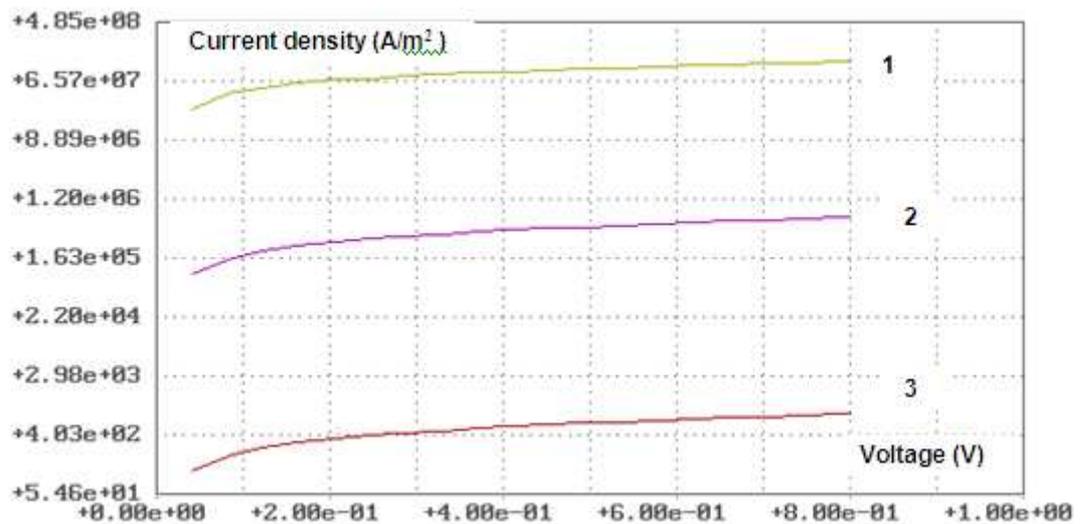


Рисунок 61. Удельные ВАХ «тонких» КП при различных размерах L : 1 – $L = L_{dq}/4$; 2 – $L = L_{dq}/4,5$; 3 – $L = L_{dq}/5$.

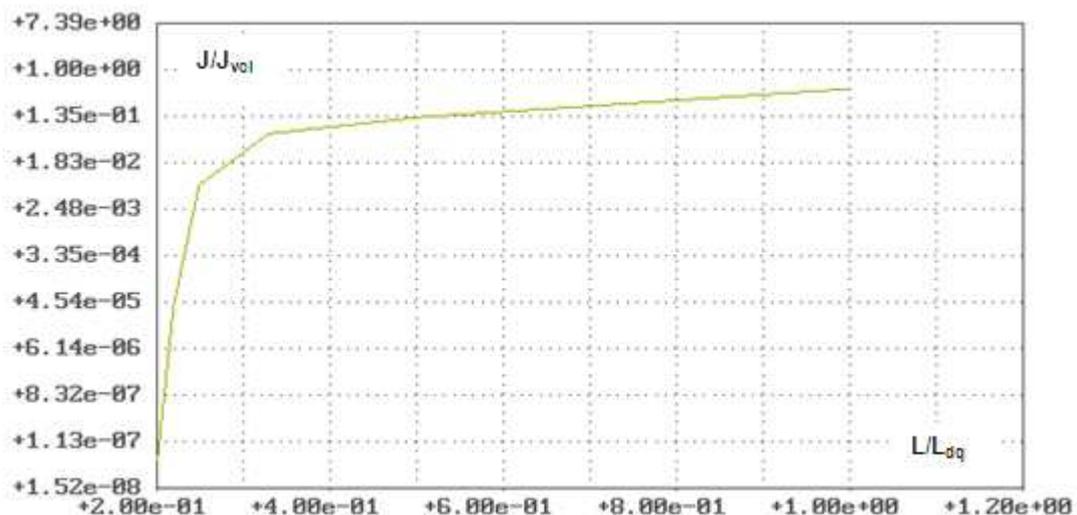


Рисунок 62. Зависимость плотности тока J в КП, выраженной в единицах плотности тока прибора с ПК из объемного материала J_{vol} , от поперечного размера ПК L в единицах длины размерного квантования L_{dq} .

Выше приведены ВАХ КП с различными поперечными размерами (см. Рисунки 54 – 58 и 61). Плотности токов и токи выражены в физических единицах

измерения, что позволяет составить представление о диапазоне возможных изменений характеристик прибора. Изменяя поперечные размеры провода или концентрацию электронов проводимости в контактах, то есть материал, из которого изготовлены контактные области, можно варьировать плотность тока в КП от практически нулевых значений до величин порядка 10^{10} А/м^2 (10^6 А/см^2) при приложенных напряжениях V_{ce} , меньших 1 В. Соответствующим образом от нуля до $10^{10} \text{ }\Omega^{-1}\text{м}^{-2}$ меняется и проводимость прибора. Заметим, что абсолютные значения плотности тока могут существенно зависеть от исходного материала, из которого изготовлен ПК. Поэтому представляют интерес кривые на Рисунке 63, отображающие удельные ВАХ КП в единицах J_{vol} , где J_{vol} бралась при $V_{ce} = 0,8 \text{ В}$.

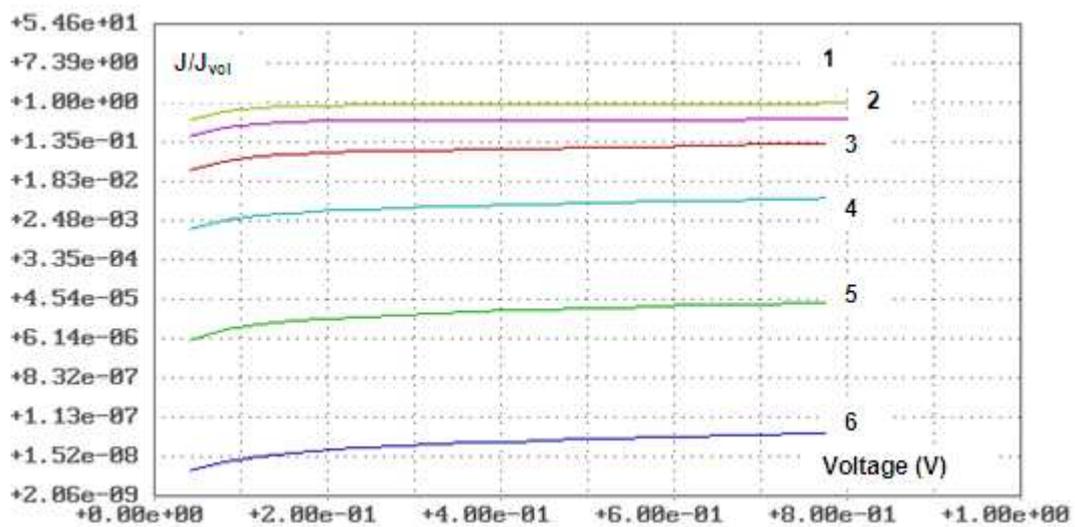


Рисунок 63. Удельные ВАХ КП в единицах плотности тока J_{vol} при различных поперечных размерах L : 1 –ПК из объемного материала; 2 – $L = L_{dq}$; 3 – $L = L_{dq}/2$; 4 – $L = L_{dq}/4$; 5 – $L = L_{dq}/4,5$; 6 – $L = L_{dq}/5$.

В отличие от проводимости обыкновенного резистора, которая линейно уменьшается с увеличением его длины, проводимость КП как функция длины ПК демонстрирует нелинейное поведение. На Рисунке 64 приведено семейство ВАХ КП с $L = L_{dq}/3$ при различных значениях длины ПК, а на Рисунке 65 – удельная проводимость этого КП при $V_{ce} = 0,8 \text{ В}$ как функция длины канала. Такое поведение проводимости для коротких проводов связано с существенной зависимостью потенциального рельефа в направлении переноса от длины прибора. При увеличении длины канала указанная зависимость ослабевает.

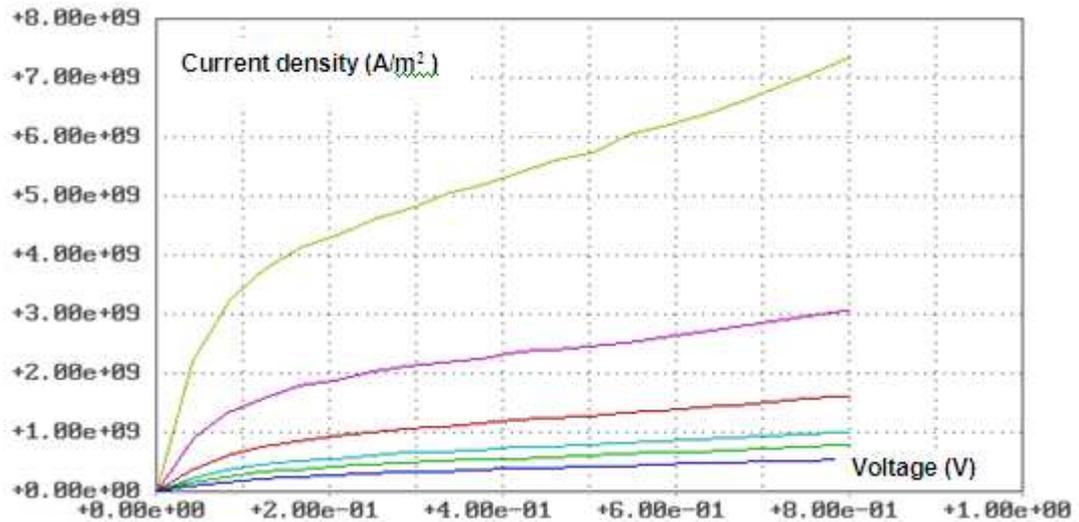


Рисунок 64. Удельные ВАХ КП с $L = L_{dq}/3$ (8,4 нм) при различных значениях длины канала (сверху вниз): 55 нм, 110 нм, 196 нм, 220 нм, 250 нм, 300 нм.

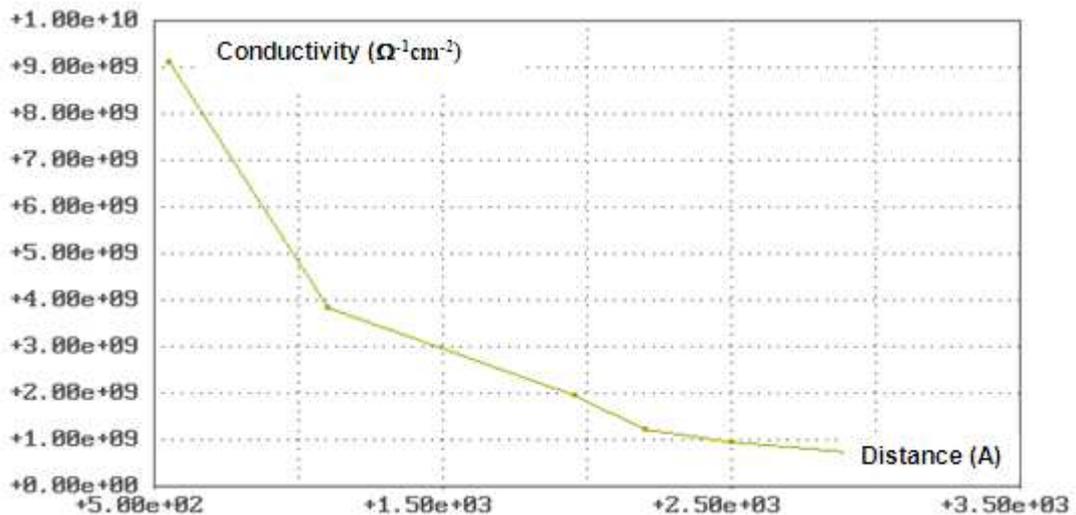


Рисунок 65. Удельная проводимость КП с $L = L_{dq}/3$ (8,4 нм) при $V_{cc} = 0,8$ В как функция длины ПК.

6.3.2. Механизмы проводимости в КП

Ток в квантовом приборе можно условно представить состоящим из компонент, связанных с тем или иным механизмом проводимости. Ясно, что эти условные компоненты (каждая отдельно) не наблюдаемы экспериментально. Однако моделирование позволяет оценить влияние различных механизмов на выходные характеристики прибора.

В соответствии с соотношением (334.1), суммарную плотность тока в КП (как и в любом другом квантовом приборе) в линейном по τ приближении можно представить в виде суммы квантовой и дрейфово-диффузионной составляющих. Эти термины нельзя считать точными, поскольку, например, дрейфово-диффузионная составляющая плотности тока существенно зависит от концентрации электронов проводимости, которая в КП в большой степени определяется поперечными размерами ПК, то есть эффектом размерного квантования. Вычислить квантовую составляющую в рамках используемой модели можно, положив $\tau = 0$. На Рисунке 66 изображена ВАХ КП с $L = L_{dq}/4$ при этом условии.

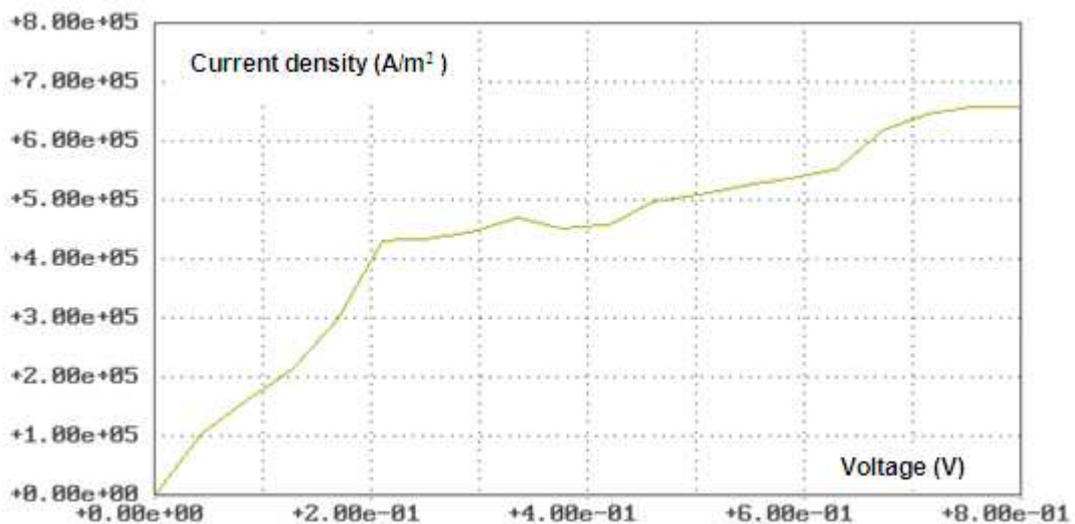


Рисунок 66. Удельная ВАХ КП с $L = L_{dq}/4$ при $\tau = 0$ и длиной ПК 220 нм.

Сравнение кривой на Рисунке 66 с характеристиками, рассчитанными при ненулевом τ (см. Рисунки 55, 57 и 61), показывает, что квантовая составляющая плотности тока в КП на три порядка ниже дрейфово-диффузионной. Естественно ожидать, что для коротких ПК вклад квантовой составляющей в плотность тока будет большим. Как видно на Рисунке 67, при уменьшении длины ПК в четыре раза плотность квантовой составляющей тока возрастает на два порядка. Однако на столько же возрастает и плотность дрейфово-диффузионного тока. Общим следует считать вывод о том, что основным механизмом проводимости в ПК является дрейфово-диффузионный. Вклад квантового тока в суммарный колеблется от десятых долей процента до десяти процентов.

Кривые на Рисунках 66 и 67 немонотонны. В отличие от N-образности суммарных ВАХ КП данная особенность объясняется чисто квантовым эффектом надбарьерного отражения. На Рисунке 68 приведена зависимость коэффициента прохождения электронов через эффективный потенциал КП в направлении переноса для провода с длиной канала 220 нм и $V_{ce} = 0$. Отчетливо видны осцилляции, связанные с надбарьерным отражением.

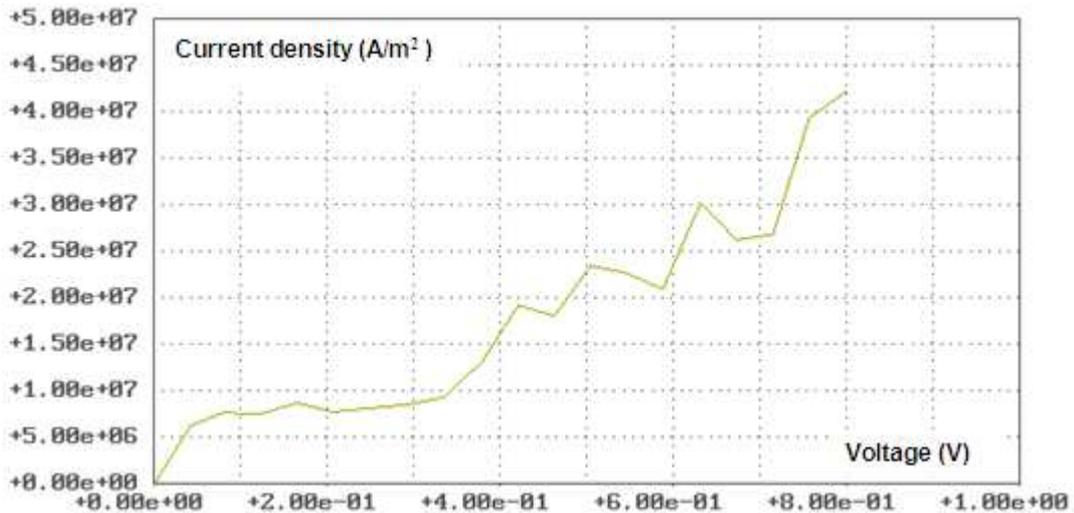


Рисунок 67. Удельная ВАХ КП с $L = L_{dq}/4$ при $\tau = 0$ и длиной ПК 55 нм.

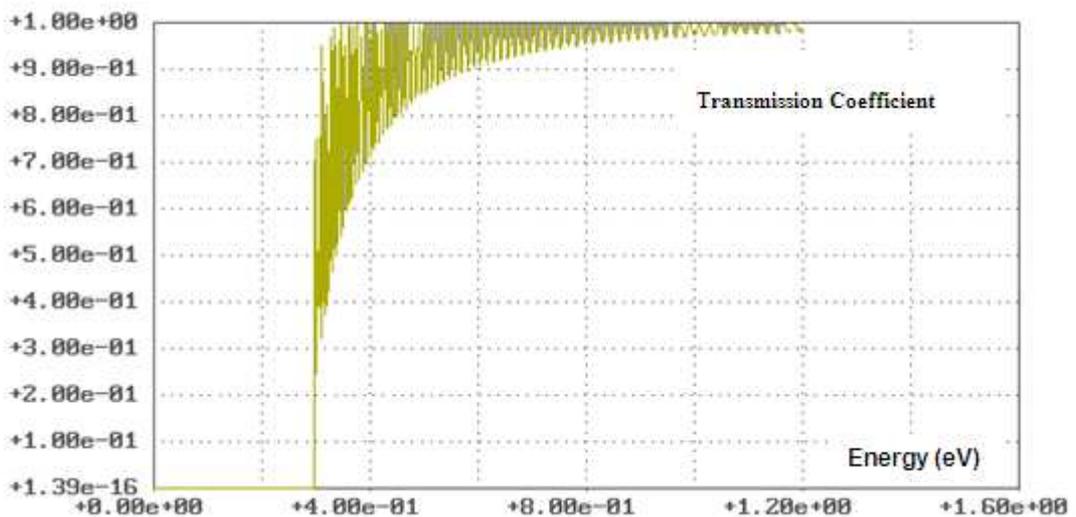


Рисунок 68. Коэффициент прохождения электронов через потенциальный рельеф КП в направлении переноса заряда ($L = L_{dq}/4$, $V_{ce} = 0$, длина ПК = 220 нм).

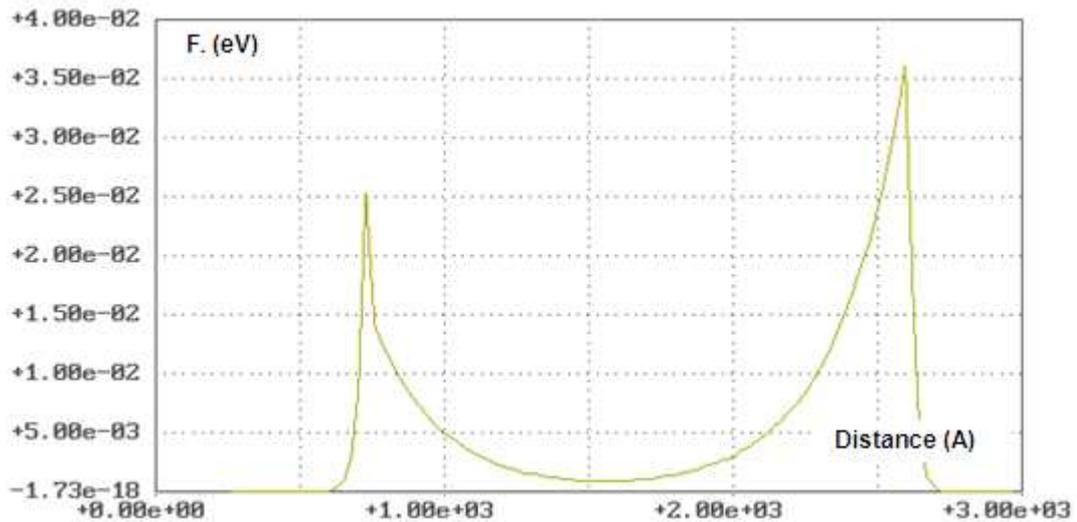


Рисунок 69. Разность химических потенциалов $F = F_1 - F_r$ в КП с $L = L_{dq}/4$ и длиной ПК 220 нм при $V_{ce} = 0,38$ В.

Согласно результатам, полученным при исследовании переноса заряда в РТС, преобладающая роль дрейфово-диффузионного механизма должна приводить к возможности нарушения локального химического равновесия между «левыми» и «правыми» электронными фазами. На Рисунке 69 демонстрируются результаты расчета разности химических потенциалов $F = F_1 - F_r$ в КП с $L = L_{dq}/4$ и длиной проводящего канала 220 нм при $V_{ce} = 0,38$ В. Как видно на рисунке F/e меньше приложенного напряжения примерно в десять раз. Это означает, что падение напряжения на КП при протекании через него тока определяется главным образом омическим вкладом, вклад неравновесных эффектов незначителен.

Величина F имеет порядок энергии тепловых флуктуаций $k_B T$ (расчеты проводились при комнатной температуре $T = 300^\circ\text{K}$) и, следовательно, неравновесность электронного газа существенно сказывается на концентрации носителей заряда в ПК и, как следствие, на плотности их тока.

Решение, представленное на Рисунке 69, подтверждает правильность выбора размеров контактных областей, как достаточных для релаксации электронных фаз к состоянию химического равновесия.

6.3.2.1. Влияние контактов на ВАХ КП

На Рисунке 57 представлены результаты расчетов плотности тока КП с $L = L_{dq}/4$ при различных значениях концентрации электронов проводимости в контактных областях. Видно, что уменьшение концентрации электронов в контактах на 25% приводит к уменьшению плотности тока прибора почти на порядок, а точнее - на 82%. Кроме того, как уже отмечалось, характерным для всех КП является слишком высокая по сравнению с обусловленной только собственной концентрацией электронов в ПК плотность тока. Эти же закономерности наблюдаются и в эксперименте [128,129,141]. Все это свидетельствует о серьезном влиянии свойств контактов на характеристики КП.

Поскольку при рассматриваемых нами изменениях концентраций в контактных областях изменения формы и высоты потенциального рельефа оказываются незначительными, то простейшее объяснение зависимости ВАХ от этой концентрации можно связать с изменением положения энергии Ферми. На Рисунке 70 видно, что при уменьшении концентрации электронов в контактах с $1,20 \cdot 10^{25} \text{ м}^{-3}$ до $9 \cdot 10^{24} \text{ м}^{-3}$ энергия Ферми понижается приблизительно на 0,05 эВ.

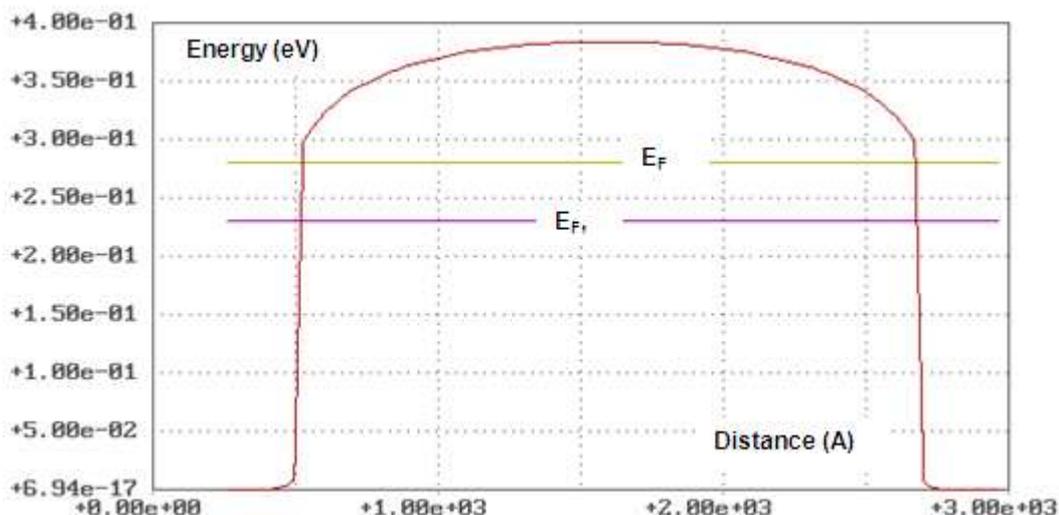


Рисунок 70. Эффективный потенциальный рельеф в направлении OX для КП с $L = L_{dq}/4$ при $V_{ce} = 0$ и положение энергии Ферми при различных концентрациях электронов проводимости в контактных областях: $1,20 \cdot 10^{25} \text{ м}^{-3}$ (верхняя прямая) и $9 \cdot 10^{24} \text{ м}^{-3}$ (нижняя прямая).

Если предположить, что избыточная проводимость КП объясняется инжекцией (испарительной эмиссией) в ПК дополнительных к собственным

электронов из контактных областей, и интенсивность этого процесса зависит от положения энергии Ферми, то становится понятной зависимость плотности тока в КП от концентрации электронов в контактах. Это предположение можно подтвердить расчетами.

При моделировании характеристик конкретных приборов мы разделяли электроны только на две фазы: «правые» и «левые». Однако для целей более детального анализа можно провести любое другое разделение, поскольку используемая модель позволяет рассматривать каждое электронное состояние как самостоятельную фазу со своим химическим потенциалом. Выбор электронных фаз для решения практических задач должен осуществляться, исходя из соображений совмещения адекватного отражения изучаемого эффекта и максимальной простоты описания.

Электронные состояния в контактных областях КП можно разделить на две качественно отличающиеся друг от друга группы [95,97,143-145]. В первую группу включим состояния с энергией E_x , большей, чем максимальное значение эффективного потенциала $U^{\text{eff}}(x)$. Будем называть их а-состояниями. Ко второй группе отнесем электронные состояния с энергией E_x , меньшей максимума $U^{\text{eff}}(x)$ и будем называть их б-состояниями. Поскольку КП – это прибор с достаточно длинным ПК, то практически все б-электроны испытывают почти полное отражение на границе между контактом и ПК. В противоположность этому, практически для всех а-электронов коэффициент прохождения через потенциальный рельеф существенно отличен от нуля. Это видно, в частности, при сопоставлении Рисунков 68 и 70. Для всех б-электронов плотность потока вероятности в направлении Ox j_x пренебрежимо мала, а для всех а-электронов – это существенная величина.

Из приведенных соображений следует, что б-электроны непосредственно не участвуют в формировании макроскопического потока в Ox направлении, который обусловлен только носителями с отличным от нуля микроскопическим потоком j_x , то есть а-электронами. Количество а-электронов в контактных областях и в ПК различается незначительно, что определяется выбранной

процедурой построения потенциала $U^{\text{eff}}(x)$. Количество a и b – электронов зависит от положения энергии Ферми, поскольку ею регулируется заполнение допустимых электронных состояний.

Испытывая стохастическое рассеяние, a - и b - электроны могут поглощать или испускать энергию и переходить из одной фазы в другую. В результате, при протекании тока через прибор может возникнуть неравновесная ситуация, соответствующая обогащению или обеднению a – фазы электронами. Явление обогащения a -фазы за счет переходов электронов из b -фазы будем называть инжекцией (или испарительной эмиссией) электронов в ПК из контактной области. Ниже будет построена модель, позволяющая изучить этот эффект.

При исследовании инжекции электронов из контактной области в ПК с целью упрощения мы будем пренебрегать различием между химическими потенциалами «левых» и «правых» электронов, то есть будем полагать, что $F_l = F_r$. Это позволит более ясно выделить рассматриваемый эффект без усложнения его второстепенными деталями. Разделим допустимые электронные состояния на две фазы a - и b – электронов с химическими потенциалами F_a и F_b соответственно. Для каждой из них определим плотности числа частиц и потока

$$n_{a,b} = \sum_{E_{a,b}} \rho_{a,b} f_{a,b}, \quad \mathbf{n}_{a,b} = \sum_{E_{a,b}} \mathbf{j}_{a,b} f_{a,b},$$

$$f_{a,b} = (1 + \exp\{(E_{a,b} - F_{a,b} + \tau(\mathbf{v}_{a,b}, \nabla F_{a,b})) / k_B T\})^{-1}.$$

Векторное выражение для плотности потока записано не случайно. Поскольку в силу отмеченных выше особенностей потоки \mathbf{j}_{bx} и, следовательно, \mathbf{n}_{bx} равны нулю, то стационарная постановка задачи переноса с учетом обмена электронов между выбранными фазами возможна только в двухмерной или трехмерной форме. Действительно, стационарные уравнения переноса имеют вид

$$\begin{cases} (\nabla, \mathbf{n}_a) = -(\gamma/\hbar)(F_a - F_b)(n_a + n_b), \\ (\nabla, \mathbf{n}_b) = -(\gamma/\hbar)(F_b - F_a)(n_a + n_b). \end{cases} \quad (392)$$

Если ограничиться одномерным рассмотрением переноса вдоль оси OX , то левая часть второго уравнения в (392) равна нулю, что автоматически означает равенство нулю и правой части, то есть химическое равновесие

$$F_a = F_b = F_{at}, \quad (393)$$

при котором никакого обмена частицами между фазами не происходит. Поэтому в области контакта мы будем решать приближенную двумерную задачу, полагая, что

$$\mathbf{j}_a = (j_{ax}, j_{ay}, 0), \mathbf{j}_b = (0, j_b, 0), \mathbf{n}_a = (I_{ax}, I_{ay}, 0), \mathbf{n}_b = (0, I_b, 0). \quad (394)$$

В результате из (392) получим

$$\begin{cases} \partial I_{ax}/\partial x + \partial I_{ay}/\partial y = -(\gamma/\hbar)(F_a - F_b)(n_a + n_b), \\ \partial I_b/\partial y = -(\gamma/\hbar)(F_b - F_a)(n_a + n_b). \end{cases} \quad (395)$$

Поскольку мы пренебрегаем различием между «левыми» и «правыми» электронами, то формально можно записать

$$I_{ax} = -\sigma_{ax} \partial F_a / \partial x, I_{ay} = -\sigma_{ay} \partial F_a / \partial y, I_b = -\sigma_b \partial F_b / \partial y, \quad (396)$$

где

$$\sigma_{ax} = (\partial F_a / \partial x)^{-1} \sum_{E_a} j_{ax} f_a, \sigma_{ay} = (\partial F_a / \partial y)^{-1} \sum_{E_a} j_{ay} f_a, \sigma_b = (\partial F_b / \partial y)^{-1} \sum_{E_b} j_b f_b.$$

Правая часть уравнений (392) и (395) описывает переходы между электронными фазами. Фигурирующий в них параметр γ определяет интенсивность этих переходов и связан с вероятностью поглощения-испускания энергии при стохастическом рассеянии. В расчетах мы полагали, что

$$\gamma = \begin{cases} \exp\{-\Delta/k_B T\}, \Delta \geq 0, \\ 1, \Delta \leq 0, \end{cases}$$

где Δ - энергетическая щель между энергией Ферми и максимальным значением потенциала $U^{\text{eff}}(x)$.

Система уравнений (395), (396), дополненная двумерным уравнением Пуассона, позволяет учесть влияние инжекции электронов из контактных областей на перенос заряда в КП. Однако для их решения необходимо сформулировать граничные условия, что является не менее важной задачей, чем написание уравнений. Область моделирования изображена на Рисунке 71. Для упрощения границы $\Gamma_{e,c}$, соответствующие омическим контактам, изображены занимающими всю границу контактной области с «внешним окружением», то есть с металлизацией, в направлении OY . Проникновение металлизации в контактную область может быть менее значительным, но это принципиально не

каждое, из которых вместе с граничными условиями (397) определяет некоторую F_{at} . Обозначим решения краевых задач для уравнений (398) и (399) через F_{ata} и F_{atb} соответственно. Эти две функции в общем случае различны и совпадают только при нулевом приложенном напряжении ($V_{ce} = 0$). Заметим, что функция F_{ata} является решением первого из уравнений (395) с нулевой правой частью и отражает поведение потенциала F_a в пределе $\gamma = 0$, то есть при отсутствии переходов между фазами и чисто омическом характере переноса. Именно этой функцией определялась бы концентрация электронов в отсутствии инжекции из контактов, а ее градиентом – плотность тока.

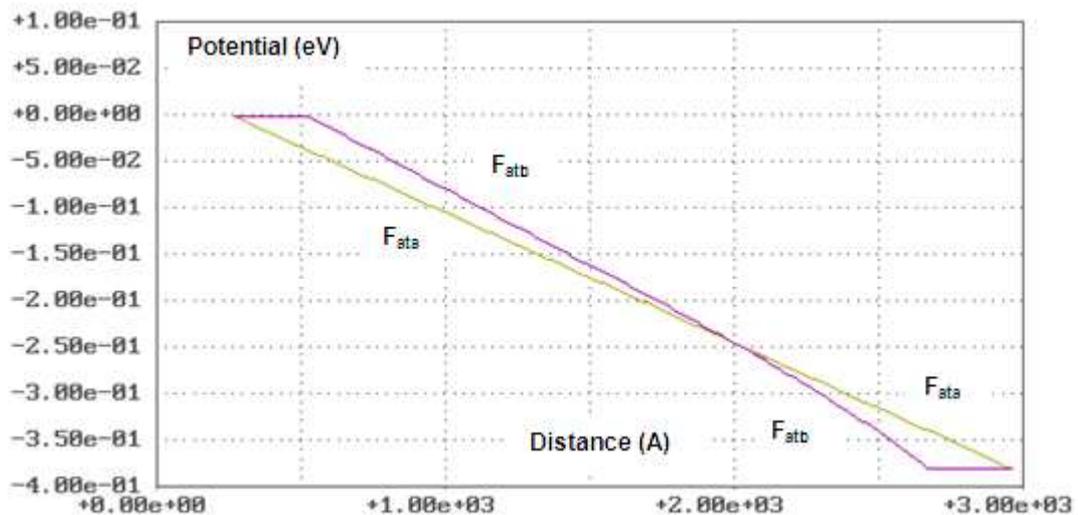


Рисунок 72. Функции $F_{ata}(x, y_0)$ и $F_{atb}(x, y_0)$ в КП с $L = L_{dq}/4$ при $V_{ce} = 0,38$.

На Рисунке 72 показано поведение F_{ata} и F_{atb} для КП с $L = L_{dq}/4$ и $V_{ce} = 0,38$ В. Хорошо видно, что в области эмиттерного контакта $F_{ata} < F_{atb}$, а в области коллекторного контакта $F_{ata} > F_{atb}$. Эти неравенства определяют знак невязки между точным решением и функциями $F_{ata,atb}$. Поскольку граничные условия для системы (395) соответствуют устойчивому аттрактору, то точное решение для F_a также будет существенно отличаться от точного решения F_b только в окрестностях источников неравновесности с характерными размерами порядка длины релаксации L_{rel} . Это означает, что вблизи эмиттерного контакта $F_a > F_{ata}$, что соответствует обогащению а-фазы носителями заряда из b-фазы, а вблизи коллекторного контакта $F_a < F_{ata}$, что соответствует обеднению а-фазы за счет переходов электронов в b-фазу. Приведенные простые соображения качественно

объясняют суть явления инжекции электронов из контакта в проводящий канал. Причем, если на одном контакте квантового провода наблюдается собственно инжекция или испарительная эмиссия, то на другом – противоположный процесс, который можно назвать конденсацией.

Представим F_a в виде

$$F_a = F_b + F_{a1},$$

тогда, согласно (395), (396) имеем уравнения

$$\begin{cases} \partial(\sigma_{ax}\partial F_{a1}/\partial x)/\partial x + \partial(\sigma_{ay}\partial F_{a1}/\partial y)/\partial y = (\gamma/\hbar)F_{a1}(n_a + n_b) + \chi, \\ \partial(\sigma_b\partial F_b/\partial y)/\partial y = -(\gamma/\hbar)F_{a1}(n_a + n_b), \end{cases} \quad (400)$$

где функцией

$$\chi = -\partial(\sigma_{ax}\partial F_b/\partial x)/\partial x - \partial(\sigma_{ay}\partial F_b/\partial y)/\partial y$$

определяются источники неравновесности. Решение первого из уравнений (400) с нулевыми граничными условиями на Γ и $\Gamma_{e,c}$ и нулевой нормальной производной на Γ_0 описывает отклонения F_a от F_b . Это уравнение релаксационного типа с источником. Сам источник χ локализован вблизи условных границ между ПК и контактами (см. Рисунок 73).

Согласно (400), решение для F_{a1} существенно отличается от нуля в окрестностях областей локализации источника χ с размерами порядка длины релаксации

$$L_r = L_0(n_a/n_b)^{1/2},$$

которая очень мала в контактных областях (порядка $10^{-3}L_0$ для провода с $L = L_{dq}/4$) и одного порядка с L_0 в проводящем канале. Эти утверждения иллюстрируются решением F_{a1} , показанным на Рисунке 74. Видно, что в большей части проводящего канала и практически всюду в контактных областях выполняется соотношение

$$F_a \approx F_b.$$

Именно этим объясняется хорошее соответствие результатов вычислений, приведенных выше, с экспериментальными результатами. Фактически, ранее мы неявно учитывали инжекцию электронов из контактов, полагая справедливым

условие (393) и решая одномерную задачу в направлении Ox , не разделяя электроны на а - и б - фазы, сдвигая тем самым решение в сторону аттрактора F_{atb} . При этом мы существенно ошибались количественно только в очень узких областях вблизи условных границ проводящего канала и контактов.

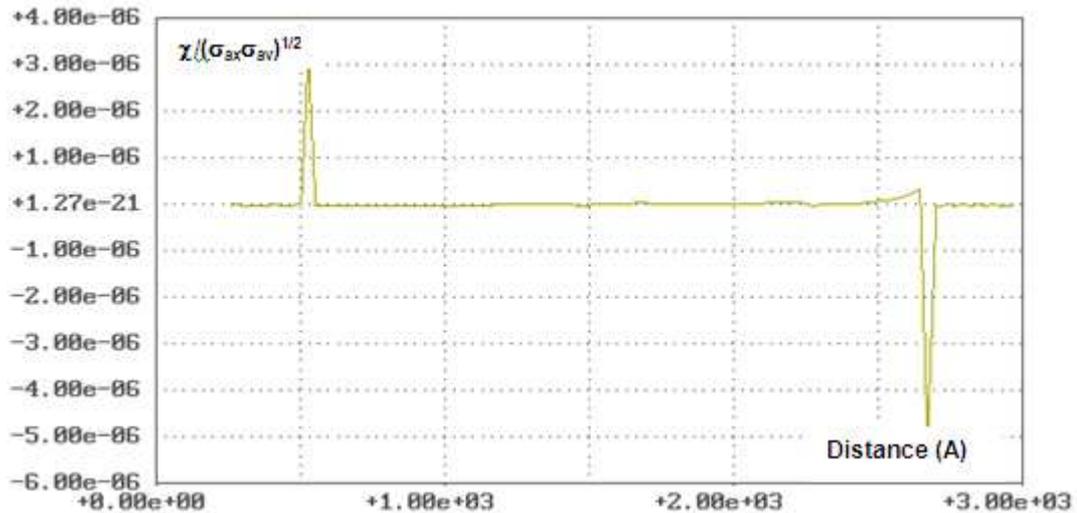


Рисунок 73. Поведение функции $\chi(x, y_0)$ в КП КП с $L = L_{dq}/4$ при $V_{ce} = 0,38$ В.

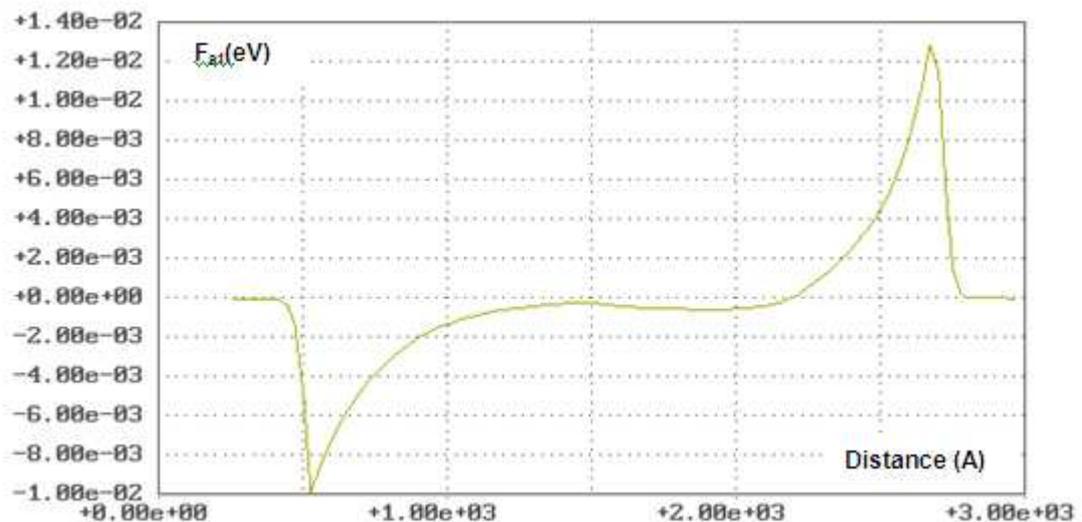


Рисунок 74. Функция $F_{a1}(x, y_0) = F_a(x, y_0) - F_b(x, y_0)$ в КП с $L = L_{dq}/4$ при $V_{ce} = 0,38$ В.

Характерная функция

$$F_{a2} = F_a - F_{ata}$$

описывает отклонение точного решения F_a от аттрактора F_{ata} . Как показано на Рисунке 75, это отклонение оказывается существенным, превышающим $k_B T$ почти всюду в квантовом проводе. В результате, как показано на Рисунке 76, большая

часть проводящего канала обогащена инжектированными из эмиттера дополнительными электронами, перешедшими из b-фазы в a-фазу. Только вблизи коллекторного контакта наблюдается значительное обеднение.

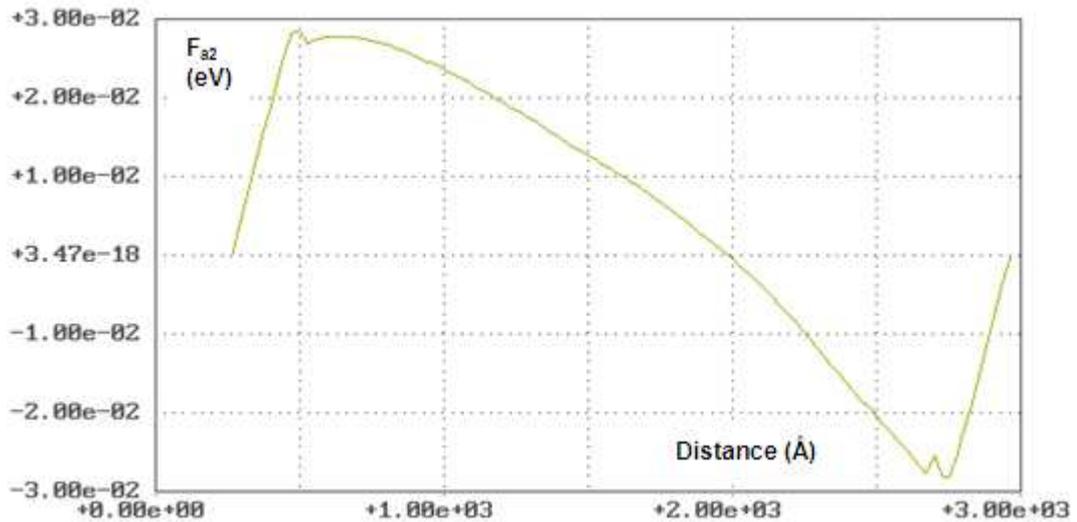


Рисунок 75. Разность $F_{a2} = F_a(x, y_0) - F_{ata}(x, y_0)$ в КП с $L = L_{dq}/4$ при $V_{ce} = 0,38$ В.

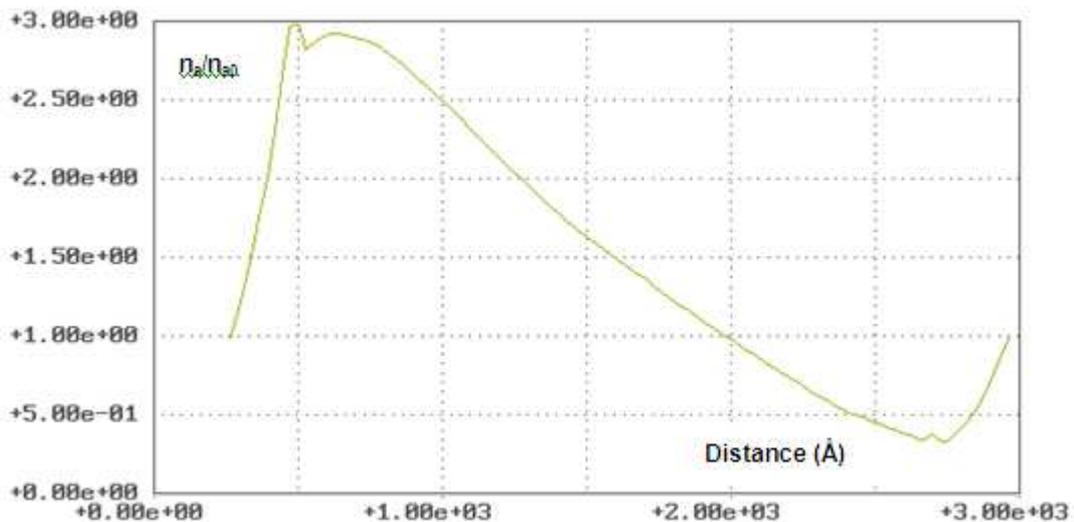


Рисунок 76. Распределение a-электронов в КП при $V_{ce} = 0,38$ В с учетом эффекта инжекции из контактов ($n_{a0} = n_a(F_a = F_{ata})$).

На Рисунке 77 представлены распределения x и y составляющих вектора плотности тока вдоль сечения $x = x_{e0}$ на Рисунке 71. Видно, что компонента J_x отлична от нуля только на части сечения, соответствующей условной границе ПК с эмиттерным контактом. Максимум J_x достигается в середине этой границы. Напротив, компонента J_y равняется нулю в середине границы ПК и контакта. Она резко убывает по абсолютной величине на этой границе и остается практически

постоянной вне ее. Рисунок 77 иллюстрирует процесс инжекции потока электронов в проводящий канал.

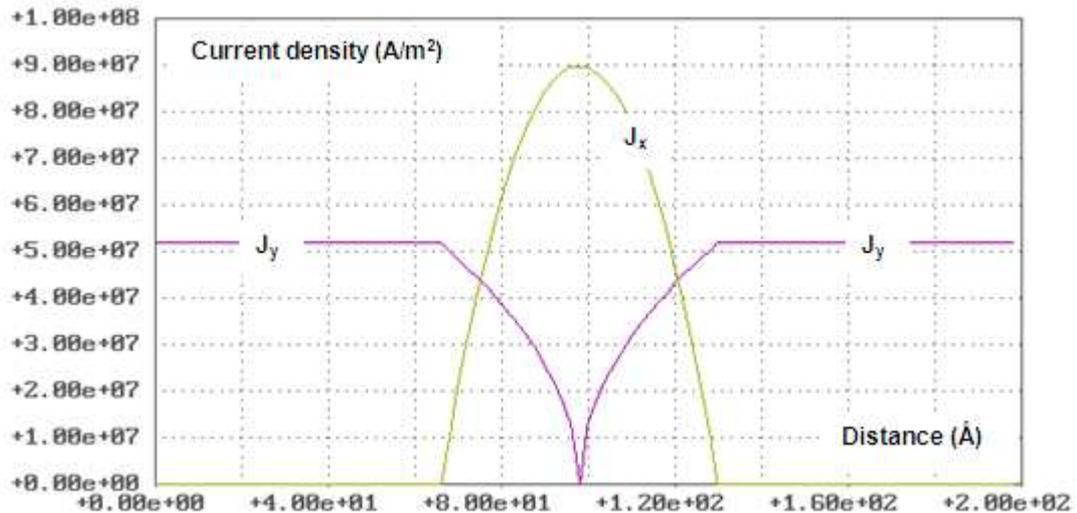


Рисунок 77. Распределение компонент плотности тока J_x и $|J_y|$ вдоль условной границы между проводящим каналом и эмиттерным контактом квантового провода.

Таким образом, мы показали, что инжекция (испарительная эмиссия) электронов из контактных областей в проводящий канал является важнейшим фактором, влияющим на проводимость квантового провода. Обычно используемая нами одномерная двухфазная модель, разделяющая электроны на «левые» и «правые», учитывает эффект инжекции в нулевом приближении сингулярной теории возмущений по параметру

$$\alpha = L_r/L_{ch},$$

где $L_{ch} = x_{c0} - x_{e0}$ – длина проводящего канала. То есть для квантовых проводов с длинным проводящим каналом решать дополнительно задачу о кинетике переходов между а - и б-фазами электронов нет необходимости. Однако для квантового провода с коротким каналом стандартного двухфазного описания может оказаться недостаточно.

6.3.2.2. Основные факторы, определяющие ВАХ КП

Полученные в этом параграфе результаты свидетельствуют о том, что квантовый провод достаточно сложный прибор, электрические характеристики

которого существенно зависят как от параметров проводящего канала, так и от свойств контактов.

Мы видели, что при одинаковых контактах плотность тока в КП с различными поперечными размерами, меньшими длины размерного квантования L_{dq} , изменяется на порядки при изменении поперечных размеров (см. Рисунки 62, 63). Для ПК из объемного материала такая зависимость не характерна, плотность тока для него практически не зависит от площади канала. Отмеченное свойство квантового провода с квантовым проводящим каналом является прямым следствием эффекта размерного квантования и обусловленной этим эффектом сильной зависимости концентрации носителей заряда в проводящем канале от его поперечных размеров.

Другим интересным свойством КП является необычное поведение удельной проводимости прибора при изменении длины ПК (см. Рисунок 65). Для КП с коротким каналом удельная проводимость сильно зависит от его длины. Это объясняется существенными изменениями потенциального рельефа в направлении переноса при изменении длины ПК. При увеличении длины ПК указанная зависимость ослабевает, и для каналов с длиной, превышающей L_{dq} на порядок и более, удельная проводимость КП практически не зависит от длины ПК.

Сильная зависимость характеристик КП от свойств контактов связана не только с эффектом инжекции, подробно описанным в предыдущем разделе. Неравновесность между фазами «левых» и «правых» электронов (см. Рисунок 69) оказывает не меньшее влияние на концентрацию носителей заряда в ПК и, как следствие, на проводимость прибора. На Рисунке 78 показано распределение концентрации электронов n в КП при $V_{ce} = 0,38$ В с учетом неравновесности между «левыми» и «правыми» электронами, но без учета эффекта инжекции из контактов. Видно, что их количество того же порядка, что и инжектированных из контактов электронов.

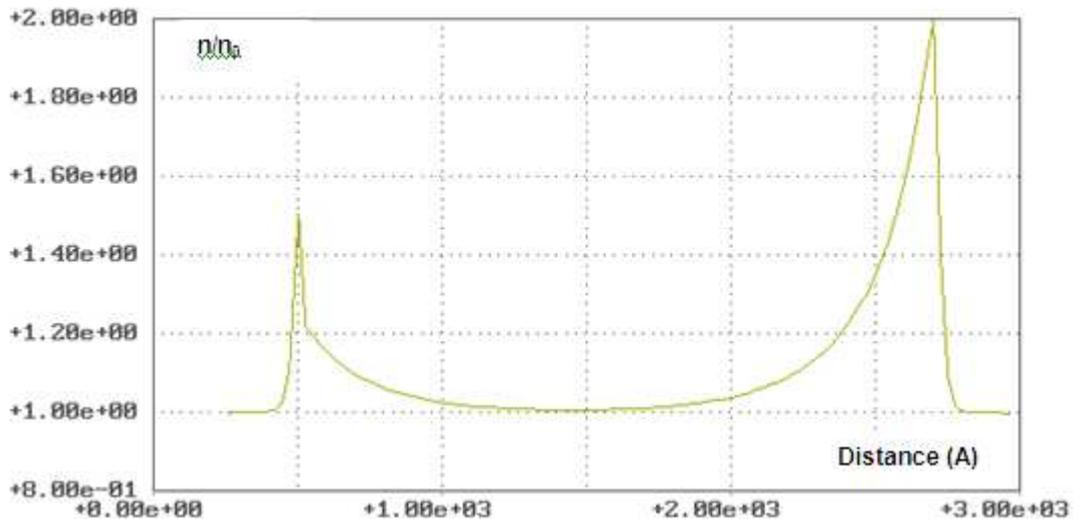


Рисунок 78. Распределение электронов в КП при $V_{ce} = 0,38$ В с учетом неравновесности между «левыми» и «правыми» электронами, но без учета эффекта инжекции из контактов ($n_0 = n(F = 0)$).

На Рисунке 79 показан суммарный вклад эффектов инжекции электронов из контактов и неравновесности между «левой» и «правой» фазами в концентрацию носителей заряда n в КП. Видно, что неравновесность электронного газа приводит к частичной компенсации обеднения коллекторной области носителями заряда, наблюдавшегося при учете только эффекта инжекции.

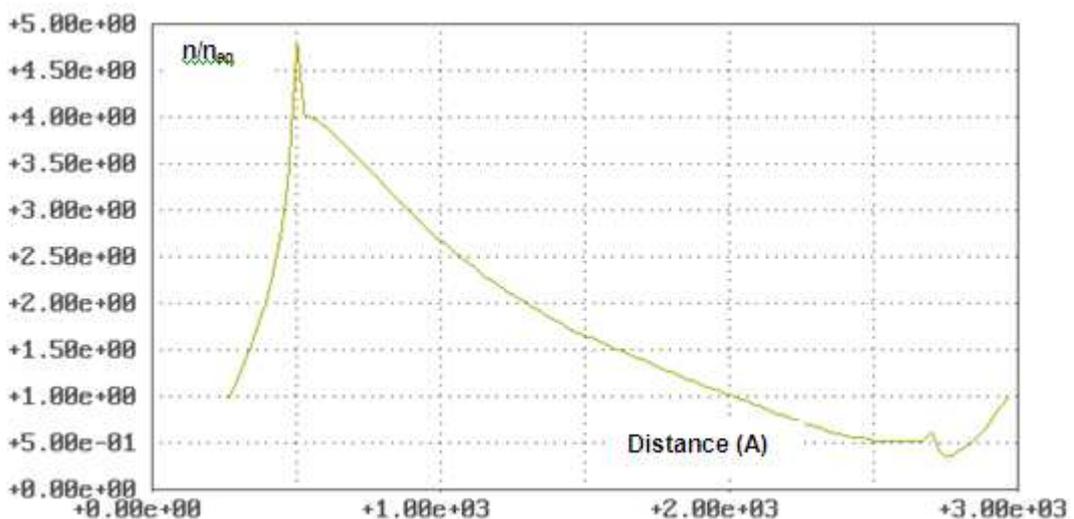


Рисунок 79. Суммарное распределение электронов в КП при $V_{ce} = 0,38$ В ($n_{eq} = n(V_{ce} = 0)$).

Исследование соответствующих уравнений показывает, что в КП неравновесность между «левой» и «правой» электронными фазами возникает только при протекании тока через прибор и обусловлена градиентами концентрации носителей заряда на границах проводящего канала и контактов. На Рисунках 69 и 78 видно, что неравновесные электроны сосредоточены в основном

в приконтактных областях. Таким образом, эффект неравновесности также как и эффект инжекции связан исключительно с наличием в квантовом проводе контактных областей, существенно отличных по своим характеристикам от проводящего канала.

Подводя итоги сказанному, можно утверждать, что основными факторами, определяющими электрические характеристики квантового провода, являются:

- поперечные размеры проводящего канала (в единицах L_{dq});
- длина проводящего канала (в единицах L_{dq});
- концентрация носителей заряда в контактных областях;
- градиенты концентрации электронов на границах контактов с проводящим каналом.

6.4. Температурные эффекты в квантовых проводах

Рассмотренные особенности переноса заряда в квантовых проводах могут приводить к интересным физическим следствиям. Поскольку высота потенциального барьера, создаваемого проводящим каналом для электронов в квантовом проводе, является функцией поперечных размеров этого канала, то от этих размеров могут зависеть электрофизические свойства провода. Кроме того, инжекция электронов из контакта в проводящий канал эквивалентна по своему физическому смыслу испарительной эмиссии. Это означает, что контактная область, из которой происходит испарение электронов, может охлаждаться в результате переноса заряда.

6.4.1. Температурная зависимость сопротивления металлических квантовых проводов

В статье [146] измерялась температурная зависимость сопротивления квантовых проводов на основе Zn наноконкомпозитов. Оказалось, что она качественно различна для «толстых» и «тонких» проводов. Так при диаметре

провода в 15 нм наблюдался типичный для металлов практически линейный рост сопротивления с ростом температуры. Однако для проводов с диаметром 9 нм и 4 нм наблюдался уже не рост, а уменьшение сопротивления при возрастании температуры. Эта аномальная температурная зависимость сопротивления более выражена для провода с диаметром 4 нм.

Данная в статье [146] теоретическая интерпретация полученных результатов не кажется убедительной. Она основывается на формуле, связывающей сопротивление квантового провода с длиной фазовой когерентности (phase - breaking length) электронов проводимости L_ϕ (см. работу [147]). При этом никак не объясняются причины качественного изменения зависимости L_ϕ от температуры при уменьшении диаметра провода, а только производится обратный пересчет L_ϕ , исходя из упомянутой формулы, на основе данных по измерению сопротивления.

Для объяснения полученных в [146] результатов исследуем в рамках предлагаемого подхода влияние поперечных размеров металлических квантовых проводов на температурную зависимость их удельного сопротивления.

Как и прежде считаем квантовый провод состоящим из эмиттерного и коллекторного контактов, соединенных прямоугольным проводящим каналом. Очевидно, что форма поперечного сечения проводящего канала может повлиять на количественные характеристики прибора, но не скажется на его качественных особенностях.

На Рисунке 80 показаны первые восемь уровней энергии электронов, возникающих в ПК за счет размерного квантования, для трех значений поперечного размера канала $L = 15$ нм, 9 нм, 4 нм. Хорошо видно, что плотность состояний электронов быстро уменьшается с уменьшением L .

На Рисунке 81 показана разница в энергиях электронов второго и первого уровней размерного квантования $\Delta E = E_2 - E_1$ для тех же значений поперечного размера ПК. При $L = 15$ нм ΔE порядка $5 \cdot 10^{-3}$ эВ плотность электронных состояний в ПК достаточно высока. При $L = 9$ нм величина $\Delta E = 1.4 \cdot 10^{-2}$ эВ, а энергетический спектр электронов разрежен в полтора раза по отношению к

проводу с $L = 15$ нм. Для провода с $L = 4$ нм значение $\Delta E = 7 \cdot 10^{-2}$ эВ, то есть энергетический спектр электронов разрежен почти в четыре раза по отношению к спектру в ПК с $L = 15$ нм.

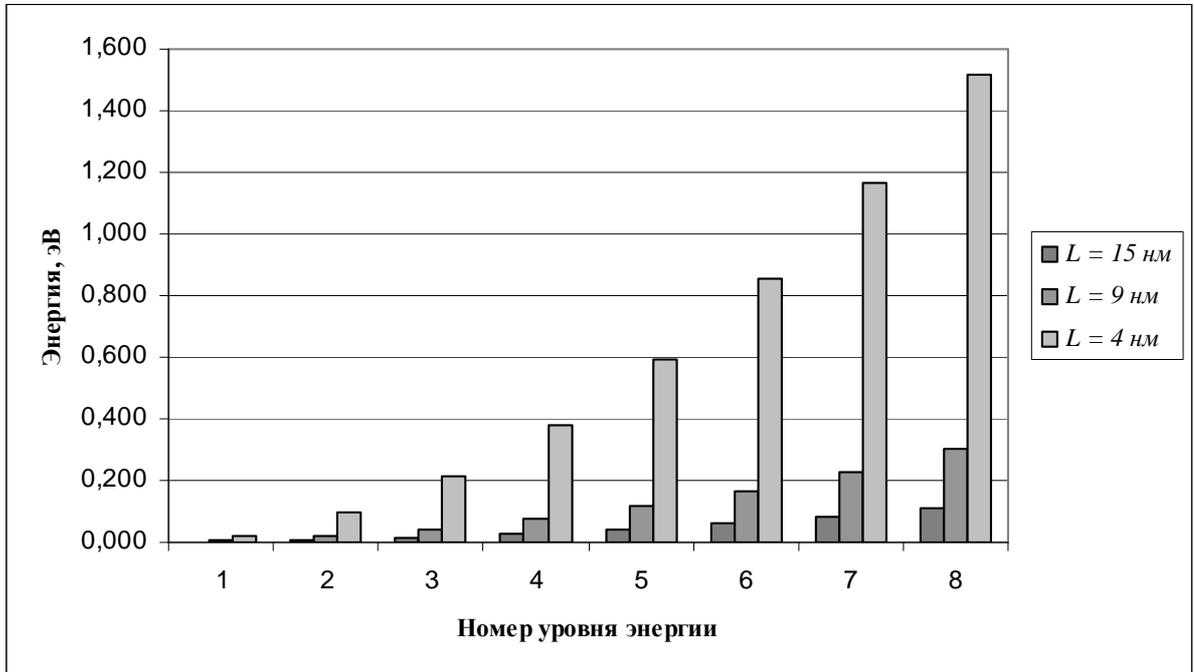


Рисунок 80. Первые восемь уровней размерного квантования энергии электронов в ПК для различных поперечных размеров проводящего канала $L = 15$ нм, 9 нм, и 4 нм.

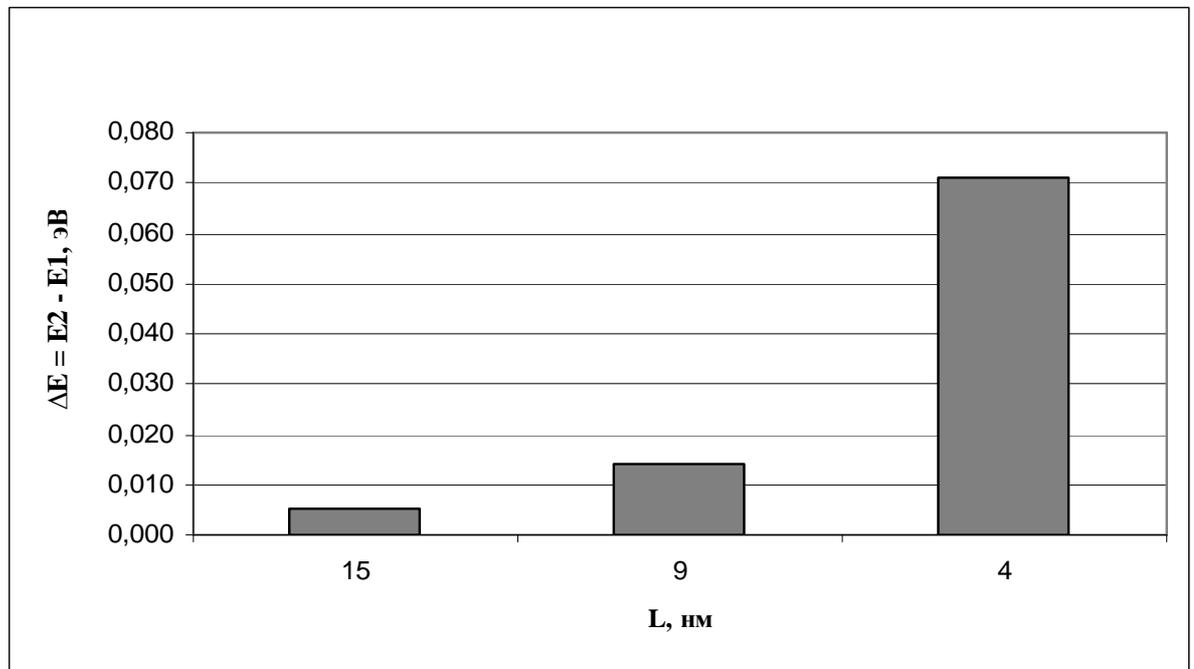


Рисунок 81. Разница между энергиями электронов на втором и первом уровнях размерного квантования в ПК с $L = 15$ нм, 9 нм и 4 нм.

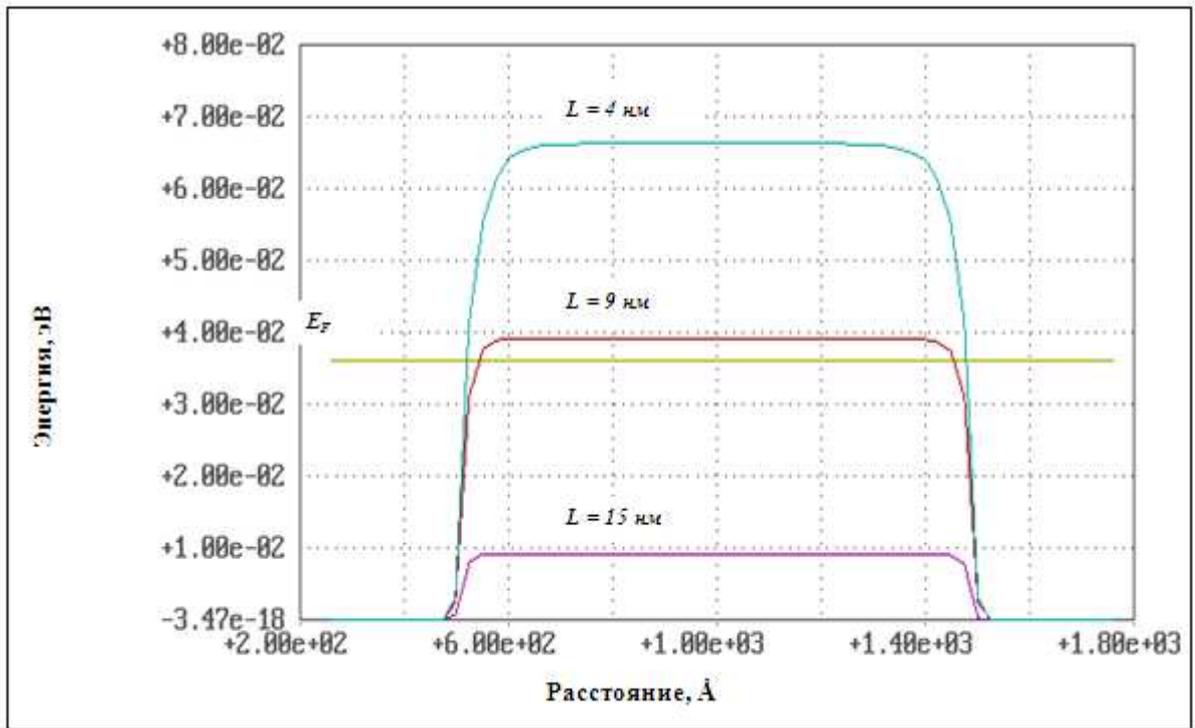


Рисунок 82. Энергия Ферми (E_F) и эффективный потенциальный рельеф для электронов проводимости в КП в направлении Ox для случаев $L = 15$ нм, 9 нм и 4 нм при нулевом смещении.

Как мы уже отмечали, количественные изменения в спектре размерного квантования энергии электронов, регулируемые размером поперечного сечения проводящего канала, приводят к качественным изменениям в эффективном потенциальном рельефе для электронов проводимости в КП. На Рисунке 82 видно, что для провода с $L = 15$ нм потенциальный барьер, создаваемый ПК между контактными областями, ниже энергии Ферми (примерно на $2.61 \cdot 10^{-2}$ эВ для рассчитываемой структуры). Назовем такие провода «толстыми». Уменьшение поперечного размера проводящего канала приводит к уменьшению плотности электронных состояний в нем. При достаточно малых L потенциальный барьер, создаваемый ПК, становится выше энергии Ферми. Такие квантовые провода назовем «тонкими». Как видно на Рисунке 82, КП с $L = 9$ нм и $L = 4$ нм являются «тонкими».

В соответствии с полученными выше результатами, для электронов в КП возможны три механизма переноса заряда:

- дрейфово-диффузионный перенос электронов с энергией, большей высоты потенциального барьера, создаваемого ПК;

- туннельный перенос через потенциальный барьер электронов с энергией, меньшей высоты этого барьера;
- инжекционный перенос электронов с энергией, меньшей высоты этого барьера.

Под инжекционным переносом понимается процесс, при котором в контакте, к которому приложено меньшее напряжение, электроны с энергией, меньшей высоты барьера, поглощают энергию тепловых флуктуаций, увеличивая свою энергию до значений, превышающих высоту барьера. Далее они перемещаются в ПК между контактами посредством дрейфа и диффузии, а затем излучают дополнительную энергию в контакте с большим приложенным напряжением. Заметим, что уменьшение энергии инжектированных электронов возможно и при переносе в ПК, однако оно ограничено снизу высотой потенциального барьера.

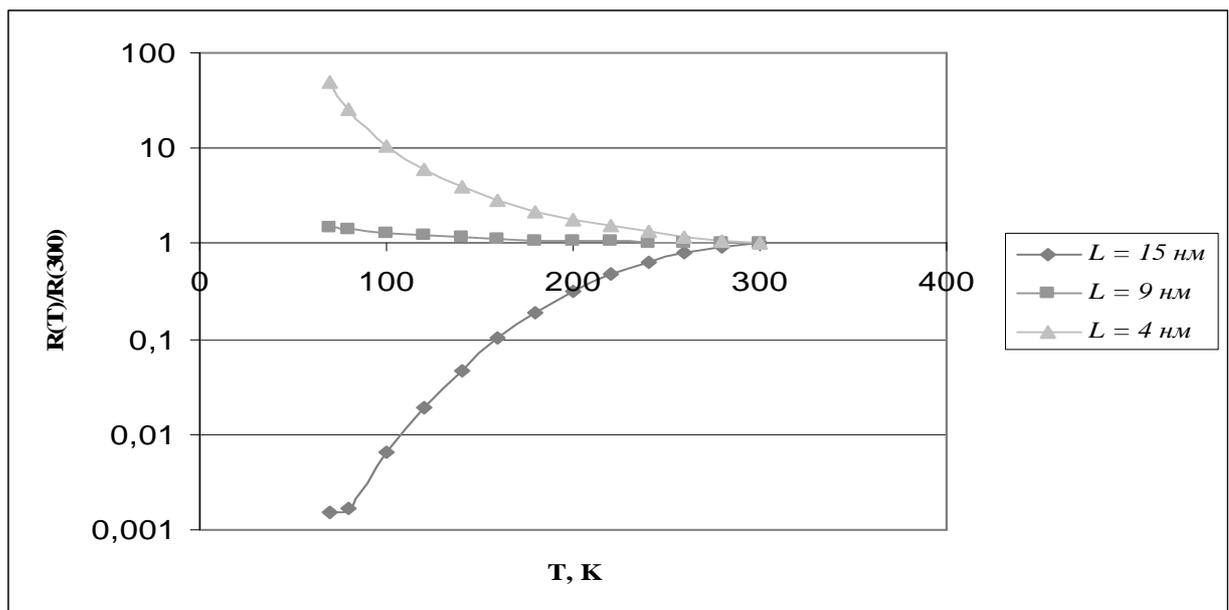


Рисунок 83. Зависимость сопротивления КП, нормированного на сопротивление при $T = 300$ К, от температуры для различных значений поперечного размера проводящего канала $L = 15$ нм, 9 нм, 4 нм.

Поскольку ширина потенциального барьера в КП, как правило, велика (100 нм для рассматриваемой структуры), туннелированием электронов через него можно пренебречь.

Если энергия Ферми электронов превышает высоту потенциального барьера более чем на среднюю энергию тепловых флуктуаций $k_B T$, то количество электронов, инжектируемых в ПК из контактов, много меньше их количества в

надбарьерной области энергий. В этом случае основной вклад в проводимость дают надбарьерные электроны, и тип проводимости КП не отличается от типа проводимости контактных областей. Поэтому зависимость сопротивления от температуры для КП с поперечным размером проводящего канала $L = 15$ нм имеет обычный, характерный для металлов вид (см. Рисунки 82 и 83). Она, главным образом, определяется температурной зависимостью времени релаксации импульса электронов.

Для «тонких» ПК, когда энергия Ферми меньше высоты эффективного потенциального барьера, основным становится инжекционный механизм проводимости. Количество собственных электронов в ПК в этом случае мало. Этим обусловлено высокое удельное сопротивление провода при низких температурах. На Рисунке 83 показана температурная зависимость сопротивлений квантовых проводов различного поперечного размера, нормированных на их значения при комнатной температуре ($T = 300^\circ\text{K}$).

Характер температурной зависимости сопротивления КП, показанной на Рисунке 83, полностью коррелирует с особенностями его потенциального рельефа, демонстрируемого на Рисунке 82. Если энергия Ферми превышает высоту потенциального барьера, создаваемого ПК, то сопротивление металлического КП при повышении температуры растет. Если энергия Ферми меньше высоты потенциального барьера, то сопротивление КП при увеличении температуры уменьшается. Эти качественные изменения в электрических характеристиках КП регулируются поперечным размером проводящего канала.

При реализации в КП инжекционного механизма проводимости количество электронов в ПК начинает существенно зависеть от температуры. Чем выше температура, тем большее количество электронов инжектируется в проводящий канал из контактов, и их концентрация в канале возрастает. Дифференциальное сопротивление ПК пропорционально

$$(\tau n)^{-1},$$

где τ – время релаксации импульса электронов, а n – их концентрация в проводящем канале. В расчетах, на основе которых получены кривые (см.

Рисунок 83), предполагалось, что время релаксации импульса электронов обратно пропорционально температуре. Для «тонких» КП уменьшение сопротивления с ростом температуры происходит за счет увеличения концентрации электронов в ПК. Этот эффект превышает обратное влияние, обусловленное температурной зависимостью времени релаксации импульса электронов.

Таким образом, в рамках многофазной модели результаты измерений температурной зависимости сопротивления металлических квантовых проводов, представленные в статье [146], можно легко объяснить особенностями переноса заряда в квантовых проводах различного поперечного размера.

6.4.2. Самоохлаждение эмиттерного контакта квантового провода

Уже несколько лет в периодической научной печати обсуждается возможность создания твердотельных охлаждающих устройств, управляемых приложенным напряжением (см. работы [100,101,139,140,148] и цитируемую в них литературу). В качестве основы функционирования такого холодильника, может быть использована описанная выше инжекция или, что то же самое, испарительная эмиссия электронов из эмиттерного контакта в проводящий канал квантового провода.

Физической основой инжекции электронов из контактных областей в ПК являются поглощение и излучение электронами энергии тепловых флуктуаций. При поглощении электроны могут переходить из состояний с энергией меньшей высоты потенциального барьера в состояния с энергией большей этой высоты. С той же вероятностью происходят и обратные переходы с излучением тепловой энергии. Если между эмиттером и коллектором КП приложено ненулевое смещение, то в приборе возникает электрический ток. Электроны, поглотившие тепловую энергию, увлекаются этим током из контакта с меньшим приложенным напряжением в проводящий канал. В ПК электроны не могут излучить тепловую энергию, поскольку в этой области отсутствуют состояния с энергией, меньшей высоты потенциального барьера. Излучение тепловой энергии и уменьшение

энергии электронов происходит только в области противоположного контакта. Таким образом, при протекании тока баланс излучения и поглощения тепловой энергии в эмиттере и коллекторе квантового провода оказывается нарушенным. При положительном смещении коллектора по отношению к эмиттеру это должно приводить к охлаждению эмиттерного контакта и разогреву коллекторного. Для оценки возможностей технического применения описанного эффекта, необходимо представлять насколько значимым может быть предсказываемое охлаждение.

Положим, что приложенное между коллектором и эмиттером напряжение V_{ce} неотрицательно. В силу симметричности конструкции смена знака напряжения означает, что эмиттер и коллектор поменялись местами. Как и ранее обозначим через Δ разницу между максимальным значением эффективного потенциала для электронов в ПК и энергией Ферми E_F при нулевом смещении. Поскольку эффект инжекции и все связанные с ним явления существенны при

$$\Delta > 0,$$

будем считать это условие выполненным.

Заметим, что при сделанных допущениях электронный газ в контактных областях вырожден. В эмиттере его теплоемкость в расчете на одну частицу определяется соотношением [109]

$$C = k_B^2 T n_e^{-2/3} (\pi/3)^{2/3} (m^*/\hbar^2),$$

где n_e – концентрация электронов в эмиттерном контакте. Для энергии электронов в этом контакте справедливо выражение

$$E(T) = C T n_e V_e = \beta T^2 n_e^{1/2} V_e,$$

где V_e – объем эмиттерного контакта и введено обозначение

$$\beta = k_B^2 (\pi/3)^{2/3} (m^*/\hbar^2).$$

Энергию, уносимую током из эмиттерного контакта в ПК можно оценить как

$$\Delta E = \Delta \int dx n_{in} \approx \Delta n_{ine} V_{ch}.$$

Здесь n_{in} – локальная концентрация инжектированных в ПК электронов, n_{ine} – средняя по ПК концентрация инжектированных электронов, V_{ch} – объем ПК. Интегрирование ведется по объему ПК.

Пусть в отсутствии тока температура прибора равна температуре окружающей среды T . При $V_{ce} \neq 0$ изменение температуры эмиттерного контакта ΔT описывается уравнением баланса энергии

$$\Delta E = E(T) - E(T - \Delta T). \quad (401)$$

Считаем, что концентрация электронов в эмиттере велика и ее изменением при протекании тока можно пренебречь. Из уравнения (401) получим для ΔT квадратное уравнение

$$\Delta T^2 - 2T\Delta T + g = 0, \quad (402)$$

где

$$g = \Delta(n_e^{2/3}/\beta)(n_{ine} V_{ch}/n_e V_e).$$

Уравнение (402) решалось численно. Для нахождения n_{ine} решалась задача о переносе заряда в КП, описанная выше.

На Рисунках 84 и 85 приведены рассчитанные зависимости температуры эмиттерного контакта

$$T_e = T - \Delta T$$

от приложенного напряжения V_{ce} при двух значениях $T = 300^0\text{K}$ и 3^0K , а также различных концентрациях электронов в эмиттере.

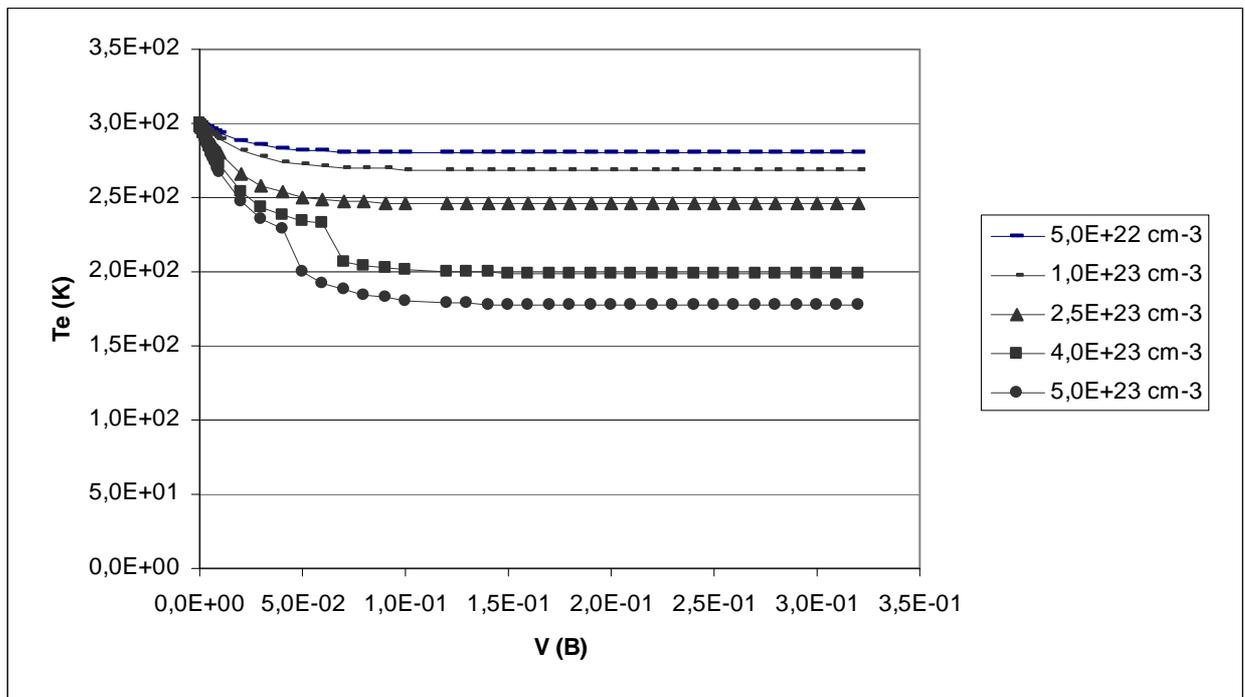


Рисунок 84. Зависимость температуры T_e от напряжения при $T = 300^0\text{K}$ и $V_e/V_{ch} = 100$.

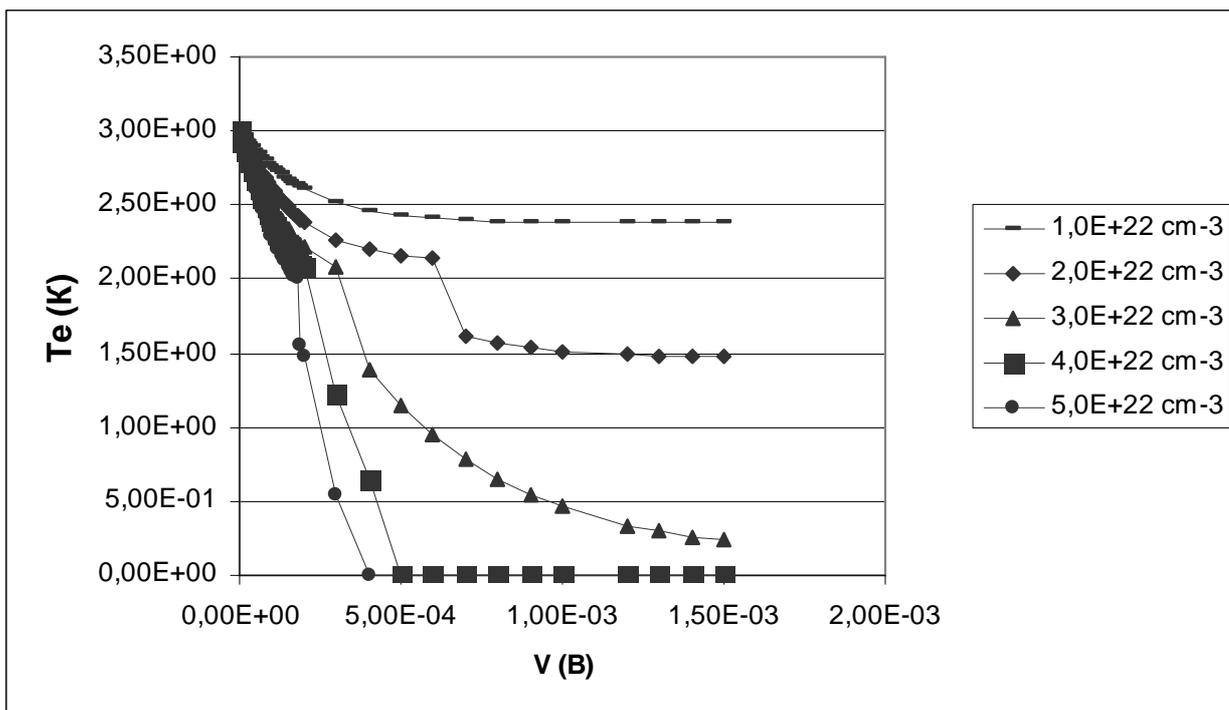


Рисунок 85. Зависимость температуры T_e от напряжения при $T = 3^0$ К и $V_e/V_{ch} = 1000$.

В расчетах полагалось, что длина ПК равна 10^{-4} см, а площадь сечения ПК варьировалась таким образом, чтобы при различных концентрациях электронов в контактах величина Δ была равна средней энергии тепловых флуктуаций kT . Во всех случаях площадь сечения ПК была порядка 10^{-12} см², а V_{ch} - порядка 10^{-16} см³.

При $T = 300^0$ К объем эмиттерного контакта выбирался в сто раз большим, чем объем проводящего канала, то есть был порядка 10^{-14} см³. Если контакт реализован на металлической пленке с толщиной 10^{-6} см, то его планарные размеры могут быть микронными (10^{-4} см). Для $T = 3^0$ К объем эмиттера считался в тысячу раз большим, чем объем ПК ($V_e = 10^{-13}$ см).

Расчеты показывают, что при соответствующем выборе материалов и размеров может быть достигнуто значительное охлаждение эмиттерного контакта квантового провода. При комнатной температуре окружающей среды оно может составлять десятки градусов (до 120^0 К при выбранных параметрах). Это позволяет надеяться, что охлаждающие устройства на основе квантовых проводов могут найти широкое применение в технике. В частности, их можно использовать для охлаждения интегральных схем (процессоров). Существенно, что температуру

охлаждения можно регулировать приложенным напряжением, то есть строить электрические схемы, управляющие тепловым режимом.

В работе [102] показано, что матрица квантовых проводов может играть роль фотоячейки в инфракрасном диапазоне длин волн. В совокупности с описанными в этом параграфе явлениями это открывает возможности для экспериментов по созданию на основе таких матриц «интеллектуальных» энергосберегающих покрытий [142,148].

6.5. Полевой транзистор на основе квантового провода

С точки зрения простейших применений в электронных схемах квантовый провод сам по себе может рассматриваться в качестве резистора. Как только что показано, возможны и другие применения этого объекта. Спектр разнообразных приборов, которые предлагается создавать и создаются на основе квантовых проводов достаточно широк (см., например, [1,2,19,23,69,70,128,138] и цитируемую в них литературу): от экзотических конструкций, использующих эффект Ааронова – Бома либо интерференцию электронных волн (интерференционные квантовые транзисторы), до обыкновенных полевых транзисторов.

В этом параграфе мы рассмотрим возможность применения квантового провода в качестве элемента квантового полевого транзистора (КПТ). Поскольку образцы такого рода приборов создаются уже достаточно давно [128,135-137], то существует возможность сопоставить рассчитанные и экспериментальные характеристики.

Функционирование КПТ основано на возможности управления потенциальным рельефом для электронов в проводящем канале КП при помощи полевого затвора. На Рисунке 86 показана одна из топологий такого прибора. Поскольку проводимость КП очень сильно зависит от потенциального рельефа, то имеет смысл предполагать, что чувствительность исток - стоковой вольт -

амперной характеристики КПТ к изменениям напряжения на затворе (крутизна) будет высокой.

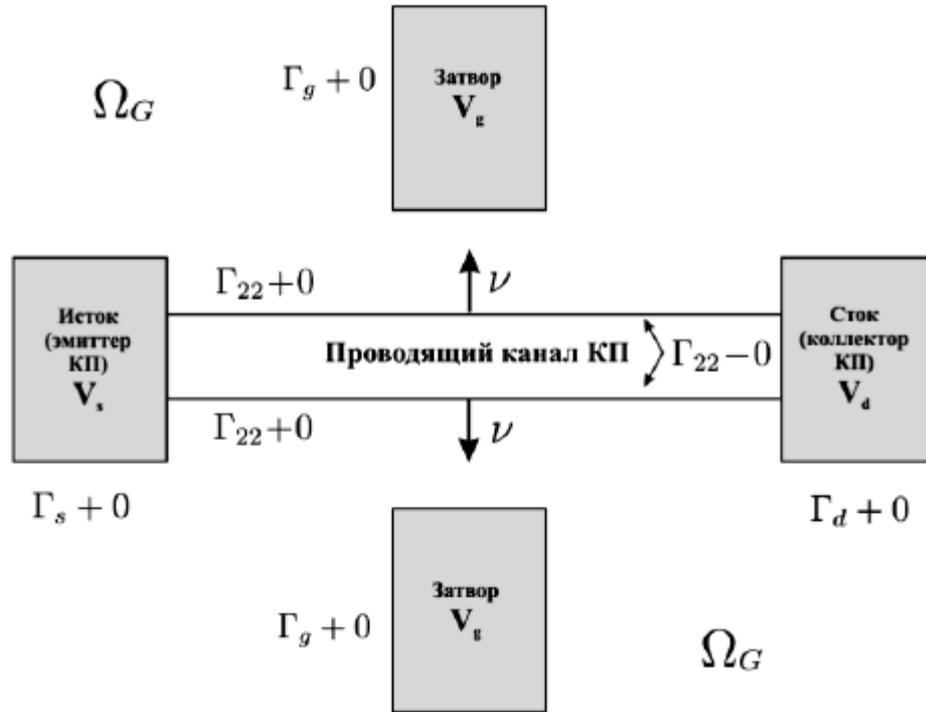


Рисунок 86. Возможная топология квантового полевого транзистора.

Истоком КППТ является эмиттерный контакт КПП, а стоком – коллекторный контакт КПП. Будем полагать, что

$$V_s = V_e, V_d = V_c$$

$$V_{ds} = V_d - V_s.$$

Напряжение на затворе будем обозначать как обычно [27] через V_g . Конструктивно затвор может быть реализован различными способами. Важно то, что между КПП и затвором находится область Ω_G , заполненная материалом с диэлектрическими свойствами и проницаемостью ϵ_1 . Границу области Ω_G с проводящим каналом КПП обозначим через Γ_{22} , а границы этой области с контактами КПП и затвором – через Γ_s , Γ_d и Γ_g соответственно. Пусть проводящий канал КПП имеет диэлектрическую проницаемость ϵ , тогда для электростатического потенциала ϕ на Γ_{22} справедливы граничные условия

$$\phi(\Gamma_{22-0}) = \phi(\Gamma_{22+0}),$$

$$\epsilon \partial \phi / \partial \nu (\Gamma_{22-0}) - \epsilon_1 \partial \phi / \partial \nu (\Gamma_{22+0}) = Q(\Gamma_{22}),$$

где $Q(\Gamma_{22})$ – поверхностный заряд на границе Γ_{22} . Кроме того, положим, что

$$\varphi(\Gamma_{s+0}) = \varphi_s, \varphi(\Gamma_{d+0}) = \varphi_d, \varphi(\Gamma_{g+0}) = \varphi_g,$$

а потенциал φ удовлетворяет внутри области Ω_G уравнению Лапласа

$$\Delta\varphi = 0.$$

Значения потенциалов $\varphi_{s,d,g}$ определяются на соответствующих границах как решения уравнения Пуассона в контактных областях и в затворе. Приближенно можно считать, что они определяются условием электронейтральности.

Внутри КП электростатический потенциал удовлетворяет, вообще говоря, уравнению Пуассона (385) с граничными условиями (387). Однако сделанные нами упрощения позволяют свести задачу для φ внутри КП к одномерной задаче (389), (391). В этом случае

$$\partial\varphi/\nu(\Gamma_{22-0}) = 0.$$

Кроме того, поверхностный заряд $Q(\Gamma_{22})$ вполне может существовать, но о его величине и пространственном распределении вряд ли можно сказать что-либо определенное. Имеющегося экспериментального и теоретического материала пока не достаточно. Поэтому для оценки влияния затвора на характеристики КПП можно положить

$$Q(\Gamma_{22}) = 0.$$

Отличие поверхностного заряда от нуля может существенно изменить характеристики прибора. Однако характер этих возможных изменений вполне понятен и аналогичен наблюдаемым у обычных полевых транзисторов [27]. В результате на Γ_{22} имеем

$$\varphi(\Gamma_{22-0}) = \varphi(\Gamma_{22+0}),$$

$$\partial\varphi/\partial\nu(\Gamma_{22+0}) = 0.$$

На Рисунке 87 показан эффективный потенциальный рельеф для электронов в КП с $L = L_{dq}$ при условии нулевого смещения между эмиттером и коллектором (сток-исток). Если на основе этого КП сформировать КПП с длиной затвора 70 нм и расстоянием между затвором и проводящим каналом 20 нм, то при приложении

к затвору отличного от нуля напряжения потенциальный рельеф существенно изменится, что демонстрируют Рисунки 88 и 89.

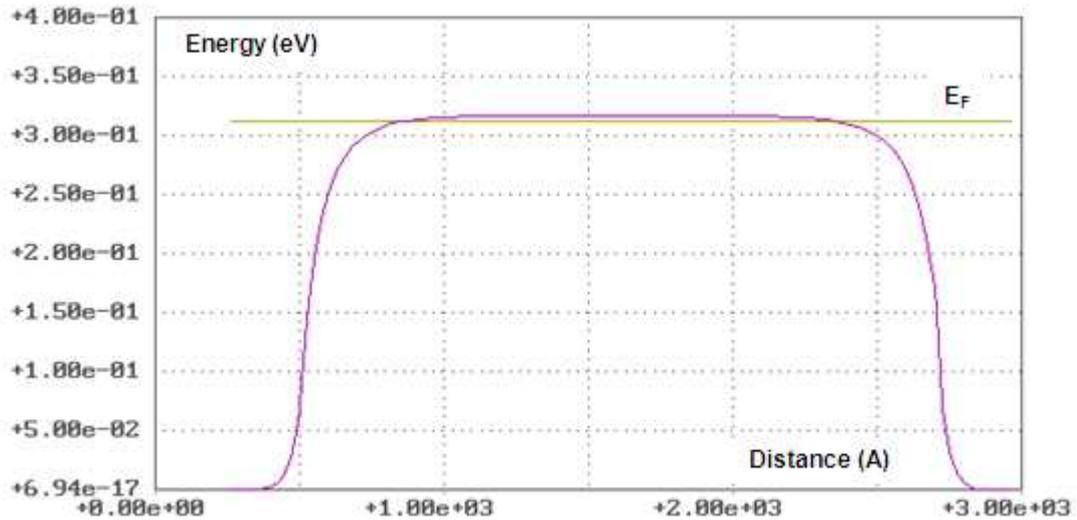


Рисунок 87. Эффективный потенциальный рельеф для электронов в КП с $L = L_{dq}$ в направлении OX при $V_{ce} = V_{ds} = 0$.

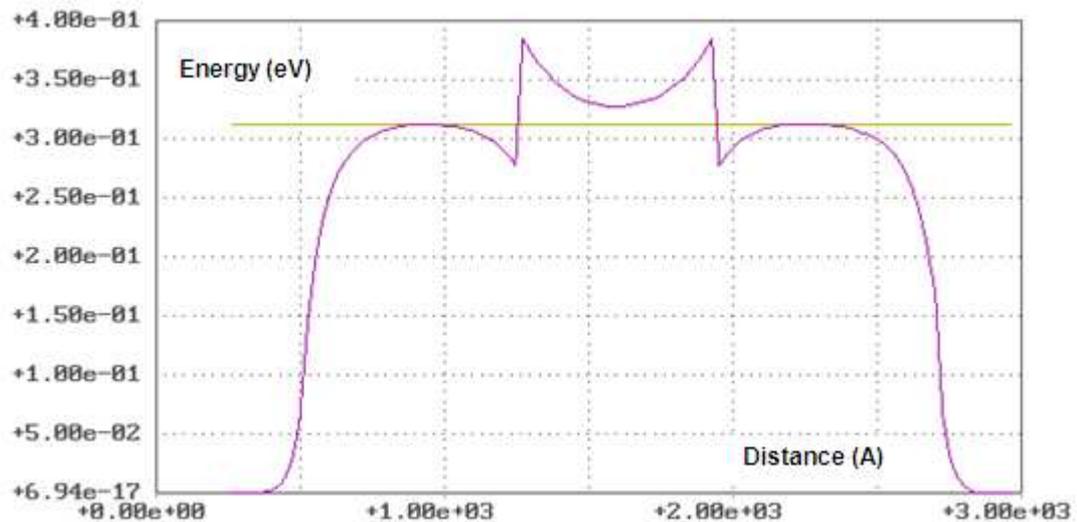


Рисунок 88. Эффективный потенциальный рельеф для электронов в КПТ на основе КП с $L = L_{dq}$ при $V_{ds} = 0$ и $V_g = -0,7$ В.

Появляющийся при $V_g < 0$ потенциальный барьер растет по высоте с увеличением абсолютной величины V_g . Это означает, что ток через структуру должен сильно зависеть от V_g , поскольку увеличение высоты потенциального барьера ведет к уменьшению коэффициента прохождения электронов через него.

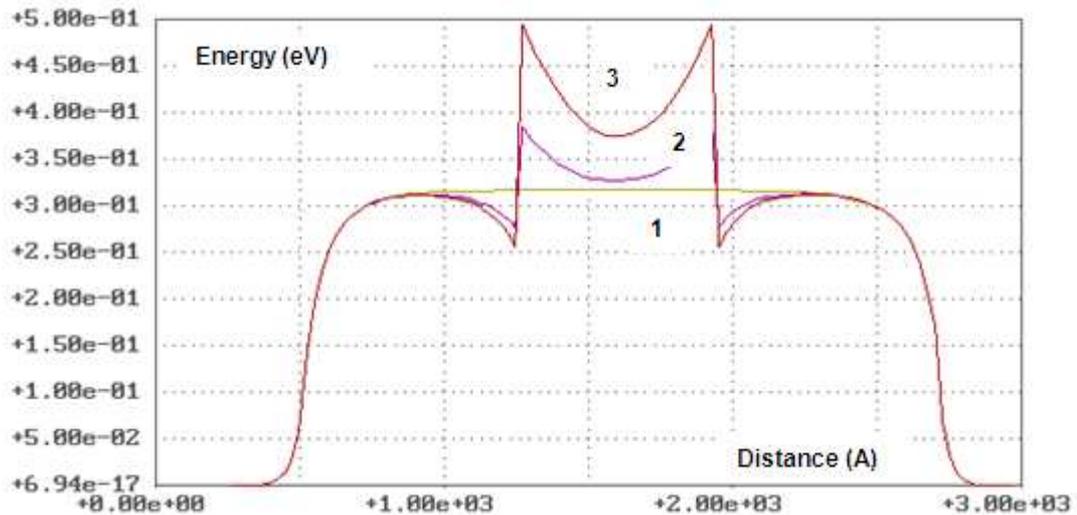


Рисунок 89. Эффективный потенциальный рельеф для электронов в КП с $L = L_{dq}$ (кривая 1) и в КПТ на основе этого КП с при $V_{ds} = 0$ и 2 - $V_g = -0,7$ В, 3 - $V_g = -1,5$ В.

Влияние приложенного напряжения V_{ds} приводит к существенным искажениям потенциального рельефа для электронов, что демонстрируется на Рисунке 90. Хорошо видны области пространственного заряда, влияющие на работу прибора.

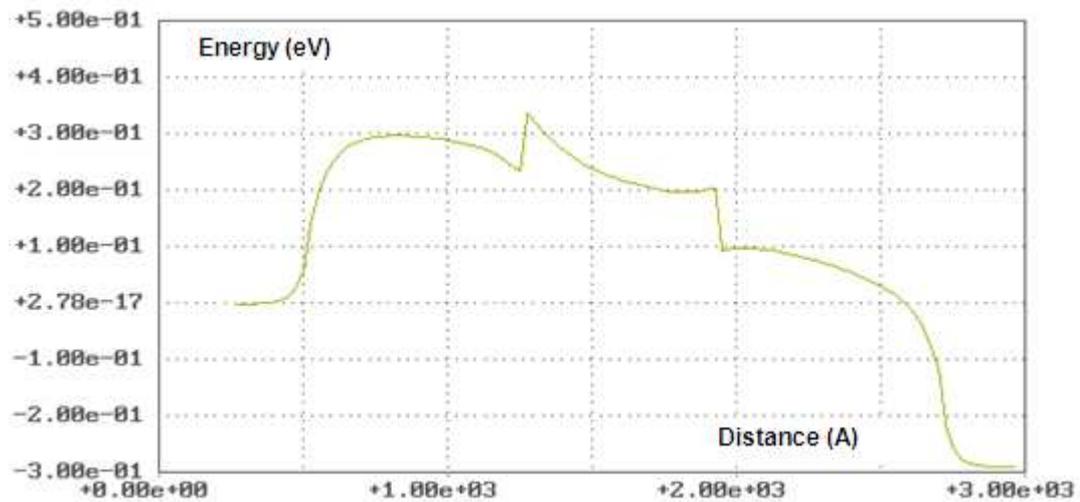


Рисунок 90. Эффективный потенциальный рельеф для электронов в КПТ при $V_{ds} = 0,29$ В и $V_g = -0,7$ В.

Очевидно, что изменения, вносимые затвором в потенциальный рельеф для электронов в проводящем канале прибора, существенно повлияют на вольт - амперные характеристики, что и демонстрируется на Рисунке 91.

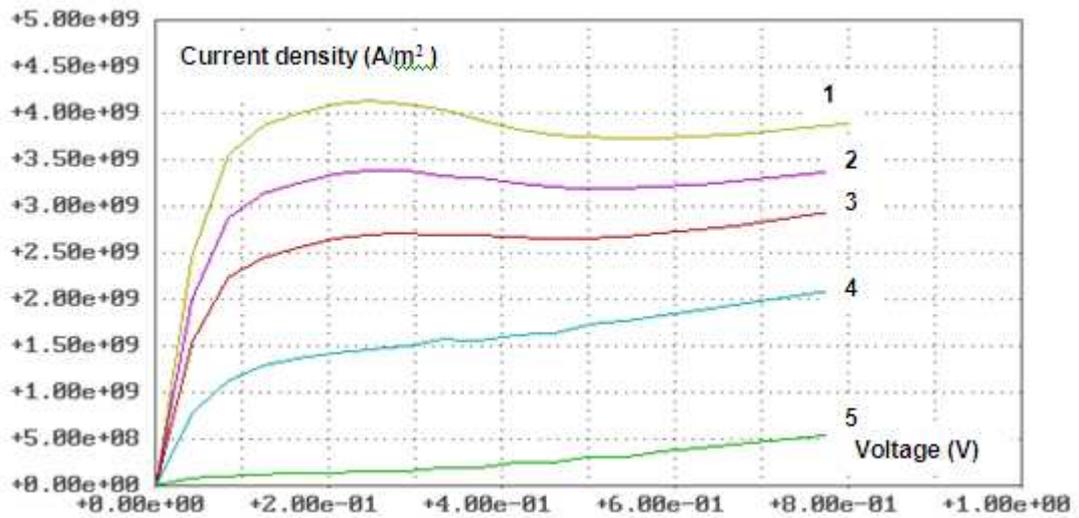


Рисунок 91. Удельные ВАХ КПТ при различных напряжениях на затворе 1 - $V_g = 0$ В, 2 - $V_g = -0,5$ В, 3 - $V_g = -0,7$ В, 4 - $V_g = -1,0$ В, 5 - $V_g = -1,5$ В.

Следует обратить внимание на качественные различия характеристик 1,2,3 и 4,5, представленных на Рисунке 91. На двух нижних ВАХ отчетливо видны изломы, характерные для чисто квантового переноса заряда и обусловленные наличием осцилляций коэффициента прохождения электронов через структуру как функции энергии (см. Рисунок 68). Такое поведение соответствует режиму запираания КПТ, когда высота потенциального барьера, создаваемого затвором становится достаточно большой.

Полученные в результате моделирования ВАХ КПТ аналогичны характеристикам обычных полевых транзисторов и подтверждают предположение о возможности полевого управления квантовыми приборами. Простейшие оценки показывают, что ВАХ КПТ обладают рекордно большой удельной крутизной g , которая составляет величины порядка 10^9 А/Вм² и на два - три порядка превышает аналогичные величины для приборов микроэлектроники. Однако собственно крутизна g_m оказывается невысокой

$$g_m \sim 10^{-5} - 10^{-7} \text{ А/В},$$

то есть на три порядка меньшей, чем у обычных полевых приборов [27]. Аналогичные выводы можно сделать, анализируя экспериментальные и расчетные работы по КПТ на основе углеродных нанотрубок и квантовых проводов из InSb (см., например, [94,128,138]).

Эти обстоятельства имеют элементарное объяснение, связанное с малой площадью сечения проводящего канала КПТ, которая не превышает величин $\sim 10^{-16} \text{ м}^2$, тогда как полевые транзисторы микроэлектроники имеют площади сечений проводящего канала не меньше 10^{-10} м^2 , то есть примерно на шесть порядков больше. Выигрыш в удельной крутизне, получаемый за счет использования квантовых эффектов, значительно повышающих чувствительность прибора к напряжению на затворе, оказывается недостаточным для компенсации малой площади сечения проводящего канала КПТ.

Расчетная характерная емкость КПТ составляет приблизительно 10^{-18} Ф . Сопротивление прибора примерно 10^6 Ом . Это позволяет надеяться на рабочие частоты порядка 1 ТГц.

6.6. Релаксационный квантовый диод

Традиционные элементы микроэлектроники не исчерпываются только полевыми приборами. Широкое применение в электронной технике имеют также биполярные диоды и транзисторы [27]. Представляет интерес конструирование приборов на основе квантовых проводов, имеющих вольт - амперные характеристики, аналогичные характеристикам биполярных приборов. В этом случае можно надеяться воспроизвести известные схемотехнические решения с использованием новой элементной базы [149] и реализовать все те преимущества нанoeлектроники, о которых говорилось в начале этой главы.

Как показывает анализ, проведенный в первых трех главах настоящей работы, с физической точки зрения, функционирование традиционных биполярных приборов основано на нарушении локального химического равновесия между электронами и дырками, возникающем при протекании тока через р-п-переход и наличии явлений электронно-дырочной генерации и рекомбинации. Результатами этой неравновесности являются и специфический вид ВАХ биполярного диода с практически односторонней проводимостью, и эффект усиления сигнала в биполярном транзисторе.

Как правило, в квантовом проводе превалирует только один тип проводимости. Поэтому, если пытаться использовать неравновесные эффекты с целью получения желаемых характеристик прибора, то речь можно вести о неравновесности между различными фазами носителей заряда, в нашем случае – электронов. Как показано выше, разные виды этой неравновесности очень сильно влияют на характеристики КП. Однако следует заметить, что согласно представленным на Рисунке 69 результатам, уровень неравновесности «правых» и «левых» электронов, выраженный в виде разности их химических потенциалов F на порядок меньше, чем падение напряжения на структуре (величина eV_{ce}). Это означает, что указанная неравновесность вносит существенный вклад в омическую составляющую падения напряжения только за счет изменения концентрации электронов. Для получения же диодной ВАХ либо эффекта усиления в транзисторе необходимо, чтобы максимальные значения F внутри прибора были порядка eV_{ce} .

Каким же образом можно получить ВАХ, аналогичную ВАХ биполярного диода, используя квантовые провода в качестве элементов конструкции? Для ответа на этот вопрос следует, в первую очередь, понять, почему в обычном квантовом проводе $F \ll eV_{ce}$? Так же как в РТС величина F в КП подчиняется уравнению (337), и для оценки F в нулевом приближении пригодна формула (338). Из нее следует, что при одной и той же плотности тока F тем больше, чем меньше в области источника неравновесности суммарная концентрация «правых» и «левых» электронов. Источники неравновесности локализованы в областях сильных градиентов концентраций электронов. В квантовом проводе - это условные границы между контактами и проводящим каналом. Концентрация электронов со стороны контактов в этом случае очень велика, что приводит к малости F по сравнению с eV_{ce} .

Приведенные рассуждения наводят на мысль о том, что если реализовать источник неравновесности «правых» и «левых» электронов в виде контакта двух проводящих каналов различных сечений, то есть надежда добиться того, чтобы значительная часть падения напряжения в структуре происходила за счет

неравновесности, что равнозначно получению диодной ВАХ. При этом не критично, что положения уровней энергии электронов в контактирующих ПК, возникающие в результате размерного квантования в поперечных переносу заряда направлениях, не будут совпадать. Изученный выше эффект инжекции электронов из контактов КП справедлив и в случае, если в качестве контакта к ПК выступает другой ПК с большим сечением. Важно только, чтобы разница энергий ΔE между нижним уровнем «тонкого» канала и ближайшим по энергии уровнем «толстого» канала была бы меньше средней энергии тепловых флуктуаций

$$\Delta E < k_B T.$$

Рассмотрим структуру, схематично изображенную на Рисунке 92, в качестве кандидата на наноэлектронный аналог биполярного диода. Будем называть релаксационным квантовым диодом (РКД) прибор, состоящий из контактов и двух проводящих каналов различного сечения, контактирующих между собой и с одним из контактов каждый. При этом подразумевается, что линейный поперечный размер каждого из ПК не больше чем L_{dq} . На рисунке через $x_{c,e}$ обозначены координаты границ, на которых задается приложенное напряжение, а через x_j координата «металлургической» границы между ПК различной толщины.

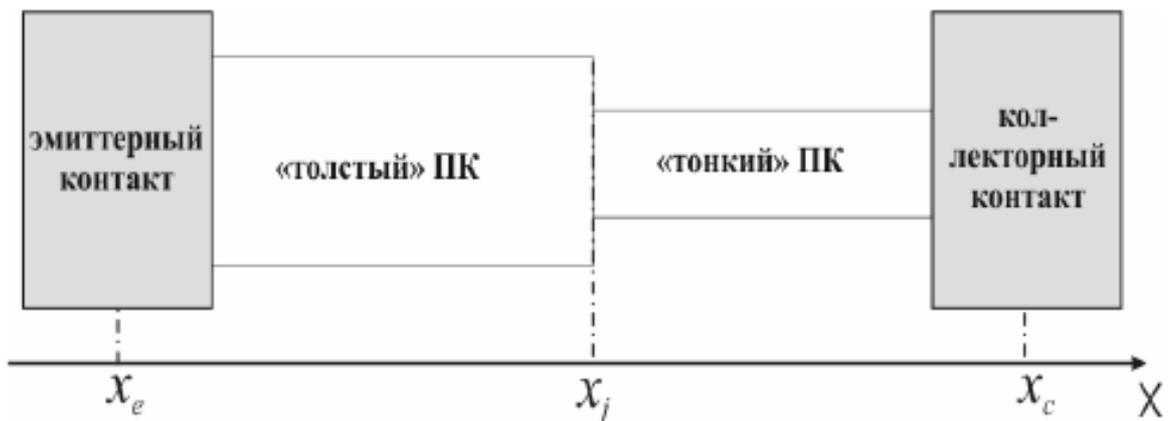


Рисунок 92. Топология релаксационного квантового диода.

Эффективный потенциальный рельеф для электронов в РКД в направлении переноса при нулевом смещении показан на Рисунке 93. В качестве проводящих каналов взяты ПК с $L = L_{dq}$ («толстый» ПК) и ПК с $L = L_{dq}/2$ («тонкий» ПК). Хорошо видно, что при переходе от «толстого» канала к «тонкому» образуется

потенциальный барьер. Именно он является основным источником неравновесности в приборе.

Аналогично соотношению (396), представим плотности потоков «левых» и «правых» электронов в РКД в виде

$$I_l = -\sigma_l dF_l/dx, I_r = -\sigma_r dF_r/dx, \quad (403)$$

где для проводимостей $\sigma_{l,r}$ справедливы формальные выражения:

$$\sigma_l = (dF_l/dx)^{-1} \sum_{Ej} j_l f_l, \sigma_r = (dF_r/dx)^{-1} \sum_{Ej} j_r f_r.$$

Считая, как и в случае РТД, что

$$F_+ = F_l + F_r, F_- = F_l - F_r,$$

получим

$$dF_+/dx = -2I/\sigma_+ - (\sigma_-/\sigma_+)dF_-/dx, \quad (404)$$

а

$$I = I_l + I_r = \text{const},$$

$$\sigma_+ = \sigma_l + \sigma_r, \sigma_- = \sigma_l - \sigma_r.$$

Простые оценки показывают, что при $x < x_j$ $\sigma_l \gg \sigma_r$, а при $x > x_j$ $\sigma_l \sim \sigma_r$. Учитывая, что

$$F_+(x_c) - F_+(x_e) = -2eV_{ce},$$

из (404) найдем приближенное выражение для ВАХ РКД

$$eV_{ce} \approx I r_{ce} + F_-(x_j), \quad (405)$$

где

$$r_{ce} = \int_{x_e}^{x_c} dx/\sigma_+$$

- омическое сопротивление прибора.

Соотношение (405) позволяет понять, какой вклад вносит неравновесность между фазами «левых» и «правых» электронов в падение напряжения в РКД. На Рисунках 94 и 95 показаны эффективный потенциальный рельеф РКД и F_- при $V_{ce} = 0,066$ В, полученные в результате численного моделирования переноса заряда через структуру. При положительных значениях V_{ce} потенциальный барьер для «левых» электронов понижается с ростом напряжения, что приводит к росту тока

через структуру. Хорошо видно, что F_e достигает максимума в области границы между «толстым» и «тонким» ПК, а абсолютное значение F_e в максимуме равняется примерно половине приложенного напряжения V_{ce} . На Рисунке 96 показаны химические потенциалы F_1 и F_r при $V_{ce} = 0,066$ В. Обратим внимание на немонотонное поведение функций внутри структуры, что согласуется с результатами работы [68].

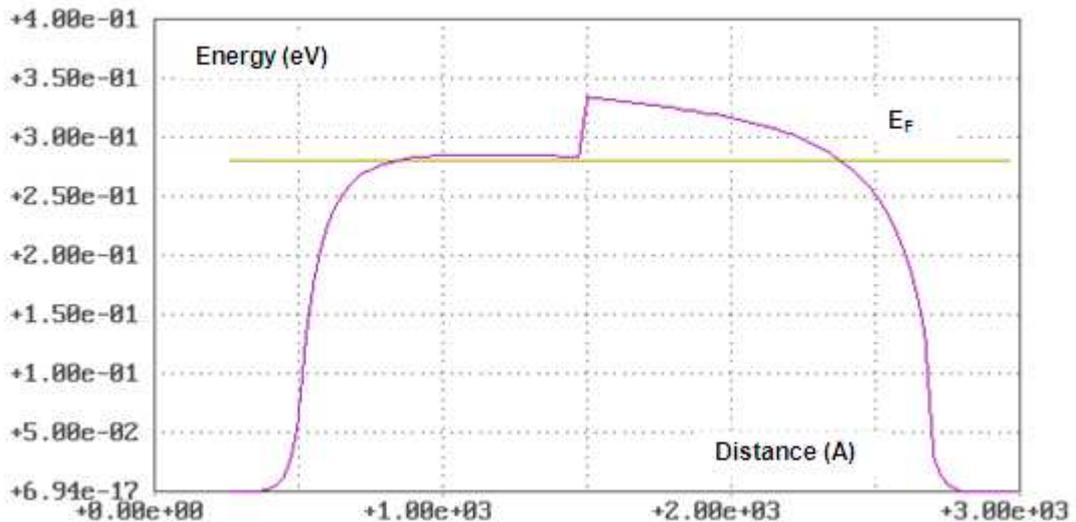


Рисунок 93. Эффективный потенциальный рельеф для электронов в РКД на основе ПК с $L = L_{dq}$ («толстый» ПК) и ПК с $L = L_{dq}/2$ («тонкий» ПК) в направлении переноса при $V_{ce} = 0$.

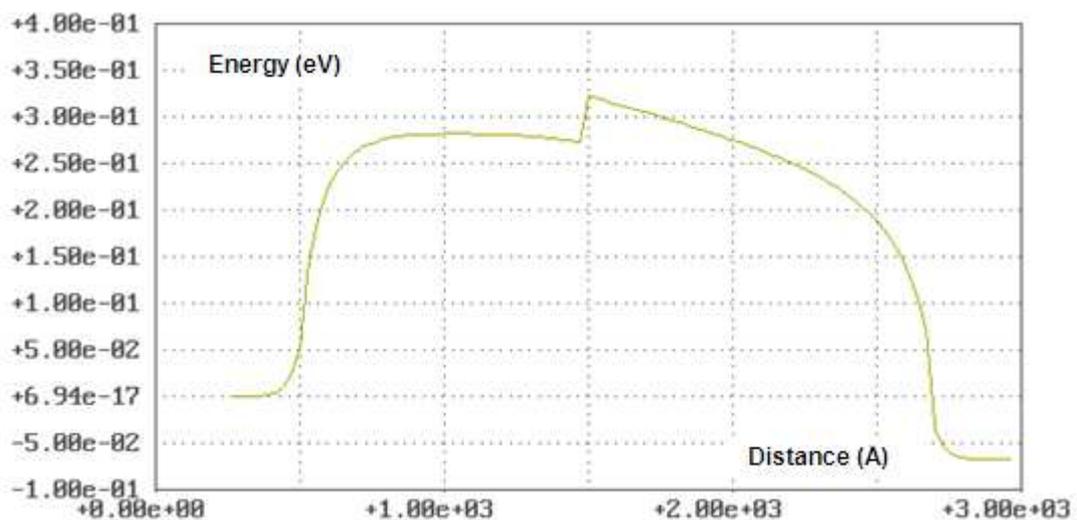


Рисунок 94. Эффективный потенциальный рельеф для электронов в РКД при $V_{ce} = 0,066$ В.

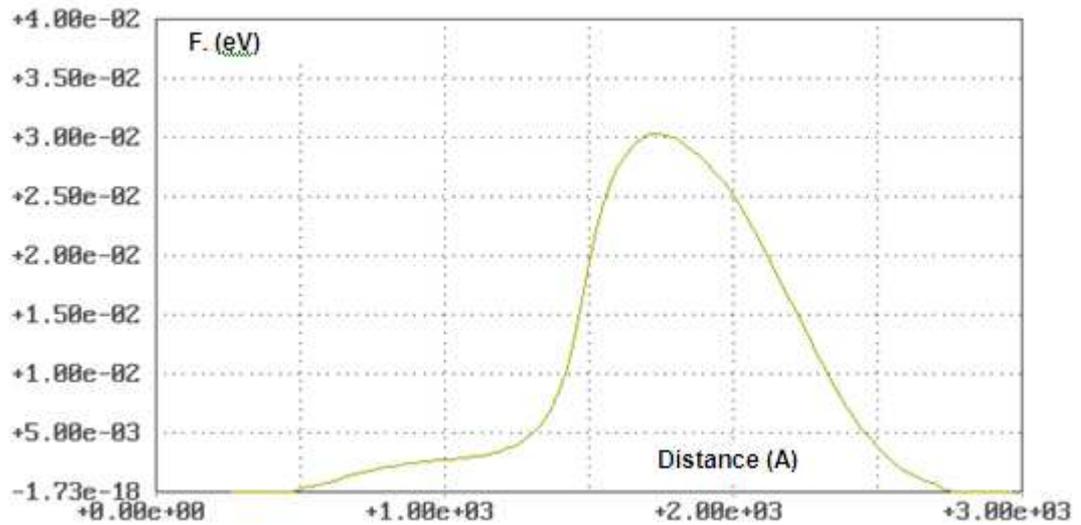


Рисунок 95. Разность химических потенциалов F в РКД при $V_{ce} = 0,066$ В.

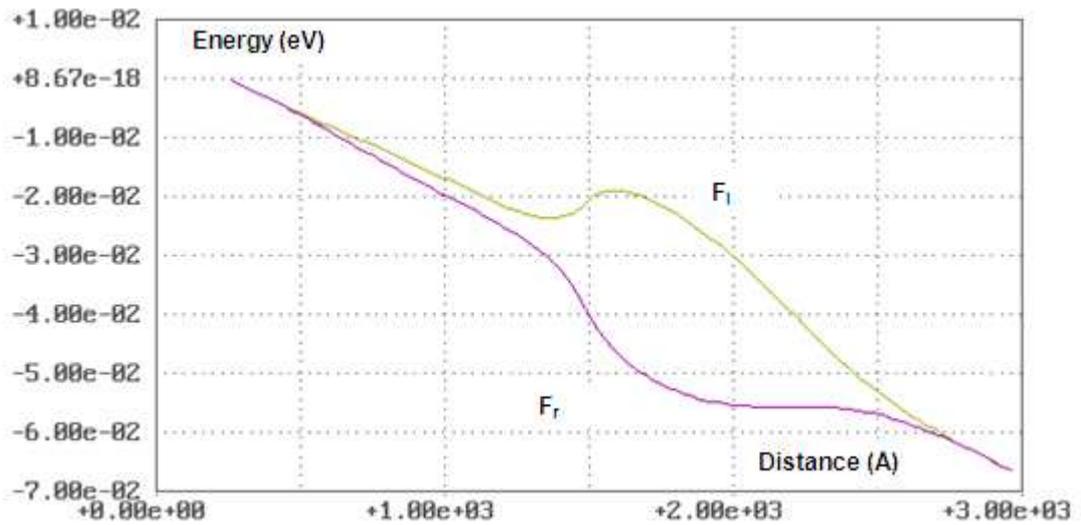


Рисунок 96. Химические потенциалы F_1 и F_r в РКД при $V_{ce} = 0,066$ В.

Если приложить к структуре напряжение другого знака, то получим потенциальные рельефы, аналогичные изображенным на Рисунках 97 и 98. Видно, что при $V_{ce} < 0$ потенциальный барьер для «левых» электронов растет, что должно приводить к эффективному увеличению сопротивления прибора. Это действительно так, в чем можно убедиться, анализируя поведение F при отрицательных смещениях, показанное на Рисунке 99 и ВАХ РКД, представленную на Рисунке 100.

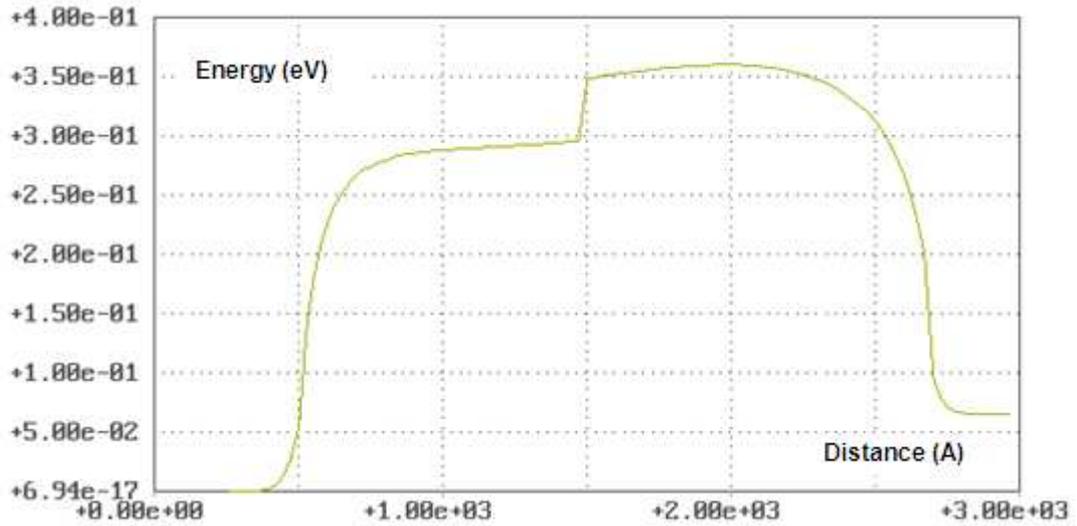


Рисунок 97. Эффективный потенциальный рельеф для электронов в РКД при $V_{cc} = -0,066$ В.

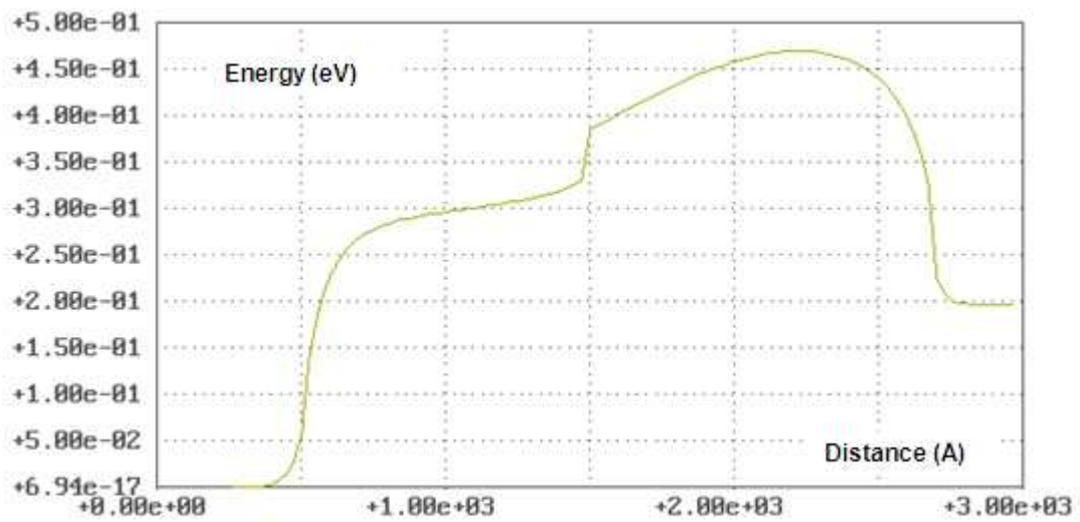


Рисунок 98. Эффективный потенциальный рельеф для электронов в РКД при $V_{cc} = -0,2$ В.

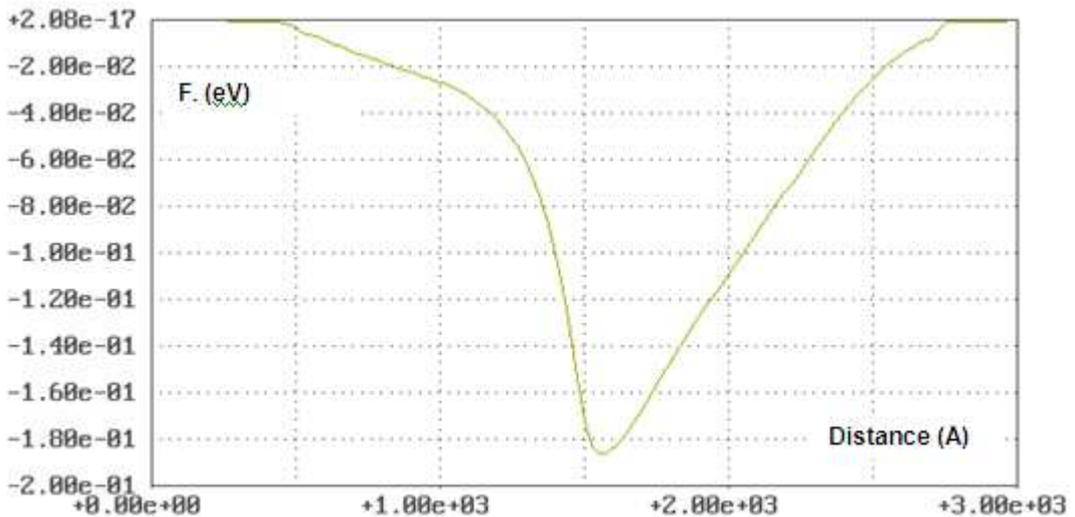


Рисунок 99. Разность химических потенциалов F в РКД при $V_{cc} = -0,2$ В.

Анализ результатов моделирования, представленных на Рисунках 99 и 101, показывает, что в РКД при $V_{ce} < 0$ вклад неравновесности F_- в падение напряжения больше, чем при $V_{ce} > 0$. В моделируемой структуре при отрицательных смещениях $F_-(x_j)$ составляет примерно 75% от V_{ce} . Очевидно, что в общем случае вклад неравновесности в ВАХ РКД на различных участках может существенно зависеть от конструктивных параметров прибора.

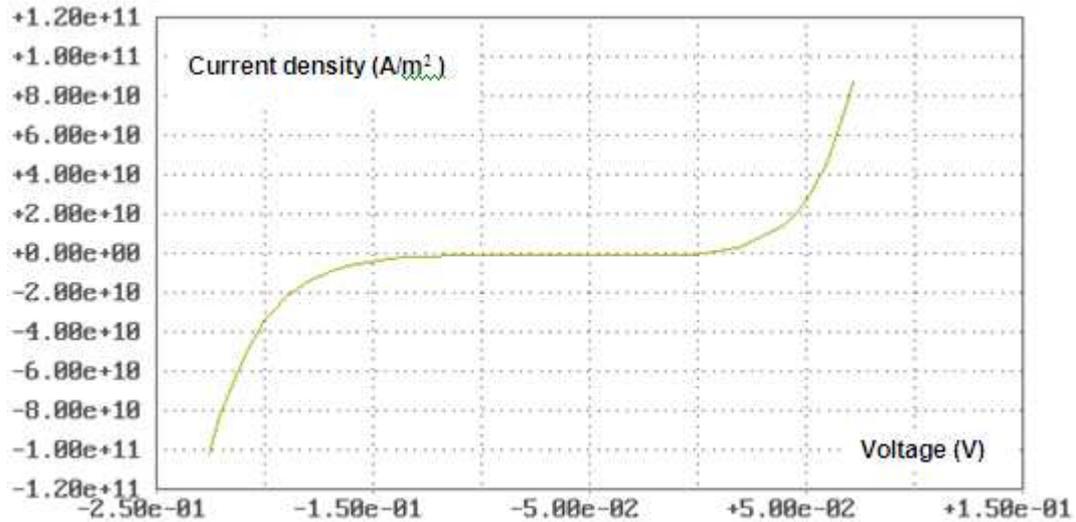


Рисунок 100. ВАХ РКД на основе ПК с $L = L_{dq}$ («толстый» ПК) и ПК с $L = L_{dq}/2$ («тонкий» ПК).

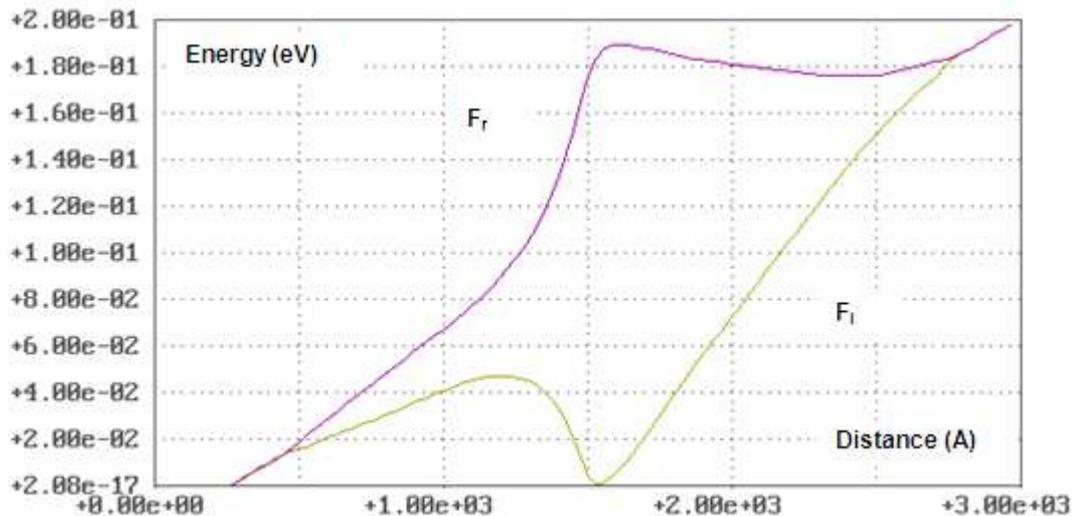


Рисунок 101. Химические потенциалы F_1 и F_r в РКД при $V_{ce} = -0,2$ В.

ВАХ РКД, представленная на Рисунке 100, полностью подтверждает сделанное в начале этого параграфа предположение о возможности реализации на основе квантовых проводов структуры, имеющей выходные характеристики, аналогичные обычному биполярному диоду. Результаты моделирования

демонстрируют, что функционирование РКД основано на неравновесных свойствах электронного газа.

Расчеты показывают, что РКД может иметь очень высокие плотности тока на прямом участке ВАХ, достигающие значений $5 \cdot 10^{10}$ А/м² ($5 \cdot 10^6$ А/см²) и выше. Для структур на основе GaAs сечение «тонкого» ПК при $L = L_{dq}/2$ порядка 10^{-16} м², то есть токи в РКД достаточно большие – микроамперы (10^{-6} А). Простые вычисления позволяют оценить удельную емкость прибора, которая оказывается порядка 10^{-3} Ф/м² [77]. Учитывая, что характерное удельное сопротивление РКТ примерно 10^{-12} Ом*м², получим для предельной частоты прибора f_{lim} оценку

$$f_{lim} \sim 10^{15} \text{ Гц.}$$

Таким образом, РКД – это электронный прибор, способный работать в диапазоне частот вплоть до тысяч ТГц.

6.7. Релаксационный квантовый транзистор

Как видно на Рисунках 95, 96, 99 и 101 область неравновесности электронного газа в РКД имеет достаточно большую протяженность. Ее размеры в несколько раз превышают длину размерного квантования L_{dq} и длину релаксации химических потенциалов L_0 в объемном материале. Это дает надежду на то, что современная технология позволит создать трехвыводное устройство, позволяющее управлять параметрами потока неравновесных электронов, изменяя ток и (или) напряжение на третьем контакте. Для этого этот третий контакт, который мы традиционно будем называть базовым, должен быть размещен между эмиттером и коллектором.

На Рисунке 102 схематично показана, удовлетворяющая сформулированному требованию структура, которую мы будем называть релаксационным квантовым транзистором (РКТ). Прибор состоит из трех контактов: эмиттерного, базового и коллекторного и трех ПК с поперечными размерами, меньшими длины размерного квантования, которые на рисунке обозначены как эмиттер, база и коллектор, а также ПК, соединяющего базу с

базовым контактом. Эмиттер и коллектор должны представлять собой ПК с поперечными размерами, большими, чем у базы, так, чтобы обеспечить концентрацию носителей заряда в базе много меньшую, чем в эмиттере и коллекторе.

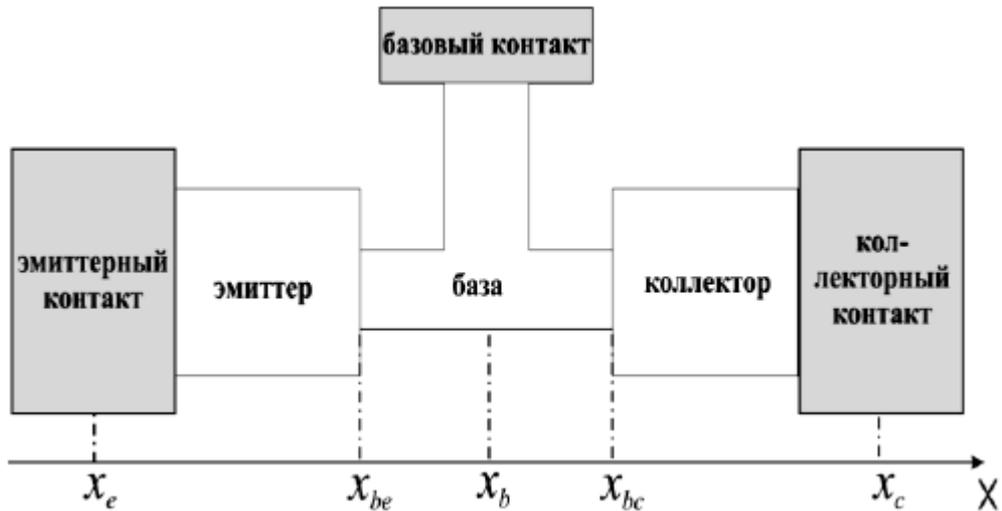


Рисунок 102. Топология релаксационного квантового транзистора.

Если длина базы

$$L_b = x_{bc} - x_{be}$$

много больше длины релаксации носителей заряда L_{rel} в ней к состоянию химического равновесия, РКТ будет представлять собой два включенных навстречу друг другу РКД с общим контактом к «тонкому» ПК, каждый из которых имеет аналогичную показанной на Рисунке 100 ВАХ диодного типа.

Если же

$$L_b < L_{rel},$$

то неравновесные электроны из эмиттера будут неравновесными и в области коллектора. Это означает, что модуляция потока этих неравновесных электронов сигналом на базе может быть усилена. На Рисунке 103 показан потенциальный рельеф РКТ при нулевых напряжениях, приложенных ко всем трем контактам.

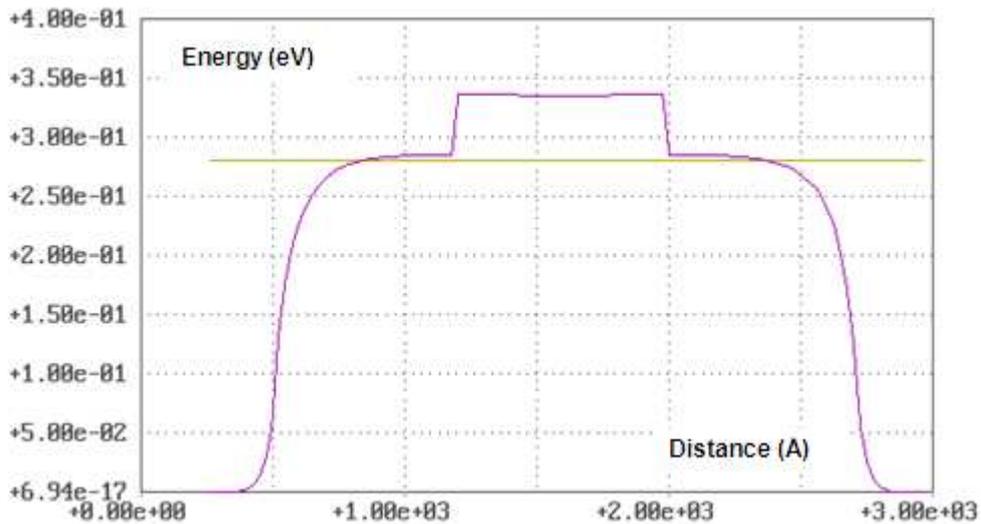


Рисунок 103. Эффективный потенциальный рельеф для электронов РКТ в направлении преноса от эмиттера к коллектору при $V_e = V_b = V_c = 0$. Поперечный размер эмиттерного и коллекторного ПК равен L_{dq} , а базового ПК – $L_{dq}/2$.

Из закона сохранения потока следует, что поток электронов, вытекающих из коллектора J_c , связан с потоком электронов, втекающих в эмиттер J_e , соотношением

$$J_c = J_e + J_b, \quad (406)$$

где: J_b – поток электронов в проводящем канале, соединяющем базу и базовый контакт. По определению коэффициент усиления прибора по току при схеме включения с общим эмиттером β равен

$$\beta = \partial J_c / \partial J_b = \alpha / (1 - \alpha), \quad (407.1)$$

где α коэффициент усиления транзистора при схеме включения с общей базой

$$\alpha = \partial J_c / \partial J_e \quad (407.2)$$

Ясно, что свойства РКТ как транзистора определяются способностью этого прибора усиливать сигнал, то есть коэффициентами α и β . Для понимания зависимости этих коэффициентов от параметров структуры произведем некоторые упрощенные аналитические оценки.

Полностью корректное описание переноса заряда в РКТ требует двумерного моделирования и применения трехфазной модели, разделяющей электроны по источнику попадания их в прибор на эмиттерные (F_e), коллекторные (F_c) и базовые (F_b). Однако качественно правильные результаты могут быть получены и при серьезных упрощениях. Поскольку основными источниками неравновесности

электронного газа являются переходы эмиттер–база (x_{eb}) и база–коллектор (x_{bc}), то можно приближенно считать, что

$$F_c = F_e = F, \quad (408)$$

и рассматривать только две электронных фазы. Тогда для плотностей потоков электронов в этих фазах имеем

$$I_b = -\sigma_b dF_b/dx, \quad I_{ec} = -(\sigma_e + \sigma_c) dF/dx, \quad (409)$$

где через I_{ec} мы обозначили суммарную плотность потока эмиттерных и коллекторных электронов

$$I_{ec} = I_e + I_c.$$

Для проводимостей $\sigma_{b,e,c}$ справедливы формальные выражения

$$\sigma_b = (dF_b/dx)^{-1} \sum_{E_b} j_b f_b, \quad \sigma_{e,c} = (dF/dx)^{-1} \sum_{E_{e,c}} j_{e,c} f_{e,c},$$

а для потоков в одномерном приближении имеют место уравнения, следующие из (380) и (409)

$$\begin{cases} dI_{ec}/dx = -\hbar^{-1}(F - F_b)(n + n_b), \\ dI_b/dx = -\hbar^{-1}(F_b - F)(n + n_b), \end{cases} \quad (410)$$

где

$$n = n_e + n_c.$$

Из соотношений (410) следует сохранение плотности суммарного потока электронов

$$d(I_{ec} + I_b)/dx = 0, \quad (411)$$

что позволяет считать суммарный поток постоянным в различных областях прибора. Будем полагать, что

$$I = I_{ec} + I_b = \begin{cases} j_e = J_e/S_e \text{ при } x < x_b \\ j_c = J_c/S_c \text{ при } x > x_b, \end{cases} \quad (412)$$

где $S_{e,c}$ – площади сечений эмиттерного и коллекторного ПК. То есть будем считать плотность потока электронов в РКТ кусочнопостоянной функцией.

Определим

$$F_+ = F + F_b, \quad F_- = F - F_b,$$

тогда

$$I_b = -(\sigma_b/2)d(F_+ - F_-)/dx, I_{ec} = -(\sigma/2)d(F_+ + F_-)/dx,$$

где

$$\sigma = \sigma_e + \sigma_c.$$

В результате, с учетом (412) получим выражение, аналогичное (404)

$$dF_+/dx = -2I/\sigma_+ - (\sigma_-/\sigma_+)dF/dx, \quad (413)$$

где

$$\sigma_+ = \sigma + \sigma_b, \quad \sigma_- = \sigma - \sigma_b.$$

Определим граничные условия для химических потенциалов следующим образом

$$F(x_e) = -eV_e, F(x_c) = -eV_c, F_b(x_b) = -eV_b, F_b(x_e) = -eV_e, F_b(x_c) = -eV_c. \quad (414)$$

Заметим, что значение F в точке x_b не фиксировано, оно совпадает с $-eV_b$ только на базовом контакте, отделенном от x_b соединительным ПК, перенос заряда в котором мы не рассматриваем. Из определения $F_{+,-}$ и выражения (413) найдем

$$\begin{cases} F_-(x_{e,c}) = 0, \\ F_+(x_e) = -2eV_e, F_+(x_c) = -2eV_c, \\ F_+(x_b) = -2eV_b + F_-(x_b). \end{cases} \quad (415)$$

Согласно (414), в точке x_b граничное условие на разность химических потенциалов F отсутствует. То есть электроны, попадающие в прибор из базового контакта, могут быть неравновесными по отношению к коллекторным и эмиттерным электронам в x_b . При этом нужно учитывать, что точка x_b - это не базовый контакт, а сопоставляемая ему точка в базовом ПК. Непосредственно на базовом контакте все электронные фазы находятся в равновесии. Последнее из условий (414) является способом учесть основные особенности функционирования прибора в рамках одномерного приближения, в то время как корректное описание требует двумерного моделирования. Ситуация вполне аналогична одномерной модели биполярного транзистора.

Для грубой оценки будем приближенно считать, что

$$(\sigma_-/\sigma_+) \sim 1 \text{ при } x < x_{eb} \text{ и } x > x_{bc},$$

$$(\sigma_-/\sigma_+) \sim -1 \text{ при } x_{eb} < x < x_{bc}.$$

При численных расчетах, результаты которых демонстрируются на рисунках, представленных в этом параграфе, эти приближения, естественно, не принимались во внимание. С учетом (412) из (414) нетрудно получить

$$\begin{cases} e(V_b - V_e) = j_e r_e + F_-(x_{eb}), \\ e(V_c - V_b) = j_c r_c - F_-(x_{bc}), \end{cases} \quad (416)$$

где

$$r_{e,c} = \int_{x_{e,b}}^{x_{b,c}} dx / \sigma_+$$

- сопротивления эмиттера и коллектора. Величины $F_-(x_{eb,bc})$ зависят от потоков $j_{e,c}$.

Исходя из (410), получим уравнение для F_- в РКТ

$$d\{((\sigma_+^2 - \sigma_-^2)/\sigma_+)dF_-/dx\}/dx = (4/\hbar)F_-(n + n_b) + 2Id(\sigma_- \sigma_+)/dx. \quad (417)$$

Источником неравновесности электронного газа является последнее слагаемое в (417). Неравновесность возможна только при ненулевом потоке электронов I и наличии градиентов величины σ_-/σ_+ . Характерная длина релаксации для меры неравновесности F_- , согласно (417), равняется

$$L_{rel} = \{(\hbar/4\gamma)(\sigma_+^2 - \sigma_-^2)/(\sigma_+(n + n_b))\}^{1/2}.$$

Легко оценить, что

$$L_{rel} \sim L_0(nn_b)^{1/2}/(n + n_b).$$

На Рисунках 104 и 105 демонстрируется характерное поведение F_- в РКТ, потенциальный рельеф которого представлен на Рисунке 103.

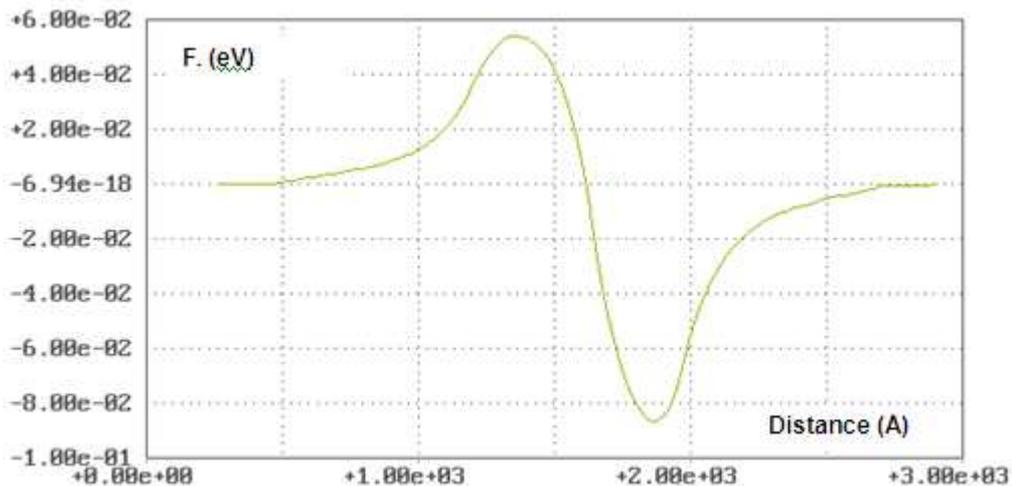


Рисунок 104. Распределение F_- в РКТ при $V_{be} = V_b - V_e = 0,06$ В и $V_{cb} = V_c - V_b = 0,1$ В.

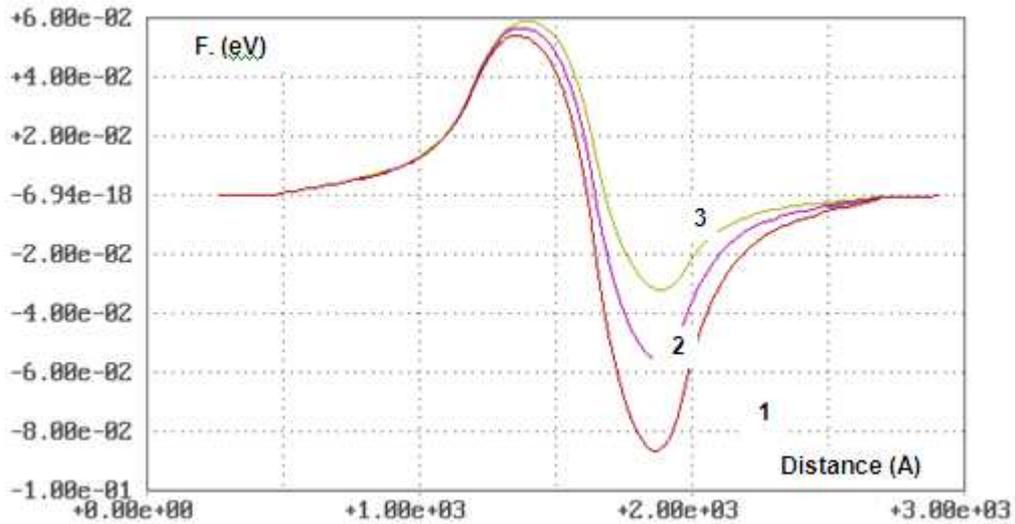


Рисунок 105. Распределение F в РКТ при $V_{be} = V_b - V_e = 0,06$ В и
 1 - $V_{cb} = V_c - V_b = 0,1$ В; 2 - $V_{cb} = V_c - V_b = 0,08$ В; 3 - $V_{cb} = V_c - V_b = 0,05$ В.

Для уравнения (417) с граничными условиями (415) можно построить приближенное решение. Если в коэффициентах уравнения (417) пренебречь зависимостью всех σ и n от x , а принимать во внимание только зависимость от координаты отношения σ_-/σ_+ , то получим

$$F(x) = \int_{x_e}^{x_c} G(x, x') f(x') dx', \quad (418)$$

где

$$G(x, x') = -(2\lambda(1 - \exp\{-2\lambda L\}))^{-1} \begin{cases} \exp\{-\lambda(x' - x)\}(1 - \exp\{-2\lambda(x - x_e)\})(1 - \exp\{-2\lambda(x_c - x')\}), & \text{при } x < x', \\ \exp\{-\lambda(x - x')\}(1 - \exp\{-2\lambda(x' - x_e)\})(1 - \exp\{-2\lambda(x_c - x)\}), & \text{при } x > x', \end{cases}$$

$$f(x) = 2I(d(\sigma_-/\sigma_+)/dx)\sigma_+/(\sigma_+^2 - \sigma_-^2), \quad \lambda = (L_{rel})^{-1}, \quad L = x_c - x_e.$$

Если, учитывая определение $f(x)$, приближенно считать, что

$$f(x) = -2j_e\chi_e\delta(x - x_{eb}) + 2j_c\chi_c\delta(x - x_{bc}),$$

то из (418) найдем

$$F(x) = -2j_e\chi_e G(x, x_{eb}) + 2j_c\chi_c G(x, x_{bc}), \quad (419)$$

где:

$$\chi_{e,c} = 2\sigma_+/(\sigma_+^2 - \sigma_-^2)|_{x=x_{e,c}}.$$

Считая, что

$$L \gg L_{rel}, x_{eb} - x_e \gg L_{rel}, x_c - x_{bc} \gg L_{rel},$$

упростим выражение для F в области $x_{eb} < x < x_{bc}$

$$F(x) = j_e L_{rel} \chi_e \exp\{-\lambda(x-x_{eb})\} - j_c L_{rel} \chi_c \exp\{-\lambda(x_{bc}-x)\}. \quad (420)$$

Поскольку коэффициенты $\chi_{e,c}$ положительны, то, как следует из (420), при $j_{e,c} > 0$ переход эмиттер - база смещен в положительном направлении, а переход база - коллектор - в отрицательном, что, собственно, и демонстрируется на Рисунках 104 и 105. Совершенно аналогично смещены в нормальном режиме функционирования р-п-переходы обычного биполярного транзистора.

Из (420) с учетом (416) нетрудно найти приближенные уравнения

$$eV_{be} = j_e a_{ee} + j_c a_{ec},$$

$$eV_{cb} = j_e a_{ce} + j_c a_{cc},$$

где

$$V_{cb} = V_c - V_b, V_{be} = V_b - V_e,$$

$$a_{ee} = r_e + \chi_e L_{rel}, a_{ec} = -\chi_c L_{rel} \exp\{-\lambda L_b\}, a_{ce} = -\chi_e L_{rel} \exp\{-\lambda L_b\}, a_{cc} = r_c + \chi_c L_{rel}.$$

Решая систему линейных уравнений, получим выражения, описывающие ВАХ прибора

$$\begin{cases} j_e = e(a_{cc} V_{be} - a_{ec} V_{cb}) / (a_{ee} a_{cc} - a_{ce} a_{ec}), \\ j_c = e(a_{ee} V_{cb} - a_{ce} V_{be}) / (a_{ee} a_{cc} - a_{ce} a_{ec}). \end{cases} \quad (421)$$

На Рисунках 106 и 107 показаны полученные в результате численного моделирования ВАХ РКТ при схемах включения с общей базой и общим эмиттером. Сравнение этих характеристик с характеристиками биполярных транзисторов [27] позволяет сделать вывод о функциональной аналогичности этих приборов.

Для нахождения коэффициента усиления α фиксируем напряжение V_{cb} и вычислим при этом условии производную тока коллектора по току эмиттера. В результате получим

$$\alpha = -(a_{ce}/a_{cc})(S_c/S_e) = (S_c/S_e) \chi_e L_{rel} \exp\{-\lambda L_b\} / (r_c + \chi_c L_{rel}). \quad (422)$$

Используя (407), находим β - коэффициент усиления РКТ по току при схеме включения с общим эмиттером

$$\beta = (S_c/S_e) \chi_e L_{rel} \exp\{-\lambda L_b\} / (r_c + \chi_c L_{rel} - (S_c/S_e) \chi_e L_{rel} \exp\{-\lambda L_b\}). \quad (423)$$

Как видим, и α , и β пропорциональны множителю

$$\exp\{-L_b/L_{rel}\}.$$

То есть, как мы и отмечали в самом начале параграфа, в случае слишком длинной базы, когда

$$L_b \gg L_{rel},$$

РКТ перестает функционировать как транзистор и превращается в два включенных навстречу друг другу диода. При длинах базы, меньших длины релаксации неравновесных электронов

$$L_b < L_{rel}, \quad (424)$$

РКТ является транзистором с достаточно хорошими усилительными свойствами, что иллюстрируется на Рисунке 108.

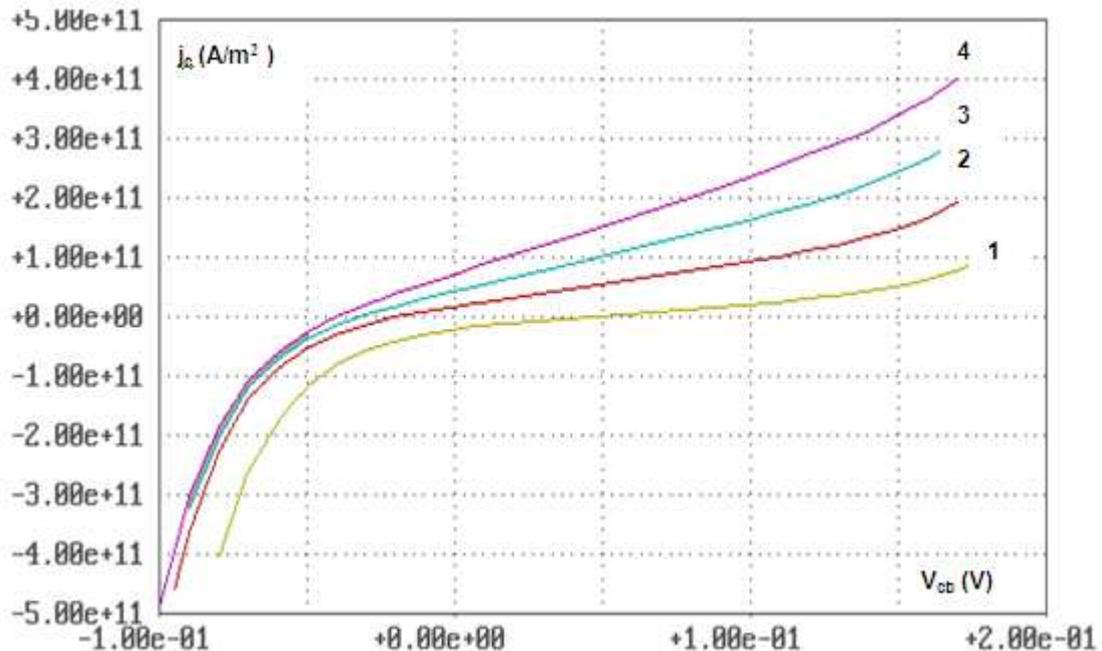


Рисунок 106. ВАХ РКТ при схеме включения с общей базой при: 1 - $V_{eb} = 0,01$ В; 2 - $V_{eb} = -0,025$ В; 3 - $V_{eb} = -0,03$ В; 4 - $V_{eb} = -0,05$ В.

Неравенство (424) является ключевым для работоспособности РКТ. Кроме того, необходимо, чтобы поперечные размеры эмиттерного и коллекторного проводящих каналов были достаточно большими для выполнения условий:

$$n_{e,c} \gg n_b,$$

но в то же время обеспечивали размерное квантование в этих проводящих каналах при выбранных рабочих температурах. Заметим, что РКТ функционирует аналогично n^+pn^+ - транзистору. Для того, чтобы эмиттер прибора инжектировал

достаточный поток неравновесных электронов в базу, необходимо потребовать, чтобы характерный размер перехода между эмиттерным и базовым ПК d_e был достаточно мал

$$d_e \ll L_{rel}.$$

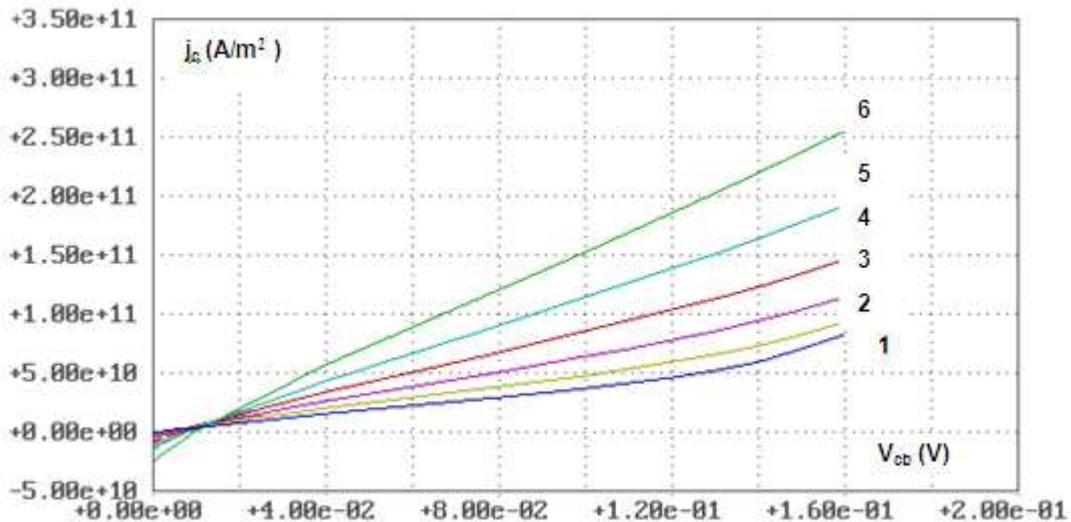


Рисунок 107. ВАХ РКТ при схеме включения с общим эмиттером при: 1 – V_{be} = 0,00 В; 2 – V_{be} = 0,01 В; 3- V_{be} = 0,02 В; 4 – V_{be} = 0,03 В; 5 – V_{be} = 0,04 В; 6 – V_{be} = 0,05 В.

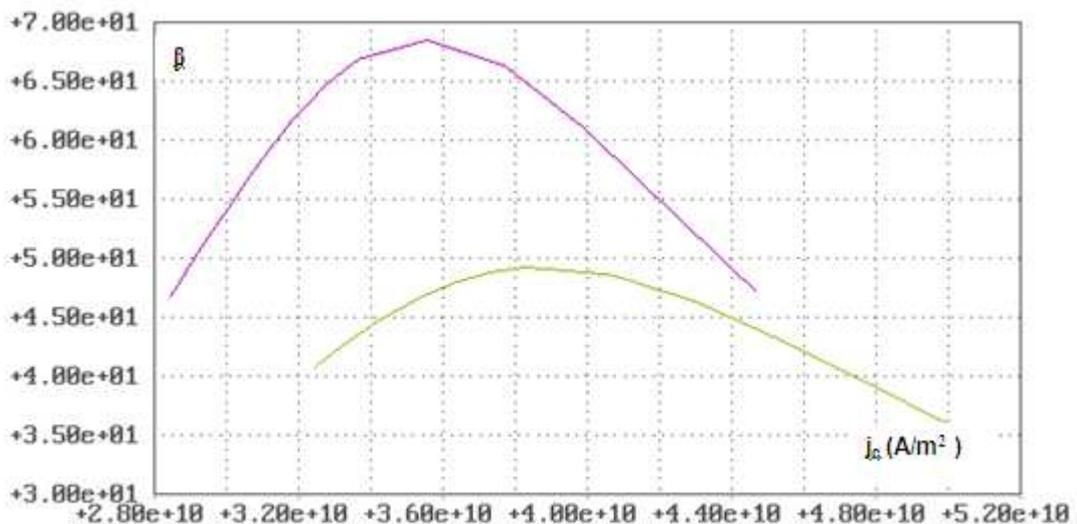


Рисунок 108. Коэффициент усиления РКТ по току при схеме включения с общим эмиттером β при L_b = 80 нм и 1 - V_{ce} = 0,06 В; 2- V_{ce} = 0,07 В.

В соответствии с результатами, представленными на Рисунках 106 и 107, РКТ обеспечивает токи коллектора в единицы и десятки микроампер. Рассматриваемый транзистор также как и РКД способен функционировать при частотах в сотни и тысячи ТГц. В работе [149] представлены расчеты характеристик некоторых схем с использованием РКТ и показано, что их характеристики могут оказаться уникальными.

Заметим, что развитые здесь приближенные методы могут быть успешно применены для анализа работы РТД, РКД и других квантовых приборов. В частности, с их помощью нетрудно получить формулы, описывающие ВАХ РКД, представленные в работе [77].

6.8. Квантовый транзистор с инжекционным затвором

Выше мы анализировали характеристики квантового полевого транзистора. Несмотря на то, что этот прибор представляет интерес для использования в наноэлектронике, некоторые его характеристики, в частности, крутизна, не могут конкурировать с аналогичными характеристиками приборов микроэлектроники. С другой стороны, при рассмотрении эффекта инжекции электронов из контактов в проводящий канал и принципов функционирования релаксационных квантовых диода и транзистора, мы убедились в том, что неравновесные явления играют очень важную роль в физике квантовых приборов и могут быть использованы при их конструировании. В этом параграфе будет рассмотрен прибор на основе квантовых проводов, конструктивно близкий к ПКТ, но в котором для управления используется не полевой, а инжекционный затвор [95].

Основная проблема ПКТ состоит в том, что управление проводимостью канала за счет изменения высоты потенциального барьера, создаваемого затвором, оказывается недостаточно эффективным. Даже характерная для квантовых проводов высокая чувствительность проводимости к параметрам потенциального рельефа не позволяет скомпенсировать потери в крутизне прибора, вызванные малой площадью поперечного сечения ПК. Существует альтернативная возможность модулировать проводимость прибора не за счет вариации параметров потенциального рельефа, а за счет изменения концентрации носителей заряда в канале. Одной из реализаций этой возможности является релаксационный квантовый транзистор, рассмотренный в предыдущем параграфе. Однако в РКТ осуществляется управление только потоком неравновесных носителей заряда и принципиально важным является создание такого потока в

области базы. Конструкция, схематично представленная на Рисунке 109, позволяет организовать обмен электронами между пересекающимися квантовыми проводами и использовать этот обмен для модуляции проводимости одного из них.

Рассматриваемый прибор будем называть квантовым транзистором с инжекционным затвором (КТИЗ). Как показано на Рисунке 109, КТИЗ представляет собой два пересекающихся под прямым углом квантовых провода, проводящие каналы которых обозначены как ПКХ и ПКУ. Для определенности будем считать, что оба провода сфабрикованы из одинакового материала, и поперечное сечение у ПКУ больше, чем у ПКХ. Следовательно, концентрация электронов в ПКУ больше, чем в ПКХ.

На Рисунках 110 и 111 показаны потенциальные рельефы для электронов в квантовых проводах с каналами ПКХ и ПКУ соответственно. Принято, что поперечный размер ПКХ $L = 3L_{dq}/4$, а ПКУ - $L = 2L_{dq}$. На рисунках видно, что область пересечения представляет для электронов, попадающих в прибор из контактов к ПКХ, потенциальную яму, а для электронов, попадающих в прибор из контактов к ПКУ, – потенциальный барьер.

Далее будем называть КП с ПКУ затвором, а КП с ПКХ – каналом рассматриваемого транзистора. Соответствие приложенных напряжений и контактов рассматриваемого прибора отображено на Рисунке 109. КТИЗ является четырехполюсником и допускает различные способы подключения во внешнюю цепь. Однако нас более всего будет интересовать работа прибора в случае, когда напряжение на затворе постоянно

$$V_{sg} = V_{dg} = V_g.$$

В КТИЗ в области пересечения проводящих каналов Ω возможен обмен электронами между ПКХ и ПКУ по механизму, полностью аналогичному инжекции электронов в проводящий канал из контактных областей. Как мы видели, эта инжекция может изменять концентрацию электронов в канале в несколько раз. На Рисунке 112 показаны ВАХ квантовых проводов с ПКХ и ПКУ, рассчитанные без учета электронных переходов в области пересечения, и ВАХ

КП с ПКХ, рассчитанная с учетом электронных переходов в области пересечения при нулевом напряжении на затворе

$$V_g = V_s = 0.$$

Видно, что инжекция электронов из ПКУ увеличивает ток в ПКХ более, чем в два раза.

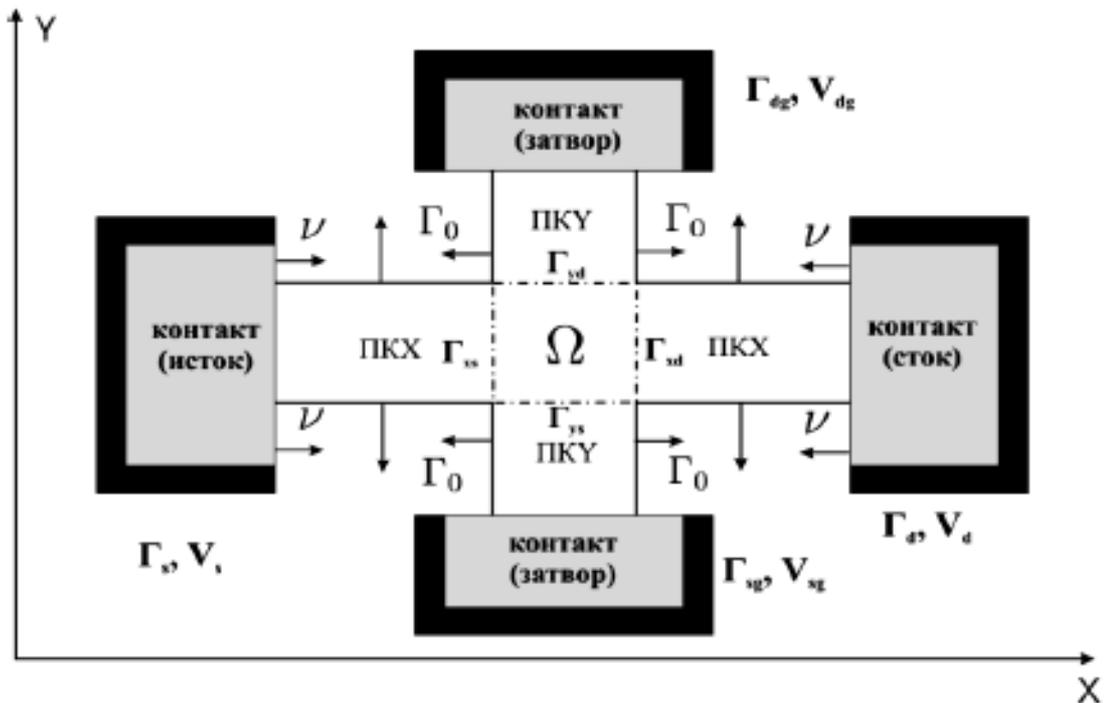


Рисунок 109. Квантовый транзистор с инжекционным затвором.

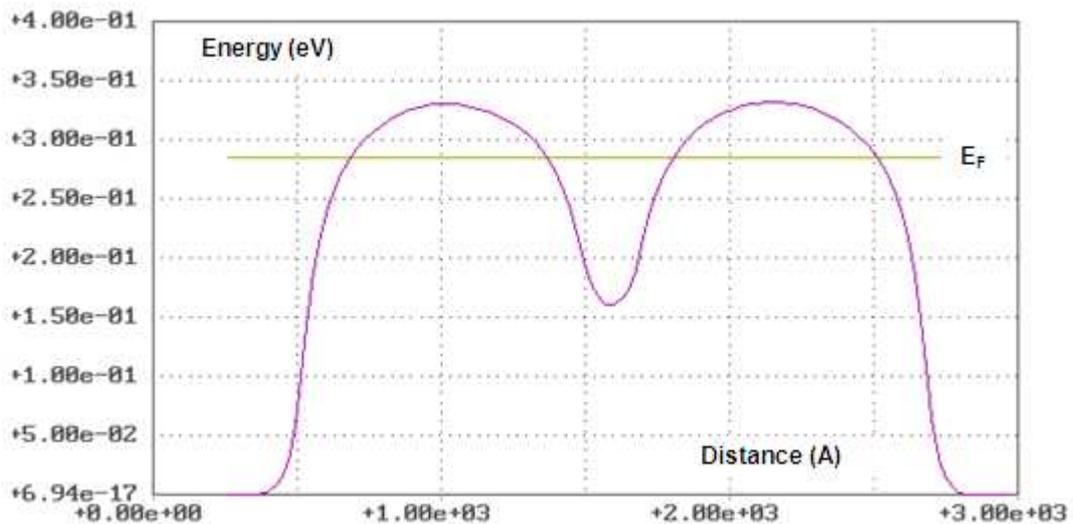


Рисунок 110. Потенциальный рельеф и энергия Ферми КП с ПКХ, имеющем поперечные размеры $L = 3L_{dq}/4$, при нулевом смещении.

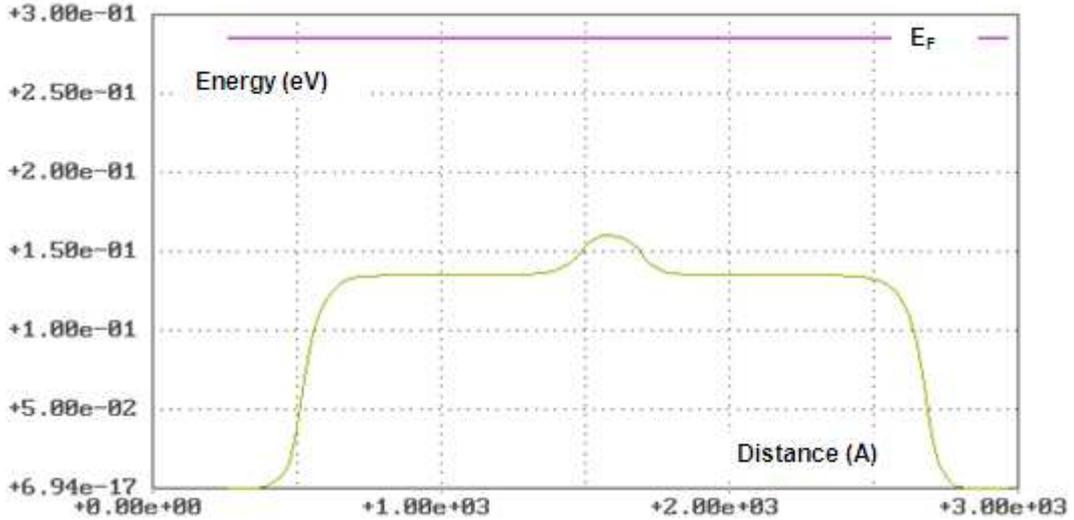


Рисунок 111. Потенциальный рельеф и энергия Ферми ПК с ПКУ, имеющем поперечные размеры $L = 2L_{dq}$, при нулевом смещении.

Модель, позволяющая учесть влияние переходов между электронами различных КП в области их пересечения Ω на перенос заряда в приборе, вполне аналогична использовавшейся ранее для описания явления инжекции электронов из контактов КП в проводящий канал [95]. Если пренебрегать различием химических потенциалов «левых» и «правых» электронов в каждом из проводов и интересоваться только неравновесностью между электронами в ПКХ и ПКУ, то получим уравнения, описывающие поведение химических потенциалов электронов, попадающих в прибор из контактов к этим каналам

$$\begin{cases} \sigma_x^{-1} \partial(\sigma_x \partial F_x / \partial x) / \partial x = \lambda_x^2 (F_x - F_y), \\ \sigma_y^{-1} \partial(\sigma_y \partial F_y / \partial y) / \partial y = \lambda_y^2 (F_y - F_x), \end{cases} \quad (425)$$

где

$$\lambda_x^2 = \hbar^{-1} \sigma_x (n_x + n_y) = L_{rx}^{-2}, \quad \lambda_y^2 = \hbar^{-1} \sigma_y (n_x + n_y) = L_{ry}^{-2},$$

$$\sigma_x = (\partial F_x / \partial x)^{-1} \Sigma_{Ex} j_x f_x, \quad \sigma_y = (\partial F_y / \partial y)^{-1} \Sigma_{Ey} j_y f_y.$$

Индексы x и y относятся к электронам, попадающим в прибор из контактов к ПКХ и ПКУ соответственно. Граничные условия для уравнений (425) выглядят следующим образом

$$\begin{cases} F_{x,y} = V_s \text{ на } \Gamma_s, F_{x,y} = V_d \text{ на } \Gamma_d, \\ F_{x,y} = V_{sg} \text{ на } \Gamma_{sg}, F_{x,y} = V_{dg} \text{ на } \Gamma_{dg}, \\ \partial F_{x,y} / \partial \mathbf{v} = 0 \text{ на } \Gamma_0. \end{cases} \quad (426)$$

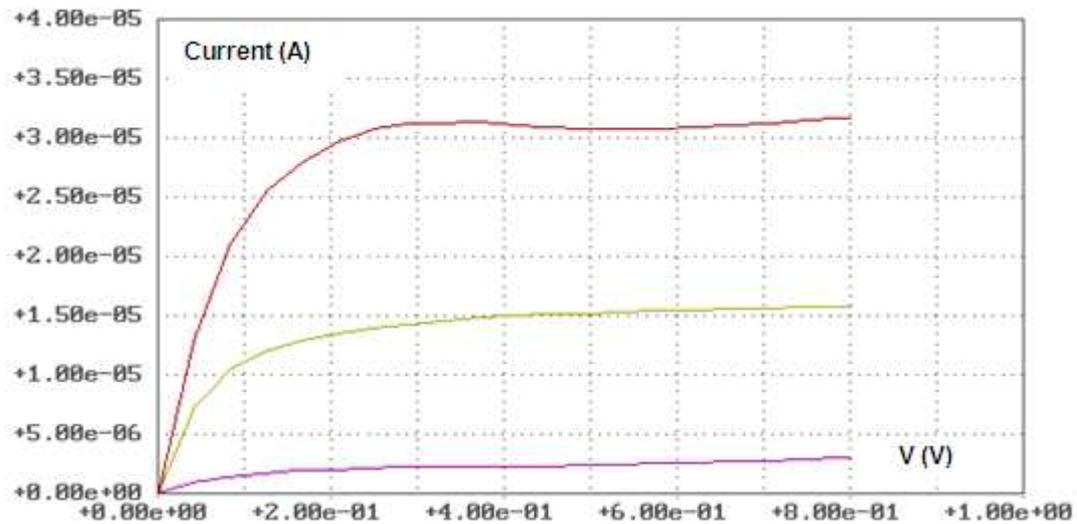


Рисунок 112. ВАХ КП с ПКХ с поперечным размером $L = 3L_{dq}/4$ (нижняя кривая, $V = V_d - V_s$) и ПКУ с поперечным размером $L = 2L_{dq}$ (верхняя кривая, $V = V_{dg} - V_{sg}$), рассчитанные без учета электронных переходов в области пересечения, и ВАХ КП с ПКХ (средняя кривая, $V = V_d - V_s$, $V_g = V_s = 0$), рассчитанная с учетом электронных переходов в области пересечения.

Для расчета статических характеристик КТИЗ задача (425), (426) решалась самосогласованным образом с одномерными уравнениями Пуассона и Шредингера, описывающими эффективный потенциал для электронов и их волновые функции в направлении переноса. На Рисунке 113 показаны ВАХ прибора, рассчитанные при условии

$$V_s = 0$$

и различных постоянных напряжений на затворе. Видно, что в этом режиме рассматриваемый четырехполюсник функционально аналогичен полевому транзистору, однако его характеристики намного лучше, чем у КПТ. В частности, для крутизны справедлива оценка

$$g_m = \partial I_d / \partial V_g \sim 10^{-3} \text{ A/V},$$

что соответствует аналогичным характеристикам обычных полевых транзисторов и на три порядка превосходит крутизну КПТ. Полученное значение g_m является рекордно большим для прибора на квантовых проводках [128,138]. Это свидетельствует об эффективности управления током стока при помощи инжекционного затвора, каковым в КТИЗ является квантовый провод с большей концентрацией носителей заряда.

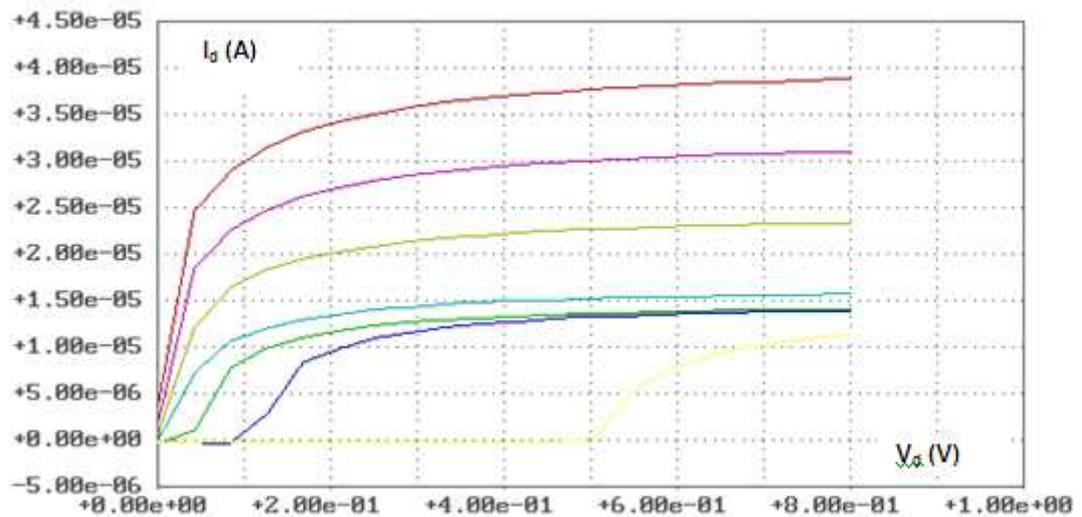


Рисунок 113. Семейство ВАХ КТИЗ при (кривые снизу вверх на рисунке): $V_g = 0,5$ В, $V_g = 0,1$ В, $V_g = 0,03$ В, $V_g = 0$ В, $V_g = -0,01$ В, $V_g = -0,02$ В, $V_g = -0,03$ В.

Важной особенностью КТИЗ является сильная зависимость его характеристик от поперечного размера затвора. Во-первых, им определяется концентрация электронов в КПУ, а значит и ток инжектированных в КПХ электронов. Во-вторых, от концентрации электронов зависят длины релаксации электронов L_{tx} и L_{ty} в области пересечения квантовых проводов Ω . Как мы видели при анализе инжекции электронов из контактов квантового провода в проводящий канал, область локализации этого процесса имеет характерный размер порядка длины релаксации. Точно также интенсивный обмен электронами между КПХ и КПУ происходит в областях с размерами порядка L_{tx} и L_{ty} вблизи границ $\Gamma_{xs,xd}$ и $\Gamma_{ys,yd}$ области их пересечения Ω . Это вполне понятно, так как, фактически, КПУ является контактом для двух отрезков КПХ, расположенных слева и справа от КПУ. Поэтому все полученные нами ранее результаты, описывающие взаимодействие контактов и проводящего канала квантового провода полностью справедливы и для КТИЗ.

Существенно, что длина релаксации L_{tx} уменьшается с ростом концентрации электронов n_y , то есть с увеличением ширины ПКУ. Это означает, что при достаточно широких ПКУ возможна ситуация, когда КТИЗ перестанет функционировать как транзистор и превратится просто в два квантовых провода с общим контактом в виде ПКУ. Для того, чтобы этого не произошло размер

области пересечения L_{cros} проводящих каналов должен быть меньше длины релаксации неравновесных электронов в направлении Ox

$$L_{\text{cros}} < L_{\text{rx}}.$$

В заключение этого параграфа заметим, что расчетная характерная емкость КТИЗ составляет $10^{-18} - 10^{-19}$ Ф, что позволяет надеяться на рабочие частоты вплоть до 100 ТГц.

6.9. Выводы

Проведенный анализ переноса заряда в квантовых проводах, а также принципов функционирования полевого квантового транзистора, релаксационных квантовых диода и транзистора, квантового транзистора с инжекционным затвором показывает, что квантовые провода и приборы на их основе являются перспективными элементами наноэлектроники. Они могут иметь выходные характеристики, аналогичные характеристикам традиционных приборов, используемых в микроэлектронике. Это позволяет сохранить в наноэлектронике наработанные за годы развития микроэлектроники схемотехнические решения. Если описанные здесь приборы удастся устойчиво создавать, то элементная база наноэлектроники будет в целом сформирована.

Расчеты показывают, что в приборах на основе квантовых проводов могут быть достигнуты очень высокие плотности токов (порядка $10^9 - 10^{11}$ А/м²) и достаточно большие токи (от 0,1 мкА до 10 мкА). При этом характерные приложенные напряжения составляют десятые и сотые доли вольт. Таким образом, характерные удельные сопротивления для рассмотренных квантовых приборов расположены в интервале от 10^{-13} Ом*м² до 10^{-10} Ом*м². Для абсолютных значений характерных сопротивлений справедливы величины 10^3 Ом – 10^6 Ом. Характерная удельная емкость активных областей приборов составляет 10^{-3} Ф/м². На основе этих данных простейшие вычисления позволяют оценить предельно достижимые для рассмотренных приборов частоты

$$f_{\text{lim}} \sim 10^{13} \text{ Гц} - 10^{16} \text{ Гц}.$$

То есть даже на нижней границе предельная частота работы приборов на основе квантовых проводов составляет десятки ТГц.

Размеры активных областей изучаемых приборов должны обеспечивать размерное квантование энергии электронов. Они существенно зависят от используемого материала и температуры. При комнатной температуре ($T = 300^{\circ}\text{K}$) для широкого класса возможных материалов наноэлектроники характерный размер активных областей приборов на основе квантовых проводов составляет несколько десятков нанометров, то есть порядка 10^{-8} м. Таким образом, использование указанных приборов в электронных схемах позволяет достигать уровня интеграции, большего, чем 10^{10} элементов на квадратном сантиметре.

Принципиально важное влияние на выходные характеристики квантовых проводов и приборов на их основе оказывают неравновесные эффекты. Использование этих эффектов при конструировании квантовых приборов позволяет добиться требуемых функциональных свойств и высокой управляемости. По-видимому, важнейшей особенностью квантовых приборов является низкая эффективность управления ими при помощи традиционных полевых затворов. Куда более перспективным представляется использование для целей управления явлений инжекции и релаксации электронного газа к состоянию термодинамического равновесия.

Для практического использования этих эффектов необходимо уметь создавать планарные объекты с контролируемыми размерами, меньшими и даже много меньшими длины релаксации неравновесных электронов L_{rel} . Эта длина зависит от используемых материалов, концентрации носителей заряда, концентрации рассеивающих центров разного рода, температуры и т.п. При уменьшении вероятности стохастического рассеяния длина L_{rel} возрастает пропорционально корню квадратному из времени релаксации импульса. Грубо можно считать, что при комнатной температуре для большинства перспективных материалов наноэлектроники L_{rel} , также как и длина размерного квантования L_{dq} , составляет несколько десятков нанометров ($L_{\text{rel}}, L_{\text{dq}} \sim 10^{-8}$ м). Умение создавать

объекты с такими и меньшими размерами является основной задачей нанотехнологии в электронике.

ЗАКЛЮЧЕНИЕ

Основные результаты и выводы.

1. Получены приближенные решения дрейфово-диффузионных уравнений переноса заряда в полупроводниковых структурах в стационарном случае в одномерном и двухмерном приближениях. Они позволяют описать функционирование полупроводниковых приборов без введения условных внутренних границ в областях р-п-переходов, а также явным образом связать ВАХ приборов с уровнем неравновесности электронного и дырочного газов - функцией

$$Y = (np - n_i^2)/n_i^2 = \exp\{(F_n - F_p)/k_B T\} - 1.$$

2. Исследовано влияние отклонения электронного и дырочного газов от состояния локального химического равновесия на ВАХ биполярного диода, биполярного транзистора и диодного тиристора. Показано, что функционирование этих приборов основано на неравновесных эффектах. Причиной возникновения неравновесности является протекание тока в направлениях, перпендикулярных р-п-переходам.

3. Разработана многофазная модель переноса заряда в мезоскопических структурах. В ней электронный газ в твердом теле представляется в виде смеси компонент, соответствующих допустимым состояниям электронов. Компоненты названы фазами электронного газа. Каждая фаза характеризуется своим химическим потенциалом. Модель позволяет в рамках единого формализма описывать квантовые и классические явления переноса заряда и, соответственно, квантовые и классические области электронных приборов.

4. Показано, что для полупроводниковых материалов характерная длина релаксации химических потенциалов электронных фаз к состоянию локального химического равновесия является величиной одного порядка с длиной размерного квантования. Это означает, что квантовые и неравновесные явления одинаково значимы для функционирования твердотельных электронных приборов с нанометровыми размерами активных областей.

5. Показано, что выражение для плотности тока многофазной модели учитывает как квантовую составляющую, обусловленную различием химических потенциалов разных фаз электронного газа, так и классическую дрейфово-диффузионную составляющую, обусловленную пространственными градиентами химических потенциалов фаз электронного газа.

6. На основе общей многофазной модели построена двухфазная самосогласованная модель переноса заряда в резонансно-туннельном диоде. Модель позволяет добиться хорошего соответствия между расчетными и измеренными ВАХ прибора, что делает возможным ее применение для оптимизации его конструкции.

7. Показано, что наличие ступеньки на участке ВАХ РТД с отрицательной дифференциальной проводимостью может быть объяснено нарушением локального химического равновесия между «левыми» и «правыми» электронами, попадающими в прибор из различных контактов, и соответствующим накоплением неравновесных электронов в области потенциальной ямы.

8. Показано, что наблюдаемые экспериментально большие значения тока долины РТД обусловлены наличием дрейфово-диффузионного тока, который с ростом приложенного напряжения выше некоторого значения начинает превышать туннельный ток.

9. Предложена модель, описывающая взаимодействие квантовой резонансно-туннельной структуры с классической подложкой. Показано, что влияние классических контактных областей на статические характеристики резонансно-туннельного диода не сводится к тривиальному эффекту нагрузочного сопротивления. Растекание тока в подложке прибора существенно искажает его выходные характеристики, меняет форму ВАХ и численные значения основных параметров ВАХ. При определенном сочетании параметров резонансно-туннельной структуры и подложки стационарные уравнения переноса заряда в подложке РТД могут иметь неединственное решение.

10. Построена двухфазная самосогласованная модель переноса заряда в квантовом проводе. Проведен анализ функционирования этого простейшего квантового прибора. Показано, что измеряемые экспериментально статические ВАХ квантового провода, удастся объяснить, если учесть неравновесные эффекты:

- инжекцию электронов из контактных областей в проводящий канал;
- нарушение локального химического равновесия между фазами «левых» и «правых» электронов, попадающих в проводящий канал из эмиттерного и коллекторного контактов, соответственно.

Эти явления могут приводить к многократному увеличению проводимости квантового провода по сравнению с собственной проводимостью.

11. Показано, что наблюдаемое экспериментально изменение температурной зависимости сопротивления металлических нанопроводов при уменьшении их поперечного сечения может быть объяснено соответствующим изменением превалирующего механизма заполнения проводящего канала носителями заряда. Для «толстых» проводов перенос заряда в канале осуществляется главным образом собственными электронами, а для «тонких» - инжектированными из эмиттерного контакта. Для «толстых» проводов температурная зависимость сопротивления определяется температурной зависимостью времени релаксации импульса, а для «тонких» - экспоненциальным ростом концентрации инжектированных электронов с ростом температуры.

12. Показано, что инжекция электронов из эмиттерного контакта в проводящий канал квантового провода может приводить к охлаждению эмиттерного контакта. Эффект может достигать десятков градусов. На его основе могут быть созданы нанохолодильники («тепловые насосы»). Температуру охлаждения можно регулировать приложенным напряжением, то есть строить электрические схемы, управляющие тепловым режимом [140,148,150,151].

13. Предложена конструкция и рассчитаны характеристики релаксационного квантового диода, являющегося функциональным аналогом биполярного диода. Функционирование прибора основано на неравновесных

эффектах, возникающих при протекании тока через n^+ - n -переход между «толстым» и «тонким» квантовым проводом.

14. Предложена конструкция и рассчитаны характеристики релаксационного квантового транзистора, являющегося функциональным аналогом биполярного транзистора. Функционирование прибора основано на использовании для усиления внешнего сигнала неравновесных электронов, инжектируемых в область базы из эмиттерного квантового провода. Для работы прибора в качестве транзистора необходимо, чтобы длина базового квантового провода была меньше длины релаксации неравновесных электронов в нем. Оценки показывают, что для квантовых проводов из InSb при комнатной температуре эта длина порядка 100 нм.

15. Для использования в цифровой наноэлектронике [152] предложен квантовый транзистор с инжекционным затвором. Его расчетные электрические характеристики аналогичны характеристикам полевого транзистора на квантовом проводе [138]. Однако управление током проводящего канала осуществляется не полевым затвором, а за счет инжекции в канал электронов из нанопровода большего, чем канал поперечного размера. Согласно проведенным расчетам, использование неравновесного эффекта инжекции должно привести к увеличению крутизны прибора на три порядка (до 10 мСм) по сравнению с полевым транзистором на квантовом проводе.

16. В диссертации показано, что целенаправленное использование неравновесных эффектов позволяет разрабатывать квантовые приборы с различными функциональными свойствами. Предложенные приборы на основе квантовых проводов являются функциональными аналогами хорошо известных элементов микроэлектроники. Их создание и применение в электронных схемах позволило бы использовать стандартные схемотехнические решения [149]. Проведенные оценки показывают, что предельные частоты рассмотренных приборов могут достигать 10^3 ТГц.

17. В мезоскопических структурах источники неравновесности электронного газа локализуются в малых окрестностях неоднородностей

потенциального рельефа, порождаемых гетерограницами и градиентами концентраций. Это позволяет сделать вывод, что наноэлектронику целесообразно строить как электронику границ и неравновесных пограничных эффектов. Наиболее перспективным материалом для такой наноэлектроники является InSb [69,118,119,152].

Перспективы дальнейших исследований.

В развитие полученных в диссертации результатов в настоящее время проводятся экспериментальные и теоретические исследования. Научной группе профессора Г.Г. Гороха, работающей в Белорусском государственном университете информатики и радиоэлектроники, удалось сформировать матрицы нанопроводов из InSb [141] в порах анодированного оксида алюминия. Проведены предварительные измерения электрических [141] и фотоэлектрических характеристик этих объектов [102]. Результаты позволяют надеяться на успешное применение матриц нанопроводов в устройствах преобразования энергии [142].

Аналогичные нанопровода были использованы американскими исследователями (Purdue University, West Lafayette, Indiana, USA, 2011 г.) при создании полевого транзистора [138], функционирующего при комнатной температуре.

В контексте освоения терагерцового диапазона длин волн [4] представляют интерес частотные свойства квантовых проводов и приборов на их основе. Они исследуются в работах [76,98,153,154]. В работе [154] рассмотрены условия возникновения релаксационной неустойчивости и отрицательного дифференциального сопротивления в квантовом проводе. Если их удастся обнаружить экспериментально, то представится возможность создания нового типа твердотельных генераторов терагерцового диапазона.

По мнению автора, наибольший интерес для наноэлектроники представляют работы по созданию планарных приборов на основе квантовых проводов [145], которые в настоящее время практически нигде не ведутся. При этом имеются все технологические возможности для таких исследований [6,7]. Применение InSb в

качестве исходного материала позволяет надеяться на создание элементной базы нанoeлектроники, функционирующей при комнатной температуре.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ

1. Долманов И.Н. Полупроводниковые приборы с резонансным туннелированием электронов / И.Н. Долманов, В.И. Толстихин, В.Г. Еленский // Зарубежная радиоэлектроника. – 1990. – № 7. – С. 66 – 89.
2. M.C. Rocco Nanotechnology Research Directions: IWGN Workshop Report. Vision for Nanotechnology R&D in the Next Decade / Edited by M.C. Rocco, R.S. Williams, P. Alivisatos // Kluwer Academic Publishers, - Dordrecht/Boston/London. – 2000. – 318 P. (Имеется перевод: Нанотехнология в ближайшем десятилетии. Пер. с англ. А.В. Хачоян // Ред. пер. с англ.: Р.А. Андриевский. - М.: Мир. – 2002. – 292 С.)
3. Frontiers in Nanoscale Science of Micron / Submicron Devices, Edited by Antti-Pekka Jauho and Eugenia V. Buzaneva // NATO ASI Series. Series E: Applied Sciences-Vol. 328. Kluwer Academic Publishers. – 1995. – 554 P.
4. Майская В. Освоение терагерцовой щели. Полупроводниковые приборы вторгаются в субмиллиметровый диапазон / В. Майская // Электроника: Наука, технология, бизнес. – 2011. – № 8. – С. 74 – 87.
5. Гаврилов О.Т. Основные ограничения макроскопического рассмотрения процессов в полупроводниковых приборах / О.Т. Гаврилов, Ю.С. Тиходеев // Электронная техника. Сер. 2. Полупроводниковые приборы. – 1973. – Вып. 1(73). – С. 61 – 71.
6. Landis S. Nano-lithography / S. Landis // London UK: ISTE Hoboken. NJ.: Wiley. – 2011.
7. Seisyan R.P. Nanolithography in Microelectronics / R.P. Seisyan // J. Tech. Phys. – 2011. - V. 81. no. 8. – P. 1 – 14.
8. Востоков Н.В. Создание наноразмерных элементов методами атомно-силовой литографии / Н.В. Востоков, Д.Г. Волгунов, В.Ф. Дряхлушин // Материалы 9-й Крымской Микроволновой Конференции. Севастополь. – 1999. – С. 3 – 4.

9. Дряхлушин В.Ф. Метод сканирующей ближнепольной оптической литографии / В.Ф. Дряхлушин, Н.В. Востоков, А.Ю. Климов, В.В. Рогов, В.И. Шашкин // *Материалы 12-й Международной Крымской Микроволновой Конференции*. Севастополь. – 2002. – С. 453 – 454.
10. Молекулярно-лучевая эпитаксия и гетероструктуры / Под редакцией Л. Ченга и К. Плога. // М.: Мир. – 1989. – 582 С.
11. Херман М. Полупроводниковые сверхрешетки. / М. Херман // М.: Мир. – 1989. – 240 С.
12. Иогансен Л. В. О возможности резонансного прохождения электронов в кристаллах через системы барьеров / Л.В. Иогансен // *ЖЭТФ*. – 1963. – Т. 45. – В. 2. – С. 207 – 213.
13. Тиходеев Ю.С. О возможности создания СВЧ твердотельных приборов, основанных на квантовых размерных эффектах / Ю.С. Тиходеев // *Электронная техника. Сер. 2. Полупроводниковые приборы*. – 1973. – В. 1 – С. 3 – 34.
14. Tsu R. Tunneling in a finite superlattice / R. Tsu, L. Esaki // *Appl. Phys. Lett.* – 1973. – V. 22. – N11. – P. 562 – 564.
15. Solner T.C.L.G. Resonant tunneling through quantum wells at frequencies up to 2.5 THz. / T.C.L.G. Solner, W.D. Goodhue, P.E. Tannenwald et al. // *Appl. Phys. Lett.* – 1983. – V. 43. – P. 588.
16. Luryi S. Frequency limit of double-barrier resonant-tunneling oscillators / S. Luryi // *Appl. Phys. Lett.* – 1985. – V. 47. – P. 490–492.
17. Шур М. Современные приборы на основе арсенида галлия /М. Шур // М.: Мир. – 1991. – 632 С.
18. Орликовский А.А. Кремниевая транзисторная наноэлектроника / А.А. Орликовский // *Известия вузов. Электроника*. – 2006. – № 5. – С. 35 – 44.
19. Лихарев К.К. О возможности создания аналоговых и цифровых интегральных схем на основе эффекта дискретного одноэлектронного туннелирования / К.К. Лихарев // *Микроэлектроника*. – 1987. – Т. 16. – В. 3. – С. 195 – 209.

20. Аверин Д.В. Когерентные колебания в туннельных переходах малых размеров / Д.В. Аверин, К.К. Лихарев // ЖЭТФ. – 1986. – Т. 90. – В. 2. – С. 733 – 743.
21. Кузьмин Л.С. Непосредственное экспериментальное наблюдение дискретного коррелированного одноэлектронного туннелирования / Л.С. Кузьмин, К.К. Лихарев // Письма в ЖЭТФ. – 1987. – Т. 45. – В. 8. – С. 389 – 390.
22. Likharev K.K. Correlated discrete transfer of single electrons in ultrasmall tunnel junctions / К.К. Likharev // IBM J. Res. Develop. – 1988. – V. 32. – N 1 – P. 144 – 158.
23. Datta S. Quantum devices. Superlattices and Microstructures / S. Datta // 1989. – V. 6. – N 1. – P. 83 – 93.
24. Килин С. Я. Квантовая информация / С. Я. Килин // Успехи Физических Наук. – 1999. – Т. 169. – С. 507—527.
25. Валиев К. А. Квантовые компьютеры: надежды и реальность / К.А. Валиев, А.А. Кокин // Ижевск: РХД. – 2004. – 320 С.
26. Бонч-Бруевич В.Л. Физика полупроводников / В.Л. Бонч-Бруевич, С.Г. Калашников // М.: «Наука». – 1990. – 688 С.
27. Зи С. Физика полупроводниковых приборов / С. Зи // М.: «Мир». – 1984. – Т. 1. – 456 С.
28. Marcowich P.A. The stationary semiconductor device equation / P.A. Marcowich // Vienna. N.Y.: Springer. – 1986.
29. Энгль В.Л. Моделирование полупроводниковых приборов / В.Л. Энгль, Х.К. Диркс, Б. Майнерцхаген // Труды Института инженеров по электротехнике и радиотехнике (ТИИЭР). – 1983. – Т. 71. – № 1. – С. 14 – 42.
30. Абрамов И.И. Моделирование физических процессов в элементах кремниевых интегральных микросхем / И.И. Абрамов // Минск. БГУ. – 1999. – 190 С.

31. Абрамов И.И. Численное моделирование элементов интегральных схем / И.И. Абрамов, В.В. Харитонов // Минск. «Вышэйшая школа». – 1990. – 224 С.
32. Мулярчик С.Г. Численное моделирование микроэлектронных структур / С.Г. Мулярчик // Минск. Издательство «Университетское». – 1989. – 368 С.
33. Стемповский А.Л. Системная среда САПР СБИС / А.Л. Стемповский, В.А Шепелев, В.А Власов. // М.: «Наука». – 1994. – 251С.
34. Киреев В. Технологии микроэлектроники. Химическое осаждение из газовой фазы / В. Киреев, А. Столяров // М.: Техносфера. – 2006. – 192 С.
35. Климонтович Ю.Л. Статистическая физика / Ю.Л. Климонтович // М.: «Наука». – 1982. – 608 С.
36. Владимиров В.В. Плазма полупроводников / В.В. Владимиров, А.Ф. Волков, Е.З. Мейлихов // М.: «Атомиздат». – 1979. – 260 С.
37. Бонч-Бруевич В.Л. Доменная электрическая неустойчивость в полупроводниках / В.Л. Бонч-Бруевич, И.П. Звягин, А.Г. Миронов // М.: «Наука». – 1972. – 414 С.
38. Мокшин А.Н. Приближенное решение уравнений переноса заряда в полупроводниковых структурах / А.Н. Мокшин, И.А. Обухов, В.К. Перес-Фернандес, С.В. Румянцев // Микроэлектроника. – 1988. – Т. 17. – В. 3. – С. 225 – 230.
39. Тамм И.Е. Основы теории электричества / И.Е. Тамм // М.: «Наука». – 1989. – 504 С.
40. Ландау Л.Д. Электродинамика сплошных сред. Теоретическая физика в 10 томах / Л.Д. Ландау, Е.М. Лифшиц // М.: «Наука». – 1982. – Т. 8. – 620 С.
41. Обухов И.А. Влияние внешнего электрического поля на перенос заряда в полупроводниках / И.А. Обухов, С.В. Крючков, М.С. Пантелеев, С.В. Румянцев // Электронная техника. Сер. 2. Полупроводниковые приборы. – 1990. – В. 5(208). – С. 31 – 42.

42. Тихонов А.Н. Уравнения математической физики / А.Н. Тихонов, А.А. Самарский // Издательство Московского Университета. – 1977. – 742 С.
43. Константинов О.В. Рекомбинационные волны в полупроводниках / О.В. Константинов, В.И. Перель // ФТТ. – 1964. – Т. 6. – № 11. – С. 3364.
44. Васильева А.Б. Асимптотические разложения решений сингулярно возмущенных уравнений / А.Б. Васильева, В.Ф. Бутузов // М.: «Наука». – 1973. – 272 С.
45. Vasil'eva A.B. The Boundary Function Method for Singular Perturbation Problems / A.B. Vasil'eva, V.F. Butuzov and L.V Kalachev // SIAM Studies in Appl. Math., SIAM, Philadelphia. – 1995.
46. Ломов С.А. Введение в общую теорию сингулярных возмущений / С.А. Ломов // М.: «Наука». – 1981. – 398 С.
47. Крючков С.В. Математическое моделирование полупроводниковых структур с S-образными ВАХ / С.В. Крючков, И.А. Обухов, С.В. Румянцев // Электронная техника. Сер. 3. Микроэлектроника. – 1989. – В. 5 (134). – С. 25 – 28.
48. Васильева А.Б. Сингулярно возмущенные системы теории полупроводниковых приборов / А.Б. Васильева, В.Г. Стельмах // ЖВМ и МФ. – 1977. – Т. 17. – № 2. – С. 339 – 348.
49. Калачев Л.В. Асимптотический анализ решения уравнения Пуассона в полупроводниках / Л.В. Калачев, С.В. Крючков, И.А. Обухов // Математическое моделирование. – 1989. – Т. 1. – № 9. – С. 129 – 140.
50. Калачев Л.В. Асимптотика решения уравнения Пуассона в модели трехмерной полупроводниковой структуры / Л.В. Калачев, И.А. Обухов // Журнал вычислительной математики и математической физики. – 1992. – Т. 31. – № 9. – С. 1509 – 1514.
51. Калачев Л.В. Асимптотика решения уравнения Пуассона для полупроводниковой структуры во внешнем поле / Л.В. Калачев, И.А. Обухов // Математическое моделирование. – 1991. – Т. 3. – № 12. – С. 78 – 85.

52. Белянин М.П. Об асимптотическом решении одной модели (p-n)-перехода / М.П. Белянин // Журнал вычислительной математики и математической физики. – 1986. – Т. 26. – № 2. – С. 306 – 311.
53. Калачев Л.В. Приближенное решение уравнения Пуассона в модели двумерной полупроводниковой структуры / Л.В. Калачев, И.А. Обухов // Вестник Моск. Ун-та. Сер. 3. Физика. Астрономия. – 1989. – Т. 30. – № 3. – С. 63 – 68.
54. Мокшин А.Н. Приближенное решение дрейфово-диффузионных уравнений переноса заряда в двумерной области / А.Н. Мокшин, И.А. Обухов, С.В. Румянцев // Электронная техника. Сер. 3. Микроэлектроника. – 1988. – В. 2(126). – С. 24 – 27.
55. Лаврентьев М.А. Методы теории функций комплексного переменного / М.А. Лаврентьев, Б.В. Шабат // М.: «Наука». – 1987. – 688 С.
56. Ленг С. Эллиптические функции / С. Ленг // М.: «Наука». – 1984. – 312 С.
57. Parrot J.E. Reformulation of basis semiconductor transport equations / J.E. Parrot // II. Solid State Electronics. – 1974. – V. 17. – P. 707 – 716.
58. Волков В.Т. Численно-асимптотический анализ переходных процессов в полупроводниках / В.Т. Волков, С.В. Крючков, И.А. Обухов, С.В. Румянцев // Журнал вычислительной математики и математической физики. – 1989. – Т. 29. – № 8. – С. 1159 – 1167.
59. Маслов В.П. Операторные методы / В.П. Маслов // М.: «Наука». – 1973. – 544 С.
60. Соколов А.А. Квантовая механика / А.А. Соколов, И.М. Тернов, В.Ч. Жуковский // М.: «Наука». – 1979. – 528 С.
61. Шокли В. Теория электрических полупроводников / В. Шокли // М.: Издательство иностранной литературы. – 1953. – 714 С.
62. Блихер А. Физика тиристоров / А. Блихер // Энергоиздат. Ленинградское отделение. – 1981. – 261С.
63. Герлах В. Тиристоры / В. Герлах // М.: Энергоатомиздат. – 1985. – 328 С.

64. Гаряинов С.А. Физические модели полупроводниковых приборов с отрицательным сопротивлением / С.А. Гаряинов, Ю.С. Тиходеев // М.: «Радио и связь». – 1997. – 276 С.
65. Landauer R. Spatial variation of currents and fields due to localized scatterers in metallic conduction / R. Landauer // IBM J. Res. Develop. – 1988. – V. 32. – N 3. – P. 306 – 316.
66. Datta S. Steady-state quantum kinetic equation / S. Datta // Phys. Rev. B. – 1989. – V. 40. – N 8. – P. 5830 – 5833.
67. Kluksdahl N.C. Self-consistent study of the resonant-tunneling diode / N.C. Kluksdahl, A.M. Krivan, D.K. Ferry, C. Ringhofer // Phys. Rev. B. – 1989. – V. 39. – N 11. – P. 7720 – 7735.
68. Mc Lennan M.J. Voltage drop in mesoscopic systems: A numerical study using a quantum kinetic equation / M. J. Mc Lennan, Y. Lee, S. Datta // Phys. Rev. B. – 1991. – V. 43. – N 17. – P. 13846 – 13884.
69. Datta S. Electronic Transport in Mesoscopic Systems / S. Datta // Cambridge University Press. UK. – 1997.
70. Абрамов И.И. Проблемы и принципы физики и моделирования приборных структур микро- и нанoeлектроники. VII. Структуры на квантовых проволоках / И.И. Абрамов // Нано и микросистемная техника. – 2009. – № 8. – С. 7 – 23.
71. Обухов И.А. Моделирование переноса заряда в мезоскопических структурах / И.А. Обухов // Изд-во «Вебер». Москва–Киeв–Минск–Севастополь. – 2005. – 226 С.
72. Обухов И.А. Проблемы моделирования электрических характеристик квантовых приборов / И.А. Обухов // Материалы 15-й Международной Крымской Микроволновой Конференции. Севастополь. – 2005. – С. 607 – 608.
73. Mizuta H. Transient quantum transport simulation based on the statistical density matrix / H. Mizuta, C.J. Goodings // J. Phys.: Condens. Matter. -1991. – V. 3. - P. 3739 - 3756.

74. Тагер А.С. Размерные квантовые эффекты в субмикронных полупроводниковых структурах и перспектива их применения в электронике СВЧ. Ч.1. Физические основы / А.С. Тагер // Электронная техника. Сер. Электроника СВЧ. - 1987. - Вып. 9. - С. 21 - 34.
75. Тавгер Б.Л. Квантовые размерные эффекты в полупроводниковых и полуметаллических пленках / Б.Л. Тавгер, В.Я. Демиховский // УФН. – 1968. – Т. 96. – В. 1. – С. 61 – 86.
76. Пашковский А.Б. Расщепление резонансных уровней двухбарьерных структур в сильном высокочастотном электрическом поле / А.Б. Пашковский // Письма в ЖЭТФ. – 2011. – Т. 93. – В. 10. – С. 620–624.
77. Обухов И.А. О возможности создания релаксационных квантовых приборов / И.А. Обухов // Письма в ЖТФ. – 1993. – Т. 19. – В. 17. – С. 12 – 16.
78. Обухов И.А. Приборы на основе квантовых проводов: перспективы и проблемы / И.А. Обухов // Материалы 6-й Международной Крымской Конференции «СВЧ -техника и телекоммуникационные технологии». Севастополь. Weber Co. – 1997. – С. 55 – 64.
79. Obukhov I.A. Electronic devices based on quantum wires: prospects and problems / I. A. Obukhov // Physics, Chemistry and Application of Nanostructures. Review and Short Notes of Nanomeeting'97 (Ed. by V.E. Borisenko, A.B. Filonov, S.V. Gaponenko, V.S. Gurin). Minsk. Belarus. – 1997. – P. 322 – 325.
80. Обухов И.А. Релаксационные квантовые приборы / И.А. Обухов // Материалы 7-й Международной Крымской Конференции «СВЧ - техника и телекоммуникационные технологии». Севастополь. Weber Co. – 1997. – С. 386 – 389.
81. Obukhov I.A. The possible application of SPM for creation of new types of quantum devices / I.A. Obukhov // Scanning Probe Microscopy – 2002 Proceedings. Nizhny Novgorod. – 2002. – P. 85 – 87.
82. Обухов И.А. О возможности применения СТМ - АСМ литографии для создания новых типов квантовых приборов / И.А. Обухов // Микросистемная техника. – 2003. – № 6. – С. 34 – 37.

83. Дианов С.А. Моделирование резонансно-туннельного диода / С.А. Дианов, Ф.Г. Искандеров, И.А. Обухов, Ю.С. Тиходеев // Электронная техника. Сер. 2. Полупроводниковые приборы. – 1991. – В. 3 (212). – С. 10 – 15.
84. Gavrilov O.T. The classical regions influence on characteristics of resonant - tunneling diode / O.T Gavrilov, S.A. Dianov, I.A. Obukhov, J.A. Matveev // International Symposium "Nanostructures: Physics and Technology". Abstracts of Invited Lectures and Contributed Papers. St. Petersburg. – 1993. – P. 56.
85. Gavrilov O.T. The classical regions influence on characteristics of resonant-tunneling diode / O.T. Gavrilov, I.A. Obukhov, S.A. Dianov // Proceeding of the Second International Conference on Nanometer Scale Science and Technology (NANO-II). Moscow. – 1993. – P. 860 – 870.
86. Obukhov I.A. Resonant - tunneling diode theory and modeling / I.A. Obukhov // Proceeding of Nanomeeting-95, Physics, Chemistry and Application of Nanostructures. Belarusian State University of Informatics and Radioelectronics. Minsk. – 1995. – P. 271 – 272.
87. Гаврилов О.Т. Флуктуационно-релаксационная неустойчивость и генерация белого шума в резонансно-туннельной структуре / О.Т. Гаврилов, И.И. Квяткевич, И.А. Обухов // Материалы 6-й Международной Крымской Конференции «СВЧ -техника и телекоммуникационные технологии». 16-19 сентября 1996 г. Севастополь. Weber Co. – 1997. – С. 308 – 311.
88. Гаврилов О.Т. Взаимодействие резонансно-туннельного диода с подложкой через распределенную границу / О.Т. Гаврилов, И.И. Квяткевич, И.А. Обухов, Ю.А. Матвеев // Письма в ЖТФ. – 1996. – Т. 22. – В.8. – С. 18 – 21.
89. Гаврилов О.Т. Влияние классических областей на характеристики резонансно-туннельного диода / О.Т. Гаврилов, И. И. Квяткевич, И.А. Обухов // Материалы 7-й Международной Крымской Конференции «СВЧ -техника и телекоммуникационные технологии». Севастополь. Weber Co. – 1997. – С. 401 – 403.

90. Обухов И. А. Моделирование статических характеристик резонансно-туннельных структур / И.А. Обухов // Материалы 10-й Международной Крымской Микроволновой Конференции. Севастополь. – 2000. – С. 423 – 427.
91. Обухов И. А. Моделирование статических характеристик резонансно-туннельных приборов / И.А. Обухов // Микросистемная техника. – 2001. – № 2. – С. 23 – 28.
92. Гаврилов О.Т. Двумерные эффекты взаимодействия резонансно-туннельных структур с подложкой / О.Т. Гаврилов, И.И. Квяткевич, И.А. Обухов // Материалы 13-й Международной Крымской Микроволновой Конференции. Севастополь. – 2003. – С. 536 – 539.
93. Obukhov I.A. Simulation of characteristics of nanoelectronics devices based on quantum wires / I.A. Obukhov, S.A. Dianov // Proceeding of the Second International Conference on Nanometer Scale Science and Technology (NANO-II). Moscow. – 1993. – P. 775 – 790.
94. Квяткевич И.И. Моделирование полевого транзистора на основе квантового провода / И.И. Квяткевич, И.А. Обухов, М.С. Чекадин // Материалы 12-й Международной Крымской Микроволновой Конференции. Севастополь. 2002. – С. 455 – 457.
95. Обухов И.А. Статические характеристики пересекающихся квантовых проводов / И.А. Обухов, И.И. Квяткевич, А.А. Лавренчук, С.В. Румянцев // Материалы 14-й Международной Крымской Микроволновой Конференции. Севастополь. – 2004. – С. 507 – 511.
96. Обухов И.А. Влияние термо-ЭДС на ВАХ квантового провода / И.А. Обухов, А.А. Лавренчук // Материалы 15-й Международной Крымской Микроволновой Конференции. Севастополь. – 2005. – С. 609 – 612.
97. Обухов И.А. Влияние контактных областей на электрические характеристики квантового провода / И.А. Обухов, И.И. Квяткевич, А.А. Лавренчук // Материалы 16-й Международной Крымской Микроволновой Конференции. Севастополь. – 2006. – С. 650 – 654.

98. Obukhov I.A. About the Frequency-Dependence of Electrical Characteristics of Quantum Devices / I.A. Obukhov // *Nanoscience and Nanotechnology*. – 2012. – V. 2. – N 5. – P. 144 – 147.
99. Обухов И.А. О температурной зависимости сопротивления металлических квантовых проводов / И.А. Обухов // *Нано- и микросистемная техника*. – 2006. – № 6. – С. 33 – 36.
100. Обухов И.А. Эффект охлаждения эмиттерного контакта квантового провода / И.А. Обухов // *Материалы 16-й Международной Крымской Микроволновой Конференции*. Севастополь. – 2006. – С. 655 – 656.
101. Обухов И.А. Самоохлаждение эмиттерного контакта квантового провода / И.А. Обухов // *Нано- и микросистемная техника*. – 2007. – № 5. – С. 21 – 23.
102. Обухов И.А. Влияние внешнего электромагнитного излучения на ВАХ квантового провода / И.А. Обухов // *Материалы 20-й Международной Крымской Микроволновой Конференции*. Севастополь. – 2010. – С. 833 – 835.
103. Вопросы квантовой теории необратимых процессов / Сборник статей под редакцией В.Л. Бонч-Бруевича. // М.: Изд. Иностранной литературы. – 1961. – 365 С.
104. Репке Г. Неравновесная статистическая механика / Г. Репке // М.: Мир. – 1990. – 320 С.
105. Ландау Л.Д. Статистическая физика. Ч. I. Теоретическая физика в 10 томах / Л.Д. Ландау, Е.М. Лифшиц // М.: Наука. – 1976. – Т. 5. – 584 С.
106. Петров Н. Современные проблемы термодинамики / Н. Петров, Й. Бранков // М.: Мир. – 1986. – 288 С.
107. Гроот С. Релятивистская кинетическая теория / С. де Гроот, В. ван Леувен, Х. ван. Верт // М.: Мир. – 1983. – 424 С.
108. Боголюбов Н.Н. Введение в теорию квантованных полей / Н.Н. Боголюбов, Д.В. Ширков // М.: Наука. – 1976. – 480 С.
109. Стратонович Р.Л. Нелинейная неравновесная термодинамика / Р.Л. Стратонович // М.: Наука. – 1985. – 480 С.

110. Capasso F. Quantum functional devices: Resonant-tunneling transistor, circuits with reduced complexity, and multiple valued logic / F. Capasso, S. Sen, F. Beltram, L.M. Lunardi, A.S. Vengurlekar, P.R. Smith, N.J. Shah, R.J. Malik, A.Y. Cho // *IEEE Trans.* - 1989. - V. ED-36. - N 10. - P. 2065 - 2082.
111. Иванов Ю.А. Исследование направлений применения резонансно-туннельного диода в интегральных схемах СВЧ диапазона / Ю.А. Иванов // *Микроэлектроника.* – 1996. – Т. 25. – № 4. – С. 249 – 258.
112. Siegel P.H. Terahertz Technology / P.H. Siegel // *IEEE Transactions on Microwave Theory and Techniques.* – 2002. – V. 50. – N 3. – P. 910 – 928.
113. Alkeev N. Factors reducing the cut-off frequency of resonant tunneling diodes / N. Alkeev, S. Averin, A. Dorofeev, N. Gladysheva // *International Journal of Microwave and Wireless Technologies.* – 2012. – V. 4(6). – P. 605 – 611.
Алкеев Н.В. Определение факторов, уменьшающих предельную частоту резонансно-туннельного диода / Н.В. Алкеев, С.В. Аверин, А.А. Дорофеев, Н.Б. Гладышева // *Материалы 22-й Международной Крымской Микроволновой Конференции.* Севастополь. – 2012. – С. 709 – 710.
114. Абрамов И.И. Проблемы и принципы физики и моделирования приборных структур микро- и нанoeлектроники. V. Резонансно-туннельные структуры / И.И. Абрамов // *Нано и микросистемная техника.* – 2007. – № 3. – С. 57 – 70.
115. Obukhov I.A. The new model of charge transport in quantum devices / I.A. Obukhov // *Proceeding of NANO-II. Part C. Moscow.* – 1993. – P. 889 – 896.
116. Борн М. Основы оптики / М. Борн, Э. Вольф // М.: Наука. – 1973. – 720 С.
117. Демиховский В.Я. Физика квантовых низкоразмерных структур / В.Я. Демиховский, Г.А. Вугальтер // Москва: Логос. - 2000. – 248 С.
118. Иванов Ю.А. Формирование ВАХ AlGaAs нанодиодов / Ю.А. Иванов, К.В. Малышев, Н.В. Федоркова // *Материалы 14-й Международной Крымской Микроволновой Конференции.* Севастополь. – 2004. – С.532–534.

119. Obukhov I.A. Some aspects of nanoelectronics development in Russia / I.A. Obukhov // WTEC Workshop Proceedings: Russian Research and Development Activities on Nanoparticles and Nanostructured Materials. S. Petersburg. – 1997. – P. 116 – 125.
120. Обухов И.А. Материалы нанозлектроники / И.А. Обухов // Материалы 7-й Международной Крымской Конференции «СВЧ -техника и телекоммуникационные технологии». Севастополь. – 1997. – С. 383 – 385.
121. Minne S.C., et al Independent parallel lithography using the AFM / S.C. Minne // J. Vac. Sci. Technol. B. – 1996. – V. 14. – P. 2456.
122. Фролов В.Д. Формирование проводящих нанообъектов на поверхности пленок α -С:Н в сканирующем туннельном микроскопе / В.Д. Фролов // Материалы 8-й Крымской Микроволновой Конференции. Севастополь. – 1998. – С. 577 – 579.
123. Ландау Л.Д. Квантовая механика. Нерелятивистская теория. Теоретическая физика в 10 томах / Л.Д. Ландау, Е.М. Лифшиц // М.: Наука. – 1974. - Т.3. – 752 С.
124. Landheer D. Role of carrier equilibrium in self-consistent calculations for double barrier resonant diodes / D. Landheer and G.C. Aers // Superlattices and Microstructures. – 1990. – V. 7. – N 1. – P.17 – 21.
125. Technology Roadmap for Nanoelectronics / Ed. by R. Compañó. // EC IST programme Future and Emerging Technologies. Second Edition. – 2000. – 104 P.
126. Smith C.G. Low-dimensional quantum devices / C.G. Smith // Rep. Prog. Phys. – 1996. – V. 59. – P. 235 – 282.
127. Lu W. Semiconductor nanowires / W. Lu, C.M. Lieber // J. Phys. D: Appl. Phys. – 2006. – V. 39. – P. R387 – R406.
128. Martel R. Single- and multi - wall carbon nanotube field effect transistors / R. Martel, T. Shcmidt, H.R. Shea, P. Avouris // Appl. Phys. Lett. – 1998. – V. 73. – N 17. – P. 2447 – 2449.
129. Tans S.J. et al. Individual single-wall carbon nanotubes as quantum wires / S.J. Tans et al. // Nature. – 1997. – V. 386. – P. 474 – 477.

130. Дьячков П.Н., Углеродные нанотрубки / П.Н. Дьячков // Природа. – 2000. – №11.
131. Лускинович П.Н. Формирование наноразмерных структур на пленке α -C:H в присутствии адсорбата / П.Н. Лускинович, В.Д. Фролов, А.Е. Шавыкин, В.Д. Хаврюченко, Е.Ф. Шека, Е.А. Никитина // Письма в ЖЭТФ. – 1995. – Т. 62. – № 11. – С. 868 – 872.
132. Молекулярно - лучевая эпитаксия и гетероструктуры / Под редакцией Л. Ченга и К. Плога. // М.: Мир. – 1989. – 582 С.
133. Волков В.А. Плазменные колебания краевых дираковских фермионов / В. А. Волков, И. В. Загороднев // Письма в ЖЭТФ. – 2013. – Т. 97. – № 7. – С. 469–472.
134. Волков В.А. Туннельная спектроскопия квазидвумерных плазмонов / В.А. Волков, Э.А. Тахтамиров, Д.Ю. Иванов, Ю.В. Дубровский, Л. Ивс, П.С Мэйн., М. Хенини, Д.К. Мауд, Дж.С. Портал, Ж.К. Маан, Д. Хилл // УФН. – 2001. – Т. 171. – № 12. – С. 1368–1370.
135. Houten H. Magnetoresistance of narrow GaAs-(Al,Ga)As heterostructures / H. van Houten, B.J. van Wees, M.G.J. Heuyman and J.P. Andre // Appl. Phys. Lett. – 1986. – V. 49. – P. 1781.
136. Thornton T.J. Mesoscopic devices / T.J. Thornton // Rep. Prog. Phys. – 1994. – V. 57. – P. 311 – 364.
137. Ferry D.K. Nanowires in nanoelectronics / D.K. Ferry // Science. - 2008. - V. 319. - P. 579 – 580.
138. Das S.R. Room Temperature Device Performance of Electrodeposited InSb Nanowire Field Effect Transistors / S.R. Das, C.J. Delker, D. Zakharov, Y.P. Chen, T.D. Sands, D.B. Janes // Appl. Phys. Lett. – 2011. – V. 98. – P. 243504-1 – 243504-3.
139. Jayasekera T. Evaporative Cooling in Semiconductor Devices [Электронный ресурс] / T. Jayasekera, K Mullen., M.A Morrison. // - 2006. – Режим доступа: <http://xxx.lanl.gov/pdf/cond-mat/0605073>.

140. Giazotto F. Thermal properties in mesoscopics: physics and applications from thermometry to refrigeration [Электронный ресурс] / F. Giazotto, T.T. Heikkila, A. Luukanen, A.M. Savin, J.P. Pekola // - 2006. - Режим доступа: <http://xxx.lanl.gov/abs/cond-mat/0508093>.
141. Горох Г.Г. Темплейтный метод формирования квантовых нанопроводов InSb с большим аспектным отношением / Г.Г. Горох, И.А. Обухов, А. А. Лозовенко, А. И. Захлебаева, Е. А. Сочнева // Материалы 23-й Международной Крымской Микроволновой Конференции. Севастополь. – 2013. – С. 820 – 823.
142. Обухов И.А. Электронные приборы на основе антимонида индия / И.А. Обухов, Г.Г. Горох // Материалы 22-й Международной Крымской Микроволновой Конференции. Севастополь. – 2012. – С. 653 – 654.
143. Обухов И.А. Неравновесные эффекты в электронных приборах / И.А. Обухов // Вебер. Москва – Киев – Минск – Севастополь. – 2010. – 303 С.
144. Обухов И.А. Особенности функционирования квантовых приборов / И.А. Обухов // Нано- и микросистемная техника. – 2009. – № 7. – С. 38 – 45.
145. Obukhov I.A. Planar One-Dimensional Quantum Devices / I.A. Obukhov // Nanoscience and Nanotechnology. – 2013. – V. 3. – N 5. – P. 115 – 122.
146. Heremans J.P. Resistance, Magnetoresistance, and Thermopower of Zinc Nanowire Composites / J.P. Heremans, C.M. Thrush, D.T. Morelli, M. C. Wu // Phys. Rev. Lett. – 2003. – V. 91. – P. 076804.
147. Beutler D.E. Localization and electron-electron interaction effects in thin Bi wires and films / D.E. Beutler and N. Giordano // Phys. Rev. B. – 1988. – V. 38. – N 1. – P. 8 - 19.
148. Ye Q.L. InSb and GaInSb nanowires for thermoelectric applications / Q.L. Ye, R. Scheffler, J. Kumari and R. Leverenz // Nanotech. – 2007. – V. 1. – P. 237 – 240.
149. Обухов И.А. Некоторые возможности квантовых интегральных схем / И.А. Обухов, В.М. Балабанов // Материалы 9-й Международной Крымской Микроволновой Конференции. Севастополь. – 1999. – С. 76 - 77.

150. Tritt T.M. Thermoelectric Materials, Phenomena, and Applications: A Bird's Eye View / T.M. Tritt and M.A. Subramanian // MRS BULLETIN. – 2006. – V. 31. – P. 188 – 198.
151. Vineis C.J. Nanostructured Thermoelectrics: Big Efficiency Gains from Small Features / C.J. Vineis, A. Shakouri, A. Majumdar and M.G. Kanatzidis // Adv. Mater. – 2010. – V. 22. – P. 3970–3980.
152. Datta S. The 21st Century Energy Efficient Logic Transistor [Электронный ресурс] / S.Datta // - 2007. - Режим доступа: <http://www.sematech.org/meetings/archives/litho/8065/pres/Key%20Note%20Summary%20Datta.pdf>.
153. Fechner A. Frequency-Dependent Electronic Transport in Quantum Wires / A. Fechner // Dissertation zur Erlangung des Doktorgrades des Fachbereichs Physik der Universität Hamburg. Germany. – 2000. – 154 P.
154. Обухов И.А. Релаксационная неустойчивость в квантовом проводе / И.А. Обухов // Материалы 23-й Международной Крымской Микроволновой Конференции. Севастополь. – 2013. – С. 869 – 872.
155. Патент Российской Федерации № 2113743 на изобретение «Генератор электромагнитных колебаний СВЧ и КВЧ диапазонов» /Патентообладатели: Квяткевич И.И., Матвеев Ю.А., Обухов И. А. и Чернявский А.А.// Приоритет изобретения 01.02.1996 г. Зарегистрирован в государственном реестре изобретений 20.06.1998 г.