

В.А. Москалюк, В.И. Тимофеев, А.В. Федяй

Сверхбыстродействующие приборы электроники

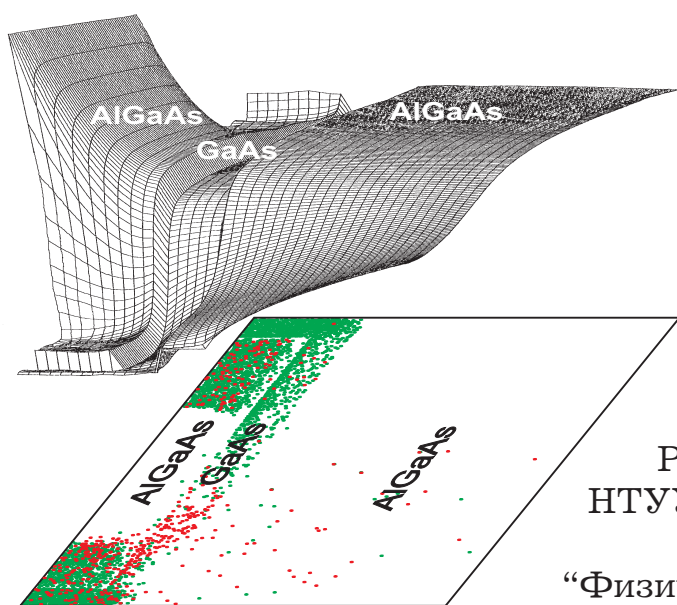
НТУУ «КПИ»
Киев 2012

В.А.Москалюк, В.И.Тимофеев, А.В.Федяй

СВЕРХБЫСТРОДЕЙСТВУЮЩИЕ

ПРИБОРЫ

ЭЛЕКТРОНИКИ



Рекомендовано Методическим советом
НТУУ «КПИ» в качестве учебного пособия
для студентов специальности
“Физическая и биомедицинская электроника”,
протокол № 9 от 24.05.2012 г.,
свидетельство НМУ № Е11/12-234

Киев 2012

УДК 621.38 : 53 (075.8)

Сверхбыстродействующие приборы электроники: Учебное пособие / Москалюк В.А., Тимофеев В.И., Федяй А.В. – К.: НТУУ «КПИ», 2012. – 479 с. илл. 226, табл. 6, библиогр. 310.

В учебном пособии рассмотрены современные и перспективные сверхбыстродействующие приборы электроники: полевые транзисторы с затвором Шоттки, гетеробиполярные транзисторы, транзисторы с высокой подвижностью электронов, мощные СВЧ транзисторы, резонансно-туннельные диоды и приборы, содержащие в активной области новые материалы (графен, нанотрубки). Анализ их работы предваряется описанием моделей переноса заряда на микроуровне без учета квантовых эффектов (уравнение Больцмана, релаксационные уравнения сохранения, метод Монте-Карло) и с учетом эффектов размерного квантования (метод огибающей волновой функций). Особое внимание уделено описанию материалов, используемых в области СВЧ и эффектов переноса горячих электронов, имеющих место при уменьшении характерных размеров приборов в сильном электрическом поле.

Предназначено для студентов специальности “Физическая и биомедицинская электроника”, может быть полезно студентам других специальностей и аспирантам вузов.

Рецензенты:

В. И. Осинский, доктор технических наук, профессор (Национальный технический университет Украины «КПИ»);

С. В. Плаксин, доктор физико-математических наук, профессор (Институт транспортных систем и технологий НАН Украины «Трансмаг»).

© В.А. Москалюк, В.И. Тимофеев, А.В. Федяй 2012

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	6
ВВЕДЕНИЕ	7
1. СРЕДСТВА МОДЕЛИРОВАНИЯ	15
1.1. Метод кинетического уравнения	15
1.1.1. Кинетическое уравнение.....	15
1.1.2. Интеграл столкновений.....	20
1.1.3. Приближение времени релаксации.....	23
1.1.4. Примеры решения кинетического уравнения.....	25
1.2. Метод Монте-Карло	31
1.3. Релаксационные уравнения	37
1.3.1. Усреднение кинетического уравнения	37
1.3.2. Релаксационные уравнения	39
1.3.3. Междолинная релаксация	41
1.4. Уравнения дрейфово-диффузионной модели	46
1.5. Квантовые модели переноса заряда	48
1.5.1. Общие положения	48
1.5.2. Метод волновых функций	53
1.6. Моделирование субмикронных структур	58
1.6.1. Задачи моделирования и типы моделей	58
1.6.2. Система моделирования субмикронных компонентов	66
Литература к гл. 1	71
2. ПОЛУПРОВОДНИКОВЫЕ МАТЕРИАЛЫ	74
2.1. Зонная структура полупроводников	74
2.1.1. Атомарные полупроводники	74
2.1.2. Бинарные полупроводники $A^{III}B^V$	79
2.1.3. Сложные соединения $A^{III}B^V$	89
2.1.4. Спектр фононов	90
2.2. Механизмы и скорости рассеяния	94
2.2.1. Фононное рассеяние	96
2.2.1.1. Общие закономерности	96
2.2.1.2. Акустическое рассеяние	98
2.2.1.3. Оптическое рассеяние	99
2.2.1.4. Междолинное рассеяние	101
2.2.2. Примесное рассеяние	103
2.3. Средние времена релаксации	106
2.3.1. Времена релаксации импульса	106
2.3.2. Времена релаксации энергии, упругость рассеяния	110
2.4. Подвижность электронов	116
Литература к гл. 2.	125
3. ГОРЯЧИЕ НОСИТЕЛИ В СИЛЬНОМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ	127
3.1. Разогрев электронного газа	127
3.2. Насыщение дрейфовой скорости	138
3.3. Динамическая проводимость	148
3.3.1. Дифференциальная проводимость	148
3.3.2. Высокочастотная проводимость	151

3.4. «Всплеск» дрейфовой скорости	156
3.5. Диффузия горячих электронов	160
Литература к гл.3	162
4. СУБМИКРОННЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ	163
4.1. Общие свойства	163
4.2. Квазидвумерная модель птШ	168
4.3. Схемные модели ПТШ	184
4.3.1. Малосигнальная модель	185
4.3.2. Шумовая модель	189
4.4. Модели ПТШ для режима большого сигнала	196
4.4.1. Субпороговый режим	197
4.4.2. Открытый затвор	202
4.4.3. Инверсный режим	205
4.5. Двумерное моделирование субмикронных ПТШ	207
4.6. Моделирование эффектов третьего измерения	212
4.6.1. Распределенные эффекты в ПТШ	212
4.6.2. Анализ ПТШ с распределенными параметрами во временной области	216
4.6.3. Моделирование размерных эффектов и технологических неоднородностей	221
Литература к главе 4	227
5. ГЕТЕРОСТРУКТУРНЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ	231
5.1. Общие свойства	231
5.2. Кинетическое моделирование ГСПТ	235
5.3. Двумерное моделирование ГСПТ	244
5.4. Гетеротранзисторы с двумя потенциальными ямами	248
5.4.1. Численное моделирование	249
5.4.2. Аналитическая модель статических характеристик	252
5.4.3. Схемные модели и шумовые параметры	255
5.5. Гетероструктурные транзисторы с квантовыми точками	258
Литература к главе 5	268
6. МОЩНЫЕ СВЧ ТРАНЗИСТОРЫ	272
6.1. Мощные субмикронные ПТШ	272
6.2. Мощные гетероструктурные полевые транзисторы	279
6.3. Гетеробиполярные транзисторы	282
6.3.1. Двумерное моделирование транзисторов	283
6.3.2. Моделирование разогретых эффектов	287
6.3.3. Схемное моделирование процессов саморазогрева	291
6.3.4. Электротепловой анализ монолитных интегральных схем на субмикронных гетеробиполярных транзисторах	298
6.3.5. Шумовые свойства субмикронных гетеробиполярных транзисторов.....	300
6.4. Мощные МОП-транзисторы	305
Литература к главе 6	313
7. РЕЗОНАНСНО-ТУННЕЛЬНЫЕ ДИОДЫ.....	318
7.1. Основные представления.....	318
7.1.1. Структурные составляющие РТД.....	318

7.1.2. Принцип работы резонансно-туннельного диода.....	319
7.1.2. Механизм токопереноса.....	322
7.2. Условно-аналитические модели.....	327
7.2.1. Метод плоских волн.....	330
7.2.2. Использование лоренцевой модельной функции.....	333
7.2.3. Полуаналитическая упрощенная модель РТД.....	337
7.2.4. Другие условно-аналитические методы.....	341
7.3. Численные модели.....	342
7.3.1. Концентрация и волновые функции.....	343
7.3.2. Влияние пространственного заряда.....	349
7.3.3. Обобщенная формула Цу-Эсаки.....	356
7.3.4. Рассеивание в квантовой яме РТД.....	359
7.3.5. Влияние вторичных факторов на характеристики РТД.....	365
7.4. Двухдолинное моделирование.....	368
7.5. Экспериментальная верификация моделей.....	377
Литература к главе 7	384
8. ПЕРСПЕКТИВНЫЕ НАНОСТРУКТУРЫ И НАНОКОМПОНЕНТЫ	390
8.1. Сравнительные характеристики наночастиц для применения в электронике	390
8.2. Наноконпоненты с использованием туннельного эффекта	394
8.3. Низкоразмерные структуры на основе графена	398
8.4. Наноконпоненты с использованием нанотрубок и нанопроволок	402
8.5 Пассивные компоненты наноинтегральных схем	408
8.6. Спинтронные структуры	418
8.7. Ячейки памяти на наноконпонентах.....	424
Литература к гл. 8	426
ПРИЛОЖЕНИЯ	430
П2. Усреднение скоростей рассеяния	430
П3. Высокочастотная проводимость	439
П4.1. Равновесное состояние контакта канал-подложка	442
П4.2. Уравнения непрерывности для канала и подложки	447
П4.3. Уравнение Пуассона для канала и подложки	450
П4.4. Уравнения сохранения энергии и импульса для канала и подложки.....	452
П4.5. Расчет малосигнальных параметров	455
П5. Алгоритм численного моделирования гетероструктурных транзисторов	459
П7.1. Вывод формулы Цу-Эсаки	468
П7.2. Конечно-разностная схема для системы уравнений огибающей волновой функции	470
П7.3. Уравнения, замыкающие конечно-разностную схему	475

ПРЕДИСЛОВИЕ

Современные тенденции развития электроники вызваны неудержимым желанием человечества повышать быстродействие компьютеров и осваивать новые информационные технологии, улучшать качество связи и передачи информации, повышать надежность локационных средств и решать все новые технические проблемы с помощью электронной техники. Такие задачи требуют соответствующих скоростей работы активных электронных компонентов, что в свою очередь вынуждает использовать для их создания как новые материалы так и все более тонкие физические эффекты. Кроме того, вполне естественным и необходимым выглядит и переход от микрометровых характерных размеров приборов к нанометровым, которые осваиваются современной технологией. При этих размерах все заметнее проявляются квантовые эффекты, которые могут быть использованы для создания принципиально новых приборов.

Учебная литература, особенно предназначенная для спецкурсов, не может оставаться в стороне от уровня и тенденций развития электроники. Предлагаемое учебное пособие представляет собой попытку отследить эти тенденции и дать будущим специалистам по микро- и наноэлектронике основы знаний, необходимых для понимания физических процессов, особенно динамических, происходящих в современных приборах. Авторы, следуя принципу «познание через моделирование», включили в настоящее пособие материал, основанный на собственном опыте моделирования. Поэтому, естественно, что перечень рассмотренных приборов не может быть полным. И только в некоторых случаях для полноты обзора использовались работы других авторов.

Авторы глубоко благодарны коллективу НИИ «Сатурн» в лице Босого В.И., Иващука А.В. и Максименко А.Г. за многолетнее творческое сотрудничество, Фалеевой Е.М., Семеновской Е.В., Овчарук М.Г. и Уваровой И.Ю. за помощь в проведении расчетов и подготовке рукописи.

Данное издание создавалось как учебное пособие по одноименному курсу для студентов, обучающихся по направлению «Микро- и наноэлектроника», но авторы надеются, что оно сможет принести пользу для гораздо большего круга читателей.

ВВЕДЕНИЕ

Главная тенденция в развитии электроники – повышение рабочих частот. Необходимое условие этого – уменьшение пролетного времени электронов и, как следствие, – уменьшение характерных размеров. Современная микроэлектроника с успехом осваивает субмикронные (< 1 мкм) и нанометровые ($< 0,1$ мкм) технологии. Не комментируя другие направления развития на рис. 1, отметим, что при этом возрастает роль междисциплинарных исследований, которые позволяют получать качественно новые соединения и структуры. Наряду с этой тенденцией остаются традиционными и направления, связанные с повышением чувствительности приемных устройств и повышением выходной мощности передающих.

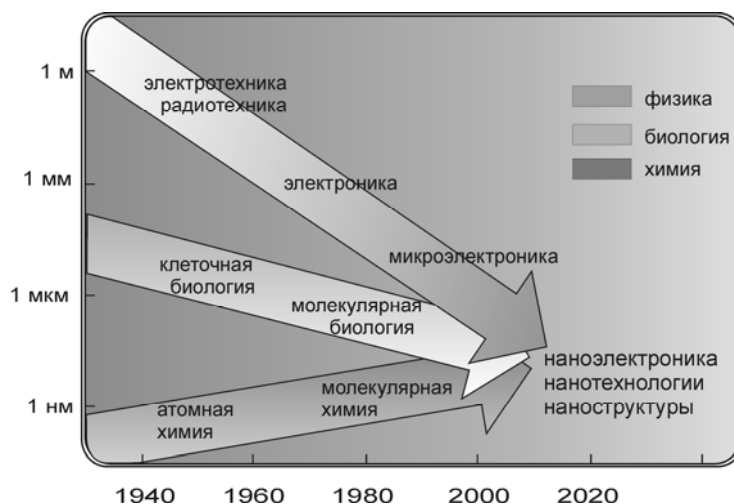


Рис.1. Тенденции развития некоторых направлений науки [1]¹

Сегодня существуют несколько направлений развития субмикронной и наноэлектроники – это поиск материалов и их соединений с принципиально новыми свойствами, совершенствование компонентов с субмикронными размерами активных областей, включая гетероструктуры, а также создание новых наноразмерных приборов, устройств и систем.

¹ 1. A. Cappy. Nanotechnology: the next industrial revolution? // The Newsletter of the French National Committee for Microelectronics Education. – 2002. – №12. – P. 4–6.

Одним из направлений повышения частотных возможностей и быстродействия электронных компонентов является использование новых полупроводниковых материалов. За последние два-три десятилетия технологическое освоение арсенида галлия существенно потеснило кремниевые приборы, которые господствовали на рынке электронной продукции, прежде всего в области сверхвысокочастотных и сверхбыстродействующих приборов. Этому способствовала не только более высокая подвижность электронов, но и некоторые особенности зонной структуры GaAs. Он, как и другие представители группы материалов $A_{III}B_V$, относится к так называемым многодолинным полупроводникам, которые имеют в зоне проводимости несколько подзон («долин»), с разными эффективными массами электронов. Это предопределяет существование эффекта отрицательной динамической проводимости (ОДП), на котором основано действие таких генераторов, как диод Ганна. С особенностью такой зонной структуры, точнее со спецификой междолинных переходов электронов, связан эффект «всплеска» дрейфовой скорости электронов в сильном электрическом поле. Этот эффект относится к так называемым квазибаллистическим явлениям и позволяет получать пикосекундный импульс тока как реакцию на включение напряжения, которое создает сильное (порядка единиц или десятков кВ/см) электрическое поле. С арсенидом галлия связано создание прежде всего полевых транзисторов с затвором в виде барьера Шоттки (ПТШ), гетероструктурных полевых транзисторов с высокой подвижностью электронов, гетеробиполярных транзисторов, монолитных интегральных схем разного назначения, и гетеролазеров. Столь широкий спектр созданных приборов обусловлен в том числе использованием тройных и четверных соединений, которые можно рассматривать как сплавы простых бинарных полупроводников $A_{III}B_V$ (AlGaAs, AlGaAsSb, InGaAs и много других). Варьируя их стехиометрический состав, можно получать материалы с прогнозируемыми параметрами.

Новые возможности предоставляет технологическое освоение в производстве электронных компонентов нитрида галлия. Характерны здесь темпы его использования для создания приборов (рис. 2), которые к тому же имеют тенденцию к увеличению, о чем можно судить по росту количества публикаций в научной литературе. При почти на порядок меньшей подвижности, чем для арсенида галлия, нитрид галлия имеет большую скорость насыщения, что позволяет получать более высокую плотность тока. А наличие существенно большей критической напряженности электрического поля предопределяет возможность создания приборов большей мощности (рис. 3).

Приборы на основе нитрида галлия в настоящее время рассматриваются как одни из самых перспективных в таких областях электроники, как светоизлучающие приборы коротковолновой области спектра, «солнечно слепые» ультрафиолетовые фотоприемники, мощные быстродействующие транзисторы для мобильных телекоммуникационных систем и других.

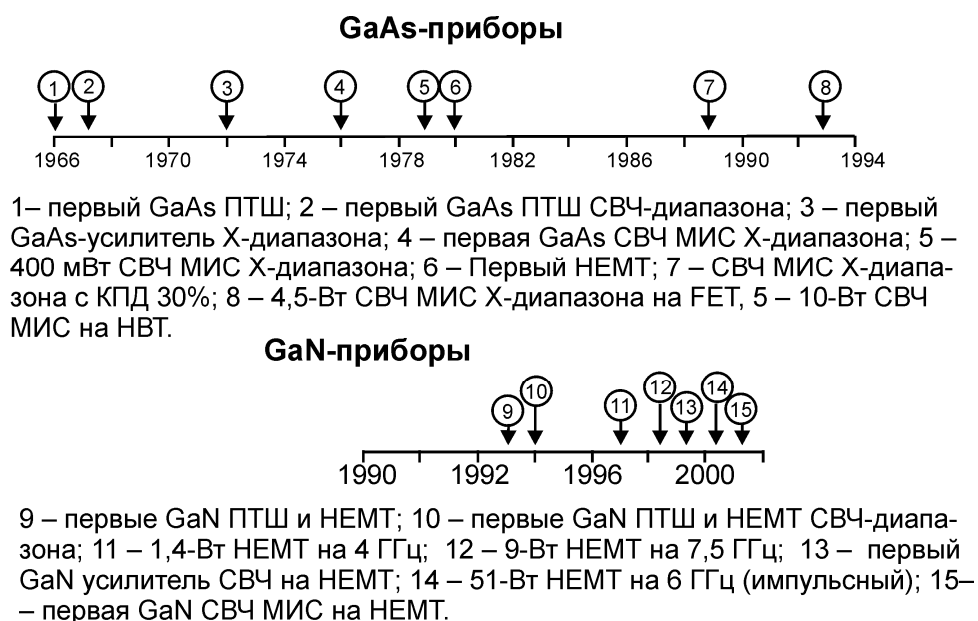


Рис. 2. Хронология создания приборов на арсениде и нитриде галлия [2]²

Нитриды III группы (GaN, InN, AlN) и тройные соединения на их основе формируют группу широкозонных материалов с уникальными свойствами, которые являются перспективными для разного применения при изготовлении приборов как оптоэлектроники, так и чисто электронного назначения. Это обусловлено в первую очередь меньшей вероятностью образования дефектов в структуре нитридов, чем в решетках арсенида и фосфидов в связи с более сильной связью атомов металлов с атомами азота, чем с атомами мышьяка и фосфора. Такая ситуация способствует повышению эксплуатационных параметров приборов на основе нитридов: долговечности, выходной мощности, рабочей частоты, термической, механической, химической, радиационной и химической стабильности, быстродействия и напряжения питания.

Развитие технологии и новые открывающиеся возможности по созданию субмикронных компонентов и интегральных схем (ИС) на их основе предусматривают усовершенствование существующих

² 2. Данилин В., Жукова Т. Транзисторы на GaN пока самый крепкий орешек // Электроника: наука, технологии, бизнес. – 2005. – Т. 26, вып. 4. – С. 20–28.

щих и создание новых компонентов, адаптированных к диапазонам сверхвысоких (СВЧ) и крайне высоких частот (КВЧ), которые могут достигать предельных значений до 500 ГГц.

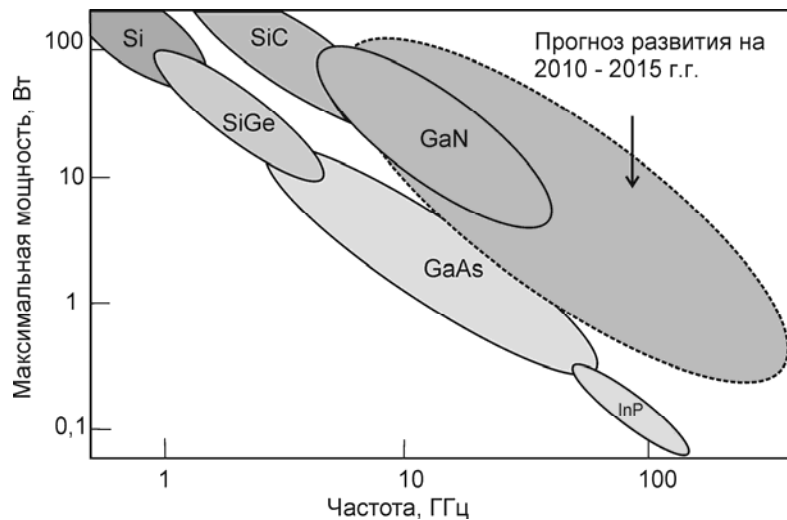


Рис. 3. Зависимость максимальной мощности от частоты для приборов на основе разных полупроводниковых материалов. Прогноз построен на основе работы [3]³ и расчетов разд. 3.

В современных электронных системах используется широкий спектр интегральных схем, выполненных на полупроводниковых соединениях $A_{III}B_V$ и $A_{II}B_{VI}$. Это связано в основном с приборами на полевых транзисторах с затвором Шоттки – ПТШ (*MESFET – Metal Semiconductor Field Effect Transistor*) и полевых транзисторах с управляющим *pn*-переходом (*Junction FET*), все большее признание получают ИС на субмикронных гетероструктурных биполярных транзисторах – ГСБТ (*HBT – Hetero Bipolar Transistor*), на транзисторах с высокой подвижностью электронов или гетероструктурных полевых транзисторах ГСПТ (*HEMT – High Electron Mobility Transistor*) и псевдоморфных гетеротранзисторах – *pHEMT* и метаморфных гетеротранзисторах *mHEMT*. Такие приборы широко используются в телекоммуникационных системах, сотовой телефонии, компьютерных сетях передачи данных, космических системах, системах локации и др. Сравнительно новыми областями применения является высокоскоростная вычислительная техника и волоконно-оптическая связь. Субмикронные ПТШ и гетероструктурные транзисторы используются, как правило, в маломощных и малошумящих цепях, а гетеробиполярные транзисторы – в мощных цепях усиления и генерации.

³ 3. Нанотехнологии в электронике / Под. ред. А.Чаплыгина. М.: Техносфера. 2005 – 488 с

В отличие от ПТШ с барьером Шоттки, в котором модуляция проводимости достигается за счет изменения толщины канала, в гетеротранзисторе модуляция проводимости осуществляется в результате изменения поверхностной плотности электронов на границе гетероперехода. Основными задачами совершенствования технологии ИС на субмикронных гетеротранзисторах является уменьшение числа дефектов, увеличение концентрации двумерного электронного газа, увеличение подвижности электронов в двумерном электронном газе, обеспечение однородности характеристик двумерного электронного газа. Совершенствование технологии гетеробиполярных транзисторов, кроме того, связано с получением тонких (менее 0,1 мкм) многослойных высоколегированных (более чем 10^{19} см⁻³) структур.

Псевдоморфные ГСПТ отличаются от обычных тем, что вместо соединения алюминия и арсенида галлия используется гетеропереход на соединении InGaAs и подложка из фосфида индия. Это приводит к увеличению разрыва зон на границе гетероперехода более чем в два раза, что способствует нейтрализации паразитного канала в слое под затвором, а также увеличению дрейфовой скорости и подвижности электронов, а значит увеличению быстродействия и предельных частот. Кроме того, *pHEMT* имеют большие уровни мощности, поскольку InP имеет большую теплопроводность, чем GaAs. В подложке из фосфида индия обеспечивается большая плотность двумерного электронного газа, что позволяет получать более высокую плотность тока. К недостаткам *pHEMT* следует отнести сложность изготовления и высокую стоимость. Последующая оптимизация структуры ГСПТ связана с появлением *mHEMT*, в которых для уменьшения рассогласования кристаллических решеток гетерослоев и арсенидгаллиевой подложки (что является главным препятствием для получения высококачественных структур) используется многослойная варизонная система буферных слоев AlGaAsSb. Развитие технологии изготовления структур на нитриде галлия способствовало появлению ГСПТ на AlGaIn/GaN, которые уступают по частотным и усилительным свойствам транзисторам на соединениях A_{III}B_V, но существенно превосходят их по плотности тока, мощности и рабочему напряжению.

Развитие современных гетеротранзисторных структур обусловлено как использованием многодолинных полупроводниковых материалов, так и достижениями технологии по созданию многослойных наноструктур. Повышение быстродействия транзисторов может быть также связано с созданием структур с субмикронными продольными и нанометровыми поперечными размерами, в том числе многослойными наноструктурами низкой

размерности [4, 5]⁴, например, гетеротранзисторами с двумя гетеропереходами. Характерные размеры затвора таких гетероструктурных транзисторов составляют 0,1-0,2 мкм при размерах высоколегированных слоев в активной области транзистора менее 0,1 мкм. Показано [6]⁵, что средние значения дрейфовой скорости в такой структуре более высоки, чем в ГСПТ с одной потенциальной ямой, при этом транзисторы с двумя гетеропереходами имеют более высокие значения тока и мощности. Сегодня достигнуты размеры гетеропереходов в двухканальном транзисторе около 3 нм на соединении InGaN [6].

Передатчики для авиационных применений и обычные радары, использующие зеркальные антенны, где требуемая выходная мощность достигается суммированием мощностей нескольких транзисторов, предъявляют строгие требования к характеристикам используемых транзисторов [7]⁶:

- значительное усиление (4...6 и более параллельно используемых транзисторов);
- стойкость к рассогласованию для предотвращения выхода транзистора из строя при настройке;
- низкий уровень паразитной генерации при рассогласовании нагрузки;
- фазовая стабильность, оказывающая огромное влияние на доплеровскую чувствительность системы в целом.

Система на фазированных антенных решетках предъявляют к транзисторам сходные требования, за исключением того, что усиление может быть ниже, поскольку в передатчике не используется сложение мощностей. Кремниевый биполярный транзистор является «рабочей лошадкой» всех современных радарных разработок с частотами до 3,5 ГГц. В типовой схеме кремниевый биполярный транзистор с усилением 9 дБ и мощностью 110 Вт «раскачивает» еще 6 таких же транзисторов для получения мощности 600 Вт. В работе [8]⁷ рассмотрен пример создания усилителя, в

⁴ 4. Yu.G. Sadofyev et. al. New approach to the independent ohmic contact formation in the structures with two parallel isotype quantum wells // J. Vac. Sci. Technol. B. – 2000. – 18(3). – P.1668–1671.

⁵ 5. Chu R. et. al. AlGaIn-GaN Double-Channel HEMTs // IEEE Transactions On Electron Devices. – 2005.–Vol. 52, №4. – P. 438–445.

⁶ 6. Liu J. et. al. DC and RF Characteristics of AlGaIn-GaN/InGaIn/GaN Double-Heterojunction HEMTs // IEEE Transactions On Electron Devices. – 2007. – Vol. 54, №1. – P. 2–10.

⁷ 7. Технология мощных СВЧ LDMOS-транзисторов для радарных передатчиков L-диапазона и авиационных применений. Компоненты и технологии № 10 2007 с.14-16.

⁸ 8. Murai T., Fujii K., Matsuno T. Power S-Band Solid-State Amplifiers for Surveillance and Traffic Control Radars/ IEEE MTT-S. 2001

котором кремниевый биполярный транзистор с усилением 7,5 дБ и мощностью 150 Вт «раскачивает» еще 4 таких же прибора для достижения выходной мощности 550 Вт в диапазоне частот от 2,7 до 2,9 ГГц.

Под влиянием рынка беспроводных коммуникаций и его требований к линейности усилительных устройств технология производства мощных *LDMOS*-транзисторов за последние годы шагнула далеко вперед. Особенно ярко это демонстрируют транзисторы для диапазона 2,1 ГГц, предназначенные для базовых станций: приборы подобного класса обладают плотностью мощности 0,9 Вт/мм при усилении 16,5 дБ и пиковой эффективности 62% [9]⁸. При этом каждое новое поколение *LDMOS*-транзисторов (*Laterally Diffused Metal Oxide Semiconductors*) обладает улучшенными характеристиками по сравнению с предыдущим. Все это позволяет сделать вывод, что *LDMOS* – достаточно надежная и отработанная технология, по многим параметрам не уступающая, а порой и превосходящая биполярную. В отличие от биполярных транзисторов, *LDMOS* – относительно молодая технология, которая преимущественно развивалась в конце 80-х – начале 90-х годов прошлого века и стала ключевой на рынке базовых станций. Благодаря более высокому усилению, лучшей эффективности и, самое главное, гораздо более высокой линейности, к середине 90-х годов *LDMOS* стала ведущей технологией на рынке устройств с высокой линейностью, полностью заменив биполярную. При развитии технологии *LDMOS* оптимизация конструкции транзистора и большинство исследований были направлены на достижение большей линейности устройства [10]⁹.

Примером развития технологий изготовления гетероструктур до размеров многослойных сверхрешеток являются современные резонансно-туннельные диоды. Недавно было продемонстрировано применение РТД для разных целей в электронике: создание цифро-аналоговых преобразователей (ЦАП), задающих генераторов, регистров сдвига и статических запоминающих устройств с произвольной выборкой (ЗУПВ) со сверхнизким потреблением энергии [11]¹⁰. РТД могут быть изготовлены для более высокопроизводи-

⁸ 9. Brech H., Brakensienk W., Burdeaux D., Burger W., Dragon C., Formicone G., Pryor B., Rice D. Record Efficiency and Gain at 2,1 GHz of High Power RF Transistor for Cellular and 3G Base Stations. IEEE IEDM. 2003.

⁹ 10. Brech H., Burger W., Dragon C., Pryor B. Voltage Optimization for State of the Art RF-LDMOS for 2,1 GHz W-CDMA Cellular Infrastructure Application. IEEE MTT-S. 2003.

¹⁰ 11. Figueiredo M. L. Ultralow voltage resonant tunnelling diode electro-absorption modulator / M.L. Figueiredo, C.N. Ironside, C.R. Stanley // Journal of Modern Optics. – 2002. – Vol. 49. – P. 939–945.

тельных ЦАП, чем другие устройства на основе комплементарных структур металл-оксид-полупроводник. Они покрывают частотный диапазон 10...100 ГГц и потребляют меньше энергии, чем существующие ЗУПВ [12]¹¹. Диапазон использования РТД в электронных устройствах СВЧ и КВЧ диапазонов постоянно расширяется. Например, в результате выполнения исследовательской программы *DAPRA's Ultra Electronics Program* (США) были созданы сумматоры в гигагерцовом диапазоне частот и показана возможность интеграции для создания базовых элементов, необходимых для сумматоров, сигнальных процессоров и цепей с многозначной логикой. Были продемонстрированы 4-битные ЦАП частотой 2 ГГц, АЦП частотой 3 ГГц, тактовые генераторы, регистры сдвига, а также статические ЗУПВ с потреблением энергии 50 нВт/бит [13]¹²

Благодаря высокой скорости процесса туннелирования резонансно-туннельные диоды имеют потенциал к повышению частотного диапазона. В 2010 году было продемонстрировано роботу GaInAs/AlAs РТД в качестве генератора сигналов частотой 1,04 ТГц и мощностью 7 мкВт [14]¹³.

¹¹ 12. Technology Roadmap for Nanoelectronics / Editor: R. Compano. – Second Edition. – Belgium, 2000. – 104 P.

¹² 13. Roco M.C. Nanotechnology Research Directions: IWGN Workshop Report. Vision for Nanotechnology R&D in the Next Decade / M.C. Roco, S. Williams, P. Alivisatos. – London: Kluwer Academic Publishers, 1999. – 318 P.

¹³ 14. Suzuki S. Fundamental oscillation of resonant tunneling diodes above 1 THz at room temperature / S. Suzuki, M. Asada, M. Teranish et.al. // Appl. Phys. Lett. – 2010. – Vol. 97. P. 242102-1–3.

1. СРЕДСТВА МОДЕЛИРОВАНИЯ

Наиболее точные количественные модели электронных процессов при наличии внешних воздействий, в том числе и динамических, могут быть созданы на основе так называемого *кинетического уравнения*. Оно в общем случае представляет собой интегро-дифференциальное уравнение, с помощью которого может быть найдена функция распределения частиц, а значит и все *средние* величины, характеризующие поведение коллектива частиц.

1.1. МЕТОД КИНЕТИЧЕСКОГО УРАВНЕНИЯ

1.1.1. Кинетическое уравнение

Для анализа неравновесных процессов в полупроводниках, происходящих, например, в сильных электрических полях, при освещении, при инжекции носителей через контакты, недостаточно энергетической функции распределения $N(E)$, которая используется для расчета равновесных концентраций. Более полно отражает состояние электронного коллектива функция $N(\vec{r}, \vec{k}, t)$, которая подобно $N(E)$ может быть представлена в виде произведения функции плотности \vec{k} -состояний $g(k) = dn_c/dV_k$ и функции $f(E(\vec{k})) = dn/dn_c$, характеризующей вероятность заполнения состояний, причем связь $E(\vec{k})$ определяется законом дисперсии для электронов в данном веществе. Функция плотности состояний может быть выбрана в виде $g(k) = 2V/(2\pi)^3$, если остаются в силе основные выводы зонной теории. Функция $f(E(\vec{k}))$ также зачастую может быть взята в виде, подобном равновесной функции Ферми-Дирака, однако входящие в нее величины приобретают несколько иной смысл.

В результате динамического воздействия происходит изменение \vec{k} -состояний электронов, а их энергия становится функцией времени – $E(\vec{k}(t))$. Пространственно неоднородную неравновесность возможно описать с помощью введения *локальной температуры* $T_e(\vec{r})$, которая в общем случае различна для электронов и дырок и к тому же не обязательно совпадает с

температурой вещества или окружающей среды. В ряде случаев, обычно при отклонении концентрации неравновесных носителей от равновесных значений, локальную неравновесность описывают с помощью введения так называемых *квазиуровней Ферми* $E_{\phi n}(\vec{r})$ для электронов и $E_{\phi p}(\vec{r})$ для дырок, которые выступают как формальные параметры функции Ферми-Дирака.

В итоге представляется возможным описывать неравновесие электронного коллектива с помощью функции распределения

$$N(\vec{r}, \vec{k}, t) = \frac{d^2 N}{dV \cdot dV_k} = \frac{g(k)}{\exp\{[E(\vec{k}(t)) - E_{\phi n}(\vec{r})] / k_B T_e(\vec{r})\} + 1}. \quad (1.1)$$

Эта функция нормируется таким образом, что, будучи проинтегрированной по всему геометрическому объёму V и по объёму k -пространства V_k , даёт полное количество частиц данного сорта в данном коллективе:

$$\int_V dV \int_{V_k} N(\vec{r}, \vec{k}, t) dV_k = N. \quad (1.2)$$

В дальнейшем ограничимся рассмотрением единичного геометрического объёма $V = 1$, выбирая функцию плотности состояний в виде $g(k) = 1/4\pi^3$ [1.1].

Знание функции распределения позволяет найти основные макроскопические характеристики коллектива. Например, интегрирование ее по всему k -пространству даёт распределение концентрации частиц в геометрическом пространстве

$$n(\vec{r}, t) = \int_{V_k} N(\vec{r}, \vec{k}, t) dV_k, \quad (1.3)$$

и соответственно распределение плотности объемного заряда $\rho(\vec{r}, t) = -en(\vec{r}, t)$.

Усреднением скорости в k -пространстве по функции распределения можно найти среднюю дрейфовую скорость направленного движения частиц

$$\vec{v}_{др}(\vec{r}, t) = \frac{1}{n(\vec{r}, t)} \int_{V_k} \vec{v} N(\vec{r}, \vec{k}, t) dV_k, \quad (1.4)$$

а значит и распределение плотности тока

$$\vec{j}(\vec{r}, t) = -en(\vec{r}, t)\vec{v}_{\text{др}}(\vec{r}, t) = -e \int_{V_k} \vec{v} N(\vec{r}, \vec{k}, t) dV_k. \quad (1.5)$$

Подобным образом можно найти локальное среднее значение $\varphi_{\text{ср}}(\vec{r}, t)$ любой физической величины, так как

$$\varphi_{\text{ср}}(\vec{r}, t) = \frac{1}{n} \int_{V_k} \varphi(\vec{k}) N(\vec{r}, \vec{k}, t) dV_k = (n\varphi)_{\text{ср}}/n. \quad (1.6)$$

Здесь необходимо обратить внимание на выбор пределов интегрирования. В последних интегралах, записанных для общего случая, интегрирование проводится по всем возможным значениям *волнового* вектора. Для электронов в периодическом поле кристаллической решетки областью неповторяющихся значений *квазиволнового* вектора является зона Бриллюэна V_B , причем интегрировать надо по состояниям, соответствующим одной энергетической зоне. Более того, иногда интегрируют по какому-либо характерному участку дисперсионной кривой, определяя концентрацию или средние значения для частиц того или иного вида. Так, например, может быть найдена концентрация и параметры электронов в Γ -, L- или X-долине зоны проводимости, «легких» или «тяжелых» дырок в валентной зоне.

Кинетическое уравнение основано на том, что в замкнутой системе количество частиц сохраняется неизменным, то есть $N = \text{const}$, то

$$dN(\vec{r}, \vec{k}, t)/dt = 0. \quad (1.7)$$

Более строго это уравнение доказывается на основе теоремы Лиувилля, рассматриваемой в статистической механике. Соотношение (1.7) служит основой для формулирования *кинетического уравнения*, которое позволяет найти изменение функции распределения при заданном воздействии на частицы. Надо заметить, что кинетическое уравнение составляют обычно отдельно для коллективов частиц разного вида (электронов определенной зоны, дырок, фононов и др.). Рассмотрим это уравнение для электронов. Так как $N(\vec{r}, \vec{k}, t)$ – функция трех переменных, две из которых векторные, то $dN = (\partial N / \partial t)dt + (\partial N / \partial \vec{r})d\vec{r} + (\partial N / \partial \vec{k})d\vec{k}$, или

$$\frac{dN}{dt} = \frac{\partial N}{\partial t} + \frac{\partial N}{\partial \vec{r}} \frac{d\vec{r}}{dt} + \frac{\partial N}{\partial \vec{k}} \frac{d\vec{k}}{dt}. \quad (1.8)$$

Здесь для краткости записи обозначено

$$\frac{\partial N}{\partial \vec{r}} = \vec{e}_x \frac{\partial N}{\partial x} + \vec{e}_y \frac{\partial N}{\partial y} + \vec{e}_z \frac{\partial N}{\partial z} = \text{grad}_r N, \quad (1.9)$$

$$\frac{\partial N}{\partial \vec{k}} = \vec{e}_x \frac{\partial N}{\partial k_x} + \vec{e}_y \frac{\partial N}{\partial k_y} + \vec{e}_z \frac{\partial N}{\partial k_z} = \text{grad}_k N. \quad (1.10)$$

Учитывая, что $d\vec{r}/dt = \vec{v}$, а $d\vec{k}/dt = \vec{F}/\hbar$, запишем (1.7) в виде

$$\frac{\partial N}{\partial t} + \left(\vec{v} \cdot \frac{\partial N}{\partial \vec{r}} \right) + \frac{1}{\hbar} \left(\vec{F} \cdot \frac{\partial N}{\partial \vec{k}} \right) = 0. \quad (1.11)$$

Силы, действующие на электроны, можно условно разделить на две группы. Это, во-первых, силы со стороны внешних полей $\vec{F} = -e(\vec{E}(\vec{r}, t) + [\vec{v} \times \vec{B}(\vec{r}, t)])$, которые изменяются в пространстве сравнительно плавно. Во-вторых, это силы, которые возникают между двумя частицами при достаточном их сближении, либо при локальных нарушениях периодичности кристаллической решетки. Несмотря на то, что такие взаимодействия имеют также электрическую природу, из-за их локальности они рассматриваются как *столкновения* или *соударения*. Оставляя в третьем слагаемом (1.11) лишь силу внешних полей, можно записать

$$\frac{\partial N}{\partial t} + \vec{v} \cdot \frac{\partial N}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial N}{\partial \vec{k}} = \left[\frac{\partial N}{\partial t} \right]_c, \quad (1.12)$$

где правая часть, называемая *интегралом столкновений*, символически учитывает влияние последних на функцию распределения. Поскольку плотность состояний $g(k)$ постоянна в k -пространстве, то, разделив все слагаемые на $g(k) = 1/4\pi^3$, получим наиболее распространенную форму записи кинетического уравнения для функции $f(\vec{r}, \vec{k}, t)$

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}} = 4\pi^3 \left[\frac{\partial N}{\partial t} \right]_c. \quad (1.13)$$

Если второе и третье слагаемое перенести в правую часть, то получим более наглядную форму записи кинетического уравнения с точки зрения его физического смысла: изменение функции распределения в данной точке возможно за счет пространственной неоднородности распределения (наличие градиентов концентрации, температуры), воздействия внешних полей и столкновений.

Поэтому иногда эти слагаемые называют диффузионным, полевым и столкновительным.

Характерным примером использования кинетического уравнения является слабо ионизированная газовая плазма или электронно-дырочная плазма в невырожденных полупроводниках. В этих случаях кинетическое уравнение называется *уравнением Больцмана*. Если столкновениями можно пренебречь, что хорошо выполняется, например, для очень разреженной газовой плазмы или для электронных потоков в вакууме, то $[\partial N/\partial t]_c = 0$ и (1.13) называется *уравнением Власова*.

Кинетическое уравнение является очень мощным средством *статистического описания* поведения больших коллективов частиц, так как получаемая в результате его решения функция распределения позволяет весьма детально анализировать физические процессы переноса носителей в электронных приборах. По своей природе кинетическое уравнение – классическое, но, придавая тем или иным величинам соответствующий квантовый смысл, можно получить так называемое *полуклассическое* описание физических процессов с учетом квантовых эффектов.

Кинетическое уравнение позволяет следить за изменением функции распределения в тех или иных условиях, а конкретный вид функции определяется статистическими свойствами коллектива частиц.

Рассмотрим, например, равновесное состояние коллектива, в котором функция распределения f_0 должна быть стационарной ($\partial f_0/\partial t = 0$) и столкновения не должны изменять ее вида ($[\partial N/\partial t]_c = 0$). В этих условиях кинетическое уравнение принимает вид

$$\vec{v} \frac{\partial f_0}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial f_0}{\partial \vec{k}} = 0. \quad (1.14)$$

Поскольку f_0 зависит от \vec{k} не явно, а через закон дисперсии $E(\vec{k})$, и учитывая, что групповая скорость электрона $\vec{v} = (\partial E/\partial \vec{k})/\hbar$, то $\partial f_0/\partial \vec{k} = (\partial f_0/\partial E) \cdot (\partial E/\partial \vec{k}) = \hbar \vec{v} (\partial f_0/\partial E)$, и уравнение преобразуется к виду

$$\vec{v} (\partial f_0/\partial \vec{r}) + \vec{F} \vec{v} (\partial f_0/\partial E) = 0. \quad (1.15)$$

Здесь интересно, что $\vec{F} \vec{v} = -e(\vec{\mathcal{E}} + [\vec{v} \times \vec{B}]) \cdot \vec{v} = -e \vec{\mathcal{E}} \vec{v}$, и поэтому магнитное поле непосредственно не влияет на вид равновесной функции распределения. Кроме того, равновесное состояние возможно при наличии статического электрического поля, но функция распределения должна быть пространственно неоднородной, причем эта неоднородность связана только с простран-

ственным распределением потенциала $V(\vec{r})$ или потенциальной энергии $U(\vec{r}) = -eV(\vec{r})$, так как $\partial f_0 / \partial \vec{r} = (\partial f_0 / \partial U)(\partial U / \partial \vec{r}) = (\partial f_0 / \partial U) \text{grad} U = -(\partial f_0 / \partial U) \vec{F}$. В этом случае (1.15) превращается в равенство $\partial f_0 / \partial E - \partial f_0 / \partial U = 0$, которое возможно, если аргументом функции распределения будет сумма $(E(\vec{k}) + U(\vec{r}))$ в отличие от (1.1), где аргументом служит только энергия $E(\vec{k})$. Эта ситуация встречается, например, когда рассматривается «искривление» энергетических зон во внешнем или контактном электрическом поле, не нарушающем равновесного состояния, и пространственное перераспределение носителей в соответствии с потенциальным рельефом [1.1]. Равенство нулю суммы диффузионного и полевого слагаемых в (1.14) означает, что возникающий из-за наличия градиента функции распределения диффузионный перенос носителей уравнивается дрейфовым движением в электрическом поле. Причем это равновесие будет локальным, выполняющимся для каждой точки пространства.

1.1.2. Интеграл столкновений

Для решения кинетического уравнения в конкретной ситуации необходимо задать распределение сил, действующих на частицы, и описать процессы соударений, которые приводят к изменению электронных состояний. В общем случае под действием локальных электрических полей, создаваемых дефектами кристаллической решетки, происходит *рассеяние* носителей заряда, то есть в процессе соударения изменяется и координата \vec{r} электрона и его волновой вектор \vec{k} . Однако, если принять, что соударения происходят мгновенно, то их результатом является изменение лишь \vec{k} -состояний, а вместе с ними и энергетических состояний, задаваемых законом дисперсии $E(\vec{k})$. Пространственное же расположение частиц при мгновенных соударениях не изменяется.

Для установления общей формы записи интеграла столкновений подсчитаем изменение общего количества частиц за счет столкновений в какой-либо точке k -пространства. Для этого обозначим вероятность перехода электрона за 1 секунду из заполненного состояния \vec{k} в свободное состояние \vec{k}' (рис. 1.1) через $w_{k\vec{k}'}$. Если же начальное состояние \vec{k} заполнено с вероятностью $f(\vec{k})$, а конечное состояние \vec{k}' свободно с вероятностью $(1 - f(\vec{k}'))$, то результирующая вероятность такого перехода будет равна $w_{k\vec{k}'} f(\vec{k})(1 - f(\vec{k}'))$. Соответственно вероятность обратных переходов

– $w_{k'k}f(\vec{k}')\left(1-f(\vec{k})\right)$. Если в окрестности точки \vec{k}' выбрать элементарный объем $dV_{k'}$, количество состояний в котором равно $dN_c = g(k')dV_{k'}$, то пропорционально им увеличится количество переходов между состоянием \vec{k} и \vec{k}' -состояниями в объеме $dV_{k'}$. В итоге полное изменение количества частиц в рассматриваемом \vec{k} -состоянии за единицу времени в результате обмена со всеми остальными состояниями можно найти интегрированием по всем \vec{k}' -состояниям

$$(\partial N/\partial t)_c = \int_{V_{k'}} g(\vec{k}') \left[w_{k'k}f(\vec{k}')\left(1-f(\vec{k})\right) - w_{kk'}f(\vec{k})\left(1-f(\vec{k}')\right) \right] dV_{k'}. \quad (1.16)$$

Отсюда с учетом постоянства функции $g(k)$ и следует представление правой части кинетического уравнения (1.13) в виде интеграла, который символически обычно обозначают $(\partial f/\partial t)_c$,

$$\begin{aligned} (\partial f/\partial t)_c &= (\partial N/\partial t)_c/g(k) = \\ &= \int_{V_{k'}} \left[w_{k'k}f(\vec{k}')\left(1-f(\vec{k})\right) - w_{kk'}f(\vec{k})\left(1-f(\vec{k}')\right) \right] dV_{k'} \end{aligned} \quad (1.17)$$

В результате кинетическое уравнение оказывается интегро-дифференциальным уравнением относительно функции распределения $f(\vec{r}, \vec{k}, t)$.

Прежде всего, отметим, что существует множество различных механизмов рассеяния и, если отдельные акты столкновений происходят случайно и независимо друг от друга, то их вероятности суммируются

$$w_{kk'}^\Sigma = \sum_i w_{kk'}^i. \quad (1.18)$$

Поэтому в правой части кинетического уравнения должна стоять сумма интегралов, соответствующих разным видам соударений.

В равновесном состоянии «приход» и «уход»

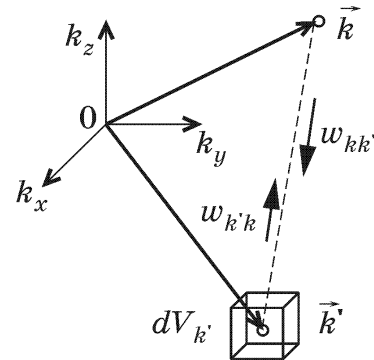


Рис. 1.1. Изображение переходов между k - и k' -состояниями. Выбором элементарного объема в окрестности k' -состояния и последующим интегрированием по всему k' -пространству достигается подсчет изменения числа частиц в выбранном k -состоянии за счет обмена со всеми остальными состояниями.

частиц для каждого \vec{k} -состояния уравниваются и подынтегральное выражение в (1.17) равно 0, поэтому

$$w_{k'k}f(\vec{k})(1 - f(\vec{k}')) = w_{kk'}f(\vec{k}')(1 - f(\vec{k})). \quad (1.19)$$

Подставляя сюда явный вид функции $f(\vec{k}) = f(E(\vec{k}))$ в виде функции Ферми-Дирака, нетрудно получить

$$w_{k'k}/w_{kk'} = \exp[(E'(k') - E(k))/k_B T]. \quad (1.20)$$

Это важное в теории квантовых переходов соотношение, называемое иногда *принципом детального равновесия*, означает в частности, что $w_{k'k} > w_{kk'}$, если $E'(k') > E(k)$, то есть более вероятными являются переходы с уменьшением энергии. Большая вероятность переходов из \vec{k}' в \vec{k} компенсируется меньшей заселенностью \vec{k}' -состояний с большей энергией, и в результате количество переходов между любой парой состояний \vec{k} и \vec{k}' «туда» и «обратно» одинаково.

Интеграл столкновений в форме (1.17) получен для вырожденного коллектива частиц. Принцип Паули учитывается в нем наличием сомножителей $(1 - f(k))$ и $(1 - f(k'))$, которые фактически запрещают переходы в уже заполненные состояния. Для невырожденных коллективов $f(k) \ll 1$ и интеграл столкновений можно представить проще

$$(\partial f/\partial t)_c = \int_{V_k} [w_{k'k}f(\vec{k}') - w_{kk'}f(\vec{k})] dV_{k'}. \quad (1.21)$$

Соответственно в равновесном состоянии

$$w_{k'k}f(\vec{k}') = w_{kk'}f(\vec{k}). \quad (1.22)$$

Но при подстановке в последнее равенство функции распределения Максвелла-Больцмана в виде $f(E(k)) = \exp[(E_{\text{ф}} - E(k))/k_B T]$ получается то же самое соотношение (1.20), что и для вырожденного коллектива. В частности, для *упругих* соударений, происходящих без изменения энергии ($E' = E$), вероятности прямых и обратных переходов равны

$$w_{kk'} = w_{k'k}. \quad (1.23)$$

Иногда именно это равенство используют как формулировку принципа детального равновесия. Однако оно является частным

случаем (1.20), пригодным лишь для переходов между состояниями с одинаковой энергией.

1.1.3. Приближение времени релаксации

Существенно упростить вид интеграла столкновений, а значит и кинетического уравнения, можно при использовании так называемого *приближения времени релаксации*. Фактически его суть состоит в допущении того, что изменение функции распределения за время dt пропорционально этой функции и, естественно, интервалу dt :

$$dN = N(\vec{r}, \vec{k}, t) dt / \tau. \quad (1.24)$$

Коэффициент пропорциональности τ , стоящий в знаменателе, имеет размерность времени и называется *временем релаксации*. По своему физическому смыслу и величине время релаксации близко к среднему времени свободного пробега в одночастичной модели дрейфа. Поэтому величина $1/\tau$ определяет вероятность того, что за единицу времени частица испытает столкновение, а вероятность столкновений за время dt равна dt/τ . С другой стороны, вероятность столкновения равна отношению числа частиц, испытавших столкновения за время dt , к общему количеству частиц dN/N . Равенство этих значений и составляет суть допущения (1.24). Оно также предполагает, что распределение электронов через некоторое время порядка τ не зависит от вида неравновесной функции распределения непосредственно перед столкновениями, частицы после столкновения “забывают” о своих первоначальных состояниях. Поэтому именно столкновения, хаотизируя движение, поддерживают термодинамическое равновесие, а всякая информация о неравновесном состоянии в результате столкновений теряется.

В соответствии с допущением (1.24) за единицу времени из \vec{k} -состояния в любые другие уходит $N(\vec{r}, \vec{k}, t)/\tau$ частиц. Но это то же самое количество, которое определяется вторым слагаемым в интеграле столкновений в форме (1.16), поэтому

$$(\partial N / \partial t)_c^- = N(\vec{r}, \vec{k}, t) / \tau = g(\vec{k}) \int w_{kk'} f(\vec{k}) (1 - f(\vec{k}')) dV_{k'}. \quad (1.25)$$

Первое слагаемое в (1.16) характеризует количество частиц, приходящих в состояние \vec{k} в результате столкновений из любых

\vec{k}' -состояний за единицу времени. В соответствии с приближением (1.24) это же значение равно N_0/τ , поэтому

$$(\partial N/\partial t)_c^+ = N_0/\tau = g(\vec{k}) \int w_{k'k} f(\vec{k}') (1 - f(\vec{k})) dV_{k'}. \quad (1.26)$$

Следует обратить внимание, что в последнем выражении стоит локально равновесная функция распределения N_0 , так как в состоянии \vec{k} частицы переходят непосредственно в результате столкновений, которые хаотизируют движение и способствуют релаксации распределения в равновесное состояние. Из (1.25) и (1.26) следует выражение для интеграла столкновений в приближении времени релаксации:

$$(\partial N/\partial t)_c = (N_0 - N)/\tau = g(\vec{k})(f_0 - f)/\tau, \quad (1.27)$$

и кинетическое уравнение (1.13) в этом приближении принимает вид

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \frac{\vec{F}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}} = - \frac{f - f_0}{\tau}, \quad (1.28)$$

где $f_0 = N_0/g(k)$ – равновесная функция распределения Ферми-Дирака.

Обоснованность приближения времени релаксации не всегда очевидна. Обычно им пользуются при относительно небольших отклонениях от положения равновесия, когда время релаксации можно считать постоянной величиной. Однако данный метод может иметь преимущества и в тех случаях, когда время релаксации является функцией E , а значения $\tau(E(k))$ могут быть предварительно определены экспериментально, либо вычислены. Последнее можно сделать на основе соотношения (1.25), если известны вероятности переходов $w_{kk'}$. Так как в (1.25) интегрирование производится по штрихованным переменным, то $f(k)$ можно вынести за знак интеграла и получить, что

$$\frac{1}{\tau(\vec{k})} = \int w_{k'k} (1 - f(\vec{k}')) dV_{k'}. \quad (1.29)$$

Для невырожденного коллектива частиц $f(k) \ll 1$ и подынтегральное выражение упрощается

$$1/\tau(\vec{k}) \equiv \lambda(\vec{k}) = \int w_{k'k} dV_{k'}. \quad (1.30)$$

Так как в дальнейшем мы встретимся с различными временами релаксации, то вместо обратного времени релаксации для того или иного вида рассеяния удобно использовать понятие так называемой *скорости рассеяния* $\lambda(\vec{k})$, которая также имеет смысл средней частоты столкновений.

1.1.4. Примеры решения кинетического уравнения

Статическая проводимость. В качестве примера использования приближения времени релаксации рассмотрим поведение функции распределения электронов в электрическом поле. При его отсутствии распределение волновых векторов и соответственно скоростей электронов будет равновероятным по всем направлениям (рис. 1.2, а) и их суммарная скорость равна 0, а функция распределения будет равновесной и симметричной относительно начала координат. Вычисление плотности тока по (1.5) через усреднение скорости по функции распределения также дает нулевой результат.

При наличии электрического поля, направленного навстречу оси x , электроны ускоряются и все точки в верхней части рис. 1.2, б,в, изображающие мгновенные значения скорости электронов, будут с течением времени равномерно смещаться в направлении действия силы поля. В момент столкновения скорость электронов становится случайной и точки на рисунке «перескакивают» в случайные места (пример соответствующего перехода обозначен стрелкой). Такие процессы приводят к двум противоположным тенденциям: под действием электрического поля функция распределения смещается вправо, а столкновения, хаотизируя движение электронов, стремятся вернуть функцию распределения в равновесное состояние. Очевидно, что в стационарном состоянии эти две тенденции уравнивают друг друга при некоторой смещенной функции $N = N_0 + \Delta N$ и $f = f_0 + \Delta f$. Усреднение скорости по (1.4) дает значение скорости дрейфа, причем вклад в значение интеграла дает только добавка $\Delta N = g(k) \cdot \Delta f$. Соответственно и величина плотности тока определяется также неравновесной добавкой

$$\vec{j}(\vec{r}) = -e \int \vec{v} \Delta N(\vec{r}, \vec{k}) dV_k = -e \int \vec{v} g(\vec{k}) \Delta f(\vec{r}, \vec{k}) dV_k. \quad (1.31)$$

Для нахождения неравновесной части функции распределения воспользуемся уравнением (1.28), которое при условии стационарности ($\partial f / \partial t = 0$), пространственной однородности поля ($\mathcal{E} = \text{const}$) и коллектива ($\partial f / \partial \vec{r} = 0$), принимает вид

$$\frac{\vec{F}}{\hbar} \cdot \frac{\partial f}{\partial \vec{k}} = -\frac{\Delta f}{\tau}. \quad (1.32)$$

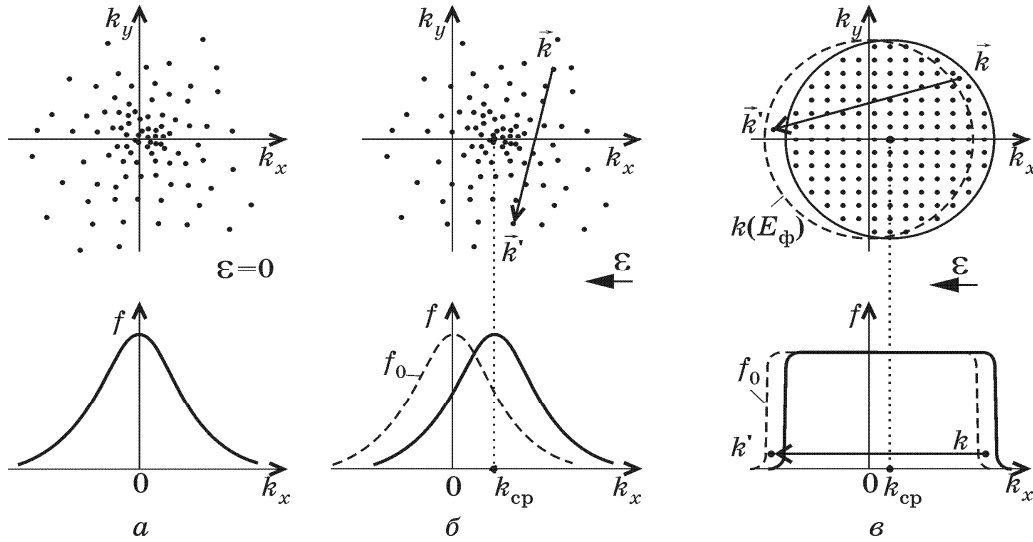


Рис. 1.2. При отсутствии электрического поля (а) равновероятное распределение волновых векторов (вверху) обуславливает симметричный характер функции распределения. В электрическом поле (б, в) смещение волновых векторов в направлении действия силы и рассеивание электронов в случайные состояния уравниваются при некоторой смещенной функции распределения. Надо обратить внимание на то, что в невырожденном коллективе (б) рассеивание вызывает переход электрона (стрелка) практически в любое состояние, а в вырожденном (в) — лишь в незаполненные состояния вблизи энергии Ферми.

Так как функция распределения f зависит от волнового вектора через закон дисперсии $E(\vec{k})$, то $\partial f / \partial \vec{k} = (\partial f / \partial E)(\partial E / \partial \vec{k}) = \hbar \vec{v} (\partial f / \partial E)$. Здесь использовано соотношение (1.22), определяющее скорость электрона из закона дисперсии. Если функция распределения не сильно отличается от равновесной ($\Delta f \ll f_0$), что наблюдается в слабых электрических полях, то в левой части (1.32) можно пренебречь членами второго порядка малости и считать, что $\partial f / \partial E \approx \partial f_0 / \partial E$. Тогда при $\vec{F} = -e\vec{\mathcal{E}}$

$$e\vec{v} \vec{\mathcal{E}} \partial f_0 / \partial E = \Delta f / \tau. \quad (1.33)$$

Отсюда и находится значение неравновесной части функции распределения в стационарном состоянии

$$\Delta f_{\text{ст}} = e\tau \vec{v} \vec{\mathcal{E}} \partial f_0 / \partial E. \quad (1.34)$$

Надо отметить, что выражение (1.33) не изменится и в том случае, если сила в (1.32) включает и силу Лоренца

$\vec{F}_л = -e[\vec{v} \times \vec{B}]$, так как $\vec{v}[\vec{v} \times \vec{B}] = 0$. Это означает, что наличие магнитного поля не оказывает непосредственного влияния на функцию распределения.

В итоге плотность тока (1.31) в слабом электрическом поле, как и ожидалось, пропорциональна напряженности поля:

$$\vec{j} = -e^2 \vec{\mathcal{E}} \int \tau v^2 g(\vec{k}) \frac{\partial f_0}{\partial E} dV_k. \quad (1.35)$$

В дальнейшем удобно перейти от интегрирования по k -пространству к интегрированию по энергии, используя соотношение $g(k)dV_k = g(E)dE$, где $g(E)$ при квадратичном законе дисперсии может быть выбрана в форме $g(E) = 4\pi(2m_a^*/h^2)^{3/2}(E - E_m)^{1/2}$ [1.1]. Кроме того, поскольку в (1.34) входит скалярное произведение $\vec{v} \vec{\mathcal{E}}$, в (1.35) вместо v^2 надо брать квадрат скорости движения в направлении вектора электрического поля. Поэтому $m^* v^2/2 = E/3$ и

$$\vec{j} = -e^2 \vec{\mathcal{E}} \int_0^\infty \frac{2\tau}{3m^*} E g(E) \frac{\partial f_0}{\partial E} dE. \quad (1.36)$$

Интегрирование здесь надо проводить по интересующей нас зоне с началом отсчета от ее дна. Если, например, зона проводимости включает несколько долин с эллипсоидальными изоэнергетическими поверхностями, различно ориентированными относительно направления вектора электрического поля, то интеграл будет состоять из суммы интегралов по каждой долине. Это позволит эффективную массу считать постоянной для каждой долины (в пределах квадратичности закона дисперсии) и вынести ее за знак интеграла. Если принять и время релаксации не зависящим от энергии, что обычно вполне допустимо в слабых электрических полях, то оставшееся подынтегральное выражение можно представить в виде $F(E)df_0$ и использовать формулу интегрирования по частям

$$\int_0^\infty F(E) df_0 = F(E) f_0 \Big|_0^\infty - \int_0^\infty f_0 dF = - \int_0^\infty f_0(E) g(E) dE.$$

Первое слагаемое обращается в 0, так как у дна зоны $g(E) = 0$, а при увеличении энергии $f_0(\infty) \rightarrow 0$. Второе слагаемое приобретает указанный вид вследствие того, что $dF/dE = g(E)$, а оно равно равновесной концентрации электронов.

В итоге при однодолинной модели зонной структуры получается прямая пропорциональность между плотностью тока и напряженностью электрического поля, известная как закон Ома

$$\vec{j} = \left(e^2 n \tau / m^* \right) \vec{\mathcal{E}} = \sigma \vec{\mathcal{E}}. \quad (1.37)$$

При наличии нескольких долин общий ток состоит из вкладов электронов каждой долины со своими параметрами τ и m^* . Иногда удается сохранить общий вид выражения (1.37), используя вместо m^* некоторую усредненную по разным долинам эффективную массу, называемую *омической*.

Нестационарная динамика дрейфа. Омическая зависимость между током и полем (1.37) получена для стационарного состояния. При изменении напряженности электрического поля наблюдаются переходные процессы, которые также можно исследовать с помощью кинетического уравнения в форме (1.28). Принимая, как и ранее, условие пространственной однородности электрического поля и функции распределения и заменяя третье слагаемое в левой части (1.28) на $\Delta f_{\text{ст}}/\tau$ согласно (1.34), получим

$$\partial(\Delta f)/\partial t + \Delta f/\tau = \Delta f_{\text{ст}}/\tau. \quad (1.38)$$

При включении электрического поля в момент времени t_0^+ (рис. 1.3) наблюдается постепенный переход из равновесного состояния ($\Delta f(t_0^+) = 0$) в стационарное состояние $\Delta f(\infty) = \Delta f_{\text{ст}}$ в соответствии с решением (1.38):

$$\Delta f(t) = \Delta f_{\text{ст}} \{1 - \exp[-(t - t_0^+)/\tau]\}. \quad (1.39)$$

При выключении поля ($\mathcal{E} = 0$) кинетическое уравнение (1.28) для неравновесной добавки Δf примет вид

$$\partial(\Delta f)/\partial t + \Delta f/\tau = 0. \quad (1.40)$$

Если выключение поля произошло в стационарном состоянии (рис. 1.3) в момент t_0^- , то есть $\Delta f(t_0^-) = \Delta f_{\text{ст}}$, то решением уравнения (1.40) будет функция

$$\Delta f(t) = \Delta f_{\text{ст}} \exp[-(t - t_0^-)/\tau], \quad (1.41)$$

которая показывает, что после выключения поля происходит *релаксация*, то есть возврат функции распределения в равновесное состояние.

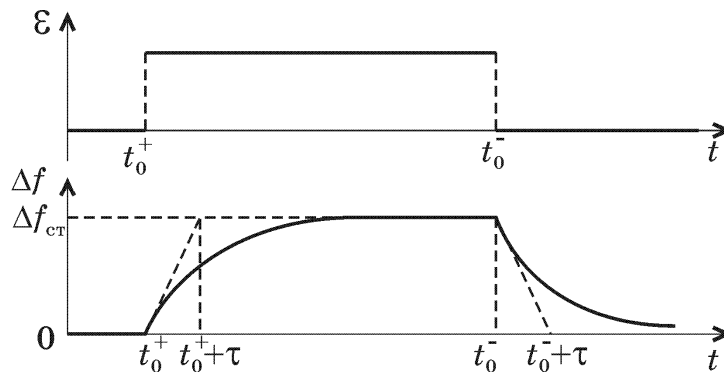


Рис. 1.3. Переходные процессы изменения неравновесной части функции распределения электронов и плотности их тока при включении и выключении электрического поля. Наклонные штриховые линии – касательные к кривым в моменты t_0^+ и t_0^- .

Подобно неравновесной добавке к функции распределения изменяются со временем дрейфовая скорость и плотность тока. Значение тока, определяемое законом Ома (1.37), достигается *не мгновенно* после включения электрического поля, так же как и исчезает после его выключения. Инерционность этого процесса определяется временем, необходимым для совершения соударений *всеми* частицами, в результате чего должна произойти хаотизация состояний, соответствующая равновесной функции распределения. Среднее значение этого времени и равно времени релаксации τ . Формально, как это видно из рис. 1.3, за время τ заканчивались бы переходные процессы, если бы их скорость оставалась такой же, как и вначале. В действительности время установления стационарного или равновесного состояния принимают равным $(3...5)\tau$. Характерные значения времени релаксации порядка $10^{-12}...10^{-14}$ с приблизительно равны времени свободного пробега в одночастичной модели дрейфа.

Следует отметить одну особенность релаксации функции распределения, характерную для вырожденных коллективов (рис. 1.2, в). В отличие от невырожденных, где вероятность заполнения состояний очень мала и возможны переходы (рис. 1.2, б) практически в любые состояния, в вырожденных коллективах все состояния внутри изоэнергетической поверхности Ферми заполнены и переходы туда запрещены принципом Паули. Поэтому при релаксации распределения в результате столкновений будут преобладать переходы из состояний \vec{k} , близких к $\vec{k}(E_\phi)$, где электроны ускоряются электрическим полем, в

состояния \vec{k}' , также близкие к $\vec{k}(E_{\text{ф}})$, но такие, где электроны тормозятся (рис. 1.2, в). Таким образом в релаксации участвуют практически лишь электроны в состояниях вблизи поверхности Ферми и соответственно в выражении (1.37) надо использовать время релаксации $\tau(E_{\text{ф}})$ именно для этой группы электронов.

Высокочастотная проводимость. Если электрическое поле, действующее на электроны, переменное и изменяется по закону $\mathcal{E}(t) = \mathcal{E}_m \exp(i\omega t)$, то в функции распределения от времени по такому же закону будет зависеть только неравновесная добавка $f(t) = f_0 + \Delta f \exp(i\omega t)$. Естественно, что это утверждение можно принять для слабого электрического поля, когда $\Delta f \ll f_0$. При тех же предположениях о пространственной однородности поля и коллектива частиц, что и в статическом поле (разд. 1.2.2), кинетическое уравнение принимает вид

$$i\omega\Delta f - \frac{e\vec{\mathcal{E}}_m}{\hbar} \frac{\partial f_0}{\partial \vec{k}} = -\frac{\Delta f}{\tau} \quad (1.42)$$

и может быть представлено в такой же форме, что и (1.32), если заменить величину τ на $\tau_{\omega} = \tau/(1 + i\omega\tau)$. Поэтому и все выводы, полученные для постоянного поля, могут быть использованы для случая переменного поля, если воспользоваться этой заменой.

Между током и напряженностью поля сохраняется пропорциональная зависимость, однако коэффициент пропорциональности, то есть удельная электропроводность σ_{ω} , становится комплексной величиной, зависящей от частоты по такому же закону, что и время релаксации τ_{ω} ,

$$\sigma_{\omega} = \frac{\sigma}{1 + i\omega\tau} = \frac{\sigma(1 - i\omega\tau)}{1 + \omega^2\tau^2} = \sigma' - i\sigma'' \quad (1.43)$$

Комплексный характер величины σ_{ω} означает, что между изменениями тока и поля имеется фазовое запаздывание на величину $\arctg \omega\tau$. Этот фазовый сдвиг обусловлен, однако, не инерционностью, связанной с массой частиц, а статистическим характером их столкновений. Проявляется этот эффект на частотах, период колебаний которых соизмерим со временем релаксации ($\omega\tau \geq 1$). При меньших частотах за время полупериода поля успевает устанавливаться значение неравновесной добавки к функции распределения, как и в статическом поле (1.34). На высоких частотах, кроме указанного фазового сдвига, за время ускоряющего полупериода поля добавка Δf не успевает достичь

этого стационарного значения, и величина плотности тока уменьшается соответственно уменьшению величины удельной электропроводности $|\sigma_\omega| = \sigma / \sqrt{1 + \omega^2 \tau^2}$.

Еще одна интересная особенность связана с мнимой частью σ_ω . Обратимся к уравнению Максвелла $\text{rot } \vec{H} = \partial \vec{D} / \partial t + \vec{j}$, которое в комплексной форме может быть представлено в виде $\text{rot } \vec{H} = (i\omega \epsilon \epsilon_0 + \sigma' - i\sigma'') \vec{E}$. Отсюда легко видеть, что мнимая часть электропроводности σ'' играет такую же роль, как и диэлектрическая проницаемость, что может заметно сказываться на оптических свойствах вещества. Однако, если первое слагаемое представляет собой ток смещения, имеющий емкостной характер, то третье слагаемое, вызванное инерционностью процессов релаксации и пропорциональное σ'' , имеет индуктивный характер.

1.2. МЕТОД МОНТЕ-КАРЛО

Метод Монте-Карло представляет собой численную реализацию решения кинетического уравнения. В общем случае с этим названием связывают многочисленную группу вычислительных методов, использующих случайные числа. Применительно к электронике этот метод, иногда называемый методом частиц [1.6], применяют для моделирования коллективных эффектов переноса зарядов в различных плазменных (твердотельных или газовых) структурах через поведение отдельных частиц: их баллистическое движение на свободном пробеге и рассеяние при столкновениях. Различают две разновидности: одночастичный и многочастичный методы Монте-Карло.

При одночастичном методе [1.3] моделируется движение одной частицы в течении промежутка времени, достаточно длинного для надежного статистического усреднения. В многочастичном методе рассматривается поведение большого количества частиц, а распределения интересующих параметров (энергии, скорости, концентрации и т.п.) находятся усреднением по ансамблю случайных реализаций этих параметров. Достаточно надежным для хорошего усреднения считается количество частиц порядка $10^4 \dots 10^5$. Очевидно, что в одночастичном методе внешние условия, то есть действующие силы, должны сохраняться постоянными на протяжении всего вычислительного эксперимента. Поэтому данная разновидность может быть использована для анализа кинетических эффектов лишь в статических полях.

Для анализа динамических процессов необходимо использовать *многочастичный* метод Монте-Карло. В нем, как отмечалось, электронная плазма в полупроводнике

представляется некоторым ансамблем модельных частиц, количество которых существенно меньше, чем реальных. Движение частиц моделируется полуклассическими уравнениями движения [1.1]:

$$\hbar d\vec{k}/dt = \vec{F}, \quad (1.44, a)$$

$$\vec{v} = \frac{1}{\hbar} \text{grad}_{\vec{k}} E(\vec{k}) = \frac{1}{\hbar} \left(\vec{e}_x \frac{\partial E}{\partial k_x} + \vec{e}_y \frac{\partial E}{\partial k_y} + \vec{e}_z \frac{\partial E}{\partial k_z} \right) \quad (1.44, б)$$

в реальном и \vec{k} -пространствах. Влияние пространственного заряда учитывают, решая согласованно с уравнениями движения уравнение Пуассона. Роль периодического электрического поля кристаллической решетки проявляется через закон дисперсии $E(\vec{k})$ с соответствующими эффективными массами. При движении частицы взаимодействуют с фононами, дефектами решетки и другими носителями, что приводит к изменению их волновых векторов и энергии. Эти эффекты учитывают как дискретные и мгновенные столкновения, в результате которых свободный пробег оканчивается рассеянием.

Основой построения алгоритма моделирования является временная диаграмма (рис. 1.4), которая представляет собой отрезки τ_i свободного движения одного электрона во внешнем поле, ограниченные моментами столкновений. Эта диаграмма накладывается на шкалу (обычно равномерную) временных интервалов T , через которые происходит перерасчет сил, действующих на частицы, и фиксируется состояние частиц.

Времена свободного пробега частиц τ_i определяются как случайные величины в соответствии с известным законом распределения по скорости рассеяния $\lambda(\vec{k})$ [1.3, 1.6]. Поскольку для разных видов соударений скорости рассеяния λ_i различны (см. разд. 2), то предварительно необходимо определить вид рассеяния также как случайную величину с вероятностью, пропорциональной скорости соответствующего вида рассеяния. В результате акта рассеяния изменяются величина и направление волнового вектора в соответствии с их законом распределения, а затем частица продолжает баллистическое движение до конца временного интервала T .

Типичный алгоритм кинетического моделирования можно представить следующим образом. Вначале задаются конфигурация прибора, профили легирования, закон дисперсии и параметры зонной структуры для данного материала, выбирается функция распределения $N(\vec{r}, \vec{k}, 0)$, то есть задается начальное распределение электронов, как пространственное, так и по \vec{k} -состояниям.

Пространственное распределение может быть задано, например, в соответствии с профилем распределения примеси. Кроме того, необходимо распределить электроны по долинам пропорционально плотности состояний в них. Поскольку для расчета сил, действующих на частицы, необходимо численно решить уравнение Пуассона, определяя потенциалы в узлах некоторой пространственной сетки, то, зная распределение частиц в пространстве, «размазывая» их тем или иным способом в некоторой локальной области, находят распределение плотности заряда. Выбирая метод численного интегрирования, для ускорения процедуры решения следует учитывать сильную разреженность матрицы соответствующей системы алгебраических уравнений.



Рис. 1.4. Характерные временные моменты в многочастичном методе Монте-Карло: шкала моментов пересчета поля и записи состояний частиц и шкала моментов актов рассеяний, между которыми моделируется свободное движение частиц.

После нахождения распределения электрического поля решаются уравнения движения поочередно для каждой частицы на протяжении ее свободного пробега τ_i , либо, если она не сталкивается на данном временном шаге, то на протяжении интервала T , а оставшаяся часть $\tau_i - T$ запоминается. Если же время свободного пробега $\tau_i < T$, то разыгрывается процедура рассеяния и находится новая траектория за время $T - \tau_i$. В момент времени T запоминаются все параметры движения частиц, то есть фиксируется распределение $N(\vec{r}, \vec{k}, T)$. Величина временного шага T должна быть достаточно большой, чтобы поле приходилось пересчитывать как можно реже, но вместе с тем и достаточно малой, чтобы не вызывать значительных численных ошибок. Обычно T выбирают таким, чтобы даже в области сильного поля частицы пролетали не более одной ячейки пространственной сетки.

Кроме параметров движения частиц на каждом временном шаге фиксируется количество частиц, поглощенных или инжектированных каждым электродом, а также вычисляется

нормальная к электроду составляющая напряженности электрического поля, что позволяет определить токи в цепи электродов.

С аппроксимации нового распределения плотности заряда, решения уравнения Пуассона и разыгрывания методом Монте-Карло времен свободного пробега начинается следующий цикл анализа движения частиц. Циклы вычислений повторяются до того момента времени, до которого интересует развитие процесса. Получающаяся последовательность распределений $N(\vec{r}, \vec{k}, nT)$ имитирует эволюцию реальной функции распределения на данном интервале времени. Одним из важнейших фрагментов алгоритма моделирования является блок статистической обработки результатов, которые могут усредняться как на каждом временном шаге, так и за более длинные интервалы времени. Следует также отметить проблему восприятия весьма больших массивов информации, получающихся в результате моделирования. Поэтому большое внимание следует уделить визуализации данных.

Для иллюстрации особенностей метода Монте-Карло и его возможностей приведем некоторые результаты моделирования физических процессов в арсенид-галлиевом полевом транзисторе с затвором в виде барьера Шоттки. Типичное представление непосредственных результатов моделирования траекторий и поля изображено на рис. 1.5 на фоне структуры транзистора в виде пространственного распределения потенциала и расположения частиц в фиксированный момент времени. Это можно сделать отдельно для разных долин, что наглядно представляет возбуждение электронов в верхние долины лишь в области сильного поля у стоковой части подзатворной области.

Подобные распределения можно сделать для скорости и энергии частиц (рис. 1.6), как в виде «индивидуальных» так и усредненных распределений. Кроме того, усреднять можно как по всей толщине транзистора, так и в отдельных слоях. При сравнении рисунков для скорости и энергии можно обратить внимание на несовпадение пространственного расположения максимумов этих кривых. Это объясняется тем, что повышение энергии сопровождается переходом электронов в верхние долины, где их эффективная масса увеличивается и соответственно скорость уменьшается.

Важнейшей особенностью многочастичного метода Монте-Карло является возможность моделирования динамических физических процессов. На рис. 1.7 представлены примеры распределений вдоль канала транзистора усредненных значений напряженности электрического поля, энергии и скорости в различные моменты времени в некотором переходном режиме. Эти распределения наряду с пространственными распределени-

ями частиц, подобными рис. 1.5 в те же временные моменты, могут дать обширную информацию о достаточно тонких динамических особенностях физических процессов в моделируемой структуре.

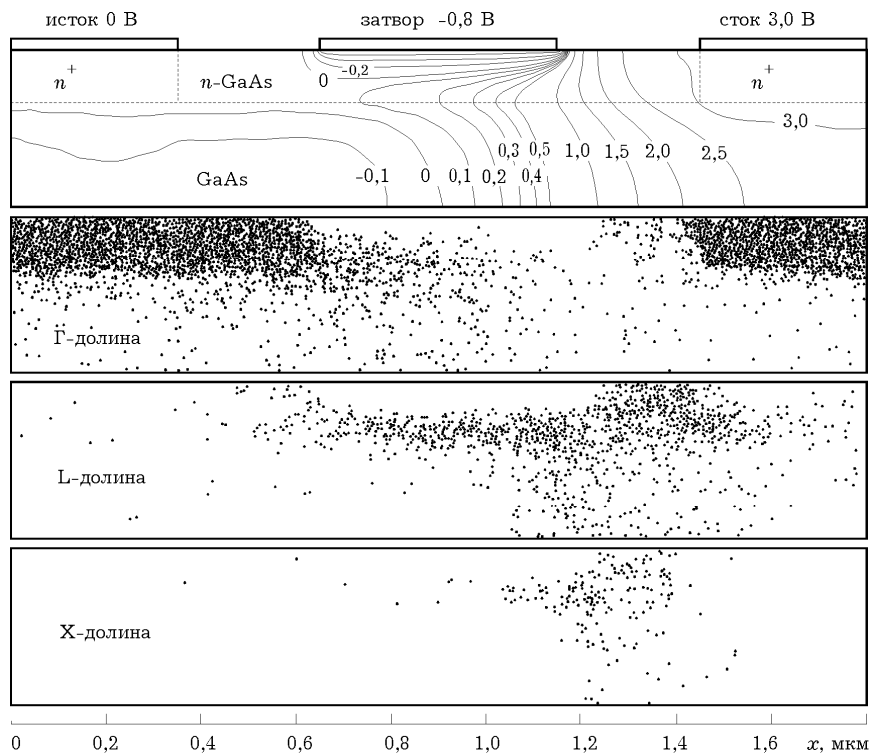


Рис. 1.5. Типичное представление распределения поля с помощью эквипотенциальных линий и распределения частиц в пространстве моделируемой структуры. Пространственное распределение частиц приведено отдельно для каждой долины. Штриховой линией обозначена нижняя граница эпитаксиального слоя с высоким легированием.

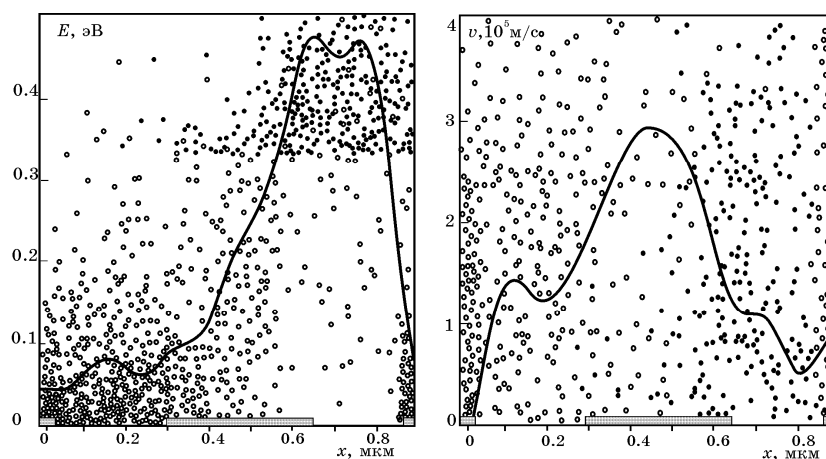


Рис. 1.6. Распределения энергии и скорости электронов вдоль канала полевого транзистора в форме «индивидуального» расположения частиц и усредненных функций. Светлые кружки соответствуют электронам нижней долины, темные – верхних.

Рассчитывая значения токов электродов при заданных на них потенциалах в установившемся режиме, можно получить

статические вольтамперные характеристики приборов. Если же на электроды подавать ступенчатые импульсы, то по переходным характеристикам, пример которых приведен на рис. 1.8, стандартными методами вычисляются частотные характеристики моделируемой структуры [1.6].

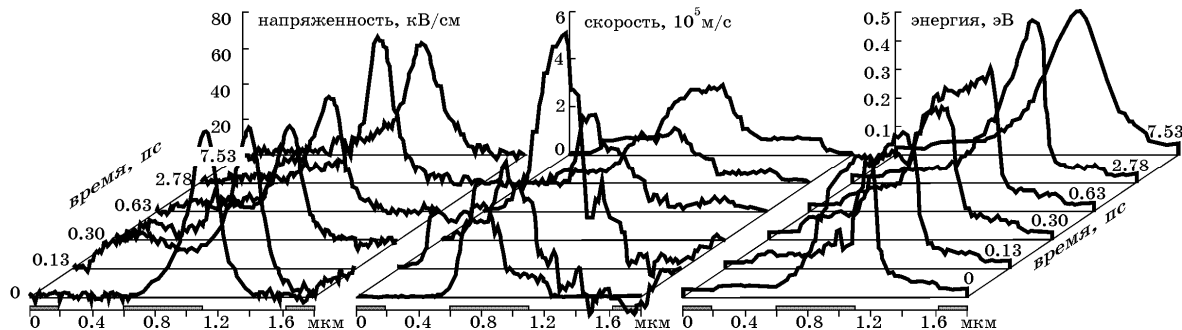


Рис. 1.7. Распределения напряженности поля, средней скорости и энергии частиц вдоль канала транзистора в различные моменты времени переходного процесса.

На представленных рисунках видна одна из характерных особенностей метода Монте-Карло: наряду с регулярными процессами переноса здесь естественным образом моделируются и шумы приборов. Это несомненное достоинство метода требует однако дополнительных вычислительных затрат, связанных с определением средних значений на фоне флуктуаций.

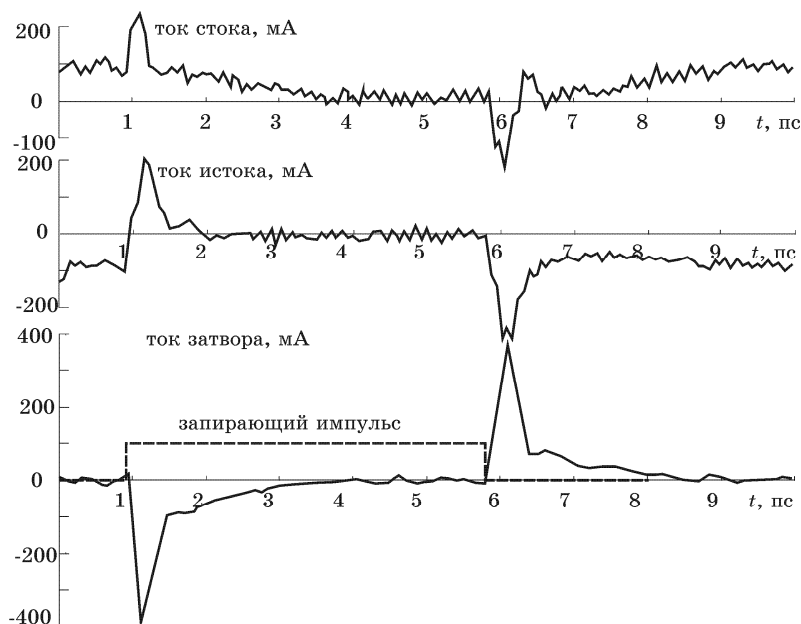


Рис. 1.8. Переходные характеристики токов стока, истока и затвора при подаче на затвор ступенчатого импульса: запирающего и открывающего.

Применительно к твердотельным структурам метод Монте-Карло особенно удобен, когда характерные размеры структур

составляют микрометры и менее. Для более крупных требуется моделировать поведение большого количества частиц на протяжении больших временных интервалов, что требует дополнительных ресурсов. Для более коротких необходимо учитывать ряд дополнительных эффектов, среди которых конечное время столкновений, квантовые эффекты и т.п.

1.3. РЕЛАКСАЦИОННЫЕ УРАВНЕНИЯ

1.3.1. Усреднение кинетического уравнения

Приближение времени релаксации, существенно упрощая кинетическое уравнение, облегчает нахождение функции распределения при известных механизмах рассеяния и значениях их скорости. Но получаемая при этом информация часто бывает избыточной, если требуется знать лишь некоторые усредненные величины или их пространственные распределения. Наиболее важными характеристиками коллектива обычно являются концентрация, скорость и энергия частиц. Более просто и экономно находить эти величины или их пространственно-временные функции не через функцию распределения, а из дифференциальных уравнений для этих величин. Такие уравнения могут быть получены непосредственно из кинетического уравнения, если провести усреднение до его решения. Для этого все слагаемые уравнения (1.12) умножаются на соответствующую усредняемую величину и интегрируются по всей или интересующей нас области k -пространства. В результате получаются дифференциальные уравнения в частных производных для средних величин, зависящих только от координат и времени. В литературе [1.5] такие уравнения встречаются в различных формах, отличающихся использованием тех или иных, зачастую достаточно тонких, допущений, связанных в основном с выбором модели (гидродинамическая, квазигидродинамическая и др.) коллектива частиц. Отчасти по этой причине такие уравнения называют *феноменологическими*, а по своей сути они являются уравнениями сохранения.

Самая простая форма подобных уравнений может быть получена, если перед усреднением первые два слагаемых (1.12), включающих частные производные по координатам и времени, свернуть в «полную» производную по этим переменным обратно тому, как это делалось в разделе 1.1.1, учитывая, что после интегрирования по k -пространству все величины не будут зависеть от \vec{k} . Если усредняемая величина $\varphi(\vec{k})$ не зависит явным

образом от \vec{r} и t , то ее можно внести под знак частных производных по этим переменным, а интеграл по выбранному объему k -пространства от этих слагаемых, умноженных на $\varphi(\vec{k})$, с учетом (1.6) примет вид:

$$\begin{aligned} \int_{V_k} \left(\frac{\partial N}{\partial t} + \vec{v} \frac{\partial N}{\partial \vec{r}} \right) \varphi(\vec{k}) d\vec{k} &= \frac{d}{dt} \int_{V_k} N(\vec{r}, \vec{k}, t) \varphi(\vec{k}) d\vec{k} = \\ &= \frac{d}{dt} [n(\vec{r}, t) \varphi_{\text{ср}}(\vec{r}, t)]. \end{aligned} \quad (1.45)$$

Третье слагаемое в кинетическом уравнении представляет собой dN/dt при фиксированных значениях \vec{r} и t . Интегрируя его по V_k и учитывая, что $d\vec{k}/dt = \vec{F}/\hbar$ не зависит от k , получим

$$\int_{V_k} (dN/dt) d\vec{k} = \int_{V_k} (d\vec{k}/dt) dN = (\vec{F}/\hbar) N|_{V_k}. \quad (1.46)$$

Если функция распределения на краях области интегрирования стремится к нулю, что часто выполняется за счет экспоненциального характера содержащихся в ней функций, то $N|_{V_k} = 0$.

Умножая перед интегрированием третье слагаемое на \vec{k} и используя формулу интегрирования по частям и (1.3), можно получить

$$\int_{V_k} \vec{k} \frac{dN}{dt} d\vec{k} = \frac{\vec{F}}{\hbar} \int_{V_k} \vec{k} dN = \frac{\vec{F}}{\hbar} \left[\vec{k} N|_{V_k} - \int_{V_k} N d\vec{k} \right] = -\frac{\vec{F}}{\hbar} n(\vec{r}, t). \quad (1.47)$$

Подобным образом преобразуется третье слагаемое, умноженное на $E(\vec{k})$ с учетом того, что $dE/d\vec{k} = \hbar \vec{v}$:

$$\begin{aligned} \int_{V_k} E(\vec{k}) \frac{dN}{dt} d\vec{k} &= \frac{\vec{F}}{\hbar} \int_{V_k} E(\vec{k}) dN = \\ &= \frac{\vec{F}}{\hbar} \left[E(\vec{k}) N|_{V_k} - \int_{V_k} N \frac{dE}{d\vec{k}} d\vec{k} \right] = -\vec{F} n(\vec{r}, t) \vec{v}_{\text{ср}}(\vec{r}, t). \end{aligned} \quad (1.48)$$

Проведение аналогичных процедур интегрирования правой части кинетического уравнения дает скорости изменения за счет столкновений соответственно концентрации $-(dn/dt)_c$, суммарного

значения волнового вектора $-(d\vec{k}_\Sigma/dt)_c$ и полной энергии $-(dE_\Sigma/dt)_c$. В итоге, подставляя вместо φ в (1.45) 1, \vec{k} или E , получим соответствующие уравнения сохранения для концентрации

$$dn/dt = (dn/dt)_c, \quad (1.49)$$

среднего значения волнового вектора

$$d(n\vec{k}_{cp})/dt - n\vec{F}/\hbar = (d\vec{k}_\Sigma/dt)_c \quad (1.50)$$

и средней энергии

$$d(nE_{cp})/dt - n\vec{F}\vec{v}_{cp} = (dE_\Sigma/dt)_c. \quad (1.51)$$

Предпоследнее уравнение с учетом $\vec{p} = \hbar\vec{k}$ обычно записывается в виде уравнения сохранения импульса

$$d(n\vec{p}_{cp})/dt - n\vec{F} = (d\vec{p}_\Sigma/dt)_c. \quad (1.52)$$

1.3.2. Релаксационные уравнения

Если использовать для кинетического уравнения приближение *времени релаксации*, то правые части (1.49), (1.51) и (1.52) можно представить в форме

$$\left(\frac{dn}{dt}\right)_c = -\int_{V_k} \frac{(N - N_0)}{\tau} d\vec{k} = -\frac{\Delta n}{\tau_n}, \quad (1.53)$$

$$\left(\frac{d\vec{p}_\Sigma}{dt}\right)_c = -\hbar \int_{V_k} \frac{\vec{k}(N - N_0)}{\tau} d\vec{k} = -\frac{n\vec{p}_{cp}}{\tau_p}, \quad (1.54)$$

$$\left(\frac{dE_\Sigma}{dt}\right)_c = -\int_{V_k} \frac{E(N - N_0)}{\tau} d\vec{k} = -\frac{n(E_{cp} - E_0)}{\tau_E}. \quad (1.55)$$

Подобно тому, как в правой части кинетического уравнения стоит неравновесная добавка к функции распределения $\Delta N = N - N_0 = g(k) \Delta f$, так и в числителе правых частей стоят отклонения от равновесных значений для концентрации $\Delta n = n - n_0$, среднего импульса ($\vec{p}_{cp0} = 0$) и средней энергии ($E_0 = 3k_B T/2$). Величины в

знаменателе называются временами релаксации *концентрации* (τ_n), *импульса* (τ_p), и *энергии* (τ_E).

Уравнение (1.52) при сохранении общего количества частиц в зоне может быть записано в форме уравнения сохранения для среднего дрейфового значения импульса $\bar{p}(\vec{r}, t)$ (индекс среднего значения опускаем)

$$d\bar{p}/dt - \vec{F} = -\bar{p}/\tau_p, \quad (1.56)$$

или для дрейфовой скорости

$$d\vec{v}/dt - \vec{F}/m^* = -\vec{v}/\tau_p, \quad (1.57)$$

которое имеет простой физический смысл: скорость направленного движения увеличивается под действием силы электрического поля, как и для свободных частиц, и уменьшается за счет рассеяния носителей, или «силы трения» $-\bar{p}/\tau_p$. В статическом электрическом поле эти силы уравновешиваются и

$$\vec{v} = -(e\tau_p/m^*)\vec{E} = -\mu\vec{E}, \quad (1.58)$$

то есть подвижность определяется средним временем релаксации импульса.

В соответствии с (1.54) это время находится через усреднение обратного времени релаксации $1/\tau = \lambda_\Sigma = \Sigma_i \lambda_i$, умноженного на волновой вектор, по функции распределения

$$\frac{\bar{p}}{\tau_p} = \frac{\hbar}{n} \int_{V_k} \vec{k} \lambda_\Sigma N(\vec{k}) d\vec{k}. \quad (1.59)$$

Наличие здесь суммарной скорости рассеяния, включающей все механизмы, присущие данному материалу, означает, что любое рассеяние, упругое или неупругое, с поглощением фонона или его испусканием, приводит в той или иной мере к потере направленного импульса и хаотизации движения.

Совершенно аналогичные соображения можно провести в связи с уравнением сохранения энергии. При неизменном общем количестве частиц в зоне уравнение (1.51) для *средней* энергии в приближении времени релаксации принимает вид (индекс среднего опускаем)

$$dE/dt - \vec{F}\vec{v} = -(E - E_0)/\tau_E. \quad (1.60)$$

Для нахождения времени релаксации энергии τ_E используется усреднение (1.55) скорости рассеяния по функции распределения, умноженной на энергию,

$$\frac{E}{\tau_E} = \frac{1}{n} \int_{V_k} E(\vec{k}) \lambda_{\Sigma} N(\vec{k}) d\vec{k}, \quad (1.61)$$

причем λ_{Σ} включает все механизмы рассеяния, кроме упругих.

Следует отметить некоторое несоответствие выражений (1.54), (1.55) с одной стороны и (1.59), (1.61) с другой. Однако это различие незначительно по самой сути приближения времени релаксации, когда отклонение от равновесного состояния невелико, а время релаксации τ – постоянно.

1.3.3. Междолинная релаксация

В многодолинных полупроводниках электроны занимают нижние состояния долин и образуют относительно обособленные коллективы со своими характерными параметрами. Поэтому усреднение кинетического уравнения можно проводить по объемам k -пространства, включающим отдельные долины. При этом, если даже общее количество электронов в зоне остается постоянным, то так называемое междолинное рассеяние может приводить к перераспределению электронов между отдельными долинами и соответственно к изменению средних значений импульса и энергии.

Рассмотрим для упрощения двухдолинную модель зоны проводимости (рис. 1.9). Нахождение уравнений сохранения для электронов i -долины предполагает интегрирование слагаемых кинетического уравнения по исходным k -состояниям только этой долины. Конечными k' -состояниями, входящими в интеграл столкновений (1.16), могут быть как внутренние состояния i -долины, так и состояния другой (других) j -долины. Учет только внутренних переходов из k - в k' -состояния (внутридолинных видов рассеяния) приведет к таким

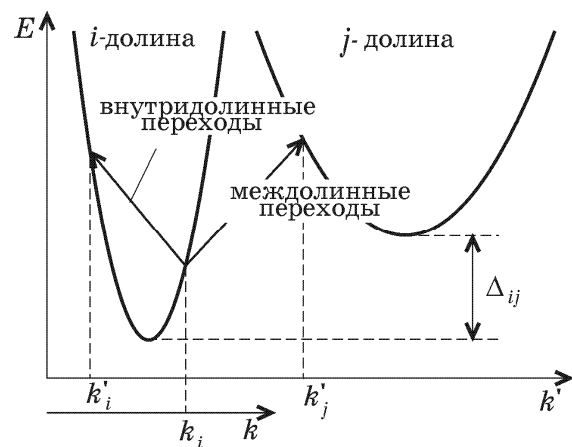


Рис. 1.9. Обозначение внутри- и междолинных переходов электронов из состояния k в k' в двухдолинной модели зонной структуры.

же по форме уравнениям сохранения, что и для однодолинной модели. Для нахождения дополнительных слагаемых, обусловленных междолинным рассеянием, в интеграле столкновений из всех k' -состояний надо оставить состояния только j -долины.

В таком случае интеграл столкновений (1.16) при процедуре усреднения, описанной в разделе 1.3.1 при $\phi = 1$, по k -пространству принимает вид

$$[dn/dt]_{c\ ij} = - \int_{V_k} \int_{V_{k'}} [w_{ij}N(k) - w_{ji}N(k')] dV_{k'} dV_k. \quad (1.62)$$

Множители $(1 - f(k))$ и $(1 - f(k'))$ в (1.16) фактически запрещают переходы в занятые состояния и для формального упрощения могут быть включены в dV_k и $dV_{k'}$, если под конечными состояниями понимать только свободные. Первое слагаемое, которое характеризует уход электронов из i -долины, с учетом (1.30) удобно представить в виде

$$- \int_{V_k} \left[\int_{V_{k'}} w_{ij} dV_{k'} \right] N(k) dV_k = - \int_{V_k} \tau_{ij}^{-1} N(k) dV_k = -n_i / \tau_{nij}. \quad (1.63)$$

Здесь τ_{nij} — время междолинной релаксации *концентрации*, усредненное по функции распределения только i -долины:

$$1/\tau_{nij} = \int_{V_k} \tau_{ij}^{-1} N(k) dV_k / \int_{V_k} N(k) dV_k. \quad (1.64)$$

Второе слагаемое, характеризующее приход электронов в i -долину, представим подобным образом, поменяв порядок интегрирования

$$\int_{V_{k'}} \left[\int_{V_k} w_{ji} dV_k \right] N(k') dV_{k'} = - \int_{V_{k'}} \tau_{ji}^{-1} N(k') dV_{k'} = n_j / \tau_{nji}. \quad (1.65)$$

Здесь τ_{nji} — также концентрационное время междолинной релаксации электронов из j -долины в i -ю, но усредненное по функции распределения j -долины:

$$1/\tau_{nji} = \int_{V_{k'}} \tau_{ji}^{-1} N(k') dV_{k'} / \int_{V_{k'}} N(k') dV_{k'}. \quad (1.66)$$

В итоге получаем уравнение сохранения для концентрации электронов в i -долине

$$dn_i/dt = - n_i / \tau_{nij} + n_j / \tau_{nji} \quad (1.67)$$

и аналогичное уравнение для j -долины

$$dn_j/dt = -n_j/\tau_{nji} + n_i/\tau_{nij}. \quad (1.68)$$

Эти уравнения представляют собой систему уравнений и естественно, что правые части у них отличаются только знаком, так как общее количество частиц в зоне остается неизменным

$$d(n_i + n_j)/dt = 0. \quad (1.69)$$

Полезно также отметить, что в стационарном состоянии ($dn_i/dt = dn_j/dt = 0$) междолинные переходы взаимно уравновешиваются и с учетом равновесной статистики [1.1]

$$\tau_{nij}/\tau_{nji} = n_i/n_j = (m_{di}^*/m_{dj}^*)^{3/2} \exp(-\Delta_{ij}/k_B T_e). \quad (1.70)$$

Подобные рассуждения можно провести и для уравнения сохранения импульса (волнового вектора). Умножая все слагаемые кинетического уравнения на \vec{k} и интегрируя по части k -пространства, содержащей состояния только i -долины, мы получим уравнение сохранения импульса в форме (1.50) или (1.52) с правой частью (1.54), в которой τ_p учитывает только внутридолинные виды рассеяния. Оставшаяся часть интеграла междолинных столкновений имеет структуру, подобную (1.62), но подынтегральное выражение умножается на \vec{k} . Первое слагаемое, характеризующее уменьшение импульса в i -долине за счет ухода из нее электронов, можно представить подобно (1.63) в виде

$$-\int_{V_k} \left[\int_{V_{k'}} w_{ij} dV_{k'} \right] \vec{k} N(k) dV_k = -\int_{V_k} \tau_{ij}^{-1} \vec{k} N(k) dV_k = -n_i \vec{k}_{i \text{ ср}} / \tau_{pij}. \quad (1.71)$$

Здесь τ_{pij} — время междолинной релаксации импульса, которое определяется усреднением

$$1/\tau_{pij} = \int_{V_k} \tau_{ij}^{-1} \vec{k} N(k) dV_k / \int_{V_k} \vec{k} N(k) dV_k. \quad (1.72)$$

Второе слагаемое, характеризующее перенос импульса электронов из j -долины в i -ю, равно 0, так как после рассеяния эти электроны принимают случайные значения волнового вектора, симметрично расположенные относительно $k = 0$. В итоге для уравнения баланса импульса в i -долине можно сохранить форму (1.52) с правой частью (1.54)

$$d(n_i \vec{p}_i)/dt - n_i \vec{F} = -n_i \vec{p}_i / \tau_{pi}, \quad (1.73)$$

но в τ_{pi} необходимо включить как внутривалинные переходы, так и межвалинные

$$1/\tau_{pi} = 1/\tau_{pii} + 1/\tau_{pij} . \quad (1.74)$$

Для j -валины уравнение сохранения импульса аналогично (1.73) при соответствующей замене индексов:

$$d(n_j \bar{p}_j)/dt - n_j \bar{F} = - n_j \bar{p}_j / \tau_{pj} , \quad (1.75)$$

с временем релаксации импульса, учитывающим как внутри- так и межвалинные переходы из j -валины в i -ю

$$1/\tau_{pj} = 1/\tau_{pjj} + 1/\tau_{pji} . \quad (1.76)$$

Уравнения (1.73) и (1.75) представляют собой систему уравнений, которая к тому же должна решаться совместно с системой (1.67) и (1.68), характеризующей заселенность валин.

Подобным образом можно провести преобразование интеграла столкновений, умноженного на E , и получить уравнение сохранения энергии для i -валины. По виду оно будет совпадать с (1.51) с правой частью в форме (1.55) со временем релаксации энергии τ_{Eii} , учитывающим только внутривалинные механизмы рассеяния. Межвалинное рассеяние приводит к появлению в правой части дополнительных слагаемых, одно из них ($-n_i E_i / \tau_{Eij}$) характеризует уменьшение суммарной энергии электронов в i -валине за единицу времени при их переходе в j -валину и связано с межвалинным временем релаксации энергии τ_{Eij} , которое определяется соотношением

$$1/\tau_{Eij} = \int_{V_k} \tau_{ij}^{-1} E(k) N(k) dV_k / \int_{V_k} E(k) N(k) dV_k . \quad (1.77)$$

Здесь обратное время релаксации τ_{ij}^{-1} находится интегрированием вероятности переходов $w_{kk'}$ по k' -состояниям j -валины

$$\tau_{ij}^{-1} = \int_{V_{k'}} w_{kk'} dk' . \quad (1.78)$$

а усреднение в (1.77) производится по k -состояниям i -валины.

При обратных переходах электроны из j -валины в i -ю приносят энергию, которую удобно представить двумя слагаемыми в правой части уравнения баланса энергии. Одно из них ($n_j E_0 / \tau_{Eji}$) пропорционально концентрации n_j и равновесной энергии E_0 , так

как электроны приходят в i -ю долину непосредственно в результате релаксирующих столкновений. Время релаксации τ_{Eji} определяется аналогично (1.77), но усреднение производится для скорости рассеяния $1/\tau_{ji}$ по состояниям j -долины.

Другое слагаемое $(n_j\Delta_{ji}/\tau_{nji})$ связано с междолинным расстоянием $\Delta_{ji} = E_j - E_i$, то есть с разным началом отсчета кинетической энергии в разных долинах. Примечательно, что скорость уменьшения этой энергии определяется концентрационным временем релаксации τ_{nji} (1.66).

В результате система уравнений, описывающих релаксацию энергии в двухдолинной модели, может быть представлена в виде:

$$\frac{d(n_i E_i)}{dt} - n_i \vec{v}_i \vec{F} = -\frac{n_i(E_i - E_0)}{\tau_{Ei}} - \frac{n_i E_0}{\tau_{Eij}} + \frac{n_j E_0}{\tau_{Eji}} + \frac{n_j \Delta_{ji}}{\tau_{nji}}, \quad (1.79)$$

$$\frac{d(n_j E_j)}{dt} - n_j \vec{v}_j \vec{F} = -\frac{n_j(E_j - E_0)}{\tau_{Ej}} - \frac{n_j E_0}{\tau_{Eji}} + \frac{n_i E_0}{\tau_{Eij}} - \frac{n_i \Delta_{ji}}{\tau_{nij}}. \quad (1.80)$$

Здесь времена релаксации τ_{Ei} и τ_{Ej} включают как внутри-, так и междолинные механизмы рассеяния:

$$1/\tau_{Ei} = 1/\tau_{Eii} + 1/\tau_{Eij}, \quad 1/\tau_{Ej} = 1/\tau_{Ejj} + 1/\tau_{Eji}. \quad (1.81)$$

Если сложить уравнения (1.79), (1.80) и разделить их на суммарную концентрацию $n = n_i + n_j$, а также ввести средние по долинам значения энергии $E = (n_i E_i + n_j E_j)/n$, скорости дрейфа $\vec{v} = (n_i \vec{v}_i + n_j \vec{v}_j)/n$ и времени релаксации энергии $1/\tau_E = [n_i(E_i - E_0)/\tau_{Ei} + n_j(E_j - E_0)/\tau_{Ej}]/n$, то можно получить уравнение сохранения энергии электрона, средней для двух долин

$$\frac{dE}{dt} - \vec{v} \vec{F} = -\frac{E - E_0}{\tau_E} + \frac{\Delta_{ji}}{n} \left(\frac{n_j}{\tau_{nji}} - \frac{n_i}{\tau_{nij}} \right). \quad (1.82)$$

Это уравнение отличается от (1.60) дополнительным слагаемым, которое связано с релаксацией избыточной энергии, возникающей при переходах электронов между неэквивалентными долинами с разным уровнем начала отсчета кинетической энергии. При переходе электрона из более высокой j -долины в i -ю его кинетическая энергия увеличивается на величину Δ_{ji} (рис. 1.9), что уменьшает скорость ее релаксации. Обратные переходы уменьшают энергию и соответственно ускоряют релаксационные процессы.

Надо обратить внимание на то, что это слагаемое проявляется только при перераспределении концентраций по долинам. В стационарном состоянии ($dn_i/dt = dn_j/dt = 0$) в соответствии с (1.67) и (1.68) слагаемые в скобках равны друг другу и эти эффекты взаимно уравниваются. Но в динамических ситуациях это явление может быть весьма заметным, так как междолинное расстояние Δ_{ji} (например, в арсениде галлия оно равно $\approx 0,3$ эВ) обычно на порядок превышает равновесное значение энергии.

Иногда общее среднее время релаксации энергии определяют как

$$1/\tau_E = (n_i/\tau_{Ei} + n_j/\tau_{Ej})/n, \quad (1.83)$$

где весовыми коэффициентами являются только концентрации. Это вполне согласуется с приближением времени релаксации (разд. 1.1.3), где время τ может быть найдено из равновесной функции распределения (1.26).

1.4. УРАВНЕНИЯ ДРЕЙФОВО-ДИФFUЗИОННОЙ МОДЕЛИ

В практике моделирования динамических свойств приборов широко используется так называемая *гидродинамическая модель*, основанная на представлении электронного коллектива в виде «несжимаемой жидкости», для которой можно принять $\text{div } \vec{v} = 0$.

Если область интегрирования V_k при усреднении кинетического уравнения содержит состояния только одной зоны, например, зоны проводимости, а межзональные генерационно-рекомбинационные переходы уравниваются друг друга, то $(dn/dt)_c = 0$ и полное число частиц в зоне сохраняется

$$\frac{dn(\vec{r}, t)}{dt} = \frac{\partial n}{\partial t} + \frac{d\vec{r}}{dt} \cdot \frac{\partial n}{\partial \vec{r}} = \frac{\partial n}{\partial t} + \vec{v} \cdot \text{grad } n = 0. \quad (1.84)$$

Если учесть тождество $\text{div}(n\vec{v}) \equiv \vec{v} \cdot \text{grad } n + n \text{div } \vec{v}$, отсюда можно получить уравнение

$$\partial n(\vec{r}, t) / \partial t + \text{div}(n\vec{v}) = 0, \quad (1.85)$$

которое (при умножении на заряд электрона) известно как *уравнение непрерывности тока* или *закон сохранения заряда*:

$$\partial \rho / \partial t + \text{div } \vec{j} = 0. \quad (1.86)$$

Однако в отличие от фундаментального уравнения непрерывности сюда входят плотности заряда и тока, создаваемые только подвижными электронами зоны проводимости.

Если равновесное состояние нарушено и скорости генерации $G_n = (dn/dt)_c^+$ и рекомбинации $R_n = (dn/dt)_c^-$ различны, то уравнение непрерывности для подвижных электронов в этих условиях обычно записывают в виде

$$\partial n(\vec{r}, t)/\partial t + \operatorname{div}(\vec{j}_n/e) = G_n - R_n. \quad (1.87)$$

Аналогичная форма уравнения может быть получена и для дырок валентной зоны

$$\partial p(\vec{r}, t)/\partial t + \operatorname{div}(\vec{j}_p/e) = G_p - R_p. \quad (1.88)$$

В общем случае электронные и дырочные токи определяются дрейфовым и диффузионным механизмами

$$\begin{aligned} \vec{j}_n &= en\mu_n\vec{E} + eD_n \operatorname{grad} n, \\ \vec{j}_p &= ep\mu_p\vec{E} - eD_p \operatorname{grad} p. \end{aligned} \quad (1.89)$$

Уравнения (1.87) и (1.88), включающие (1.89), называют *уравнениями непрерывности для подвижных носителей*. Совместно с уравнением Пуассона

$$\nabla^2 V = -\rho(\vec{r})/\varepsilon\varepsilon_0 = e(\Delta n - \Delta p)/\varepsilon\varepsilon_0 \quad (1.90)$$

и описанием механизмов генерации и рекомбинации они представляют систему уравнений для анализа поведения неравновесных носителей в полупроводниковой плазме. Если скорости генерации и рекомбинации представить в виде (1.53), то времена релаксации τ_n и τ_p называют *временами жизни* неравновесных носителей.

Система уравнений дрейфово-диффузионной модели позволяет рассчитать распределения концентрации носителей и электрического поля, а по ним значения скорости носителей, распределение плотности тока и соответственно значения тока электродов прибора. Для сильных электрических полей такая модель имеет ограниченные возможности. В некоторых случаях можно использовать в том или ином виде аппроксимации зависимостей подвижности и коэффициента диффузии от напряженности электрического поля. Но по своей сути эта система является «локальной» моделью токопереноса, в которой скорость

носителей определяется локальным значением напряженности поля. Она не может описывать эффекты всплеска дрейфовой скорости, эффекты накопления носителей из-за изменения их энергии и эффективной массы при переходе в верхние долины.

Модели традиционных электронных компонентов относятся в основном к классу дрейфово-диффузионных. С их помощью можно достаточно точно рассчитывать характеристики устройств традиционной микроэлектроники. Однако при выводе уравнений этих моделей учитывалось, что характерные размеры компонентов превышают длину свободного пробега в сотни и более раз. Если же это не выполняется, то сами понятия дрейфа и диффузии теряют смысл, использование моделей становится необоснованным, а их результаты перестают согласовываться с экспериментальными данными.

Все понятия, которые используются в дрейфово-диффузионных моделях, имеют интуитивно понятный физический смысл, ведь в рамках указанных моделей волновые свойства электрона учтены эффективно: взаимодействие с периодическим полем кристалла учтено в эффективной массе и зонной структуре; наличие полупроводника и принцип тождественности электронов - в использовании квантовой статистики вместо классической и так далее. Это значительно упрощает понимание и построение моделей, ведь намного проще представлять электрон как частицу, чем как объект, которому свойственны как волновые, так и корпускулярные свойства.

Таким образом в микроэлектронике для решения практических задач обычно удается избежать необходимости использования аппарата двух самых сложных разделов современной физики твердого тела: квантовой теории и теории рассеивания. И это невзирая на то, что волновые свойства электрона, как и процессы взаимодействия с различными дефектами решетки (рассеивание), в конечном итоге являются определяющими с точки зрения применения конкретных материалов для тех или иных целей.

1.5. КВАНТОВЫЕ МОДЕЛИ ПЕРЕНОСА ЗАРЯДА

1.5.1. Общие положения

Метод кинетического уравнения, строго говоря, применим лишь для систем, которые состоят из классических частиц. Как известно, в традиционной твердотельной электронике для описания процессов переноса заряда обычно используется одноэлектронное приближение и метод эффективной массы. С помощью последнего в левой части уравнения (1.13) квантовая природа электронов учитывается через соответствующую

модификацию функции распределения. В целом же, движение электронов считается классическим. Интеграл столкновений в правой части указанного уравнения учитывает квантовую природу электронов [1.8]. Таким образом, при использовании кинетического уравнения считается, что электроны между столкновениями двигаются как классические частицы; а рассеиваются как квантовые частицы. По этим причинам при использовании кинетического уравнения его считают "полуклассическим", так как оно все же частично учитывает квантовую природу электронного ансамбля.

Кинетическое уравнение Больцмана является достаточным и очень мощным средством моделирования неравновесных электронных процессов в существующих компонентах твердотельной электроники. Однако возникает вопрос, какие квантовые эффекты можно "потерять", считая движение электронов "классическим" в промежутках между столкновениями. Один из таких эффектов - туннельный эффект, который заключается в конечной вероятности прохождения потенциального барьера. Из него также следует возможность найти электрон в так называемой "запрещенной зоне", если говорить в терминах зонной теории. Такие эффекты могут существенно влиять на параметры конкретного электронного компонента, если его характерные размеры меньше определенных характерных длин, которыми характеризуется электронный коллектив. Такими длинами являются:

1. Длина волны де Бройля электрона λ_{Φ} , связанная с его волновым вектором k соотношением: $\lambda_{\Phi} = 2\pi/k$. При низкой температуре большинство электронов, участвующих в процессах переноса, находятся вблизи уровня Ферми. Поэтому волновой вектор k выбирают равным волновому вектору Ферми: $k = k_{\Phi}$. Условно считают, что значение длины волны де Бройля в типичных полупроводниковых структурах $\lambda_{\Phi} = 35$ нм [1.9]. В конкретном случае λ_{Φ} может отличаться от этого значения.

2. Средняя длина релаксации импульса L_p определяется через время релаксации импульса: $L_p = \tau_p \cdot v_{\Phi}$, v_{Φ} - скорость Ферми, которая определяется за формулой $v_{\Phi} = \hbar k_{\Phi} / m$, где k_{Φ} - волновой вектор Ферми. При типичной длине волны Ферми $\lambda_{\Phi} = 35$ нм скорость Ферми приблизительно равна $v_{\Phi} = 3 \cdot 10^5$ м/с. Считая, что время релаксации импульса $\tau_p = 0,1$ пс, типичное значение длины релаксации импульса L_p составляет около 0,3 мкм.

3. Длина релаксации фазы L_{Φ} определяется как средний путь, который должен пройти электрон для того, чтобы избавиться от так называемой «фазовой памяти». Последнее понятие означает буквально сохранение детерминированной связи между фазой волновой функции электрона в разные моменты времени и в

разных точках пространства. Роль величины L_{ϕ} заключается в том, что для наблюдения интерференционных эффектов она должна иметь значение больше размеров системы. Понятия фазовой релаксации, хаотизации фазы и фазовой когерентности не связаны непосредственно с актами упругих или неупругих столкновений, которые не обязательно приводят к хаотизации фазы. И наоборот, рандомизация фазы может не сопровождаться изменением импульса или энергии. Иногда, однако, акт рассеивания сопровождается хаотизацией фазы. Примером может быть электрон-электронное рассеивание. Однако в большинстве случаев хаотизация фазы не влияет на величину длины свободного пробега, поскольку акты хаотизации по большей части упруги. Для понимания общей причины хаотизации фазы, можно пользоваться общим правилом: флуктуационные центры рассеивания вызывают фазовую релаксацию, а стационарные центры рассеивания - нет.

Упомянутые длины используются для определения понятия макроскопической, микроскопической и мезоскопической системы.

*Микроскопическая система*¹ - это система, геометрические размеры которой меньше хотя бы одного из параметров λ_{ϕ} , L_p , L_{ϕ} . Если же значения всех трех параметров значительно меньше геометрических размеров системы, то она является *макроскопической*. Именно к таким системам может быть применен метод кинетического уравнения. Системы, которые не могут быть отнесены ни к микроскопическим, ни к макроскопическим называются *мезоскопическими* (от греч. "mesys" - средний, промежуточный). Таким образом, мезоскопические системы - это системы, геометрические размеры которых меньше наибольшего из параметров λ_{ϕ} , L_p , L_{ϕ} и больше наименьшего из них. Следует прибавить, что мезо- и микроскопические системы считаются такими, если соответствующие условия выполняются хотя бы для одного пространственного направления. Например, многослойные полупроводниковые гетероструктуры (сверхрешетки, резонансно-туннельные диоды, транзисторы с высокой подвижностью электронов (HEMT)) относятся к мезоскопическим системам.

Понятно, что четких границ между этими типами систем не существует. Чаще всего электронные структуры, имеющие размеры от нескольких десятков нанометров до единиц микрометров классифицируются как мезоскопические, иногда их называют *субмикронными*. Конкретные цифры определяются

¹ Понятие «микроскопичность» тут не привязывается к геометрическим размерам соответствующих систем (хотя их типичные размеры составляют менее десятков нм и обычно их называют «наноскопическими»), а лишь указывает на наличие в таких системах квантовых эффектов, которые не могут быть учтены в рамках полуклассического подхода.

интенсивностью процессов рассеивания и хаотизации, а также параметрами электронного коллектива в каждом конкретном случае.

Для микроскопических систем метод кинетического уравнения по большей части не может быть применен принципиально. Будучи применимым к мезоскопическим системам, метод кинетического уравнения Больцмана дает погрешность, которая возрастает с уменьшением размеров системы. По этим причинам возникает необходимость использования другого формализма для описания микроскопических и мезоскопических систем. Действующими на сегодня методами, которые частично или полностью выполняют функцию метода кинетического уравнения являются [1.10-1.12]:

- метод неравновесных функций Грина;
- метод функционала плотности;
- метод функции распределения Вигнера;

Поскольку для моделирования микроскопических структур, которые здесь рассматриваются, использование кинетических моделей не является эффективным, в дальнейшем остановимся на феноменологических.

Широкое использование приобрели феноменологические модели электронных процессов, такие как: метод волновых функций, метод туннельного гамильтониана и другие специфические методы, которые разрабатываются для моделирования конкретного прибора. Отчасти они наиболее эффективны, однако их применимость ограничивается узким кругом структур.

Качественные представления о процессах в структурах, в которых длина фазовой когерентности L_{ϕ} превышает геометрические размеры, были разработаны в 1957 году Ландауэром [1.13] и развивались вместе с Буттикером, что привело к появлению феноменологического формализма Ландауэра-Буттикера, который является системой чисто утилитарных представлений о явлениях в микроскопических системах, которые входят в состав макроскопической системы. В 1992 году формула Ландауэра была строго выведена из представлений кинетики с учетом межэлектронного взаимодействия [1.14]. Историческая роль Ландауэра заключалась в том, что он первым предложил рассматривать задачу о нахождении тока через малые объекты в виде задачи на прохождение и отражение электронных волн (англ. "*scattering problem*" - задача на "рассеивание") в отличие от, скажем, формализма Кубо, где задача о нахождении тока рассматривалась как задача нахождения линейного отклика на прикладываемое напряжение.

Невзирая на почтенный возраст отмеченного формализма, он стал основой для феноменологических моделей резонансно-

туннельных диодов (созданы Тсу и Эсаки в 1973 г.), одноэлектронных транзисторов (созданы Лихаревим в 1987 г.), углеродных нанотрубок (созданы Иджимой в 1991 г.) и некоторых других перспективных электронных компонентов, в том числе - устройств молекулярной электроники [1.15].

Этот формализм является интуитивно понятным качественным представлением об электронных процессах в сверхмалых системах. Однако, путем модификации и обобщения можно получить даже эффективные количественные модели. Хорошими примерами могут быть: «комбинированная» модель резонансно-туннельного диода [1.16, 1.17, 1.18], «ортодоксальная» теория одноэлектронного туннелирования [1.19]. и др., которые однозначно более эффективны, а иногда могут быть точнее, чем *ab initio*² модели.

Заметим, что метод построения теории Ландауэра-Буттикера не является типичным для отечественной науки в целом и физики в частности, которая тяготеет к получению представления о явлениях в конкретной системе путем обоснованных упрощений общей теории. Конкретные теории обычно строились так, что имели стойкую логическую связь с общей теорией. В отличие от этих подходов западная школа часто рассматривает конкретное явление для системы, которая приведена к максимально упрощенному виду (пределы нулевой температуры, отсутствие рассеивания, бесконечные длины релаксации, отсутствие электрон-электронных взаимодействий), который все еще допускает существование эффектов или явлений, закономерности которых мы желаем изучить.

Классическим примером структур низкой размерности являются полупроводниковые гетероструктуры. Если в гетероструктуре характерный размер лишь в одном пространственном измерении соизмерим с λ_F , L_p , или L_F , то это значит, что такая структура содержит тонкие пленки разных полупроводников толщиной порядка нескольких нанометров. Если в силу конструкции напряжение прикладывается перпендикулярно к плоскостям раздела наноразмерных слоев, то говорят, что в таких конструкциях реализуется *поперечный транспорт*, а сами структуры называются структурами с поперечным (иногда вертикальным) электронным транспортом. Примером таких структур являются резонансно-туннельные диоды, сверхрешетки, разнообразные устройства функциональной электроники на их основе.

Проблема, с которой имеет дело наноэлектроника при моделировании структур с поперечным транспортом, типична для

² *ab initio* (лат. «от начала») – решение задачи из первых основополагающих принципов без привлечения эмпирических допущений.

нанoeлектроники в целом. Представления об электронном транспорте в «старых» устройствах, не могут отображать реальные процессы, которые происходят в «новых» устройствах. Поэтому нанoeлектроника не может формироваться как «надстройка» микроэлектроники, и с этой точки зрения является революционным, а не эволюционным этапом развития электроники.

Формировать принципы теоретической нанoeлектроники приходится на том же фундаменте, на котором была построена микроэлектроника: многочастичная квантовая теория в широком смысле этого понятия, которое включает в себя спецразделы: статистическую физику для описания равновесных систем, квантовую кинетику - для неравновесных; теорию рассеивания и некоторые другие разделы теоретической физики.

Однако, сегодня большинство моделей наноструктур являются феноменологическими и не выводятся строго из положений теоретической физики. Они не являются строгими даже в той мере, в которой такими являются полуэмпирические модели устройств традиционной электроники. Их возникновение объясняется отставанием теоретической нанoeлектроники от экспериментальной. В то время как экспериментальные образцы уже изготовлены и функционируют, выявляя нужные свойства, адекватные *ab-initio* модели процессов остаются редкостью; отчасти проблемой является даже качественное моделирование.

1.5.2. Метод волновых функций

Обобщение метода эффективной массы. Движение электронов в гетероструктурах можно описать многочастичным уравнением Шредингера, которое рядом аппроксимаций сводится к системе независимых между собой уравнений для каждого из электронов [1.20]:

$$\hat{H}\psi = E\psi, \quad (1.91)$$

где ψ - одноэлектронная волновая функция, E - энергия электрона, \hat{H} - оператор Гамильтона в одноэлектронном приближении

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}) + V(\mathbf{r}), \quad (1.92)$$

в котором m - масса покоя электрона, \hbar - постоянная Планка, $V(\mathbf{r})$ - потенциальная энергия электрона в поле всех других электронов и ионов, которая не изменяется при смещении на вектор трансляции \mathbf{a} : $V(\mathbf{r} + \mathbf{a}) = V(\mathbf{r})$, $U(\mathbf{r})$ - непериодическая часть потенциальной энергии электрона, которая обусловлена внешними

полями, в частности разрывами экстремумов зон в гетероструктурах, прикладываемым напряжением и объемным зарядом.

Для решения уравнения Шредингера в *массивном* полупроводнике существует простой метод, который позволяет избавиться от рассмотрения члена $V(\mathbf{r})$. Это так называемый *метод эффективной массы*. Он основывается на том, что вблизи экстремумов зон уравнения (1.91) с гамильтонианом (1.92) можно представить уравнением (1.91) с гамильтонианом :

$$\hat{H} = -\frac{\hbar^2}{2m^*} \Delta + U(\mathbf{r}), \quad (1.93)$$

Функции, являющиеся решением уравнения Шредингера с гамильтонианом (1.93), корректно называть «*оггибающими волновой функции*», а не просто «волновыми функциями». Такое название обусловлено тем, что при $U(\mathbf{r}) = \text{const}$ решением (1.91) являются функции Блоха: $\psi(\mathbf{r}) = u_{nk}(\mathbf{r})e^{i\mathbf{k}\mathbf{r}}$, где функция $u_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r} + \mathbf{a})$ повторяет периодичность решетки и существенно изменяется в течение периода решетки. Множитель $e^{i\mathbf{k}\mathbf{r}}$ вблизи экстремумов зон изменяется на периоде решетки медленно, потому он фактически огибает всю волновую функцию. Если $U(\mathbf{r})$ медленно изменяется на периоде решетки, то можно показать, что это практически не повлияет на $u_{nk}(\mathbf{r})$; зато изменение отобразится на втором члене [1.21]. В множитель $e^{i\mathbf{k}\mathbf{r}}$ можно ввести зависимость волнового вектора от координаты $\mathbf{k}(\mathbf{r})$ или же заменить $e^{i\mathbf{k}\mathbf{r}}$ на медленно изменяющуюся $\chi(\mathbf{r})$: $\psi(\mathbf{r}) = u_{nk}(\mathbf{r})e^{i\mathbf{k}(\mathbf{r})\mathbf{r}} = u_{nk}(\mathbf{r})\chi(\mathbf{r})$. Функция $\chi(\mathbf{r})$ фактически и является искомой в уравнении (1.91) с гамильтонианом (1.93); естественно, что ее тоже называют «оггибающей». Дальше, однако, будем обозначать оггибающую волновой функции через $\psi(\mathbf{r})$, если не будет указано иное.

Метод эффективной массы используется в традиционной электронике в частности для расчета локализованных состояний электрона в кристалле [1.6]. Его с некоторыми модификациями используют и для решения уравнения Шредингера в гетероструктурах.

Если считать, что $U(\mathbf{r})$ можно представить в виде:

$$U(\mathbf{r}) = U_x(x) + U_y(y) + U_z(z),$$

то уравнение (1.93) распадется на три независимых уравнения:

$$\hat{H}_x \psi_x = E_x \psi_x, \quad (1.94)$$

$$\hat{H}_y \psi_y = E_y \psi_y, \quad (1.95)$$

$$\hat{H}_z \psi_z = E_z \psi_z, \quad (1.96)$$

а его решение можно будет искать в виде: $\Psi = \Psi_x \Psi_y \Psi_z$; $E = E_x + E_y + E_z$ [1.22]. Здесь $\hat{H}_z = -\frac{\hbar^2}{2m^*} \frac{d^2}{dz^2} + U(z)$; аналогично определяются \hat{H}_x и \hat{H}_y .

Рассмотрим структуру с поперечным транспортом, который содержит гетерограницы, которые заданы уравнениями $z = \text{const}$. Согласно приведенного определения структур с поперечным транспортом, составляющие $U_x(x)$ и $U_y(y)$ можно соответствующим выбором начала отсчета потенциальной энергии приравнять к нулю. Поэтому уравнение (1.94) и (1.95) будут иметь тривиальные решения в виде плоских волн:

$$\Psi_{\parallel} \sim \Psi_x \Psi_y e^{i(k_x x + k_y y)} \equiv e^{i(\mathbf{k}_{\parallel} \mathbf{r})}.$$

В направлении $0z$ следует решать уравнение Шредингера (1.96), которое в раскрытом виде записывается :

$$-\frac{\hbar^2}{2m^*} \frac{d^2 \Psi_z}{dz^2} + U(z) = E_z \Psi_z, \quad (1.97)$$

В дальнейшем нас будет интересовать лишь составляющая огибающей Ψ_z , которую мы будем обозначать, опуская индекс, через ψ .

Расчет электронного тока. Из обычного уравнения Шредингера можно получить закон сохранения плотности потока вероятности, который определяется по формуле $\mathbf{J} = \frac{i\hbar}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*)$ [1.8]. Умножив ее на заряд электрона, получим плотность создаваемого движением электрона электрического тока, для которого также автоматически выполняется уравнение непрерывности.

Граничные условия. Рассмотрим процессы, которые происходят вблизи границы двух материалов А и Б, находящейся в плоскости $z = a$. Известно, что потенциальная энергия электрона зоны проводимости испытывает резкое изменение на гетерогранице, в первом приближении - скачкообразное. Различные примеры гетерограниц изображены на рис. 1.9. Кроме того, скачкообразное изменение испытывает также эффективная масса, которая в материалах А и Б в общем случае разная.

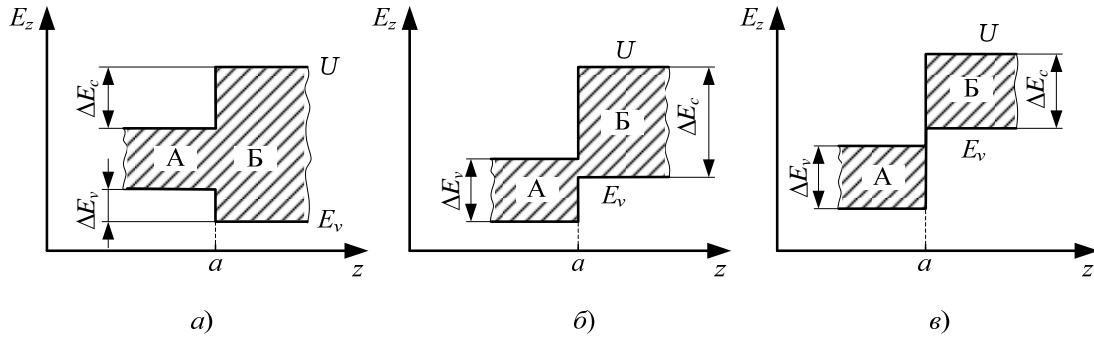


Рис. 1.9. Типы разрывов границ запрещенной зоны (заштрихована) для гетеропереходів: *а)* переход I типа (А-Б = GaAs-AlGaAs; GaSb-AlSb; GaAs-GaP; InGaAs-InAlAs; InGaAs-InP); *б)* переход II типа (А-Б = InP-Al_{0.48}In_{0.52}As; InP-InSb; BeTe-ZnSe; GaInP-GaAsP; Si-SiGe); *в)* переход III типа (А-Б = InAs-GaSb)

Рассмотрим электрон в зоне проводимости вблизи гетерограницы GaAs/Al_{0.33}Ga_{0.67}As в плоскости $z = a$. На границе этой гетеропары происходит движение электронов проводимости в состояниях Γ -долины, которая является самой низкой для обоих соединений. Именно для такого случая проводятся дальнейшие выкладки.

Запишем уравнение Шредингера для огибающей волновой функции в окрестности плоскости $z = a$ слева и справа:

$$-\frac{\hbar^2}{2m_A^*} \frac{d^2\psi}{dz^2} + U_A = E_z\psi \quad (z = a^+) \quad (1.98)$$

$$-\frac{\hbar^2}{2m_B^*} \frac{d^2\psi}{dz^2} + U_B = E_z\psi \quad (z = a^-), \quad (1.99)$$

где индексами «А» и «Б» обозначены соответствующие величины в соединениях «А» и «Б».

Обозначим решения уравнений (1.98) и (1.99) через ψ_A и ψ_B соответственно. Считается, что на гетерогранице периодическая составляющая волновой функции $u_{nk}(\mathbf{r})$ и ее производная непрерывны [1.23]. Тогда из требования непрерывности волновой функции следует непрерывность ее огибающей, и должны положить:

$$\psi_A(a) = \psi_B(a). \quad (1.100)$$

Запишем выражения для z -проекции потоков плотности вероятности в огибающей волновой функции в плоскостях $z = a^+$ и $z = a^-$, J_A и J_B соответственно:

$$J_A = \frac{i\hbar}{2m_A} \left(\psi_A^* \frac{d\psi_A}{dz} - \psi_A \frac{d\psi_A^*}{dz} \right), \quad (1.101)$$

$$J_B = \frac{i\hbar}{2m_B} \left(\psi_B^* \frac{d\psi_B}{dz} - \psi_B \frac{d\psi_B^*}{dz} \right). \quad (1.102)$$

Выполнение фундаментального закона сохранения плотности потока вероятности в квантовой механике следует из обычного уравнения Шредингера. Данный закон автоматически выполняется для уравнения Шредингера с обычной, постоянной массой электрона. Но в плоскости точки $z = a$, в соответствии с принятой моделью, эффективная масса испытывает скачкообразное изменение. Поэтому для выполнения закона сохранения необходимо принудительно наложить условие:

$$J_A = J_B, \quad (1.103)$$

из чего следует, что

$$\frac{1}{m_A^*} \frac{d\psi_A}{dz} = \frac{1}{m_B^*} \frac{d\psi_B}{dz}. \quad (1.104)$$

Уравнения (1.100) и (1.104) представляют собой общие требования к огибающей волновой функции и ее производной на гетерогранице, и являются хорошими приближениями для большинства гетерограниц. Их можно записать в неявном, однако компактном виде, внося под первую производную в (1.97) эффективную массу:

$$-\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m^*(z)} \frac{d\psi_z}{dz} + U(z) = E_z \psi_z. \quad (1.105)$$

Уравнение (1.105) широко применяется при моделировании электронного транспорта в гетероструктурах; его также применяют, когда эффективная масса изменяется плавно. Последнего эффекта можно достичь, например, изменяя молярную долю x в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ от нуля до 33% на протяжении нескольких нанометров. Такие структуры называются «варизонными». Они могут быть изготовлены современными технологиями:

- молекулярно-лучевой эпитаксии (англ. аббревиатура - *MBE*);
- химического испарения металлоорганических соединений из газовой фазы (*MOCVD*).

Уравнения (1.105) часто называют *уравнением Шредингера с эффективной массой*. Оно является ключевым компонентом так называемого «метода огибающей волновой функции», которая применяется для моделирования гетероструктур. Название этого

метода является синонимом *метода эффективной массы*. В отечественной литературе часто опускают слово «огibaющая» и говорят просто «метод волновых функций».

1.6. МОДЕЛИРОВАНИЕ СУБМИКРОННЫХ СТРУКТУР

1.6.1. Задачи моделирования и типы моделей

Задачи, решаемые в инженерной практике с помощью моделей, зависят, прежде всего, от круга пользователей. Удобно, хоть в определенной мере и условно, выделить три их основных типа. К первым относятся разработчики СВЧ устройств, которые используют модель как завершённый программный продукт, зачастую встраиваемый в систему автоматизированного проектирования (САПР), не требующий изменения или подстройки каких-либо внутренних параметров модели и поэтому предельно формализованный и хорошо протестированный в заданной области изменения исходных физико-топологических и электрических параметров. По внутренней структуре это могут быть модели, основанные на простых аналитических соотношениях, полученных из исходных физических уравнений, либо на аппроксимации моделируемых параметров численными или аналитическими методами. Основное требование к таким моделям – быстрый и адекватный (при заданной точности) расчет той или иной системы параметров, характерных для данного прибора. Особым требованием является минимальный набор входных данных, так как разработчики устройств зачастую ведут проектирование в условиях недостаточной информации о параметрах используемых приборов.

Иного рода задачи могут решаться с помощью моделирования разработчиками приборов. К типичным конструкторским задачам относится анализ связей между исходными топологическими, физическими и электрическими данными и выходными параметрами приборов. Как правило, этот анализ представляет интерес не сам по себе, а для выявления оптимальных (по выбранным критериям) или допустимых исходных параметров приборов. Для этого модель обычно используется многократно с целенаправленным перебором входных данных, который выполняется либо вручную, либо осуществляется с помощью стандартных или специализированных процедур оптимизации. Необходимым атрибутом этого процесса является анализ чувствительности результата (целевой функции) к вариантам тех или иных входных данных или параметров. Это позволяет выявлять наиболее сильные связи и проводить поиск более целенаправленно.

В качестве примеров задач, которые возникают перед разработчиками, например, полевых транзисторов с барьером Шоттки и гетеротранзисторов, можно привести следующие:

- выбор профилей легирования, в том числе допустимых концентраций и вида примесей в подложке и буферном слое;
- исследование влияния сопротивлений омических контактов на параметры приборов;
- поиск оптимального сопротивления металлизации затвора и его размеров;
- оценка влияния паразитных элементов конструкции прибора на его параметры;
- поиск и прогнозирование оптимальных режимов работы с точки зрения усилительных и шумовых свойств;
- анализ статистических и регулярных неоднородностей топологии прибора;
- роль паразитных примесей в материале (медленных ловушек, поверхностных состояний) и т.п.

Этот перечень, который постоянно пополняется, в принципе не может быть полным. Решение технологических задач и задач конструирования приборов требуют хорошего понимания физических процессов, происходящих в отдельных областях прибора, умения получить и обработать результаты моделирования, грамотно их интерпретировать. Широкий спектр задач, которые решаются при этом конструкторами и технологами, стимулирует разработку большого спектра моделей различного уровня, отображающих те или иные процессы в приборах, так как обычно создание «полных» моделей, которые бы «все учитывали» невозможно. Для проверки результатов моделирования, вызывающих сомнения, или для получения информации о немоделируемых процессах полезно создание тестовых структур.

Для адекватного моделирования физических процессов и создания моделей для систем автоматизированного проектирования работа должна проходить в тесном взаимодействии с разработчиками приборов и технологами, так как процесс моделирования должен соотноситься с экспериментальными данными, полученными при испытании приборов и тестовых структур. При этом совершенствование моделей с точки зрения физической сущности и численных процедур, с одной стороны, направлено на адекватное описание физических процессов, а с другой, связано с выбором методов расчета, системы допущений и формированием алгоритмов с учетом анализа разнообразных физических факторов, выделения главных и игнорирования второстепенных.

Физико-топологические модели. Удобнее всего классификацию физико-топологических моделей провести на примере моделей субмикронного полевого транзистора с барьером Шоттки,

одного из наиболее типичных СВЧ приборов. При этом приведенная классификация не претендует на полноту и строгость в соблюдении классификационного признака, а отражает лишь сложившуюся в практике моделирования ситуацию.

Среди большого количества различных моделей [1.24] полевых транзисторов с затвором Шоттки (ПТШ) можно выделить четыре типа, отличающихся различным уровнем моделирования.

Аналитические модели основаны на соотношениях, полученных в результате приближенного анализа основных уравнений переноса электронов и уравнений поля. Типичными являются модели [1.24-1.26], использующие допущения о резкой границе обедненной области под затвором и ее полном обеднении, об одномерном характере переноса заряда в канале и стационарном характере дрейфового движения. Последнее допущение не позволяет использовать эти модели для субмикронных транзисторов. Подходящим выбором аппроксимации поле-скоростной характеристики $v(\mathcal{E})$ можно расширить предел моделирования до длин затвора 0,5...0,7 мкм, однако точность моделирования остается невысокой. В аналитических моделях обычно рассчитываются параметры схемной модели обычно с сосредоточенными параметрами, а затем уже определяются Y - или S -параметры и шумовые параметры. Последние рассчитываются с помощью включаемых в схемную модель эквивалентных шумовых источников. Физическая информативность моделей этого уровня невелика, однако, они являются экономичными с точки зрения вычислительных затрат.

Квазидвумерные модели [1.27-1.29] основаны на численном решении системы одномерных дифференциальных уравнений, описывающих процессы вдоль канала, а информация об обедненной области в приближении резкой или плавной [1.29] границы включается в систему уравнений. Принципиальным отличием этого уровня моделей является то, что в исходную систему входят дифференциальные уравнения для импульса и энергии электронов, позволяющие моделировать динамику нестационарного дрейфового движения (эффект всплеска скорости). Это дает возможность применять эти модели для субмикронных транзисторов с длиной затвора до 0,2 мкм. Время моделирования при малосигнальном анализе в частотной области достаточно невелико, что позволяет получать функциональные зависимости схемных параметров от электрического режима в широком диапазоне питающих напряжений. Это делает возможным их использование для режима большого сигнала во временной области [1.30] на высокопроизводительных компьютерах. Важной особенностью моделей данного уровня является весьма высокая информативность о физических процессах в транзисторе,

так как в процессе численного интегрирования исходной системы рассчитываются средние значения таких физических величин как скорость, энергия, напряженность электрического поля и т.п. Анализ шумовых параметров транзисторов проводится, как в аналитических моделях, с помощью схемной модели, включающей эквивалентные шумовые источники тока или напряжения. Однако особенности квазидвумерных моделей позволяют определять реакцию выходного напряжения на флуктуации скорости в канале. Такой анализ требует заметных затрат времени, но обеспечивает весьма высокую точность моделирования шумовых параметров транзистора. Надо отметить также, что модели такого уровня могут использоваться для транзисторов, имеющих простейшие профили легирования.

Двумерные модели в традиционном смысле основаны на численном двумерном интегрировании системы фундаментальных (макроскопических) уравнений, включающей уравнение Пуассона, уравнение непрерывности и уравнение динамики носителей. Большинство существующих моделей, например [1.31], построены в так называемом дрейфово-диффузионном приближении, не позволяющим анализировать нестационарную динамику носителей, существенную для субмикронных транзисторов. Часто такие модели обозначают термином «локальная» модель, при этом предполагается, что в каждой точке пространства скорость носителей следует за изменением напряженности электрического поля. Результатом моделирования физической части модели являются временные и пространственные распределения средних значений тока, электрического поля, скорости и т.п., поэтому модели данного уровня отличаются высокой информативностью. Обычно моделирование проводится во временной области, и его результатом являются переходные характеристики, которые при необходимости могут быть пересчитаны в параметры схемной модели. Шумовой анализ может быть проведен с помощью локальных шумовых источников или температуры, рассчитываемых для каждого элемента двумерной сетки, однако большие временные затраты требуют дальнейшего совершенствования этой процедуры. Адекватность двумерных моделей может быть существенным образом улучшена при переходе к «нелокальной» модели путем дополнения системы релаксационными уравнениями: уравнениями сохранения энергии, импульса и частиц (концентрации).

Кинетические модели [1.32] относятся к двумерным лишь по организации области моделирования. Их принципиальное отличие от предыдущих, являющихся детерминированными, состоит в использовании кинетического уравнения Больцмана, решаемого численно совместно с уравнением Пуассона методом Монте-Карло. Являясь статистическими, кинетические модели позволяют

анализировать движение ансамбля моделируемых частиц, как в координатном пространстве, так и в пространстве квазиимпульсов. Используя те или иные модели зонной структуры, существующие для данного материала механизмы рассеяния через единичные акты столкновений, учитывая при необходимости некоторые квантовые эффекты, кинетические модели позволяют проследить «реальную» эволюцию всего коллектива носителей при заданном законе изменения напряжений на электродах через динамику каждой отдельной частицы. На этапе обработки результатов моделирования по желанию можно отслеживать как полную функцию распределения $f(\vec{r}, \vec{k}, t)$, так и производить выборки и соответствующие усреднения в \vec{r} - и \vec{k} -пространствах, анализируя особенности поведения и вклад отдельных групп электронов, например, влияние процессов в подложке, переходы электронов между долинами и т.п. частные эффекты. Шумовой анализ не представляет принципиальных трудностей, так как при расчете токов через электроды флуктуации проявляются «естественным» образом. Все это позволяет считать кинетические модели моделями самого высокого уровня, которые физически полностью информативны. Характерное время моделирования зависит от требуемой выходной информации.

Произведя анализ моделей разного уровня с точки зрения их использования в проектировании можно отметить, что аналитические модели для субмикронных транзисторов имеют ограниченное применение и применяются в тех случаях, когда требуется высокое быстродействие и можно удовлетвориться невысокой точностью моделирования. Для транзисторов со сложным профилем легирования пригодны практически лишь двумерные детерминированные и кинетические модели, в которых информация о реальном профиле вводится в качестве исходных данных. Эти же модели можно использовать при проектировании устройств дискретной схмотехники, так как они позволяют моделировать реакцию транзистора на заданные электрические воздействия. Более удобно это делать для кинетических моделей, так как они обычно используют прямые методы интегрирования уравнения Пуассона. В принципе для этой цели возможно создание и квазидвумерных моделей, однако они особенно удобны для малосигнального анализа, заметно выигрывая у двумерных моделей по быстродействию, которое существенно при создании систем проектирования аналоговых транзисторных устройств.

Сравнивая между собой двумерные детерминированные и кинетические модели, надо отметить, что последние, проигрывая в быстродействии, имеют преимущества в информативности, в точности моделирования (особенно специфики субмикронных приборов и шумовых свойств). Это делает их незаменимыми для

исследования процессов в транзисторах, для конструкторско-технологического проектирования приборов, в том числе приборов с гетероструктурным каналом, транзисторов с проницаемой базой и других приборов со сложной топологической конфигурацией. С другой стороны, применение кинетических моделей для современных субмикронных высоколегированных биполярных структур с характерными размерами слоев около 0,1 мкм и необходимости усреднения характеристик в условиях больших градиентов функций может привести к росту вычислительных погрешностей и дополнительным вычислительным затратам. В системах автоматизированного проектирования кинетические модели и двумерные модели могут иметь ограниченное применение из-за большого времени моделирования. Для этих целей с успехом можно использовать квазидвумерные модели, дающие приемлемую точность и сравнительно высокое быстродействие. К достоинствам этих моделей следует отнести также простоту учета эффектов третьего измерения [1.33-1.36]: топологических неоднородностей и распределенных эффектов.

Схемные модели. Расчет выходных характеристик ПТШ, таких как матрицы проводимости и рассеяния, коэффициент усиления, матрица спектральных плотностей мощности шума, минимальный коэффициент шума и др., связан с анализом реакций в виде токов и напряжений на электродах транзистора на некоторые тестирующие воздействия. Отклики на электродах в виде временных зависимостей, либо в виде приращений токов и напряжений получаются на основе физико-топологической модели соответствующего уровня и схемной модели.

В зависимости от уровней входных воздействий возможны различные режимы работы ПТШ – от линейного (малосигнальный режим), когда изменения входного сигнала происходят в малой окрестности рабочей точки по электрическому режиму, до существенно нелинейного, когда амплитуда входного сигнала может быть сравнимой с напряжениями питания (режим большого сигнала).

И в первом и во втором случае возможно представление транзистора в виде схемной модели. Эта модель является некоторой электрической схемой, содержащей пассивные элементы и зависимые источники тока (напряжения), отражающие реакцию отдельных областей структуры транзистора на входные тестирующие воздействия.

Как правило, такая схема включает в себя элементы, описывающие активную область транзистора (проводимость канала, зависимый от входного напряжения источник тока с управляющим параметром, соответствующим крутизне вольт-амперной характеристики, а также входную и проходную емкость); омические сопротивления пассивных участков канала и

омических контактов стока и истока; сопротивление металлизации затвора; паразитные емкостные и индуктивные элементы внешних областей структуры транзистора. Как правило, последняя группа параметров может быть найдена из эксперимента, электродинамического анализа, либо рассчитана по экспериментальным параметрам матрицы рассеяния с привлечением процедуры оптимизации.

Выходные характеристики в виде, например, матрицы проводимости транзистора, рассчитываются известными методами анализа электронных схем. При этом транзистор представляется в виде эквивалентного четырехполюсника с соответствующими проводимостями и зависимыми источниками тока.

В режиме большого сигнала описанная схемная модель транзистора содержит нелинейные элементы (например, диоды). Анализ модели в этом случае заключается в применении методов, справедливых для динамических систем.

Необходимо отметить, что анализ транзистора на основе схемной модели (в виде некоторой «эквивалентной» схемы) обладает рядом недостатков, к главным из которых можно отнести неоднозначность такого описания (существует некоторое множество «эквивалентных схем», удовлетворительно описывающих выходные характеристики), а также трудности идентификации «внешних» по отношению к активной области транзистора параметров схемной модели, влияние которых усиливается по мере расширения рабочего диапазона частот.

Строго говоря, представление транзистора в виде схемной модели справедливо в той мере, насколько справедливо одномерное физико-топологическое моделирование структуры транзистора (когда в канале несущественны поперечные составляющие электрического поля). В случае, когда продольные и поперечные размеры канала транзистора сравнимы, описание транзистора в виде схемной модели с сосредоточенными параметрами обладает большой погрешностью.

Существенным достоинством схемного моделирования реакции транзистора на входное тестирующее воздействие является возможность надежной идентификации параметров схемной модели, поскольку имеется возможность получения спектра экспериментальных зависимостей, не связанных с диапазоном сверхвысоких частот. При необходимости те или иные компоненты схемной модели можно рассматривать как частотнозависимые.

В общем случае при анализе «двумерных» структур с применением двумерных физико-топологических моделей используется представление транзистора в виде некоторых четырехполюсников, на вход которых подается тестирующее воздействие в виде «единичной» ступенчатой временной функции тока или напряжения, а затем по известной реакции на выходе (переходной

характеристике) рассчитываются частотные характеристики схемы с использованием процедуры быстрого преобразования Фурье.

Последний подход, как правило, требует больших вычислительных затрат, однако позволяет получить наиболее адекватные временные зависимости на выходе. Учет «внешних» паразитных элементов приводит к увеличению времени установления выходных величин, т.е. использование моделей высокого уровня не избавляет от проблем идентификации паразитных элементов и требует развития методов электродинамического анализа этих элементов.

Одним из важнейших аспектов моделирования малошумящих устройств является разработка шумовых моделей, позволяющих описывать источники шума различной природы и содержащих алгоритмы анализа шумовых характеристик в системе волновых шумовых параметров, удобной при анализе СВЧ многополюсников, а также позволяющих рассчитывать минимальный коэффициент шума и оптимальные импедансы входа и выхода. Расчет минимального коэффициента шума транзистора связан с совместным использованием сигнальных и шумовых параметров (матрицы рассеяния и матрицы спектральных плотностей мощности шума), а также с оптимальным выбором питающих напряжений. Оптимизация параметров структуры транзистора приводит к целевой функции, представляющей собой поверхность с множеством локальных экстремумов, и требует разработки специальных процедур минимизации.

Формальные модели. Наибольшие трудности и затраты вызывает моделирование во временной области параметров ПТШ, работающего в режиме большого сигнала. Непосредственное использование двумерных физико-топологических моделей, позволяющих рассчитывать временные характеристики для анализа СВЧ устройств средствами систем автоматизированного проектирования, неоправдано из-за больших вычислительных затрат. Как правило, в САПР используются аппроксимации зависимостей нелинейных элементов от электрического режима, либо временных или частотных характеристик. Такой подход позволяет для заданной топологии прибора и выбранного рабочего режима построить формальные модели указанных зависимостей в виде аппроксимаций (например, степенным рядом или полиномом), что существенно повышает эффективность анализа и не требует обращения к физическим моделям при расчете устройств.

Аппроксимация может быть одно- или многомерной, непрерывной или кусочной. Возможно сочетание линейных и нелинейных аппроксимаций отдельных зависимостей или характеристик.

Создание библиотеки формальных моделей связано на предварительном этапе с надежной идентификацией параметров схемной модели или реакции (отклика) на выходе. Эта задача может решаться либо путем проведения измерений, либо путем многократного счета по физико-топологическим моделям различного уровня. Количество расчетных точек аппроксимируемых зависимостей зависит от уровня входного сигнала (степени нелинейности), а также требований к точности аппроксимирующих полиномов.

Наиболее перспективными для построения формальных аппроксимирующих процедур являются квазидвумерные физико-топологические модели. Такие модели, во-первых, не требуют значительных вычислительных затрат, а во-вторых, используют приближение схемной модели, что позволяет рассчитывать необходимые зависимости схемных параметров в заданном рабочем диапазоне, легко тестируемые по низкочастотным измерениям в малосигнальном приближении.

Особый класс формальных моделей представляют так называемые макромоделю, когда устройство в целом или прибор представляются в виде некоторой электрической схемы, содержащей идеальные электрические компоненты, описывающие управляющие, передаточные, импедансные и другие характеристики моделируемого устройства или прибора. Как правило, формальная схемная модель (макромодель) содержит существенно меньше компонентов, чем схемная модель, отражающая физическую сущность процессов в моделируемых устройствах и приборах. Макромоделю требуют минимальных вычислительных затрат при анализе и легко встраиваются в САПР.

Для моделирования цифровых схем используют логические формальные модели, не использующие электрических величин и физических констант, а отражающие функционирование на логическом уровне.

Учет в формальных моделях внешних возмущающих факторов (температуры, радиации, шумов и т.п.) осуществляется путем введения в модель дополнительных аппроксимирующих зависимостей или аналитических функций, учитывающих соответствующее воздействие.

1.6.2. Система моделирования субмикронных компонентов

Создание новых компонентов связано с разработкой и верификацией различных математических моделей для описания физических процессов в активных полупроводниковых структурах.

При этом возможны различные подходы при моделировании,

применение моделей субмикронных структур различного уровня, их комплексное взаимодействие и обмен данными. Задача заключается в создании целостной *системы моделирования субмикронных приборов* и анализе возможных методов моделирования [1.37] с учетом их адаптации к объекту моделирования [1.38, 1.39].

Основные уравнения математических моделей (см. разд. 1) основываются на анализе уравнений для нахождения функции распределения (кинетическое уравнение Больцмана), уравнений сохранения частиц, импульса и энергии с учетом перераспределения носителей между долинами для их дальнейшего использования для описания физических процессов в субмикронных приборах на основе многодолинных полупроводников, включая арсенид галлия, фосфид индия, нитрид галлия и др.

В настоящее время в области моделирования полупроводниковых СВЧ приборов сложилась ситуация, которая характеризуется, с одной стороны, большим разнообразием моделей и, с другой, широким спектром задач, требующих их использования. В связи с этим возникает необходимость систематизации моделей и предъявляемых к ним требований. Актуальным является круг вопросов, связанных с созданием системы моделирования твердотельных СВЧ приборов, в частности субмикронных полевых транзисторов с барьером Шоттки, гетероструктурных транзисторов и в последующем маломощных и мощных устройств на их основе.

При численном моделировании устройств, содержащих субмикронные компоненты, целесообразным представляется подход, заключающийся в построении иерархического ряда математических моделей, предназначенных для задач описания физических эффектов и процессов, совершенствования технологии и конструкций приборов, схемотехнического анализа и расчета характеристик устройств при различных уровнях входного воздействия и режимах работы. При этом модели могут отличаться не только по своему назначению, но и методами, уровнем строгости, принятыми допущениями и требуемыми вычислительными затратами. Наличие всего иерархического ряда моделей позволяет, например, использовать результаты физического моделирования для определения параметров и аппроксимаций для более простых моделей, а последние, в свою очередь, могут задавать хорошее начальное приближение для постановки задач проектирования и исследовательских задач в субмикронных структурах.

Ядром системы моделирования может являться блок моделей приборов, отражающий специфику данной системы (рис. 1.10). Все многообразие моделей целесообразно выстроить в определенный иерархический ряд. Исходные данные для моделирования

сосредоточены в библиотеке материалов и включают информацию, получаемую непосредственно из электронных и фононных законов дисперсии и параметров кристаллической решетки данного вещества.

Отдельной задачей является идентификация паразитных параметров полупроводниковых структур, влияние которых становится все более существенным при уменьшении размеров активных областей до субмикронных и наноразмеров. Задача решается путем моделирования и измерений характеристик тестовых структур.

К верхнему ряду относятся модели физических процессов, в результате которых определяются поле-скоростные характеристики, заселенность долин, времена релаксации и т.п., а также модели отдельных фрагментов приборов, таких как модели омических контактов, слоев металлизации, барьеров Шоттки, гетеропереходов и т.д.

В результате моделирования на данном уровне определяются электрофизические константы материалов, а также вспомогательные параметры приборов, не входящие в активную область прибора и поэтому обычно не моделируемые на приборном уровне. Полученная здесь информация совместно с топологическими и электрическими данными является исходной для физико-топологического моделирования приборов.

Среди существующих физико-топологических моделей приборов в свою очередь можно выделить модели различного уровня и установить определенную классификацию. Но все они, хоть и в разной степени, позволяют получать физическую информацию о процессах, происходящих в приборе, через распределения носителей, полей, скоростей, токов и т.п. величин в пространстве и быть может во времени. Эта информация представляет интерес для понимания процессов в приборе и, кроме того, используется для нахождения схемных электрических параметров прибора или его отдельных фрагментов. Характеристика используемой схемной модели (в виде схемы с сосредоточенными параметрами и управляемыми источниками, распределенной модели, в виде многополюсника) зависит как от типа прибора, так и особенностей его использования (частотный диапазон, величина входного сигнала, область применения и т.п.).

Для дальнейшего использования в системах проектирования устройств зачастую удобно использовать формальные модели приборов в виде числовых таблиц или функций, аппроксимирующих различного рода зависимости, полученные в процессе физико-топологического моделирования.

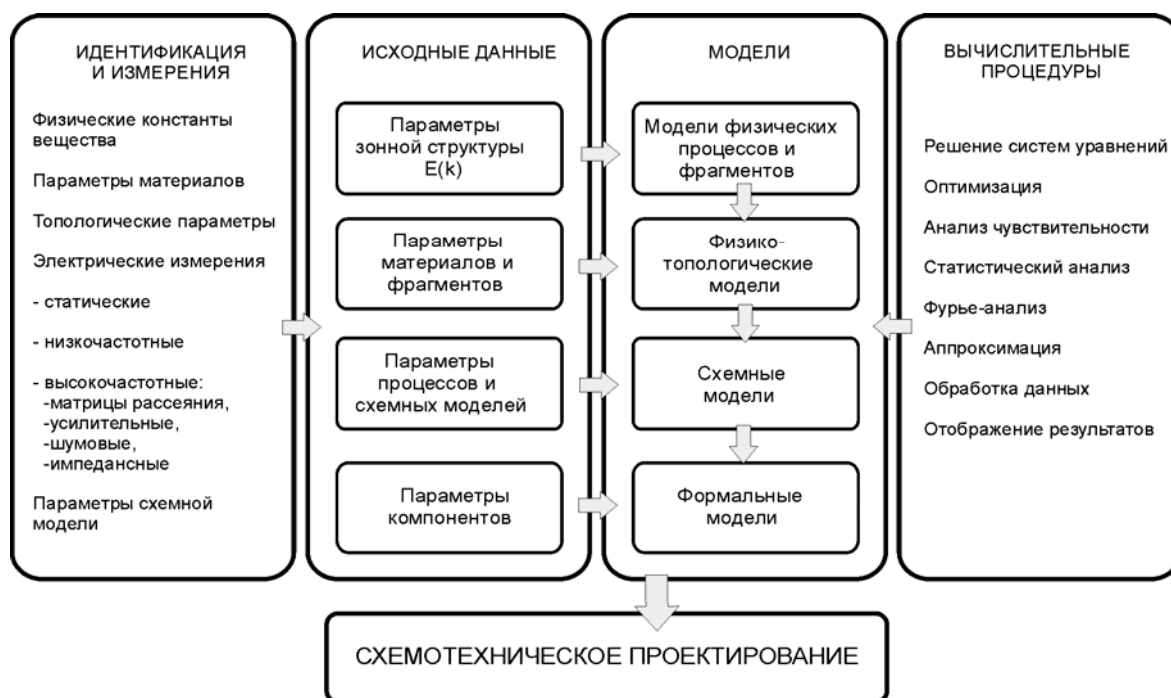


Рис. 1.10. Структурная схема системы моделирования субмикронных приборов

Важнейшей составной частью, определяющей в значительной степени спектр решаемых задач, является блок процедур. Он представляет из себя набор программ, реализующий типичные методы, используемые в практике моделирования и проектирования. Среди них могут быть как стандартные, универсальные, так и специализированные, предусмотренные для решения вполне определенных задач. Особо необходимо отметить задачу отображения результатов моделирования. Кроме удобства, графическое представление результатов является необходимостью при анализе больших массивов информации, которые встречаются, например, при моделировании многочастичным методом Монте-Карло, а также в другого рода статистических моделях, при двумерном моделировании физических процессов.

Функционирование всей системы определяется организацией работы отдельных программ. Основные требования к ней: гибкость в формировании маршрута моделирования в зависимости от решаемой задачи, передача данных между отдельными уровнями моделей и интерактивный ввод данных, удобное представление результатов. Интерактивный ввод данных, во-первых, обеспечивает возможность использовать блок моделей с любого уровня и, во-вторых, вводить дополнительные данные, не моделируемые на предыдущем уровне.

Необходимым дополнением к системе моделирования являются измерения. Подключение этого блока к машинной системе затруднительно, так как среди перечисленных типичных для приборов СВЧ видов измерений лишь некоторые могут

проводиться автоматизировано. Поэтому функционирование данного блока осуществляется автономно, что не исключает использования отдельных машинных программ для обработки результатов измерений. Основная роль измерений сводится к созданию информационного обеспечения для системы моделирования, то есть предполагает получение исходных данных для моделей того или иного уровня или получение дополнительных результатов, не моделируемых на более высоком уровне. Достаточно обширные измерения необходимы также для верификации моделей.

Проведенный анализ относится прежде всего к физической части моделей и не затрагивает моделирование схемных параметров транзисторов. Опыт показывает, что схемные параметры, являясь интегральными, не столь чувствительны к погрешностям моделирования деталей физических процессов. При выборе модели надо также учесть, что, во-первых, паразитные элементы транзистора не включаются в область моделирования и требуются дополнительные способы их учета при использовании физико-топологических моделей. Во-вторых, для расчета Y - или S -параметров на низких частотах требуется очень длительный (порядка сотен пикосекунд или нескольких тысяч временных шагов кинетической модели) период наблюдения переходного процесса. Это, безусловно, ограничивает применение последних при проектировании транзисторных устройств, когда требуется многократное обращение к модели.

Необходимым является совместное использование моделей различного уровня при проектировании. Таких вариантов может быть предложено много в зависимости от целей проектирования, имеющейся вычислительной техники и др. Одним из простейших является предварительное определение с помощью кинетических или двумерных моделей параметров для более простых моделей (квазидвумерных, аналитических, аппроксимационных) с последующим их использованием для целей анализа или оптимизации. Подобные варианты могут быть использованы и при конструкторско-технологической разработке транзисторов, когда очень высокая физическая информативность кинетических и двумерных моделей может оказаться избыточной.

Для адекватного анализа физических процессов и оптимизации характеристик необходимым является создание иерархического ряда математических моделей субмикронных приборов различного уровня и определение функциональных особенностей их применимости к задачам адаптивного проектирования.

Литература к главе 1

- 1.1. Москалюк В. А., Синекон Ю. С., Кассинг Р. Физика электронных процессов: Учеб. пособие. – Ч. I: Электронные состояния. – К.: УкрИНТЭИ, 2001. – 148 с.
- 1.2. Зеегер К. Физика полупроводников: Пер. с англ. – М.: Мир, 1977. – 616 с.
- 1.3. Иващенко В. М., Митин В. В. Моделирование кинетических явлений в полупроводниках. Метод Монте-Карло. – К.: Наук. думка, 1990. – 192 с.
- 1.4. Бонч-Бруевич В.Л., Калашников С.Г. Физика полупроводников. Учеб. пособие для вузов – 2-е изд., перераб. и доп. – М.: Наука. Гл. ред. физ.-мат. лит. 1990.–688 с., ил.-ISBN 5-02-014032-5.
- 1.5. Шур М. Современные приборы на основе арсенида галлия: Пер. с англ. – М.: Мир, 1991. – 632 с.
- 1.6. Хокни Р., Иствуд Дж. Численное моделирование методом частиц: Пер. с англ. – М.: Мир, 1987. – 640 с.
- 1.7. Прохоров Э. Д., Белецкий Н. И. Полупроводниковые материалы для приборов с междолинным переносом. – Х.: Вища шк.; Изд. при Харьк. ун-те, 1982. – 144 с.
- 1.8. Москалюк В.О. Фізика електронних процесів. Динамічні процеси. Київ: “Політехніка”, 2004. – 180 с.
- 1.9. *Fransson J.* Non-orthogonality and electron correlations in nanotransport. Spin- and time-dependent currents. Dissertation for the Degree of Doctor of Philosophy in Physics presented at Uppsala University. Uppsala, Sweden, 2002. – 101 P.
- 1.10. *Sun J.P Mains R.K., Haddad G.I.* “Resonant tunneling diodes: models and properties”, Proc. of IEEE, vol. 86, pp. 641-661, 1998.
- 1.11. Абрамов И.И. Проблемы и принципы физики и моделирования приборных структур микро- и наноэлектроники. V. Резонансно-туннельные структуры // Нано- и микросистемная техника. – 2007. – №3. – с. 57–70.
- 1.12. Федяй А.В. Моделі резонансно-тунельних діодів (аналітичний огляд). Кінетичні моделі // Н.-техн. Ж. «Електроніка і зв'язь» – 2007. - №5. – С. 11–16.
- 1.13. *Landauer R.* Spatial variation of currents and fields due to localized scatterers in metallic conductor. // IBM Journal of Research and Development. – 1957. – Vol. 1, 233.
- 1.14. *Meir Y., Wingreen, N.S.* Landauer formula for the current through an interacting electron region // Phys. Rev. Lett. – 1992. – Vol. 68. – P. 2512–2515.
- 1.15. *Hirose K., Ono Tomoya, Fujimoto Y., Tsukamoto S.* First-principle calculation in real-space formalism // London: Imperial College Press. – 2005. – 253 p.
- 1.16. *Pinaud O.* Transient simulation of resonant-tunneling diode // J. Appl. Phys. – 2002. – Vol. 92, No. 4. – P. 1987–1994.
- 1.17. Абрамов И.И., Гончаренко И.А., Коломейцева Н.В. Комбинированная модель резонансно-туннельного диода. //

Физика и техника полупроводников. – 2005. – Том 39, Вып. 9. – с. 1138–1145.

1.18. Москалюк В.О., Федяй А.В., Ярошенко О.Ю. Прикладна програма для моделювання переносу заряду в квантоворозмірних гетероструктурах з графічним інтерфейсом користувача // Н.-техн. журнал «Електроніка і зв'язь», тематический випуск «Електроніка і нанотехнології». – 2011. – №1. – С. 48-53.

1.19. Likharev K. K. Single-electron devices and their applications // Proc. Of IEEE. – 1999. – Vol. 87. – P. 606-632.

1.20. Шалимова К.В. Физика полупроводников. Учебное пособие для вузов. – М.: Энергоатомиздат, 1985. – 392 с.

1.21. John H. Davis. The physics of low-dimensional semiconductors: an introduction. – Cambridge University Press, 2005. – 438 p.

1.22. Ландау Л.Д., Лифшиц Е.М. Теоретическая физика: учеб. пособ. для вузов в 10 томах. Том 3. Квантовая механика (нерелятивистская теория). – 5-е изд. стереот. – М. Физматлит, 2002. – 808 с.

1.23. Bastard G. Wave mechanics applied to semiconductor heterostructures. – Halsted Press, New York, 1988. – 357 p.

1.24. Петров Г.В., Толстой А.И. Основные направления в моделировании субмикронных полевых транзисторов с затвором Шоттки // Изв. ВУЗов. Сер. Радиоэлектроника. – 1986. – Т. 29. – №10. – С.28 – 42.

1.25. Pucel R.A., Haus H.A., Statz H. Signal and noise property of GaAs microwave field-effect transistors. Adv. in electronics and Electron Physics.–1975.– N-Y.; Acad. Press. – Vol. 38.– P.195 – 265.

1.26. Sugeta T., Ida M., Uchida M. Microwave performance GaAs-Shottky barrier gate FET's.– Rev. Electron Common. Lab.–1975.–Vol. 23, №11/12.–P.1182–1192.

1.27. Carnez B., Cappy A., Kaszynski A., et al. Modeling of a submicrometer gate-field transistors including effects of nonstationary Electron dynamics // J. Appl. Phys. – 1980.–Vol.51, №1.–P.784–790.

1.28. Кальфа А.А., Пашковский А.Б., Тагер А.С. Исследование вольтамперных характеристик полевых транзисторов с субмикронным затвором // Электронная техника. Сер. I, Электроника СВЧ. – 1984. – Вып. 4 (364). – С. 27–30.

1.29. Москалюк В.А., Тимофеев В.И. Моделирование процессов переноса электронов в канале и подложке субмикронного полевого транзистора // Электронная техника. Сер. Электроника СВЧ. – 1990.–Вып. 2 (426).– С. 22–27.

1.30. Гарбер Г.З. Моделирование работы полевых транзисторов с субмикронным затвором Шоттки на арсениде галлия // Электронная техника. Сер. 2, Полупроводниковые приборы. – 1985. – Вып. 2. – С. 103–107.

1.31. Мартынов Я.Б. Двумерная численная модель статических и высокочастотных характеристик полевого транзистора с затвором Шоттки // Электронная техника. Сер. I, Электроника СВЧ. – 1984.–Вып.4 (364).–С.29–32.

- 1.32. Москалюк В.А., Минаков В.В., Касиян В.Т. Программа моделирования полупроводниковых приборов субмикронных размеров многочастичным методом Монте-Карло // Электронная техника. Сер. I, Электроника СВЧ. – 1988. – Вып. 8 (412). – С. 71–73.
- 1.33. Москалюк В.А., Тимофеев В.И., Шовкун И.Д. Распределенная модель СВЧ полевого транзистора // Изв. вузов. Сер. Радиоэлектроника. – 1986. – Т. 29. – № 6. – С. 87–89.
- 1.34. Москалюк В.А., Тимофеев В.И. Моделирование СВЧ полевого транзистора с топологическими неоднородностями // Автоматизация проектирования в электронике. – 1985. – Вып. 32. – С. 92–96.
- 1.35. Иващук А.В., Тимофеев В.И. Вплив морфології омичних контактів на надвисокочастотні параметри польових транзисторів // Оптоэлектроника и полупроводниковая техника. – 2000. – №35. – С.157–161.
- 1.36. Иващук А.В., Матуленіс А., Тимофеев В.И. Вплив морфології омичних контактів на шуми гарячих електронів в транзисторних структурах // Электроника и связь. – 1999. – №7. – С.96–99.
- 1.37. Гусельников Н.А., Москалюк В.А., Тимофеев В.И. О создании системы моделирования твердотельных СВЧ-приборов // Автоматизация проектирования в электронике. – 1993. – Вып. 48. С. 15–21.
- 1.38. Сигорский В.П. Моделирование электронных компонентов в системах автоматизированного проектирования // Изв. вузов. Сер. Радиоэлектроника. – 1986. – Т. 29. – №6. – С. 3–15.
- 1.39. Сигорский В.П. Проблемная адаптация в системах автоматизированного проектирования // Изв. вузов. Сер. Радиоэлектроника. – 1988. – Т.31. – №6. – С. 5–22.

2. ПОЛУПРОВОДНИКОВЫЕ МАТЕРИАЛЫ

2.1. ЗОННАЯ СТРУКТУРА ПОЛУПРОВОДНИКОВ

Полупроводниковые свойства имеют простые вещества: бор В, углерод (графит) С, кремний Si, германий Ge, олово (серое) Sn, фосфор Р, мышьяк As, сурьма Sb, сера S, селен (красный) Se, теллур Те, йод J. К полупроводникам относятся также бинарные соединения типа A_xB_{8-x} , где А – элемент группы x , а В – элемент группы $(8-x)$. Наиболее распространены и применимы из них – вещества типа A_3B_5 , к которым относятся соединения азота (нитриды), фосфора (фосфиды), мышьяка (арсениды), сурьмы (антимониды) с элементами третьей группы (бора, алюминия, галлия, индия). Перспективными являются соединения типа A_2B_6 . Полупроводниковые свойства проявляют тройные и более сложные соединения, а также некоторые органические вещества.

2.1.1. Атомарные полупроводники

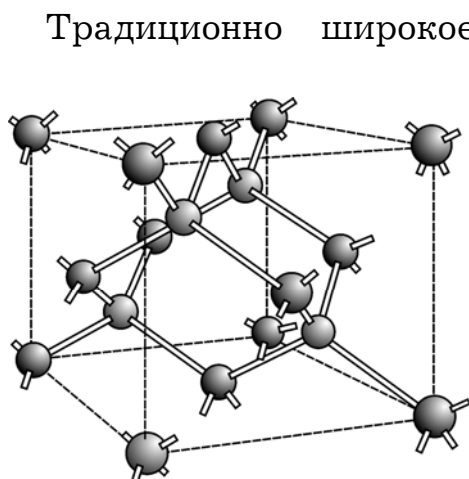


Рис. 2.1. Кристаллическая структура алмаза, состоящая из двух гранецентрированных кубических подрешеток.

Традиционно широкое применение в электронике имеют атомарные полупроводниковые вещества (германий и кремний). Структура верхних оболочек этих атомов, а также других атомов этой подгруппы IV группы имеет вид ns^2np^2 . Из-за близкого энергетического расположения s - и p -уровней эти атомы являются четырехвалентными. Объединяясь в кристалл при помощи ковалентных связей, они образуют решетку такую же, как и у алмаза. Она содержит две подрешетки из гранецентрированных кубических ячеек, сдвинутых друг относительно друга на четверть диагонали куба (рис. 2.1). В таком кристалле каждый атом связан с четырьмя ближайшими атомами. В

каждой связи участвуют два электрона, по одному от каждого атома, что является характерным для ковалентной связи.

Рассмотрим особенности зонной структуры наиболее распространенных полупроводниковых материалов. Нужно отметить, что в разных источниках вид и параметры дисперсионных кривых иногда заметно отличаются, что может объясняться как различием в методах вычисления этих кривых, так и разной степенью их детализации. Не останавливаясь на этих отличиях, рассмотрим только наиболее существенные особенности законов дисперсии.

Периодичность закона дисперсии электронов в кристалле дает возможность его изображения в пределах только первой зоны Бриллюэна. Для решетки типа алмаза она имеет вид октаэдра со срезанными вершинами (рис. 2.2). Тут же показано, как такими фигурами можно заполнить все k -пространство. Вместо проекций волнового вектора на основные направления, заданные

числовым набором кристаллографических индексов, часто используют другую систему обозначений. В ней характерные точки и линии симметрии в зоне Бриллюэна обозначаются буквами, а закон дисперсии задается в виде одномерной функции вдоль тех или иных направлений между двумя точками. Центр зоны Бриллюэна принято обозначать буквой Γ , направления эквивалентные $[100]$ — Δ , а соответствующие им шесть точек на поверхности зоны — X . Восемь эквивалентных направлений $[111]$ обозначают буквой — Λ , а точки их пересечения с поверхностью — L .

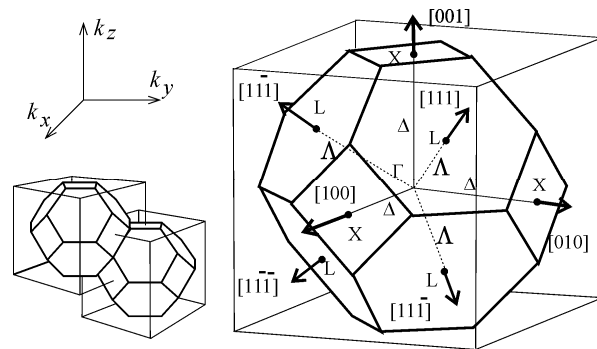


Рис. 2.2. Форма первой зоны Бриллюэна для кристаллической решетки типа алмаза и обозначение характерных точек и направлений симметрии [2.1]

Кремний. Вид дисперсионных кривых кремния для двух направлений волнового вектора приведен на рис. 2.3.

Волновые функции электронов и соответствующие им дисперсионные кривые для валентной зоны образуются из волновых функций электронов в атомарных s - и p -состояниях в результате sp^3 -гибридизации. Им соответствуют три кривые на рис. 2.3. Одна из кривых лежит у нижнего края валентной зоны и на рисунке не показана. Максимум энергии валентной зоны вырожден и находится в центре зоны Бриллюэна для двух ветвей, для третьей за счет спин-орбитального взаимодействия он

опускается на 0,035 эВ. Разная кривизна этих ветвей в максимуме дает основание говорить о разной эффективной массе подвижных носителей: для “легких” дырок $m_{p,л}^* \approx 0,16m$, для “тяжелых” $m_{p,т}^* \approx 0,5m$. Для вычисления эффективной плотности состояний у края валентной зоны удобно пользоваться некоторой усредненной массой:

$$m_{pd}^* = \left(m_{p,л}^{*3/2} + m_{p,т}^{*3/2} \right)^{2/3} \approx 0,56m.$$

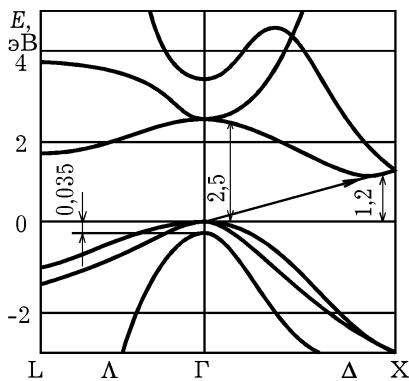


Рис. 2.3. Дисперсионные кривые в зоне проводимости и валентной зоне кремния в Λ - и Δ -направлениях. Расстояние между абсолютными экстремумами кривых соответствует запрещенной зоне.

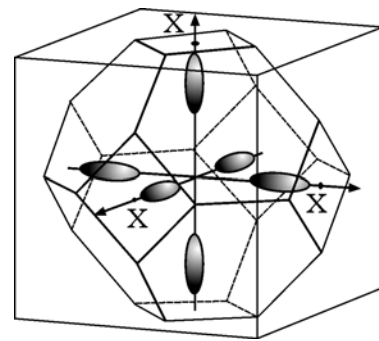


Рис. 2.4. Расположение изоэнергетических поверхностей, соответствующих Δ -долинам у дна зоны проводимости кремния, в первой зоне Бриллюэна.

Структура зоны проводимости еще более сложная, чем валентной. Абсолютный минимум в зоне проводимости имеет одна из ветвей в Δ -направлении на расстоянии 0,855 отрезка Γ -X. Принимая во внимания симметрию зоны Бриллюэна (рис. 2.2), можно считать, что этот минимум состоит из шести эквивалентных минимумов, или, как часто их называют, *долин*. На рис. 2.4 эти долины изображены при помощи изоэнергетических поверхностей, которые имеют форму эллипсоида вращения. Так как полуоси эллипсоида характеризуют компоненты тензора эффективной массы, удобно говорить о *продольной* и *поперечной* эффективных массах электронов, которые для кремния соответственно имеют значения: $m_l^* \approx 0,98m$ и $m_t^* \approx 0,19m$. При определении удельной электропроводности пользуются некоторой усредненной скалярной величиной эффективной массы, которую называют *омической* и обратную величину которой находят как среднее арифметическое из обратных значений эффективных масс в трех направлениях.

Поэтому $m_c^* = 3m_t^*m_l / (2m_l^* + m_t^*) \approx 0,26m$. Наличие шести эквивалентных долин у дна зоны проводимости означает шестикратное увеличение количества состояний. Это учитывают, сохраняя неизменным вид выражения для функции распределения плотности состояний и используя эффективную массу плотности состояний, как среднее геометрическое из компонентов тензора эффективной массы, умноженное на коэффициент $M^{2/3}$. Для кремния $M = 6$, поэтому $m_d^* = 6^{2/3}(m_t^{*2}m_l^*)^{1/3} \approx 1,08m$. Ширина запрещенной зоны у кремния различна в разных точках зоны Бриллюэна. В центре зоны при $k = 0$ она равняется $\approx 2,5$ эВ. Расстояние между абсолютными экстремумами, т.е. для кремния расстояние от максимума валентной зоны в точке Γ до минимума зоны проводимости в Δ -направлении равно $\approx 1,2$ эВ при абсолютном нуле и $\approx 1,1$ эВ при комнатной температуре, уменьшаясь с ростом температуры. В области комнатных температур иногда считают, что $\Delta E(T) = \Delta E(0) + \alpha T$, где температурный коэффициент ширины запрещенной зоны $\alpha = -4 \cdot 10^{-4}$ эВ/К.

С точки зрения формы запрещенной зоны кремний относят к так называемым *непрямозонным* полупроводникам. Это значит, что получив энергию, равную минимальному расстоянию между зонами, валентный электрон может осуществить переход в зону проводимости, если меняется также и значение его квазиимпульса. Такой переход, обозначенный стрелкой на рис. 2.3, называется *непрямым*. Прямой переход осуществляется без изменения квазиимпульса и требует большей энергии. Важным параметром с точки зрения эмиссионных и контактных свойств является *энергия электронного сродства*, которая определяется как расстояние от дна зоны проводимости до уровня вакуума и для кремния равна 4,05 эВ.

Германий. Законы дисперсии и структура энергетических зон германия во многом подобны кремнию. Валентная зона (рис. 2.5) имеет вырожденный максимум в центре зоны Бриллюэна и характеризуется наличием “легких” и “тяжелых” дырок ($m_{pl}^* = 0,04m$, $m_{pt}^* = 0,34m$), что соответствует эффективной массе плотности состояний $m_{pd}^* = (m_{pl}^{*3/2} + m_{pt}^{*3/2})^{2/3} \approx 0,36m$. Третья ветвь отстает от края зоны на величину энергии спин-орбитального взаимодействия, которая равна 0,28 эВ.

Наиболее существенное отличие от кремния состоит в том, что абсолютный минимум зоны проводимости лежит на границе первой зоны Бриллюэна в точке L . Соответствующие этому минимуму

изоэнергетические поверхности изображены на рис. 2.6 и имеют вид эллипсоидов вращения. Длинные оси этих эллипсоидов расположены в Δ -направлениях симметрично под углом $\arcsin(\sqrt{2}/\sqrt{3})$ относительно Δ -направления. Получить значение омической эффективной массы в этом направлении через продольные и поперечные эффективные массы можно, используя соответствующее превращение тензора эффективной массы: $m_c^* = 3m_t^*m_l^* / (2m_l^* + m_t^*)$. Анизотропия эффективной массы у германия выражена еще больше ($m_l^* = 1,58m$, $m_t^* = 0,08m$), поэтому скалярное значение *омической* эффективной массы $m_c^* \approx 3m_t^*/2 \approx 0,12m$. Эффективная масса *плотности состояний* электронов в зоне проводимости $m_d^* = 4^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 0,56m$ учитывает наличие восьми эквивалентных долин, каждая из которых лежит лишь наполовину в первой зоне Бриллюэна (рис. 2.6), поэтому $M = 8/2 = 4$.

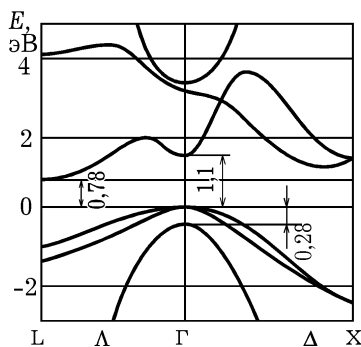


Рис. 2.5. Дисперсионные кривые в зоне проводимости и валентной зоне германия в Δ - и Δ - направлениях. Ширина запрещенной зоны германия – расстояние между абсолютными экстремумами кривых.

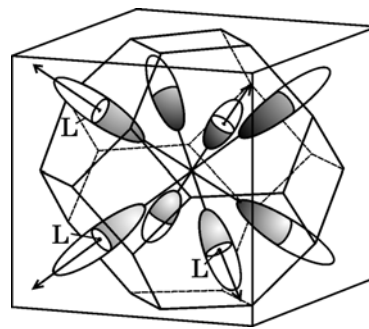


Рис. 2.6. Изоэнергетические поверхности, соответствующие L-долинам у дна зоны проводимости германия. Эллипсоиды вращения лежат лишь наполовину в первой зоне Бриллюэна.

Ширина запрещенной зоны германия в центре зоны Бриллюэна 1,1 эВ. Расстояние между абсолютными экстремумами энергетических зон равно 0,78 эВ при нулевой температуре и 0,66 эВ при комнатной. Температурный коэффициент ширины запрещенной зоны практически такой же, как и у кремния, энергия электронного сродства у германия 4,0 эВ.

Соединение $\text{Si}_x\text{Ge}_{1-x}$ является сплавом кремния и германия с их процентным содержанием, которое задается параметром x . Кристаллическая решетка, имея вид решетки алмаза, содержит большое количество дефектов, вызванных случайным расположением атомов кремния и германия. Структура

энергетических зон имеет общие черты, свойственные и кремнию и германию. Дисперсионные кривые в зоне проводимости содержат те же характерные долины, однако их расположение зависит от состава соединения. Естественно, что при крайних значениях $x = 0$ и $x = 1$ зонная структура приобретает соответственно вид для чистого германия или кремния. При увеличении содержания кремния положения долин изменяются, как указано на рис. 2.7: быстро возрастает уровень дна Γ - и L -долин и относительно плавно X -долины. Для $x = 0,15$ положения X - и L -долин выравниваются и ширина запрещенной зоны составляет 0,86 эВ. При последующем увеличении части кремния она постепенно приближается к значению 1,11 эВ.

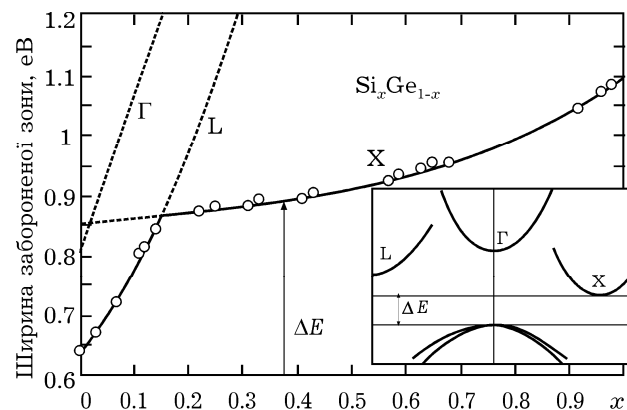


Рис. 2.7. Энергетическое положение долин в зоне проводимости $\text{Si}_x\text{Ge}_{1-x}$ отсчитанное от потолка валентной зоны, в зависимости от доли кремния

В [2.2] приведены аппроксимации, которые определяют положение соответствующих долин относительно потолка валентной зоны:

$$E_X = a_1 + b_1x + c_1x^2, \text{ де } a_1 = 0,86, b_1 = -0,0658, c_1 = 0,318;$$

$$E_L = a_2 + b_2x + c_2x^2, \text{ де } a_2 = 0,635, b_2 = 1,505, c_2 = 0,1205;$$

$$E_\Gamma = a_3 + b_3x, \text{ де } a_3 = 0,805, b_3 = 2,615.$$

2.1.2. Бинарные полупроводники $\text{A}^{\text{III}}\text{B}^{\text{V}}$

Разнообразие свойств соединений типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$ обуславливает их широкое использование в приборах и устройствах разного технического назначения. Соединения типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$ являются

ближайшими электронными аналогами кремния и германия. Они образуются в результате взаимодействия элементов III-б подгруппы периодической таблицы (бора, алюминия, галлия, индия) с элементами V-б подгруппы (азотом, фосфором, мышьяком и сурьмой). Висмут и таллий не образуют соединений данного ряда. Соединения $A^{III}B^V$ принято классифицировать по металлоидному элементу. Соответственно, различают нитриды, фосфида, арсениды и антимониды.

Для соединений $A^{III}B^V$ характерен особый тип химической связи, так называемый донорно-акцепторный. Из четырех ковалентных связей, которыми каждый атом встраивается в решетку, три образуются обобществлением валентных электронов атомов A^{III} , а четвертая связь осуществляется неразделенной парой валентных атомов B^V . Образование этой связи соответствует энергетически выгодному переходу электронов от атома B^V в энергетический избыток, общий для донора (атомов B^V) и акцептора (атома A^{III}). В каждой ковалентной связи максимум электронной плотности смещен в сторону атома с большей электроотрицательностью, то есть электронная плотность смещена к узлам решетки, где находятся атомы B^V . Благодаря такой поляризации связей атомы A^{III} приобретают некоторый эффективный положительный заряд, а атомы B^V – отрицательный. Величина этого эффективного заряда определяет степень ионности соединения, которая естественно изменяется при переходе от одного соединения к другому [2.3] в соответствии с положением химических элементов в периодической таблице Д. И. Менделеева.

Полупроводниковые соединения $A^{III}B^V$ образуют гомологический ряд, в котором наблюдается закономерное изменение многих свойств при изменении атомных номеров компонентов. Эти закономерности можно проследить с помощью таблицы 2.1.

Точки плавления лежат выше соответствующих температур плавления элементов, из которых состоит соединение, за исключением антимонида индия, температура которого (525°C) лежит между температурой плавления сурьмы (630°C) и индия (157°C).

Внутри каждой группы соединений наблюдается уменьшение ширины запрещенной зоны с ростом суммарного атомного номера и атомных масс элементов, входящих в соединение. Это объясняется более сильным размытием «электронных облаков» ковалентных связей по мере движения вниз по периодической таблице, благодаря чему усиливается металлический характер связи. Одновременно с ослаблением

ковалентных химических связей снижается температура плавления и твердость материала, то есть полупроводник становится более пластичным [2.5].

Таблица 2.1. Свойства и параметры полупроводников типа $A^{III}B^V$ [2.4]

Соединение	Период решетки, $\times 10$, нм	Плотность, $Mг/м^3$	Температура плавления, $^{\circ}C$	Твердость ¹⁾	$A \times 10^6, K^{-1}$ 2)	Ширина запрещенной зоны, эВ	Подвижность электронов, $м^2/(В \times с)$	Подвижность дырок, $м^2/(В \times с)$	Показатель преломления при $h\nu = \Delta E$	Диэлектрическая проницаемость ³⁾
$BN_{(куб)}$	3,615	3,49	3000	10	-	6,0	-	-	2,1	7,1
AlN	3,110 4,975	3,28	2400	7	6,1	5,88	-	-	2,2	9,1
GaN	3,186 5,176	6,11	1700	-	5,65	3,40	0,03	-	2,4	12,2
InN	3,540 5,704	6,91	1100	-	-	1,95	-	-	2,9	-
AlP	5,463	2,37	2000	5,5	4,2	2,45	0,008	0,003	3,0	9,8
GaP	5,451	4,07	1467	5	5,9	2,26	0,019	0,012	3,45	11,1
InP	5,869	4,78	1070	-	4,6	1,35	0,46	0,015	3,45	12,4
AlAs	5,661	3,60	1770	5	5,2	2,16	0,02	-	3,2	10,1
GaAs	5,653	5,32	1237	4,5	6,4	1,43	0,8	0,045	3,65	13,1
InAs	6,058	5,67	942	4	5,3	0,36	0,95 3,3	0,046	3,52	14,6
AlSb	6,136	4,28	1060	4,8	4,2	1,58	0,02	0,055	3,4	14,4
GaSb	6,096	5,65	710	4,5	6,2	0,72	0,4	0,14	3,8	15,7
InSb	6,479	5,78	525	3,8	4,9	0,18	7,8	0,075	4,0	17,7

¹⁾ Твердость по минералогической шкале.

Это иллюстрируется рис. 2.8, где для кубической структуры кристаллической решетки представлены поверхности, соответствующие положениям дна зоны проводимости в точках симметрии Γ , L и X, построенные по данным, приведенным в работе [2.6]. Следует обратить внимание, что взаимное расположение Γ -, L- и X-долин может изменяться в зависимости от химического состава. Прямозонными полупроводниками будут соединения индия и галлия (кроме GaP). Соединения алюминия и фосфида галлия имеют дно зоны проводимости в точке X.

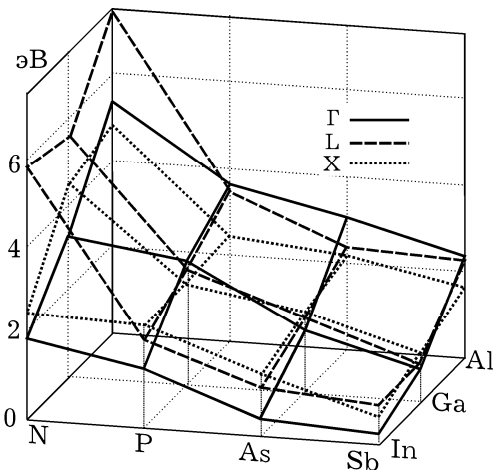


Рис. 2.8. Зависимости ширины запрещенной зоны в разных точках симметрии (Γ , L и X) зоны Бриллюэна от состава соединения.

За исключением нитридов все соединения $A^{III}B^V$ кристаллизуются в решетках (рис. 2.9) кубического типа (цинковой обманки, сфалерита). Для гранецентрированной кубической решетки типа цинковой обманки, как и алмаза, первая зона Бриллюэна имеет вид октаэдра со срезанными вершинами (рис. 2.9). Ее симметрия позволяет задавать закон дисперсии только в пределах $1/64$ части зоны Бриллюэна.

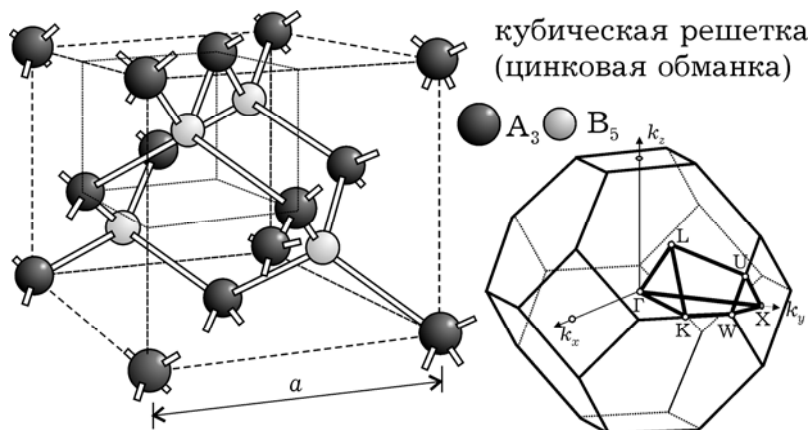


Рис. 2.9. Структура кубической кристаллической решетки типа цинковой обманки и форма первой зоны Бриллюэна. Выделен минимальный элемент симметрии – $1/64$ часть зоны

Для нитридов характерна также структура гексогонального типа (вюрцит) (рис. 2.10). В решетках обоих типов каждый атом элемента III группы находится в окружении тетраэдра четырех атомов элемента V группы и наоборот. Структура вюрцита в отличие от структуры алмаза не имеет центра симметрии. Эта особенность приводит к отличию в свойствах поверхностей $[111]$ и $[1\bar{1}0]$, полностью состоящих из разноименных атомов. Разное поведение граней сказывается при травлении, окислении, и при выращивании кристаллов.

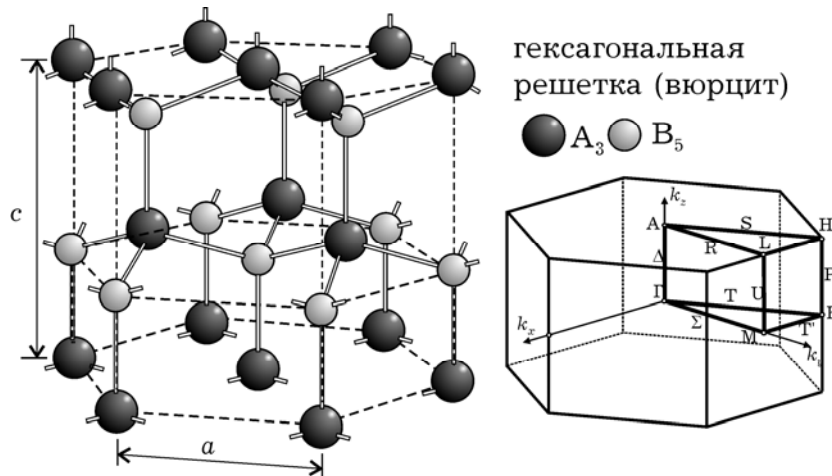


Рис. 2.10. Структура гексагональной кристаллической решетки типа вюрцита и форма первой зоны Бриллюэна. Выделен наименьший элемент симметрии – $1/24$ часть зоны

Рассмотрим особенности зонной структуры некоторых наиболее перспективных с точки зрения быстродействия полупроводниковых материалов GaAs, GaN и InP [2.6, 2.7].

Арсенид галлия. GaAs является наиболее распространенным полупроводником из группы соединений A_3B_5 .

В структуре энергетических зон атомарных полупроводников IV группы и соединений $A^{III}B^V$ много общего. Так, например, характер дисперсионных кривых в валентной зоне арсенида галлия качественно аналогичен германию и кремнию. В центре зоны (рис. 2.11) находятся максимумы двух ветвей, соответствующих “легким” ($m_{рл}^* \approx 0,1m$) и “тяжелым” ($m_{рт}^* \approx 0,5m$) дыркам. Третья ветвь за счет спин-орбитального взаимодействия лежит ниже на 0,33 эВ.

В отличие от германия и кремния большинство соединений A_3B_5 являются *прямозонными* полупроводниками, т.е. абсолютный минимум зоны проводимости лежит в центре зоны Бриллюэна (точка Γ). Для GaAs расстояние между экстремумами составляет при нулевой температуре 1,5 эВ, уменьшаясь до 1,45 эВ при комнатной. Эффективная масса электрона вблизи абсолютного экстремума (в Γ -долине) является скалярной величиной ($m^* \approx 0,067m$), так как соответствующие изоэнергетические поверхности являются практически сферическими.

Одной из наиболее ярких особенностей арсенида галлия является наличие верхней долины на высоте $\approx 0,3$ эВ от края зоны проводимости в точке L. Изоэнергетические поверхности у дна этой долины представляют собой эллипсоиды вращения, еще более вытянутые, чем у германия. Это говорит о большом различии

продольной и поперечной эффективных масс, которые имеют соответственно значения: $m_l^* = 1,9m$, $m_t^* = 0,075m$.

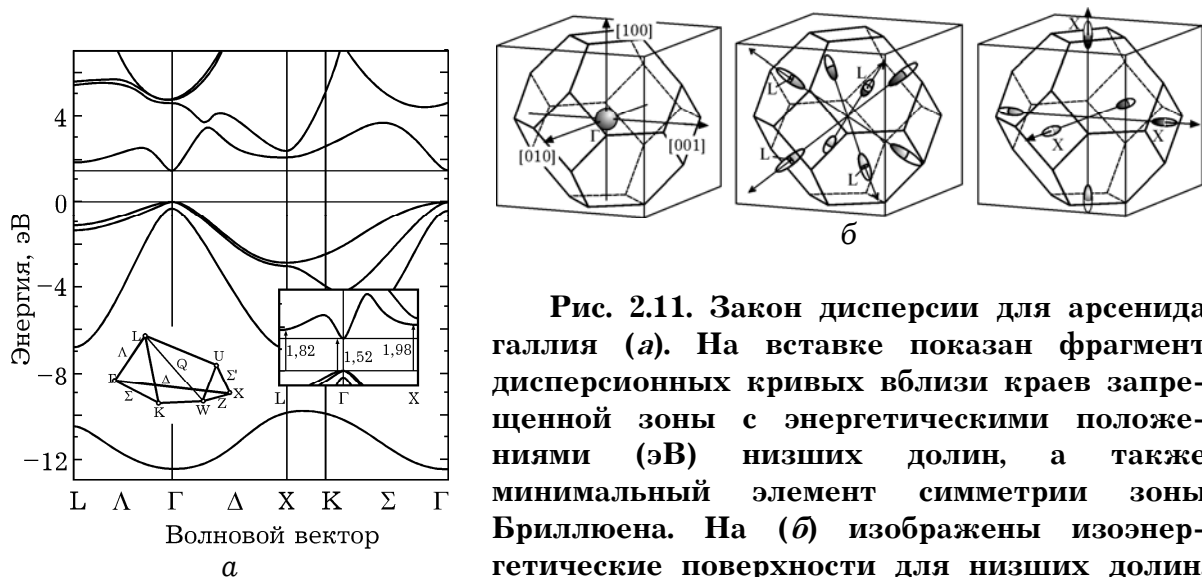


Рис. 2.11. Закон дисперсии для арсенида галлия (а). На вставке показан фрагмент дисперсионных кривых вблизи краев запрещенной зоны с энергетическими положениями (эВ) низших долин, а также минимальный элемент симметрии зоны Бриллюена. На (б) изображены изоэнергетические поверхности для низших долин. Γ -долина практически изотропна

Так же, как и в германии, рассчитываются омическая эффективная масса $m_c^* \approx 3m_t^*/2 \approx 0,11m$ и эффективная масса плотности состояний $m_d^* = 4^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 0,56m$. С этой особенностью Γ - и L-долин связан эффект возникновения отрицательной дифференциальной подвижности. В зоне проводимости GaAs есть и более высокие долины, например, на расстоянии $\approx 0,5$ эВ в X-направлении, но заметное влияние на электрические свойства они имеют только при наличии очень сильного электрического поля. Энергия электронного сродства составляет 4,07.

Нитриды III группы («тринитриды») периодической системы Менделеева могут иметь кристаллическую структуру сфалерита, вюрцита и каменной соли.

Нитрид галлия имеет две модификации кристаллической структуры: кубическую, подобную решетке для арсенида галлия, или решетки типа «цинковой обманки», и гексагональную типа «вюрцита». Для кубических решеток взяты параметры зонной структуры [2.6, 2.7], для которой в зоне проводимости в точках симметрии зоны Бриллюена имеются Γ -, X- и L-долины (рис. 2.12, а). Дну зоны проводимости соответствует центральная Γ -долина с изотропной эффективной массой $m_\Gamma^* = 0,13m$. Следующие шесть X-долин находятся на энергетическом расстоянии $\Delta_{\Gamma X} = 1,3$ эВ на

краю зоны Бриллюэна. Им соответствует заметная анизотропия эффективных масс $m_l^*/m_t^* = 0,5/0,3$. При этом омическая эффективная масса $m_c^* = 0,35m$, а эффективная масса плотности состояний с учетом количества долин ($M_X = 6/2$) $m_d^* = M^{2/3} \sqrt[3]{m_t^{*2} m_l^*} = 0,76m$. Более высокие L-долины находятся на 2,3 эВ выше, чем Γ -долина.

Для гексагональной структуры кристаллической решетки также характерно, что минимум зоны проводимости находится в центре зоны Бриллюэна (рис. 2.12, б), где эффективная масса является изотропной и составляет $(0,19...0,21)m$ [2.8]. Но кроме этого минимума, в центре зоны есть еще одна Γ -долина, выше первой на $(1,9...2,2)$ эВ. Однако, относительно положений верхних долин в научной литературе нет единства: некоторые авторы считают, что следующей является долина, расположенная между точками L и M на высоте от 1,2 до 2,0 эВ, другие такой считают A-долину – $(1,3...2,1)$ эВ [2.9]. Практически отсутствует информация относительно эффективной массы в верхних долинах, иногда ее принимают равной обычной массе. Изотропная эффективная масса для верхней Γ -долины составляет $(0,9...1)m$ [2.10]. Важным параметром с точки зрения эмиссионных и контактных свойств является энергия электронного сродства, которая определяется как расстояние от дна зоны проводимости до уровня вакуума и для нитрида галлия равна 4,1 эВ [2.9].

Нитрид алюминия (AlN), так же, как и нитрид галлия, кристаллизуется как в структуру цинковой обманки, так и в гексагональную структуру типа вюрцита.

Кубический нитрид алюминия в отличие от нитрида галлия является непрямозонным полупроводником (рис. 2.12, в), то есть абсолютный минимум зоны проводимости лежит не в центре зоны Бриллюэна, а соответствует кристаллографическому направлению $\langle 100 \rangle$ (X-долина) на расстоянии 4,9 эВ от максимума валентной зоны [2.10].

Ширина запрещенной зоны, которая соответствует центральной Γ -долине, составляет около 6 эВ, в ней эффективная масса изотропна и равна $0,4m$. На энергетическом расстоянии $\Delta_{\Gamma L}$ около 4,0 эВ находятся еще 4 эквивалентных L-долины. Анизотропия эффективной массы в них выражена практически так же, как и для GaN ($m_l^* = 0,53m$, $m_t^* = 0,35m$) [2.7], поэтому скалярное значение омической эффективной массы $m_c^* = 0,4m$. Эффективная масса плотности состояний электронов в зоне

проводимости $m_d^* = 4^{2/3} (m_t^* m_l^*)^{1/3} \approx 0,79m$. Энергия электронного сродства составляет лишь 0,6 эВ [2.9].

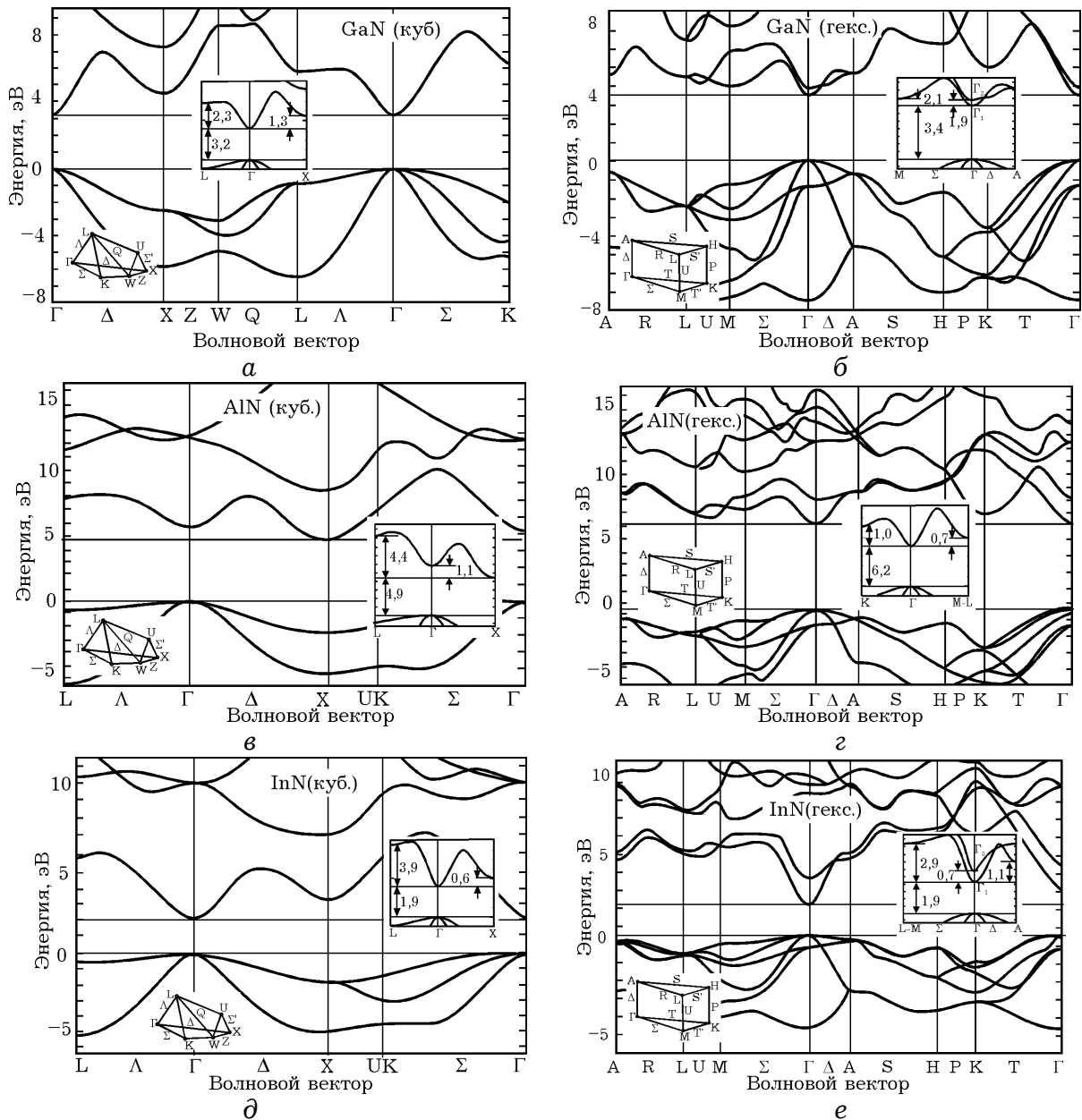


Рис. 2.12. Закон дисперсии для нитридов с кубической (а, в, д) и гексагональной (б, г, е) структурой кристаллической решетки. На вставках изображены фрагменты дисперсионных кривых у края запрещенной зоны с параметрами, которые использовались при расчетах, а также минимальный элемент симметрии зоны Бриллюэна с характерными направлениями и точками симметрии

Нитрид алюминия типа *вюрцита* имеет закон дисперсии (рис. 2.12, г) в общих чертах подобный гексагональному нитриду галлия и является прямозонным полупроводником, в отличие от своей кубической модификации.

Еще одной особенностью нитрида алюминия является его ширина запрещенной зоны, которая составляет в центре зоны Бриллюэна 6,2 эВ (Г-долина) [2.7]. Нитрид алюминия является самым широкозонным полупроводником из соединений $A^{III}B^V$. На энергетическом расстоянии 0,7 эВ находится следующая L-М-долина. Учитывая, что для вюрцитной модификации AlN $m_l^* = 0,35m$, $m_t^* = 0,29m$, эффективная масса плотности состояний в одной долине равна 0,17m [2.7]. Также на краю зоны Бриллюэна, но в Δ -направлении на расстоянии 1,0 эВ лежат шесть эквивалентных К-долин с эффективной массой плотности электронных состояний с учетом всех долин 1,02m. Энергия электронного сродства составляет лишь 0,6 эВ [2.9]. Принимая во внимание то, что нитрид алюминия вызывает интерес своей шириной запрещенной зоны и исключительными электронными свойствами сравнительно недавно, потому и сведения относительно зонной структуры нитрида алюминия в литературе часто различаются.

Нитрид индия типа цинковой обманки является прямозонным полупроводником и по структуре своей дисперсионной кривой (рис. 2.12, д) подобен кубическому GaN. Для кубического нитрида индия характерна в зоне проводимости в точках симметрии зоны Бриллюэна последовательность Г-, X- и L-долин.

Абсолютный минимум зоны проводимости лежит в центре зоны Бриллюэна (точка Г). Для InN расстояние между экстремумами составляет при нулевой температуре 2,0 эВ, уменьшаясь до 1,9 эВ при комнатной. Эффективная масса электрона вблизи абсолютного экстремума в Г-долине является скалярной величиной ($m^* \approx 0,12m$) [2.9].

Следующая X-долина находится на энергетическом расстоянии $\Delta_{ГХ} = 0,57$ эВ в середине зоны Бриллюэна. Ей соответствуют значения продольной и поперечной эффективных масс: $m_l^* = 0,48m$, $m_t^* = 0,27m$ [2.6]. При этом омиическая эффективная масса $m_c^* = 0,32m$, эффективная масса плотности состояний с учетом количества долин $m_d^* = 0,52m$.

Структура зоны проводимости для *гексагонального* нитрида индия несколько сложнее (рис. 2.12, е) и еще более неопределена. В [2.9] приводится ширина запрещенной зоны, которая в зависимости от температуры принимает значение около 2 эВ. По [2.10] это значение около (0,65...0,7) эВ. Дну зоны проводимости соответствует центральная Γ_1 -долина с практически изотропной эффективной массой, которая равна (0,04...0,07)m. Как и для

гексагонального GaN, из-за отсутствия точных экспериментальных данных, литературные источники расходятся относительно параметров, а соответственно и последовательности долин в зоне проводимости InN типа вюрцита. Одни авторы считают следующей A-долину, расположенную в Δ -направлении на энергетическом расстоянии $(0,7...2,7)$ эВ. Другие считают таковой вторую Γ_2 -долину, расположенную также в центре зоны Бриллюэна выше на $(1,1...2,6)$ эВ [2.8, 2.9]. Им соответствует значительная анизотропия эффективных масс $m_l^*/m_t^* = 0,84/0,16$, поэтому скалярное значение омической эффективной массы $m_c^* \approx 3m_t^*/2 \approx 0,24m$, а эффективная масса плотности состояний в одной долине $m_d^* = 0,28m$. Отмечают и наличие высших долин в зоне проводимости InN, лежащих между точками L и M на высоте $\Delta_{\Gamma ML} = (2,9...3,9)$ эВ [2.9], с эффективной массой плотности состояний с учетом всех долин $1,08m$. Энергия электронного сродства для данного материала неизвестна.

Фосфид индия имеет закон дисперсии (рис. 2.13), в общих чертах подобный арсениду галлия. Ширина запрещенной зоны составляет 1,34 эВ. В зоне проводимости самой низкой является центральная Γ -долина с практически изотропной эффективной массой, которая равна $(0,064...0,08)m$ [2.6, 2.9]. На энергетическом расстоянии 0,59 эВ находятся восемь L-долин на краю зоны Бриллюэна с эффективной массой плотности состояний в одной долине $0,25m$, или с учетом всех долин $0,63m$. Также на краю зоны Бриллюэна, но в Δ -направлении на расстоянии 0,85 эВ лежат шесть X-долин с эффективными массами плотности электронных состояний соответственно $0,32m$ и $0,66m$. Энергия электронного сродства составляет 4,38 эВ.

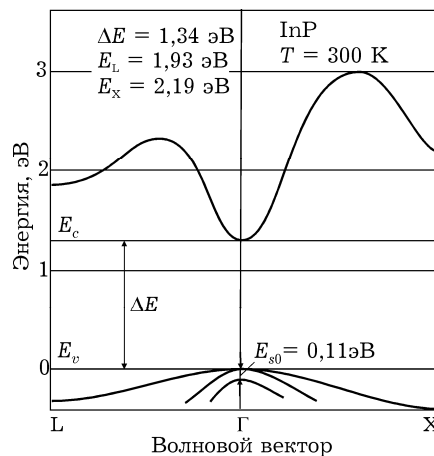


Рис. 2.13. Закон дисперсии для фосфида индия (реконструировано из [2.9]). Характерной является ΓLX -последовательность долин в зоне проводимости

2.1.3. Сложные соединения $A^{III}B^V$

Широкое использование в электронике нашли тройные соединения материалов A^{III} и B^V , например, InGaAs, AlGaAs, InAsP и тому подобные. Не рассматривая всего разнообразия этих соединений (достаточно полный обзор их свойств сделан в работах [2.6, 2.7]), некоторые тенденции определим на примере $Al_xGa_{1-x}As$. Это соединение можно рассматривать как сплав бинарных арсенидов Al и Ga, доля которых определяется параметром x . При возрастании x от 0 до 1 свойства соединения постепенно изменяются от таковых для чистого GaAs к чистому AlAs. Прежде всего нужно отметить изменение зонной структуры (рис. 2.14).

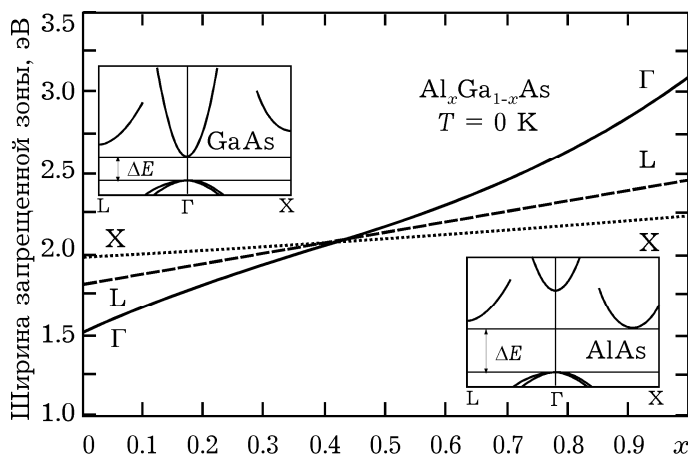


Рис. 2.14. Энергетическое положение долин в зоне проводимости $Al_xGa_{1-x}As$, отсчитанное от потолка валентной зоны, в зависимости от части алюминия. На вставках изображены эскизы законов дисперсии для чистых материалов

В отличие от арсенида галлия арсенид алюминия является непрямозонным полупроводником, в котором самой низкой является X-долина на расстоянии 2,17 эВ от валентной зоны. Точнее, ее минимум находится на некотором расстоянии от края зоны Бриллюэна в Δ -направлении. При увеличении доли алюминия положения долин изменяются: они поднимаются, но по-разному, так что последовательность ΓLX - в GaAs изменяется на последовательность $XL\Gamma$ - в AlAs. Интересной особенностью соединения является то, что при $x = 0,45$ положения всех долин совпадают.

В [2.10] приведены аппроксимации, которые определяют изменение положения соответствующих долин относительно потолка валентной зоны от части алюминия:

$$\begin{aligned}
 E_{\Gamma}(x) &= 1,424 + 1,247x \quad \text{для } 0 < x < 0,45, \\
 E_{\Gamma}(x) &= 1,424 + 1,247x + 1,147(x - 0,45)^2 \quad \text{для } 0,45 < x < 1,0, \\
 E_L(x) &= 1,708 + 0,642x, \\
 E_X(x) &= 1,900 + 0,125x + 0,143x^2.
 \end{aligned}$$

Здесь нужно обратить внимание на то, что эти данные приведены для $T = 300$ К, а на рис. 2.14 для $T = 0$ К. Обычно для температурной зависимости ширины запрещенной зоны для разных долин ($i = \Gamma, L, X$) используют соотношение [2.6]

$$\Delta E_i(T) = \Delta E_i(0) - \frac{\alpha_i T^2}{T + \beta_i},$$

где α_i и β_i – так называемые *Varshni*-параметры, которые служат для аппроксимации экспериментальных зависимостей.

С изменением доли алюминия изменяется и кривизна дисперсионных кривых, которая определяет эффективную массу.

Таблица 2.2. Линейная аппроксимация значений эффективной массы в $\text{Al}_x\text{Ga}_{1-x}\text{As}$ [2.12].

эффективная масса	плотности состояний	омическая
m_{Γ}^*/m_0	$0,067 + 0,083x$	$0,067 + 0,083x$
m_L^*/m_0	$0,56 + 0,1x$	$0,11 + 0,03x$
m_X^*/m_0	$0,85 - 0,14x$	$0,32 - 0,06x$

Иногда для параметров сложных соединений используют квадратичные аппроксимации. Например, для ширины запрещенной зоны соединения бинарных соединений А и В пригодна билинейная форма

$$\Delta E(A_x B_{1-x}) = x\Delta E(A) + (1-x)\Delta E(B) - x(1-x)C,$$

где параметр C находится из результатов эксперимента.

Подобные соотношения используют и для более сложных, четверных, соединений, которые также можно рассматривать как сплавы простых бинарных соединений

2.1.4. Спектр фононов

На электронные свойства полупроводников в наибольшей степени, кроме зонной структуры, влияют особенности взаимодействия электронов с тепловыми колебаниями кристаллической решетки.

Во многих вопросах, связанных с тепловыми свойствами твердых тел, оправдывает себя модель Дебая, согласно которой тепловое возбуждение вызывает не колебания отдельных атомов, случайные и независимые друг от друга, а волновые процессы смещения атомов около положения равновесия, распространяющиеся в кристаллической решетке как едином целом.

Волновые колебания решетки, как и другие, квантуются, а кванты этих волн, *фононы*, рассматриваются как квазичастицы, перемещающиеся в кристалле со скоростью, соответствующей групповой скорости волновых процессов и в первом приближении равной скорости звука.

Спектр частот фононов определяется периодичностью структуры и параметрами решетки и описывается *законом дисперсии*. Как и для электронных волн в кристалле этот закон представляется в виде функций $\omega(\vec{k}_\phi)$ или $E(\vec{k}_\phi)$, где \vec{k}_ϕ – волновой (точнее квазиволновой) вектор фононов. Эта функция, как и для электронного закона дисперсии, является четной и периодической функцией с периодом $2\pi/a$ (для одномерного кристалла), поэтому достаточно определить его на половине зоны Бриллюэна.

Простейшая модель распространения колебаний в одномерной цепочке связанных упругими силами одинаковых атомов с периодом, равным a , дает закон дисперсии в виде [2.13]:

$$\omega(k_\phi) = \omega_m \sin(k_\phi a/2), \quad (2.1)$$

где $\omega_m = 2v_{зв}/a$ – предельная частота фононов (рис. 2.15, а), которая зависит от скорости распространения звука в твердом теле и соответствует самой короткой длине волны $\lambda_{\min} = 2a$, при которой соседние атомы колеблются в противофазе. Такие колебания соответствуют стоячей волне, групповая скорость которой $v_{гр} = d\omega/dk_\phi$ равна нулю. Иногда эту предельную частоту характеризуют так называемой *температурой Дебая* θ_D , определяемой из соотношения $\hbar\omega_m = k_B\theta_D$. Для типичных полупроводников период решетки составляет величины порядка 10^{-10} м, скорость звука $\sim 10^3$ м/с, поэтому предельные частоты ω_m порядка $10^{12} \dots 10^{13}$ Гц лежат в ультразвуковой области, а температуры Дебая составляют несколько сотен кельвинов. Для самых низких частот закон дисперсии линеен: $\omega \approx \omega_m k_\phi a/2$, поэтому фазовая и групповая скорости одинаковы и равны скорости звука. Такие колебания и соответственно фононы называют *акустическими*.

Для кристаллов, элементарная ячейка которых состоит из двух атомов, закон дисперсии (рис. 2.15, б, в) имеет существенные отличия – в нем появляется так называемая *оптическая ветвь*. Понять её происхождение можно с помощью сравнительно простой модели (рис. 2.15, б). К линейной цепочке из одинаковых атомов, расположенных на расстоянии $a' = a/2$, применим закон дисперсии (2.1), который при замене a' на a дает значение ω_m в

точках $k_{\phi} = \pm\pi/a' = \pm 2\pi/a$. Но этот интервал в два раза превышает размеры зоны Бриллюэна ($-\pi/a \leq k_{\phi} \leq \pi/a$) для кристалла с периодом a . Поэтому после приведения закона дисперсии к первой зоне Бриллюэна, как это делают для электронных волн [2.1], он принимает вид (рис. 2.15, в), где оптическая ветвь является продолжением акустической, смещенным на расстояние $2\pi/a$.

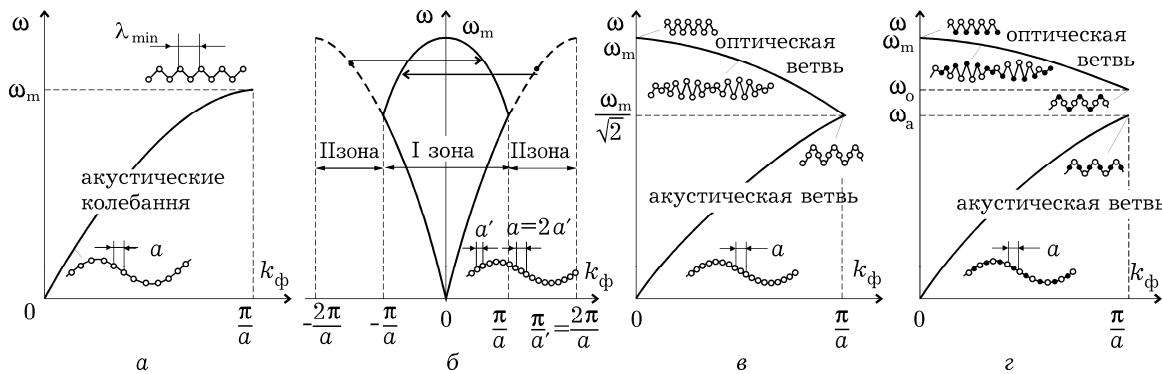


Рис. 2.15. Законы дисперсии волн смещения в одномерной модели кристалла – цепочке атомов: с одноатомным базисом (а), с базисом из двух одинаковых атомов (б, в) и двух разных атомов (г). Показано (б) образование оптической ветви для двухатомного базиса путем приведения верхней части “акустической” ветви к первой зоне Бриллюэна (масштаб по горизонтальной оси уменьшен). На фрагментах для удобства изображения продольные смещения атомов представлены как поперечные

Акустические колебания распространяются в цепочке атомов как в едином целом, а самые коротковолновые имеют частоту $\omega_m/\sqrt{2}$ (как в точке $k_{\phi} = \pi/2a$ на рис. 2.15, а), когда на одном периоде решетки укладывается половина длины волны и $\lambda = \lambda_{\min} = 2a$. Эта точка вырождена, она соответствует также нижней частоте оптической ветви, где с увеличением частоты увеличивается и длина волны, а соседние атомы одной ячейки колеблются в противофазе.

В принципе такая же картина наблюдается и в цепочке с периодом, состоящим из двух разноименных атомов, но с разной «жесткостью» связей слева и справа от атома. Из-за этого в точке $k_{\phi} = \pi/a$ вырождение снимается (рис. 2.15, г), так как акустические и оптические колебания имеют разные частоты и в спектре фононов образуется разрыв. Эти самые коротковолновые фононы, как акустические, так и оптические, иногда называют *междолинными*.

В рассмотренной одномерной модели возможны лишь продольные виды (*моды*) колебаний, хоть на приведенных рисунках они представлены как поперечные. Это сделано

исключительно для удобства изображения. В трехмерном кристалле кроме продольных имеется и две моды поперечных колебаний, причем в кубическом кристалле обе поперечные моды имеют одинаковый закон дисперсии. На рис. 2.16 изображены дисперсионные кривые для фононов в наиболее распространенных полупроводниковых материалах: кремнии, германии, арсениде галлия [2.14] и нитриде галлия [2.15]. Общий вид этих кривых для продольных мод (LA – акустическая, LO – оптическая) вполне соответствует описанной выше модели, только закон дисперсии описывается функцией, более сложной, чем (2.1).

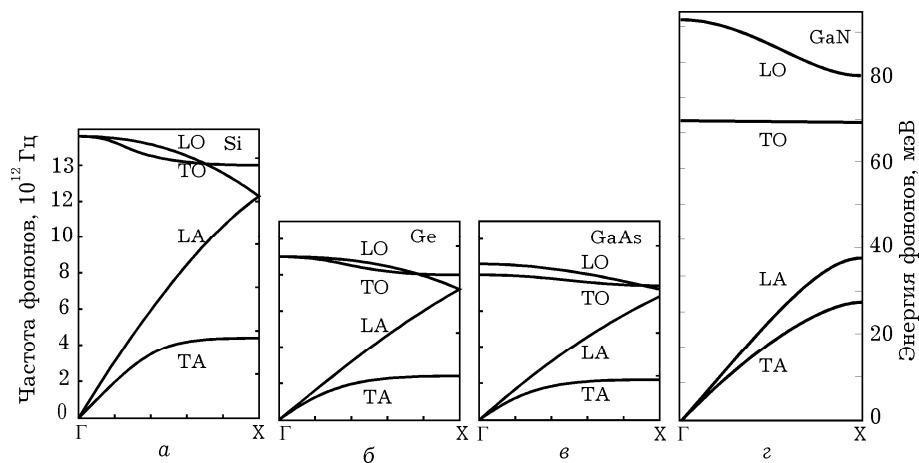


Рис. 2.16. Дисперсионные кривые фононов для кремния (*а*), германия (*б*), арсенида галлия (*в*) и кубического нитрида галлия (*г*) в кристаллографическом направлении [100] (L – продольные, Т – поперечные, А – акустические и О – оптические моды). На правых шкалах отложены энергии фононов, на левых – их частоты

В кремнии и германии, элементарные ячейки которых состоят из одинаковых атомов, ветви продольных LA- и LO-мод вырождены на краю зоны Бриллюэна, а энергия междолинных фононов равна $\sim 0,05$ эВ и $\sim 0,03$ эВ соответственно. Такое заметное различие связано прежде всего с разной массой атомов Si и Ge.

В арсениде галлия вырождение отсутствует, но небольшое отличие масс атомов галлия и мышьяка приводит и к небольшому различию частот оптической и акустической мод: $\omega_o/\omega_a = \sqrt{M_{As}/M_{Ga}}$. В других бинарных соединениях $A^{III}B^V$ из-за большего различия атомных масс подрешеток это отношение может быть существенно большим.

Так, например, в нитриде галлия (*г*) граничная частота оптических колебаний превышает частоту акустических более чем вдвое ($M_{Ga}/M_N \approx 5$).

На рис. 2.16. изображены также поперечные акустические (ТА) и оптические (ТО) ветви. В неполярных кристаллах германия и кремния поперечная и продольная оптические моды вырождены при $k_{\phi} = 0$. В бинарных соединениях, где оптические колебания вызывают образование дипольного электрического момента, эти моды имеют разные частоты, а при $k_{\phi} = 0$ выполняется соотношение

$$\omega_{LO}/\omega_{TO} = \sqrt{\epsilon_{ст}/\epsilon_{\infty}}, \quad (2.2)$$

где $\epsilon_{ст}$ и ϵ_{∞} — соответственно статическая и высокочастотная диэлектрическая проницаемость.

Характерной особенностью фононов, кроме их спектра, является распределение по энергиям. Среднее число фононов $N_{\phi}(\omega)$ с заданной энергией (частотой) определяется как для бесспиновых частиц статистикой Бозе-Эйнштейна и представляется формулой Планка:

$$N_{\phi}(\omega) = 1/[\exp(\hbar\omega/k_B T) - 1]. \quad (2.3)$$

Зависимость числа фононов от их энергии (рис. 2.17) показывает, что при $\hbar\omega > k_B T$ число фононов $N_{\phi} \ll 1$. Это условие обычно хорошо выполняется для оптических фононов при комнатных температурах. Для длинноволновых акустических фононов с частотами $\hbar\omega \ll k_B T$ можно приближенно считать, что $N_{\phi}(\omega) \approx k_B T / \hbar\omega$, то есть их количество пропорционально температуре.

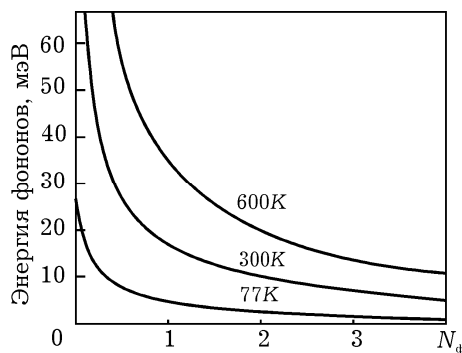


Рис. 2.17. Зависимость между средним числом фононов данной частоты и их энергией при различных температурах.

2.2. МЕХАНИЗМЫ И СКОРОСТИ РАССЕЯНИЯ

Многообразие и особенности эффектов переноса заряда в твердом теле зависит не только от конфигурации и величины внешних сил, но и от физических особенностей конкретных

механизмов соударений, приводящих к рассеянию направленных потоков носителей. Эти процессы изучаются в разделе физики твердого тела, называемом *теорией рассеяния*. Здесь рассмотрим характерные особенности некоторых важнейших механизмов рассеяния и принципы их учета в интеграле столкновений.

Прежде всего надо помнить, что в *идеальной* кристаллической решетке поведение электронов описывается волновыми функциями в виде волн Блоха, которые распространяются в кристалле в виде бегущих волн с постоянной вдоль кристалла амплитудой. Появление отраженных волн, то есть рассеяние электронов, возможно лишь при нарушении идеальной периодичности поля кристалла, что возможно при существовании каких-либо дефектов решетки. Нарушения периодичности могут быть вызваны двумя причинами. Первая из них связана со стационарными, локализованными в пространстве структурными дефектами – примесями, границами, различного рода дислокациями. Вторая – вызвана тепловыми колебаниями ионов в узлах кристаллической решетки относительно положения равновесия. Несмотря на то, что эти колебания представляют собой волновые возбуждения периодической структуры решетки, зависящие от времени и пространственно не локализованные, рассеяние электронов на них рассматривают как столкновения с квазичастицами, *фононами*, количество которых к тому же может изменяться в процессе столкновения. К подобным дефектам относятся также возбуждения в энергетическом спектре электронов (экситоны, плазмоны), при определенных условиях заметным оказывается также электрон-электронное и электрон-дырочное рассеяние. Но наиболее типичными являются примесное рассеяние и различные виды фононного рассеяния.

Эффективность соударений часто вместо времени релаксации τ оценивается обратной величиной, которая называется *скоростью рассеяния* и имеет смысл среднего количества актов рассеяния данного вида за единицу времени. Ее определяют непосредственным интегрированием вероятности переходов по всем конечным состояниям

$$\lambda(k) = 1 / \tau(k) = \int w_{kk'} dV_{k'}. \quad (2.4)$$

Именно эти величины используют при моделировании методом Монте-Карло для определения вероятности того или иного вида рассеивания. При наличии разных механизмов их скорости рассеивания суммируют

$$\lambda_{\Sigma} = \sum_i \lambda_i.$$

Это правило выполняется, если разные акты рассеивания происходят независимо друг от друга.

2.2.1. Фононное рассеяние

2.2.1.1. Общие закономерности

При взаимодействии носителя заряда с тепловыми колебаниями кристаллической решетки его энергия может увеличиваться или уменьшаться на величину кванта колебаний решетки, то есть энергии фонона. В этом случае говорят о *поглощении* или *возбуждении* фонона носителем заряда. Многофононные процессы маловероятны по сравнению с однофононными. Поэтому закон сохранения энергии при фононном рассеянии имеет вид:

$$E' - E = \pm \hbar\omega. \quad (2.5)$$

Кроме этого должен выполняться закон сохранения импульса (квазиимпульса), который удобнее выразить через волновые вектора:

$$\vec{k}' - \vec{k} = \pm \vec{k}_{\text{ф}}. \quad (2.6)$$

В приведенных формулах штрихами обозначены энергия и волновой вектор электрона после рассеяния, а знаки “±” относятся соответственно к поглощению фонона и к его возбуждению.

Квантовомеханический анализ процессов перехода между электронными состояниями дает выражения для вероятностей рассеяния с поглощением фонона электроном и с его возбуждением

$$w_{kk'}^{\pm} = w(\vec{k}_{\text{ф}}) (N_{\text{ф}} + 1/2 \mp 1/2) \delta(E' - E \mp \hbar\omega), \quad (2.7)$$

где $w(\vec{k}_{\text{ф}})$ – некоторая функция, зависящая от конкретного вида фононного рассеяния, δ -функция учитывает закон сохранения энергии, верхний знак относится к поглощению фонона, а нижний к его испусканию.

Пропорциональность вероятности числу фононов $(N_{\text{ф}} + 1/2 \mp 1/2)$ говорит о том, что скорость рассеяния с возбуждением фононов в $(N_{\text{ф}} + 1)/N_{\text{ф}}$ раз больше, чем при их поглощении. Однако это правило проявляется в полной мере, во-первых, при достаточно больших энергиях электрона. Если же $E < \hbar\omega$, то может произойти рассеяние только с поглощением

фонона. Во-вторых, это различие существенно при обычных температурах для оптических фононов, когда $\hbar\omega > k_B T$ и $N_\phi \ll 1$. Для акустических, особенно длинноволновых ($k_\phi \approx 0$), фононов $N_\phi \gg 1$ и скорости рассеяния с поглощением и возбуждением практически одинаковы.

При каждом акте соударения энергия электрона уменьшается или увеличивается на энергию фонона $\hbar\omega$. Однако в среднем за счет преобладания процессов возбуждения энергия электронов будет уменьшаться, а кристаллическая решетка нагреваться. Средняя энергия, отдаваемая электроном за одно столкновение, может быть определена как энергия одного фонона, умноженная на отношение разности вероятностей с излучением и поглощением к полной вероятности, т.е. их сумме:

$$\Delta E_1 = \hbar\omega \frac{w_{kk'}^- - w_{kk'}^+}{w_{kk'}^- + w_{kk'}^+}. \quad (2.8)$$

Отношение этой величины к энергии электрона определяет в среднем часть передаваемой им решетке энергии, то есть степень упругости фононных соударений, и называется *коэффициентом упругости*. Вводя понятие электронной температуры через соотношение $E = 3k_B T_e/2$ и используя выражения (2.7) для вероятностей рассеяния и (2.3) получим:

$$\zeta = \frac{\Delta E_1}{E} = \frac{\hbar\omega}{E} \frac{1}{2N_\phi + 1} = \frac{2\hbar\omega}{3k_B T_e} \operatorname{th} \frac{\hbar\omega}{2k_B T}. \quad (2.9)$$

При высоких температурах или малых энергиях фононов, когда $\hbar\omega \ll k_B T$ (при замене $\operatorname{th} x \approx x$), коэффициент упругости принимает значение

$$\zeta \approx (\hbar\omega)^2 / 3(k_B)^2 T T_e \ll 1, \quad (2.10)$$

и такие соударения можно считать практически упругими.

Энергия оптических фононов может быть соизмерима с величиной $k_B T_e$, поэтому их взаимодействие с электронами является существенно неупругим. Однако с ростом энергии электронов или электронной температуры упругость соударений возрастает так, что коэффициент упругости уменьшается обратно пропорционально T_e .

2.2.1.2. Акустическое рассеяние

При акустическом рассеянии основную роль играет взаимодействие электрона с продольными колебаниями решетки, которые вызывают изменения межатомных расстояний. Волновые процессы изменения параметра решетки a приводят к соответствующим изменениям запрещенной зоны. Поскольку длина волны акустических колебаний особенно при малых значениях k_{ϕ} гораздо больше межатомного расстояния, то здесь вполне применима макроскопическая модель искривления энергетических зон (разд. 3.6). В этом случае можно говорить о локальном изменении энергии электрона в зоне проводимости, которое пропорционально продольной деформации кристалла, то есть относительному изменению межатомного расстояния

$$\Delta E = D_a (\Delta a/a). \quad (2.11)$$

Коэффициент пропорциональности D_a (размерность эВ) называется константой *акустического деформационного потенциала*, а данный вид рассеяния – рассеянием на акустическом деформационном потенциале. Это рассеяние связано с возникновением локальных электрических полей, образующихся за счет изменения взаимодействия между атомами при продольных деформациях решетки и вызывающих изменения траектории движения электрона.

Поперечные акустические фононы связаны с деформациями сдвига атомных слоев и практически не изменяют межатомных расстояний, но могут вызвать пьезоэлектрический эффект. Однако *пьезоэлектрическое рассеяние* в типичных ситуациях намного слабее рассеяния на продольных акустических фононах.

Из совместного решения уравнений баланса энергии и импульса следует, что в рассеивании принимают участие лишь длинноволновые акустические фононы, энергия которых заметно меньше энергии электронов, которую в простейшем случае можно считать равной тепловой – $3k_B T_e/2$. Поэтому поглощение или испускание фонона практически не изменяет энергии электрона и рассеяние можно считать упругим ($\zeta \ll 1$).

При дрейфовых процессах переноса носителей часто принимают, что функция распределения фононов (2.3) остается равновесной, а при $\hbar\omega \ll k_B T$ количество фононов $N_{\phi} + 1 \approx N_{\phi} \approx k_B T/\hbar\omega \gg 1$ пропорционально температуре. Этот факт свидетельствует о том, что вероятности рассеяния с излучением и поглощением фонона в этих условиях практически одинаковы. Это позволяет относительно просто провести вычисление интеграла

(2.4) с учетом (2.7) и получить выражение для скорости акустического рассеяния [2.16]:

$$\lambda_a = \frac{(2m^*)^{3/2} D_a^2 k_B T}{2\pi \rho v_{зв}^2 \hbar^4} \sqrt{E}, \quad (2.12)$$

где m^* – эффективная масса плотности состояний, D_a – деформационный акустический потенциал, $v_{зв}$ – скорость звука, соответствующая наклону акустической ветви продольных колебаний при $k_f \approx 0$, ρ – плотность вещества.

Характерно, что время релаксации $\tau = 1/\lambda$ уменьшается с возрастанием энергии электрона обратно пропорционально $E^{1/2}$ или его скорости. В вырожденных электронных коллективах, например, в металлах энергия электронов определяется энергией Ферми и практически не зависит от температуры, поэтому удельная электропроводность $\sigma = en\mu$ падает с ростом температуры как $1/T$ из-за уменьшения подвижности $\mu = e\tau/m^*$. В невырожденных полупроводниках в слабых электрических полях в энергии электрона преобладает вклад теплового движения, $E = 3k_B T/2$ и подвижность изменяется $\sim T^{-3/2}$.

2.2.1.3. Оптическое рассеяние

Для оптических колебаний кристаллической решетки характерен пологий ход дисперсионной кривой (см. рис. 2.16), поэтому часто принимают, что частота оптических фононов ω_0 не зависит от волнового вектора и $\hbar\omega_0 = k_B \Theta_D$. Их энергия по порядку величины сравнима с тепловой энергией носителей, поэтому рассеяние на оптических фононах считается *неупругим*. Следовательно, условия рассеяния с излучением фононов и с их поглощением существенно различаются. Как следует из (2.5), поглощение фонона возможно даже при самых низких энергиях (температурах), а для испускания фонона энергия электрона должна быть не ниже энергии фонона: $E \geq \hbar\omega_0$. С другой стороны, пропорциональность вероятности рассеяния количеству фононов $(N_f + 1/2 \mp 1/2)$ приводит к различию скоростей рассеяния: вероятность испускания может быть существенно больше вероятности поглощения, если $N_f \ll 1$ (конечно, если энергия электрона достаточна для испускания фонона).

Взаимодействие с оптическими фононами в атомарных полупроводниках с чисто ковалентными (гомеополярными)

связями и в бинарных соединениях с заметной долей полярных (гетерополярных) связей имеет различные механизмы.

В атомарных полупроводниках противофазные продольные колебания атомов двух подрешеток, характерные для оптической ветви, не приводят к изменению ширины запрещенной зоны, как это было при акустических колебаниях, но изменяют потенциальную энергию взаимодействия между соседними атомами. В этих условиях изменяется и энергия электрона, причем пропорционально смещению: $\Delta E = D_0 \Delta x$. Коэффициент пропорциональности D_0 в этом случае имеет размерность эВ/см и называется константой *оптического деформационного потенциала*, а данный вид рассеяния называют рассеянием на оптическом деформационном потенциале или *неполярным оптическим рассеянием*. Для расчета скорости такого рассеяния можно воспользоваться выражением [2.16]

$$\lambda_o^\pm = \frac{(m^*)^{3/2} D_o^2}{\sqrt{2}\pi r \hbar^3 \omega_o} (N_\phi + 1/2 \mp 1/2) \operatorname{Re} \sqrt{E \pm \hbar \omega_o}. \quad (2.13)$$

Здесь, как и ранее, верхний знак относится к поглощению фонона, а нижний к испусканию. Оператор Re дает нулевую скорость испускания при $E \leq \hbar \omega_o$, а при поглощении фонона он лишен смысла

В бинарных полупроводниках, таких как, арсенид галлия, встречные продольные колебания подрешеток атомов разной полярности приводят к образованию поляризационных электрических полей, которые и являются причиной так называемого *полярного оптического рассеяния*. Его скорость рассеяния может быть определена по формуле [2.16]:

$$\lambda_{\text{по}}^\pm = \frac{e^2 \sqrt{m^*} \omega_o (N_\phi + 1/2 \mp 1/2)}{4\sqrt{2}\pi \epsilon_o \hbar \sqrt{E}} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_{\text{ст}}} \right) \ln \left| \frac{\sqrt{E} + \operatorname{Re} \sqrt{E \pm \hbar \omega_o}}{\sqrt{E} - \operatorname{Re} \sqrt{E \pm \hbar \omega_o}} \right|. \quad (2.14)$$

Наиболее примечательной особенностью полярного оптического рассеяния является зависимость скорости рассеяния от энергии электрона: $\lambda_{\text{по}} \sim E^{-1/2}$, если не принимать во внимание относительно медленную логарифмическую зависимость в (2.14). С увеличением энергии электрона уменьшается вероятность рассеяния, что способствует накоплению энергии и еще меньшему её рассеянию. Такая ситуация наблюдается, если полярное оптическое рассеяние является преобладающим, и приводит к неустойчивому процессу, который называется *полярным*

оптическим убеганием и который приводит к быстрому разогреву электронного газа (разд. 3).

2.2.1.4. Междолинное рассеяние

Результатом междолинного рассеяния является переход электрона из одной долины дисперсионной кривой в другую. Для атомарных полупроводников (германий, кремний) он обычно происходит между эквивалентными долинами, например, для кремния (рис. 2.18, а) из Δ -долины в направлении $[100]$ также в Δ -долины эквивалентных направлений $[010]$, $[001]$ и т.п. В электрическом поле такое рассеяние может привести к перераспределению заселенности долин электронами, которая в равновесном состоянии одинакова. Для бинарных соединений A_3B_5 более характерен переход между неэквивалентными долинами. Например, для GaAs типично рассеяние (рис. 2.18, б) из центральной Γ -долины в боковые L-долины, расположенные в направлениях, эквивалентных $[111]$, или даже в лежащие выше X-долины.

Практически все междолинные переходы сопровождаются изменениями волнового вектора, соизмеримыми с размерами зоны Бриллюэна, и поэтому возможны лишь при участии фононов с большими значениями волновых векторов. Такие фононы называются *междолинными* и соответствуют участкам фононных дисперсионных кривых у границ зоны. Их энергия обычно выбирается средней между крайними значениями энергий акустических и оптических фононов: около 0,05 эВ для Si и 0,03 эВ для Ge и GaAs (см. рис. 2.16). Эти значения соизмеримы с тепловой энергией и рассеяние с поглощением или излучением междолинного фонона будет существенно неупругим.

Кроме того, для переходов между неэквивалентными долинами необходимо учитывать изменение начала отсчета энергии. Если электрон рассеивается из i -долины в j -долину, то закон сохранения энергии (2.5) принимает вид

$$E_j - E_i = \pm \hbar \omega_{ij} - \Delta_{ij}, \quad (2.15)$$

где $\hbar \omega_{ij}$ — энергия междолинного фонона, $\Delta_{ij} = E_{cj} - E_{ci}$ — междолинное расстояние, отсчитываемое (с учетом знака) от дна j -долины до дна i -долины. Значения энергий E_i и E_j отсчитываются от от минимумов соответствующих долин.

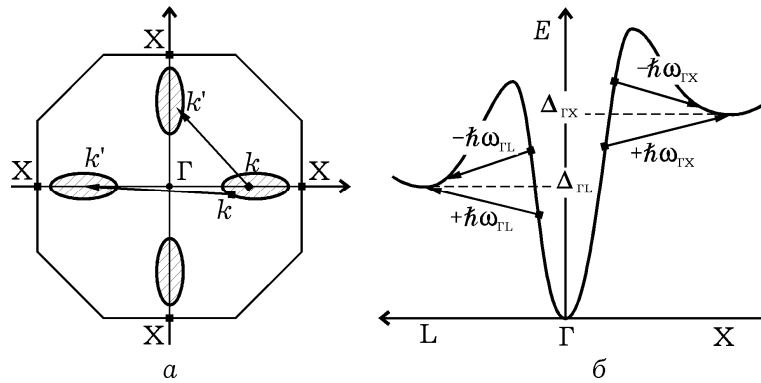


Рис. 2.18. Примеры переходов при междолинном рассеивании между эквивалентными долинами в кремнии (а) и между неэквивалентными долинами в арсениде галлия (б)

Скорости междолинного рассеяния могут быть рассчитаны по формуле, похожей по структуре на (2.13) для неполярного оптического рассеяния [2.16]:

$$\lambda_{\text{мд}}^{\pm} = \frac{Z_j (m_j^*)^{3/2} D_{ij}^2}{\sqrt{2\pi} \rho \hbar^3 \omega_{ij}} (N_{\text{ф}} + 1/2 \mp 1/2) \text{Re} \sqrt{E_i \pm \hbar\omega_{ij} - \Delta_{ij}}, \quad (2.16)$$

где Z_j – число долин, в которые может произойти рассеяние; m_j^* – эффективная масса плотности состояний в j -долине; D_{ij} – константа междолинного взаимодействия, имеющая значения порядка $10^8 \dots 10^9$ эВ/см; $N_{\text{ф}}$ – число фононов с энергией $\hbar\omega_{ij}$, определяемое по (2.3). Оператор Re делает невозможным рассеяние при недостаточной энергии. Если рассеяние происходит между эквивалентными долинами, то вместо Z_j подставляется число, на единицу меньше их количества, и естественно $\Delta_{ij} = 0$. Для долин, лежащих у края зоны Бриллюэна, их число надо уменьшать вдвое.

Количественное представление о скорости различных видов рассеяния можно получить из рис. 2.19, где изображены зависимости скоростей рассеяния от энергии электронов в Γ - и L -долинах арсенида галлия. Следует обратить внимание на “пороговый” характер кривых для оптических и междолинных видов рассеяния, которые требуют перехода электрона в более высокое энергетическое состояние.

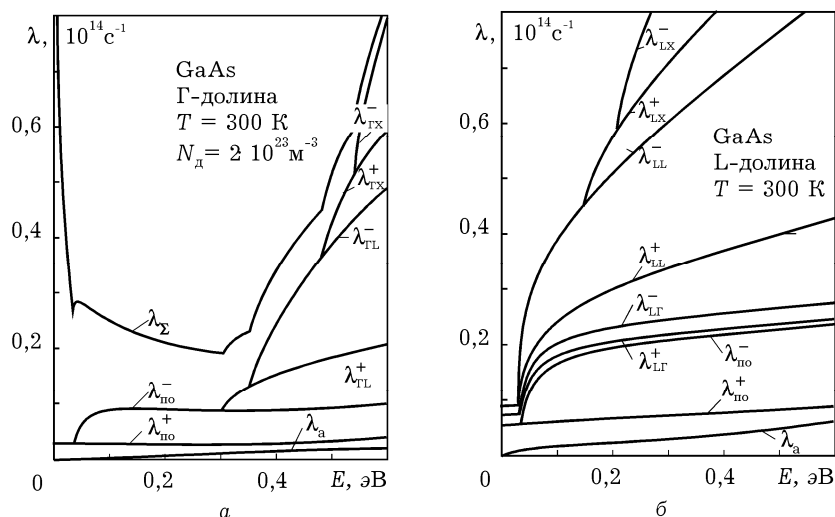


Рис. 2.19. Диаграмма зависимостей скорости рассеивания для разных видов от энергии электрона в Γ -долине (а) и L-долине (б). График для каждого вида суммируется с предыдущими. Верхняя кривая λ_{Σ} суммарной скорости включает также примесное рассеяние.

2.2.2. Примесное рассеяние

В типичных случаях примеси, вводимые в полупроводник, являются *ионизированными* и вызываемое ими рассеяние состоит в отклонении траектории электрона электростатическим полем иона. Подобная задача рассматривалась Резерфордом при исследовании рассеяния α -частиц. Электрон или дырка в кулоновском поле примесного иона движутся по гиперболическим траекториям (рис. 2.20). Так как масса иона примеси на 3..4 порядка превышает массу электрона, то передачи энергии в результате их взаимодействия практически не происходит и такое соударение будет практически *упругим*. Соответственно не изменяется значение скорости электрона, а только ее направление, и $|\vec{k}'| = |\vec{k}|$. Угол рассеяния θ между векторами \vec{k}' и \vec{k} тем больше, чем меньше значение скорости электрона или дырки и так называемое *прицельное расстояние* b (рис. 2.20).

Применительно к носителям в твердом теле модель Резерфорда дает неправильный результат при больших прицельных расстояниях: при малых углах рассеяния интеграл (2.4) расходится, что связано с медленностью убывания кулоновского потенциала. Это можно устранить, если учесть поля, создаваемые другими ионами примеси и носителями заряда. Их поля в той или иной мере экранируют поле иона примеси и “обрезают” действие

кулоновского потенциала, ослабляя его действие прежде всего на больших расстояниях.

В модели Конуэлл-Вайскопфа [2.16] область действия поля иона примеси ограничивают половиной среднего расстояния между ионами (рис. 2.21, а), так что максимальное прицельное расстояние

$$b_{\max} = a/2 = 1/2 \sqrt[3]{N_d}. \quad (2.17)$$

Это выражение получается из условия, что одному атому примеси в среднем соответствует кубический объем с ребром, равным a . Иногда в литературе можно встретить несколько отличный результат, который получается, если считать, что этот элементарный объем представляет сферу радиуса a . В итоге учет экранирования рассеивающего поля полями соседних атомов примеси дает для скорости рассеяния

$$\lambda_d = \frac{v}{b_{\max}} \left(\frac{U_i}{2E} \right)^2 \ln \left(1 + \frac{2E}{U_i} \right), \quad (2.18)$$

где $E = m^* v^2/2$ – кинетическая, а $U_i = e^2/4\pi\epsilon\epsilon_0 b_{\max}$ – потенциальная энергия электрона в поле иона на расстоянии b_{\max} .

В модели Брукса-Херринга дальное действие рассеивающего поля иона примеси ограничивается его экранированием полем носителей заряда [2.16]. При этом распределение потенциала в окрестности донорного иона (рис. 2.21, б) описывается выражением

$$V(r) = (e/4\pi\epsilon\epsilon_0 r) \exp(-r/r_s), \quad (2.19)$$

где $r_s = (\epsilon\epsilon_0 k_B T / e^2 n)^{1/2}$ – так называемый радиус экранирования, а n – концентрация экранирующих зарядов (электронов), и в формуле (2.18) изменяется последний множитель:

$$\lambda_d = \frac{v}{b_{\max}} \left(\frac{U_i}{2E} \right)^2 \left[\ln(1 + \eta) - \frac{\eta}{1 + \eta} \right]. \quad (2.20)$$

Здесь $\eta = (2kr_s)^2$, а при квадратичном законе дисперсии $E = \hbar^2 k^2 / 2m^*$ и $\eta = 8m^* E r_s^2 / \hbar^2$.

Надо отметить, что выражение (2.19) справедливо для максвелловского распределения. Но иногда его используют в случаях, когда средняя энергия хаотического движения электронов заметно превышает величину $k_B T$, подставляя в

выражение для радиуса экранирования электронную температуру T_e .

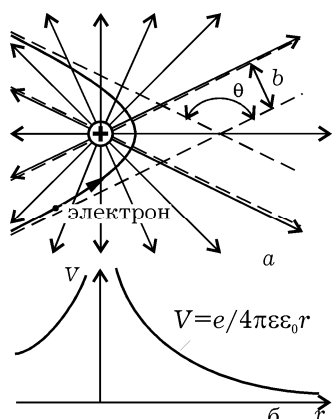


Рис. 2.20. Траектории движения электрона в кулоновском поле иона донорной примеси.

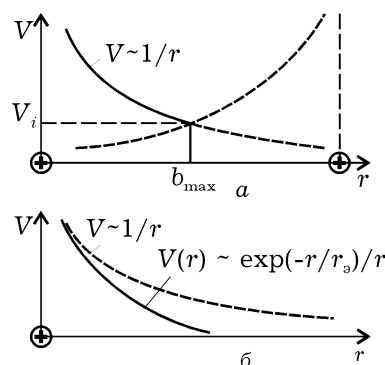


Рис. 2.21. Ограничение потенциальной функции иона примеси при экранировании рассеивающего поля: *a* – соседними ионами примеси (модель Конуэлл-Вайскопфа) и *б* – подвижными носителями заряда (модель Брукса-Херринга).

Обе модели подтверждают известный результат о температурной зависимости подвижности. Если считать, что энергия носителей соответствует тепловой ($E \sim k_B T$, $v \sim T^{1/2}$), и пренебречь медленной логарифмической зависимостью в (2.18), то получим, что скорость примесного рассеяния пропорциональна концентрации примеси $N_{пр}$ и обратно пропорциональна $T^{3/2}$. Модель Конуэлл-Вайскопфа при очень низких температурах дает сильно завышенный результат (формально $\lambda \rightarrow \infty$ при $T \rightarrow 0$), если не учитывать эффекты вымораживания примеси.

Среди других видов рассеяния, которые могут играть заметную роль при определенных условиях, надо назвать *электрон-электронное рассеяние*, которое проявляется при высоких концентрациях примеси. При межэлектронном взаимодействии полный импульс и энергия коллектива сохраняются неизменными, лишь перераспределяясь между частицами, хаотизируя их движение и сглаживая функцию распределения.

В последнее время широко используются трех- и четырехкомпонентные полупроводниковые материалы, например, $Al_x Ga_{1-x} As$, $In_x Ga_{1-x} As_y P_{1-y}$. Их можно рассматривать как сплавы бинарных соединений $A_3 B_5$, в которых даже при идеальной структуре кристаллической решетки расположение разных атомов одной группы имеет случайный характер. В таких

соединениях может оказаться существенным *рассеяние на сплавном потенциале*, вызванном хаотическими отклонениями потенциала кристаллической решетки от идеальной периодической функции. Величина сплавного потенциала зависит от разности энергий электронного сродства бинарных компонентов и пропорциональна произведению их концентраций.

С позиций теории рассеяния могут быть рассмотрены и взаимодействия, вызывающие межзонные переходы: генерацию носителей под действием света, соударений, сильных электрических полей и т.п. и их рекомбинацию с излучением квантов света, фононов. Однако времена релаксации для этих процессов обычно на несколько порядков больше, чем для внутризонных, поэтому генерационно-рекомбинационные явления относят к медленным и рассматривают отдельно

2.3. СРЕДНИЕ ВРЕМЕНА РЕЛАКСАЦИИ

2.3.1. Времена релаксации импульса

Рассмотренные в предыдущих разделах выражения для скоростей различных механизмов рассеяния характеризуют частоты тех или иных видов столкновений как *единичных актов* при заданных параметрах носителей. Времена релаксации концентрации, импульса и энергии определяют скорости протекания переходных процессов в коллективе и по своей сути являются характеристиками, *усредненными* по всей совокупности параметров носителей, то есть по функции их распределения. Из-за сложностей интегрирования получение аналитических соотношений затруднено при использовании распределения Ферми-Дирака, поэтому для усреднения обычно выбирается функция распределения Максвелла-Больцмана, в которой для учета неравновесности коллектива вместо температуры решетки T используется электронная температура T_e . Перенося относительно громоздкие процедуры интегрирования в приложение П2, приведем лишь окончательные результаты для времен релаксации импульса.

Примесное рассеяние, как уже отмечалось, является практически упругим и, не изменяя общей энергии носителей, не участвует в междолинных переходах и релаксации энергии. В то же время, эффективно хаотизируя движение носителей, оно вносит существенный (особенно при низких температурах, когда роль фононов невелика) вклад в релаксацию импульса.

Среднее время релаксации импульса находится из интеграла (1.46), в котором скорость рассеяния $\lambda_{\text{пр}}$ согласно (2.19) и (2.20) выбирается пропорциональной $E^{-3/2}$, а сомножители в квадратных скобках считаются постоянными. На основе соотношения $N(k)dV_k = N(E)dE$ при квадратичном законе дисперсии производится переход от интегрирования по k -пространству к интегрированию по энергии. Выбирая функцию распределения $N(E) = g(E) \cdot f(E)$ максвелловской ($f(E) = \exp[(E_{\text{ф}} - E)/k_{\text{Б}}T_e]$) и пределы интегрирования от $E = 0$ (дно зоны) до ∞ , получим для времени релаксации импульса

$$\frac{1}{\tau_{\text{р д}}} = \sqrt{\frac{8\pi}{m^*}} \left(\frac{e^2}{4\pi\epsilon\epsilon_0} \right)^2 \left[\ln(1 + \eta) - \frac{\eta}{1 + \eta} \right] N_{\text{д}} (k_{\text{Б}}T_e)^{-3/2}. \quad (2.21)$$

Здесь принято также, что импульс в левой части (1.46), а также величина η в (2.19) и (2.19) вычисляются для энергии, равной средней энергии теплового движения, то есть $p = \sqrt{3m^*k_{\text{Б}}T_e}$, и для модели Брукса-Херринга

$$\eta = (2kr_{\text{э}})^2 = 12\epsilon\epsilon_0 m^* (k_{\text{Б}}T_e)^2 / \hbar^2 e^2 N_{\text{д}}. \quad (2.22)$$

При использовании модели Конуэлл-Вайскопфа, не учитывающей экранирование примесного центра носителями, значение η задается выражением

$$\eta = 6\pi\epsilon\epsilon_0 k_{\text{Б}}T_e / e^2 N_{\text{д}}^{1/3}, \quad (2.23)$$

а в (2.21) опускается второе слагаемое в квадратных скобках.

Выражение (2.21) хорошо описывает наблюдающуюся зависимость подвижности от температуры, если примесное рассеяние является преобладающим. Для более точного совпадения с экспериментом можно использовать в (2.21) поправочный коэффициент (см. приложение П2), зависящий как от способа усреднения, так и от формы единичного объема, занимаемого в среднем одним атомом примеси.

Для **акустического рассеяния** характерна линейная зависимость скорости рассеяния от температуры кристаллической решетки T , что связано с уменьшением длины свободного пробега при увеличении количества фононов, и от скорости носителей, пропорциональной $E^{1/2}$. При этом процедура усреднения не представляет сложности (см. приложение П2). Если выбрать

функцию распределения максвелловской: $f(E) = \exp(-E/k_B T_e)$, то обратное время релаксации импульса будет практически совпадать со скоростью акустического рассеяния (2.12) при замене $E = 3k_B T_e/2$:

$$\frac{1}{\tau_{\text{пак}}} = \frac{\sqrt{3} m^{*3/2} D_a^2}{\pi \hbar^4 \rho v_{\text{ЗВ}}^2} k_B T \sqrt{k_B T_e}. \quad (2.24)$$

Другие варианты усреднения, приведенные в приложении, дают результат, отличающийся от последнего выражения коэффициентом пропорциональности, близким к 1.

Оптическое рассеяние является неупругим, при каждом взаимодействии такого рода энергия электрона изменяется на величину $\pm \hbar \omega_0$. Так как эта величина оптического фонона может быть соизмерима с энергией носителей, то скорости рассеяния с поглощением фонона и с его излучением могут заметно различаться. Усреднение скоростей рассеяния (2.13) на *оптическом деформационном потенциале* по функции распределения Максвелла-Больцмана с электронной температурой T_e дает для времен релаксации импульса выражение (см. формулу (П2.18) приложения П2)

$$\frac{1}{\tau_{\text{по}}^{\pm}} = \left(\frac{m^*}{2\pi} \right)^{3/2} \frac{D_0^2}{\rho \hbar^2} \frac{e^{\mp(t-t_e)}}{(k_B T_e)^{1/2} \text{sh } t} K_1(t_e), \quad (2.25)$$

здесь обозначено $t_e = \hbar \omega_0 / 2k_B T_e$, $t = \hbar \omega_0 / 2k_B T$, K_1 — модифицированная функция Бесселя второго рода первого порядка.

Общее время релаксации импульса находится через сумму обратных времен для поглощения и возбуждения:

$$\frac{1}{\tau_{\text{по}}} = \frac{1}{\tau_{\text{по}}^+} + \frac{1}{\tau_{\text{по}}^-} = 2 \left(\frac{m^*}{2\pi} \right)^{3/2} \frac{D_0^2}{\rho \hbar^2} \frac{\text{ch}(t-t_e)}{(k_B T_e)^{1/2} \text{sh } t} K_1(t_e). \quad (2.26)$$

В предположении “теплых” электронов $T_e \approx T$ и $\text{ch}(t-t_e) \approx 1$.

Для **полярного оптического рассеяния** усреднение скорости рассеяния (2.14) по распределению Максвелла-Больцмана с температурой T_e дает подобные результаты (см. приложение П2), отличающиеся коэффициентом пропорциональности и на единицу меньшим порядком функции Бесселя: для обратных времен при поглощении и излучении

$$\frac{1}{\tau_{p\text{ по}}^{\pm}} = \frac{\alpha\omega_0}{\sqrt{2\pi}} \frac{t_e^{1/2}}{\text{sh } t} \exp[\mp(t - t_e)] K_0(t_e), \quad (2.27)$$

и для суммарного времени

$$\frac{1}{\tau_{p\text{ по}}} = \alpha\omega_0 \sqrt{\frac{2}{\pi}} \frac{\text{ch}(t - t_e)}{\text{sh } t} t_e^{1/2} K_0(t_e). \quad (2.28)$$

При этом для “теплых” электронов ($T_e \approx T$) $\text{ch}(t - t_e) \approx 1$ и

$$\frac{1}{\tau_{p\text{ по}}} = \alpha\omega_0 \sqrt{\frac{2}{\pi}} \frac{t_e^{1/2}}{\text{sh } t} K_0(t_e). \quad (2.29)$$

В последних выражениях стоит так называемая *полярная константа*

$$\alpha = \frac{e^2 \sqrt{m^*}}{4\sqrt{2} \pi \varepsilon_0 \hbar (\hbar\omega_0)^{1/2}} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_{\text{ст}}} \right). \quad (2.30)$$

Междолинное рассеяние также является неупругим, поскольку в каждом акте рассеяния поглощается или излучается *междолинный фонон* с энергией $\hbar\omega_{ij}$, близкой к энергии оптических фононов. При рассмотрении междолинного рассеяния в отличие от других механизмов, усреднение кинетического уравнения производится не по всей зоне Бриллюэна, а по характерным состояниям, соответствующим отдельным долинам. Поэтому даже при сохранении общего количества частиц во всей зоне их концентрация может перераспределяться между отдельными долинами. Соответственно инерционность этого процесса определяется концентрационным временем релаксации τ_n , которое находится, как и время релаксации импульса τ_p (см. приложение), усреднением скорости междолинного рассеяния (2.16) по функции распределения. Используя для этого, как и в предыдущих случаях, максвелловскую функцию с электронной температурой T_e , получим для времен релаксации τ_{ij}^{\pm} с поглощением (верхний знак) и испусканием (нижний знак) междолинного фонона $\hbar\omega_{ij}$ при переходе электрона из i -долины в j -долину:

$$\frac{1}{\tau_{pij}^{\pm}} = \left(\frac{m_j^*}{2\pi} \right)^{3/2} \frac{Z_j D_{ij}^2}{\rho \hbar^3 \omega_{ij}} \frac{e^{-\delta_{ij}/2}}{\text{sh } t} |\delta_{ij}| \sqrt{k_B T_e} e^{\mp(t-t_e)} K_1(|\delta_{ij}|/2), \quad (2.31)$$

где $\delta_{ij} = \Delta_{ij}/k_B T_e = (E_{cj} - E_{ci})/k_B T_e$ – нормированное расстояние от дна j -долины до дна i -долины, то есть при переходе электрона из нижней долины в верхнюю $\delta_{ij} > 0$, а величины t и t_e определяются, как и при оптическом рассеянии, но в них надо подставлять энергию *междолинного* фонона. Суммирование обратных времен релаксации с излучением и поглощением фононов дает значения обратных времен релаксации концентрации и импульса:

$$\frac{1}{\tau_{nij}} = \frac{1}{\tau_{pij}} = \left(\frac{m_j^*}{2\pi} \right)^2 \frac{2Z_j D_{ij}^2}{\rho \hbar^3 \omega_{ij}} \frac{|\delta_{ij}| e^{-\frac{\delta_{ij}}{2}}}{\text{sh } t} \text{ch}(t - t_e) \sqrt{k_B T_e} K_1(|\delta_{ij}|/2). \quad (2.32)$$

При использовании последних соотношений надо обратить внимание на то, что, во-первых, в эти соотношения входит эффективная масса плотности состояний, которая обычно учитывает количество долин, и в таком случае величину Z_j надо исключить. Во-вторых, выражения (2.31) и (2.32) получены в предположении, что энергия междолинного фонона намного меньше междолинного расстояния ($\hbar \omega_{ij} \ll \Delta_{ij}$). Это вполне оправдано для большинства практически важных случаев, когда междолинное расстояние заметно превышает величину $k_B T$. Например, для GaAs энергия междолинных фононов около 0,03 эВ, а расстояние между ближайшими долинами $\Delta_{GL} \approx 0,33$ эВ. Более общие выражения приведены в приложении.

Полученные результаты пригодны для рассеяния между неэквивалентными долинами. Для анализа рассеяния *между эквивалентными долинами* ($\Delta_{ij} = 0$) можно использовать выражения для скорости рассеяния (2.13) и последующие соотношения для времен релаксации при неполярном оптическом рассеянии с соответствующей заменой обозначений.

2.3.2. Времена релаксации энергии, упругость рассеяния

Взаимодействие электрона с примесными атомами является практически полностью упругим и не вызывает релаксации

энергии. Поэтому считают, что для примесного рассеяния $1/\tau_{\text{Eпр}} = 0$.

Для определения времен релаксации энергии при *фононных* видах рассеяния можно воспользоваться соотношением (1.42), где во избежание громоздких процедур усреднения по функции распределения удобно использовать выражение для средней скорости потерь энергии одного электрона за счет столкновений в следующей простой форме

$$(dE/dt)_c = \hbar\omega_{\text{ф}} / \tau_{p\text{ф}}^+ - \hbar\omega_{\text{ф}} / \tau_{p\text{ф}}^- \quad (2.33)$$

Смысл этого соотношения в следующем: так как $1/\tau_p$ представляет собой среднее количество столкновений за 1 с и после каждого столкновении энергия электрона изменяется на величину энергии фонона $\hbar\omega_{\text{ф}}$, то первое слагаемое соответствует средней энергии, приобретаемой электроном за 1 с за счет поглощения фононов, а второе – энергии, теряемой при их излучении. Приравнявая (2.33) правой части (1.42), получим выражение:

$$\frac{1}{\tau_E} = \frac{\hbar\omega_{\text{ф}}(1/\tau_{p\text{ф}}^+ - 1/\tau_{p\text{ф}}^-)}{E - E_0}, \quad (2.34)$$

которое можно использовать для определения времен релаксации энергии по найденным ранее временам релаксации импульса при излучении и поглощении фононов.

Однако в большинстве расчетов удобнее пользоваться не временем τ_E , а отношением τ_E/τ_p , которое имеет простой физический смысл. Предварительно отметим, что рассеяние с излучением фононов преобладает над рассеянием с их поглощением, поэтому $1/\tau_{p\text{ф}}^+ > 1/\tau_{p\text{ф}}^-$ и в числителе (2.34) стоит средняя энергия, передаваемая электроном при фононном рассеянии кристаллической решетке за 1 секунду, а в знаменателе избыточная средняя энергия электрона, которая в принципе может быть им отдана в процессе релаксации. Умножение (2.34) на τ_p приводит к соотношению

$$\frac{\tau_p}{\tau_E} = \frac{\hbar\omega_{\text{ф}}}{E - E_0} \frac{1/\tau_{p\text{ф}}^+ - 1/\tau_{p\text{ф}}^-}{1/\tau_{p\text{ф}}^+ + 1/\tau_{p\text{ф}}^-}, \quad (2.35)$$

которое характеризует среднюю часть энергии, теряемой электроном за время τ_p , то есть фактически за одно соударение. По структуре и содержанию оно соответствует выражению (2.9), которое характеризует степень упругости соударений как единичных актов и которое содержит вероятности переходов $w_{kk'}^+$ и $w_{kk'}^-$, отличающиеся только количеством фононов, участвующих в рассеянии с их поглощением и излучением. Это отличие в $w_{kk'}^\pm$ определяется множителем $(N_\phi + 1/2 \pm 1/2) = e^{\pm t}/2\text{sh}t$, где $t = \hbar\omega_\phi / 2k_B T$. В выражении (2.35) вместо $w_{kk'}^\pm$ стоят обратные времена релаксации $1/\tau_{p\phi}^\pm$, которые, как это видно из (2.25) и (2.27), в условиях разогрева электронов для оптических видов рассеяния отличаются еще и знаком в множителе $\exp(\pm t_e)$, включающем электронную температуру через параметр $t_e = \hbar\omega_\phi / 2k_B T_e$. Поэтому второй сомножитель в (2.35), будет в отличие от (2.9), равен $\text{th}(t - t_e)$, а само отношение примет вид

$$\frac{\tau_p}{\tau_E} = \frac{4}{3} \frac{t t_e}{t - t_e} \text{th}(t - t_e). \quad (2.36)$$

При междолинном рассеянии, когда энергия электрона изменяется не только на величину энергии фонона, но и на величину междолинного расстояния $\pm\Delta_{ij}$, аргументом функции Бесселя в (2.31) является параметр $|\delta_{ij}/2 \pm t_e|$ (см. приложение П2) и соответствующее выражение будет более сложным. Однако для наиболее типичных случаев, когда $\hbar\omega_{ij} \ll \Delta_{ij}$, для $1/\tau_{ij}^\pm$ приближенно справедливо (2.31) и отношение τ_p/τ_E , то есть упругость междолинного рассеяния, также можно определять из (2.36).

Это выражение может быть упрощено для двух типичных случаев: для “теплых” электронов, когда $T_e \approx T$, $t_e \approx t$ и $\text{th}x \approx x$:

$$\tau_p/\tau_E \approx 4t t_e / 3, \quad (2.37)$$

и для “горячих” ($T_e \gg T$, $t_e \ll t$)

$$\tau_p/\tau_E \approx 4t_e \text{th}t / 3. \quad (2.38)$$

Важным свойством отношения времен релаксации импульса и энергии является практически линейная зависимость от обратной электронной температуры, что предопределяет специфику

протекания некоторых физических процессов в сильном электрическом поле (разд. 3). Однако это характерно только для *оптических и междолинных* фононов, энергия которых из-за узости спектра принимается постоянной.

При участии в рассеянии акустических фононов наблюдается несколько иная ситуация. Во-первых, из-за их малой энергии и соответственно высокой упругости этого рассеяния иногда пренебрегают акустическим механизмом релаксации энергии. Но при отсутствии оптического рассеяния именно акустический механизм вызывает передачу энергии от электронов кристаллической решетке, обеспечивая термодинамическое равновесие. Во-вторых, упругость акустического рассеяния практически не зависит от электронной температуры, так как с ростом T_e возрастает и энергия фононов, участвующих в рассеянии. Эта особенность связана с тем, что спектр акустических колебаний достаточно широкий и, как отмечалось в п. 2.2.1.2, в рассеянии принимают участие низкочастотные фононы, волновой вектор которых лежит в пределах $|k_\phi| \leq 2|k|$. При квадратичном законе дисперсии электронов $k = \sqrt{2m^*E/\hbar^2}$ и соответственно $|k_\phi| \leq 2\sqrt{2m^*E/\hbar^2}$. При малых частотах фононный закон дисперсии (2.1) можно считать линейным, т. е. $\omega_\phi = v_{зв}|k_\phi|$, поэтому энергия таких фононов $\hbar\omega_\phi \leq v_{зв}(8m^*E)^{1/2}$. Если считать, что в рассеянии участвуют электроны со средней энергией $E = 3k_B T_e/2$, то для акустических фононов $t_e, t \ll 1$ и выражение (2.37) для коэффициента упругости принимает вид

$$\frac{\tau_p}{\tau_E} = \frac{4m^* v_{зв}^2}{k_B T}. \quad (2.39)$$

Таким образом, в отличие от оптических механизмов упругость акустического рассеяния не зависит от электронной температуры. Подстановка типичных значений ($m^* \approx 0,1m$, $v_{зв} \approx 10^3$ м/с, $T = 300$ К) показывает, что отношение τ_E/τ_p , которое приближенно характеризует количество необходимых для релаксации энергии соударений, имеет порядок $10^2 \dots 10^3$.

Релаксация распределения импульса происходит практически за одно соударение, то есть за время порядка τ_p . В этих условиях равновесие между решеткой и электронами устанавливается за счет повышения T_e (разд. 3). Так, в электрическом поле с напряженностью порядка 10^3 В/м при

комнатной температуре $T_e \approx 350$ К, а в более сильных полях это превышение может быть многократным. В то же время для оптического рассеяния при $T_e = 350$ К, $T = 300$ К и $\hbar\omega_0 = 0,03$ эВ время релаксации $\tau_{E_0} = 2,5\tau_{p_0}$, что объясняется его большей неупругостью по сравнению с акустическим рассеянием.

Различие времен релаксации импульса и энергии определяет некоторые **особенности переходных процессов** при включении и выключении электрического поля. Решая уравнение сохранения импульса (1.57) в предположении постоянства времени релаксации можно получить, при включении электрического поля в момент времени t_0^+ (рис. 2.22, а) происходит постепенный переход из равновесного состояния ($v(t_0^+) = 0$) в стационарное состояние $v(\infty) = \mu\mathcal{E}$ по закону

$$v(t) = \mu\mathcal{E} \{1 - \exp[-(t - t_0^+)/\tau_p]\}. \quad (2.40)$$

После выключения поля ($\mathcal{E} = 0$) в момент t_0^- происходит релаксация из стационарного состояния, которая согласно решению уравнения (1.44) описывается функцией

$$v(t) = \mu\mathcal{E} \exp[-(t - t_0^-)/\tau_p]\}. \quad (2.41)$$

Она показывает (рис. 2.22, б), что после выключения поля происходит возврат в равновесное состояние. Инерционность этих переходных процессов определяется временем релаксации импульса τ_p , которое по величине практически совпадает с временем τ , входящим в кинетическое уравнение.

Иная ситуация наблюдается для энергии. Решением уравнения (1.47) при включении поля будет функция

$$E(t) = E_0 + e\tau_E\mu\mathcal{E}^2 \left[1 - \frac{\tau_E}{\tau_E - \tau_p} e^{-(t-t_0^+)/\tau_E} + \frac{\tau_p}{\tau_E - \tau_p} e^{-(t-t_0^+)/\tau_p} \right], \quad (2.42)$$

а при выключении

$$E(t) = E_0 + e\tau_E\mu\mathcal{E}^2 \exp[-(t - t_0^-)/\tau_E]. \quad (2.43)$$

Характерная особенность этих функций (рис. 5.10, в) состоит в том, что длительность переходных процессов определяется в

основном временем τ_E , большим, чем τ_p . Некоторое влияние релаксационного процесса с временной константой τ_p заметно при включении поля и проявляется в более медленном начале нарастания энергии. Это отображается третьим слагаемым в скобках (2.42), вклад которого уменьшается при усилении неравенства $\tau_E > \tau_p$. При выключении поля релаксация энергии до равновесного значения E_0 происходит по экспоненциальному закону (2.43) с характерной временной константой τ_E .

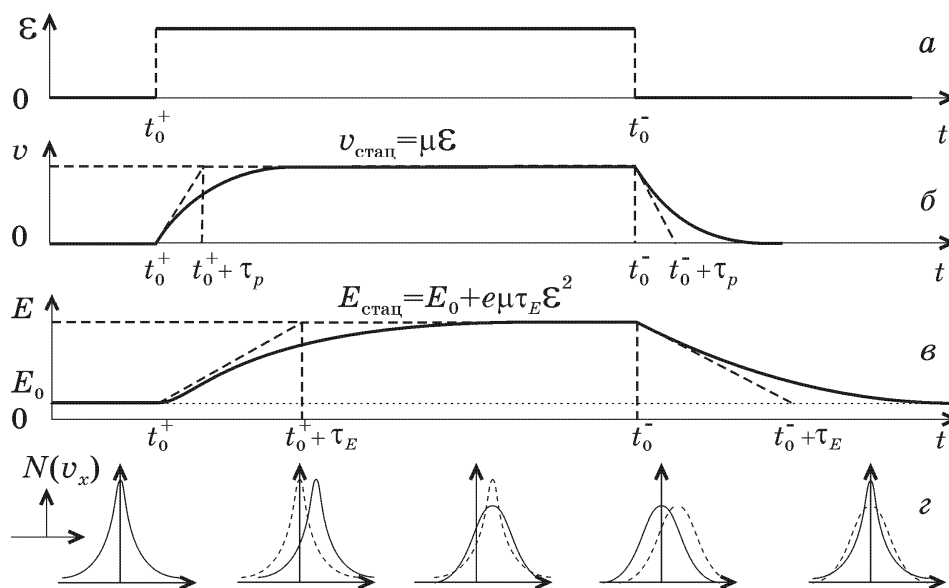


Рис. 2.22. Временные диаграммы переходных процессов, происходящих при включении и выключении электрического поля (а). Установление стационарных состояний при включении и релаксация при выключении для скорости (б) и энергии (в) происходят с существенно разными скоростями. В нижнем ряду (г) символически представлены изображения функции распределения для соответствующих моментов времени.

Результатом таких переходных процессов является эволюция функции распределения $N(v_x)$, качественно изображенная на рис. 2.22, г. При включении поля она за время порядка τ_p из равновесного состояния смещается так, что среднее значение стационарной скорости устанавливается равным скорости дрейфа $v = \mu\epsilon$. Но полностью стационарное состояние наступает лишь за время порядка $\tau_E > \tau_p$, когда происходит расширение функции распределения, означающее увеличение средней энергии электрона. При выключении поля быстрее происходит релаксация дрейфовой скорости до нулевого значения и функция становится симметричной, оставаясь широкой. И лишь через время порядка τ_E она сужается, что говорит о уменьшении энергии до равновесного значения E_0 и возврате в равновесное состояние.

2.4. ПОДВИЖНОСТЬ ЭЛЕКТРОНОВ

Результаты теории рассеяния находят наиболее важное применение, пожалуй, в вопросе о расчете подвижности. Именно особенности различных механизмов рассеяния и детали зонной структуры определяют характерные электрические свойства тех или иных полупроводников и динамику электронных процессов.

Напомним, что подвижность $\mu = e\tau_p/m_c^*$ определяется средним временем релаксации импульса τ_p и *омической* эффективной массой m_c^* . Среднее время релаксации импульса обычно находят из суммы обратных времен релаксации для отдельных механизмов рассеяния, хотя точность этого правила, основанного на принципе независимости различных видов столкновений, иногда подвергается сомнению. Удельный вес того или иного вида рассеяния определяет влияние внешних факторов на подвижность. В данном разделе рассматривается подвижность в слабых электрических полях, когда температура электронного газа T_e практически равна температуре кристаллической решетки T .

Кремний. Из фононных механизмов рассеяния для кремния характерны акустическое и неполярное оптическое. Акустическое рассеяние, строго говоря, является анизотропным, так как существует взаимодействие электронов и с продольными и с поперечными колебаниями кристаллической решетки. Но преобладающим является взаимодействие с продольными колебаниями. Поэтому для расчета времени релаксации импульса при акустическом рассеянии можно с хорошей точностью использовать выражения (2.24) и (2.39), полученные в изотропном приближении, но со значением скорости распространения продольных колебаний $v_{зв} = v_l = 9,0 \cdot 10^3$ м/с и соответствующим деформационным потенциалом $D_a = 6$ эВ. Диэлектрическая проницаемость и плотность вещества соответственно равны $\epsilon = 11,9$ и $\rho = 2,33$ г/см³. Надо также иметь в виду, что в (2.24) необходимо использовать значение эффективной массы плотности состояний, определяемой через компоненты тензора эффективной массы ($m_l^* = 0,98m$, $m_t^* = 0,19m$) с учетом количества эквивалентных долин ($M = 6$), соответствующих дну зоны проводимости кремния: $m_d^* = 6^{2/3}(m_t^* m_l^*)^{1/3} = 1,08m$. В (2.39) для расчета времени релаксации энергии более пригодной будет *омическая* эффективная масса $m_c^* = 3m_t^* m_l^* / (2m_l^* + m_t^*) = 0,26m$.

Оптическое рассеяние в кремнии является неполярным и происходит на оптическом деформационном потенциале. При наличии разных долин такое рассеяние рассматривается фактически как эквивалентное междолинное. Так как нижние долины кремния расположены в зоне Бриллюэна в точках Δ , имеющих значение $k = 0,85(2\pi/a)$, то возможными оказываются два типа переходов (рис. 2.23).

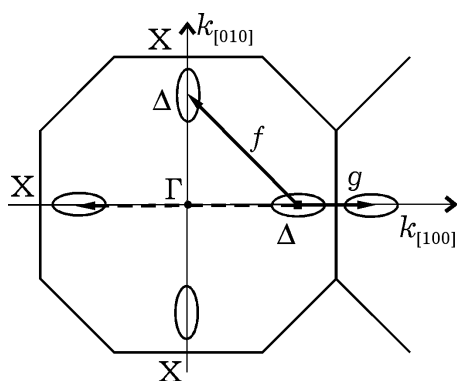


Рис. 2.23. Виды переходов при оптическом рассеянии в кремнии: f – переходы между перпендикулярными долинами, g – переходы между соосными долинами (штриховая линия), которые удобнее рассматривать, как переходы в соседнюю зону Бриллюэна.

Переходы между соосными долинами (штриховая линия) удобнее с учетом периодичности закона дисперсии рассматривать как так называемые g -переходы (сплошная линия) в соседнюю зону Бриллюэна с изменением волнового вектора на $\Delta k = 0,29(2\pi/a)$. Если обратиться к дисперсионной кривой фононов для этого направления (см. рис. 2.16) и учесть, что возможно взаимодействие электрона только с продольными оптическими колебаниями, то видно, что такое изменение волнового вектора электрону могут обеспечить фононы с энергией $\hbar\omega_{01} \approx 0,062$ эВ.

Второй возможный тип оптического рассеяния вызывает так называемые f -переходы (рис. 2.23) между долинами, лежащими на перпендикулярных осях. Из рассмотрения соответствующих фононных дисперсионных кривых (на рис. 2.16 не приведены) следует, что такие переходы могут происходить при участии двух типов фононов с энергиями $\hbar\omega_{02} \approx 0,059$ эВ и $\hbar\omega_{03} \approx 0,047$ эВ. Этим трем видам рассеяния соответствуют оптические деформационные потенциалы: $D_{01} = 7,5 \cdot 10^8$ эВ/см, $D_{02} = 2,0 \cdot 10^8$ эВ/см и $D_{03} = 4,3 \cdot 10^8$ эВ/см [2.16]. Для упрощения расчетов без большой потери точности все эти три вида рассеяния можно рассматривать как один с некоторыми усредненными эквивалентными параметрами, например, $\hbar\omega_0 = 0,05$ эВ и $D_0 = 5,5 \cdot 10^8$ эВ/см. Эффективная масса плотности состояний при этом не включает количество долин $m_{d1}^* = (m_t^* m_l^*)^{1/3}$, но время

релаксации умножается на количество долин, в которые происходит рассеяние, равное 5.

Следует также иметь в виду, что оптические виды рассеяния происходят преимущественно с излучением фонона, поэтому для расчета времени релаксации можно использовать выражение (2.25) для $1/\tau_{po}^-$, а для скорости потерь энергии – второе слагаемое (2.33) для излучения фонона. Для оптического рассеяния понятие времени релаксации энергии, строго говоря, не может быть введено, так как в этом случае скорость потерь энергии электрона не пропорциональна отклонению его энергии от равновесного значения. Но для оценок можно использовать соотношение (2.33).

Для анализа примесного рассеяния пригодно выражение (2.21). Основной проблемой при этом является выбор того или иного приближения. Напомним, что в модели Брукса-Херринга фактором, ограничивающим дальное действие поля иона примеси, принимается экранирование кулоновского поля подвижными носителями заряда. В модели Конуэлл-Вайскопфа таким фактором является наличие соседних примесных атомов. В качестве некоторого критерия для выбора той или иной модели может служить соотношение между средним межпримесным расстоянием и радиусом экранирования. Так, при концентрации примеси 10^{21} м^{-3} и комнатной температуре $T = 300 \text{ К}$, радиус экранирования $r_s = 0,13 \text{ мкм}$, а максимальное прицельное расстояние $b_m = 0,05 \text{ мкм}$. Соотношение $b_m < r_s$ говорит о том, что более сильным фактором является ограничение рассеяния соседними ионами примеси и более предпочтительна модель Конуэлл-Вайскопфа. Приближение Брукса-Херринга обосновано в случае существенного возрастания роли экранирования подвижными носителями, что наблюдается при повышении концентрации примеси и (или) при низких температурах, когда

$$k_B T < e^2 N_d^{1/3} / 4\epsilon\epsilon_0. \quad (2.44)$$

В области низких температур надо также принимать во внимание “вымораживание” примеси.

Результаты расчетов времен релаксации, поясняющие вклад различных механизмов рассеяния в результирующее значение подвижности, представлены на рис. 2.24, а. Преобладающим, кроме области низких температур, является рассеяние импульса на фононах. Влияние оптических фононов проявляется лишь при повышении температуры до значений около 100 К. При низких температурах возрастает роль примесного рассеяния, причем при самых низких температурах – на нейтральных примесных атомах.

Итогом такого анализа могут служить зависимости подвижности от температуры (рис. 2.24, б) при различных концентрациях примеси. В логарифмических координатах температурная зависимость подвижности при самых низких температурах спрямляется (при отсутствии примеси) и ее наклон соответствует степенной зависимости $T^{-3/2}$, характерной для акустического рассеяния. При повышении температуры эта линейность исчезает, что говорит о преобладании оптического рассеяния. На самом крутом участке (при температуре около 200 К) наклон соответствует зависимости $\mu \sim T^{-2,4}$, что хорошо соответствует экспериментальной зависимости, представленной на том же рисунке и взятой из книги [2.14].

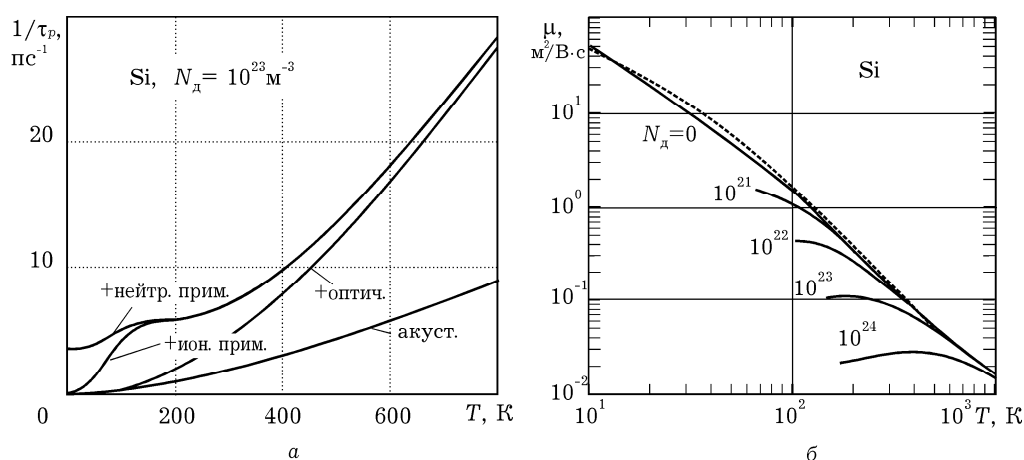


Рис. 2.24. Температурные зависимости обратных времен релаксации импульса при фиксированной концентрации примеси (а) и подвижности при различных концентрациях примеси (б) для кремния. Штриховая линия – экспериментальная. Графики обратных времен релаксации для каждого вида суммируются с предыдущими

Наличие примеси снижает подвижность тем больше, чем выше ее концентрация. Это проявляется сильнее при уменьшении температуры, когда фоновые механизмы рассеяния ослабевают. Заметное отличие от экспериментальных значений (рис. 2.25) наблюдается лишь при больших концентрациях примеси, когда начинается вырождение электронного газа, не учитывающееся в представленных расчетах.

Германий. Для германия характерны те же механизмы рассеяния, что и для кремния. Акустическое рассеяние также является преобладающим в области низких температур (при отсутствии примеси). Для расчета времени акустической релаксации использовались значения акустического деформационного потенциала $D_a = 8$ эВ, скорости звука $v_{зв} = 5,4 \cdot 10^3$ м/с,

плотности $\rho = 5,3 \text{ г/см}^3$, диэлектрической проницаемости $\varepsilon = 16,0$ и эффективной массы плотности состояний $m_d^* = 0,56m$.

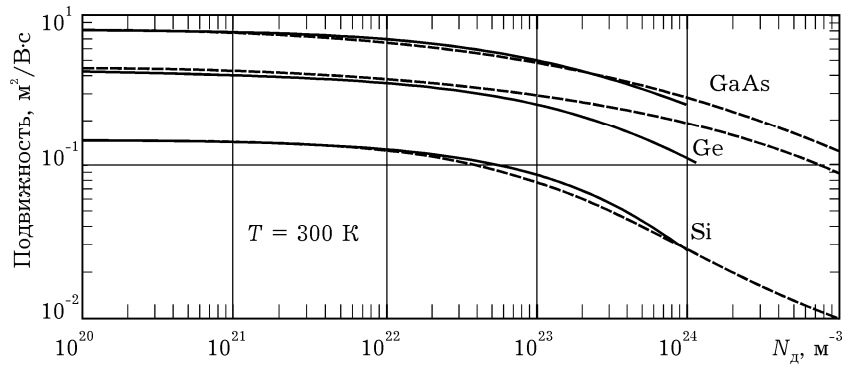


Рис. 2.25. Концентрационные зависимости подвижности для кремния, германия и арсенида галлия при комнатной температуре. Штриховые линии — экспериментальные [2.14]

Оптическое рассеяние происходит на оптическом деформационном потенциале и приводит к переходам электронов в эквивалентные долины. Основное отличие от кремния связано с особенностью закона дисперсии. Как известно, дно зоны проводимости находится в Λ -направлении на краю зоны Бриллюэна в точках L (рис. 2.5). Это соответствует наличию восьми половин или четырех эквивалентных долин с сильной анизотропией эффективной массы $m_t^*/m_l^* = 0,08/1,58$. При этом, как и в кремнии, возможны различные типы переходов электронов между долинами, но в среднем их можно смоделировать как междолинное эквивалентное рассеяние на оптическом деформационном потенциале при $D_0 = 5,5 \cdot 10^8 \text{ эВ/см}$ и участии оптических фононов с энергией $\hbar\omega_0 = 0,034 \text{ эВ}$. Для анализа примесного рассеяния здесь так же, как и в кремнии, предпочтительнее модель Конуэлл-Вайскопфа.

Результаты расчета обратных времен релаксации, показывающие вклад различных механизмов рассеяния, представлены на рис. 2.26, а. При самых низких температурах преобладающим является примесное рассеяние, как на нейтральных, так и на ионизированных атомах примеси. Оно, естественно, зависит от концентрации примеси и ослабевает при уменьшении N_d и увеличении температуры, что наглядно видно на рис. 2.26, б, представляющем температурную зависимость подвижности.

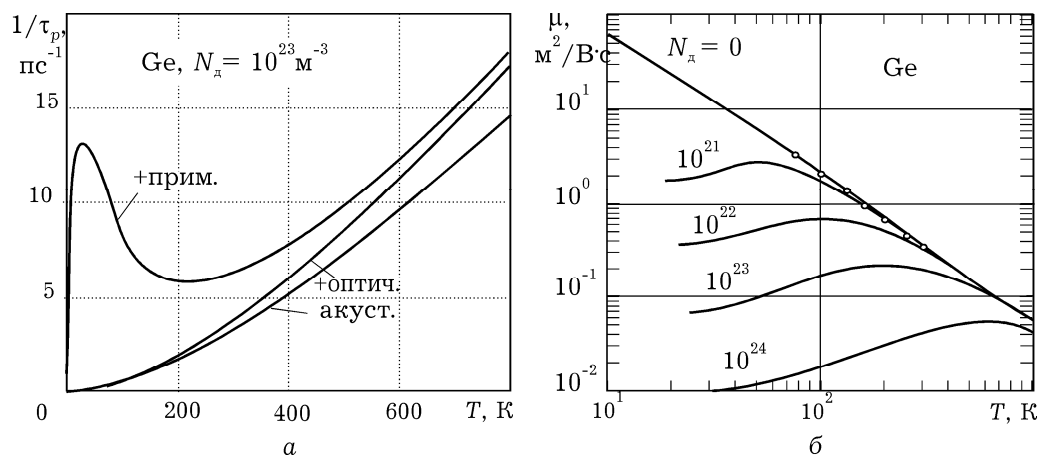


Рис. 2.26. Такие же зависимости, как и на рис. 2.24 для германия. Экспериментальные [2.9] значения для $N_d = 0$ обозначены точками

При возрастании температуры возрастает роль фоновых механизмов рассеяния. Но в сравнении с кремнием здесь меньший вес имеет оптическое рассеяние. Это проявляется в более пологом спаде подвижности с максимальным наклоном кривой $\mu(T)$ в логарифмических координатах, соответствующим степенной зависимости $\sim T^{-1,6}$.

Как и для кремния, но еще заметнее, рассчитанные при $T = 300$ К (рис. 2.25) значения $\mu(N_d)$ расходятся с экспериментальными при концентрациях примеси порядка 10^{24} м^{-3} . Это по всей вероятности связано с меньшей энергией ионизации атомов примеси в германии и соответственно с еще большим концентрационным вырождением электронов в зоне проводимости. Также неоправданно высокой оказалась роль рассеяния на нейтральной примеси, что особенно проявляется при высоких концентрациях примеси.

Арсенид галлия. Одна из главных особенностей рассеяния в GaAs связана с заметным вкладом ионной связи между атомами Ga и As. Это приводит к тому, что из фоновых механизмов основную роль играет полярное оптическое рассеяние с участием оптических фононов с энергией $\hbar\omega_0 = 0,035$ эВ. Результаты расчета скорости этого вида рассеяния (2.14) и соответствующего времени релаксации (2.26) существенно зависят от точности определения близких значений статической и высокочастотной диэлектрических проницаемостей $\epsilon_{ст}$ и ϵ_∞ , которые связаны (2.2) со значениями частот ω_{LO} и ω_{TO} продольной и поперечной мод фононного спектра (рис. 2.16, в) при $k_\phi = 0$. Здесь (рис. 5.15) приведены результаты для значения $\omega_{LO}/\omega_{TO} = (\epsilon_{ст}/\epsilon_\infty)^{1/2} = 1,15$. Этот вид рассеяния проявляется при температурах выше 100 К и

при комнатной температуре является преобладающим (при отсутствии примесного рассеяния).

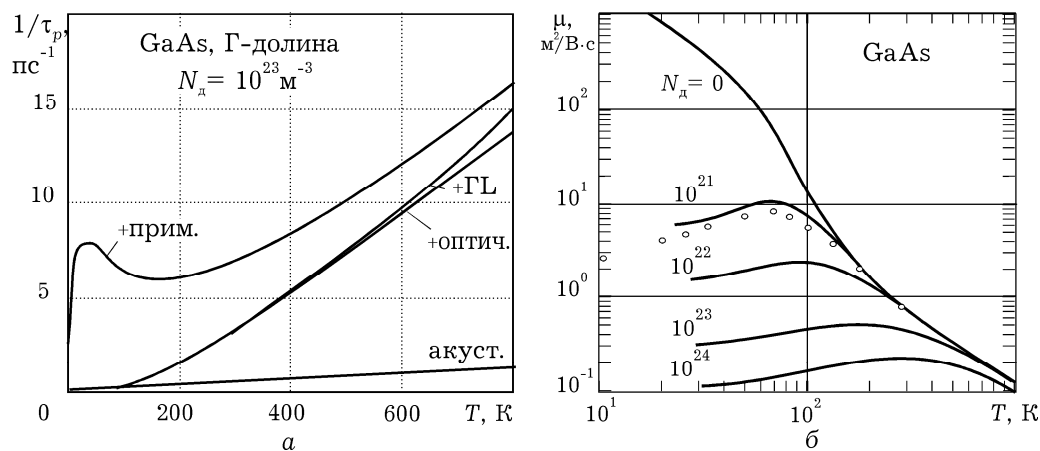


Рис. 2.27. Такие же зависимости, как и на рис. 2.24 для арсенида галлия. Точками обозначены экспериментальные значения для $N_d = 10^{21} \text{ м}^{-3}$

Второе существенное отличие связано с особенностью закона дисперсии GaAs и со структурой его долин [2.9]. Кроме нижней Γ -долины у дна зоны проводимости с изотропной эффективной массой $m_\Gamma^* = 0,063m$ имеются и более высокие: восемь L-долин (как у германия) на краю зоны Бриллюэна ($M_\Gamma = 4$) на расстоянии по оси энергий $\Delta_{\Gamma L} = 0,33$ эВ с сильно анизотропной эффективной массой $m_t^*/m_l^* = 0,075/1,9$, а также на расстоянии $\Delta_{\Gamma X} = 0,52$ эВ шесть X-долин, как у кремния, но не в точках Δ , а на краю зоны Бриллюэна ($M_X = 3$), с эффективными массами $m_t^*/m_l^* = 0,23/1,3$. Присутствие L-долин начинает проявляться при температурах выше ~ 600 К: взаимодействие электронов с оптическими фононами с энергией $\hbar\omega_{ij} = 0,026$ эВ приводит к междолинному рассеянию, вызывающему их переходы из состояний Γ -долины в состояния L-долины. Для расчета времен релаксации этого вида рассеяния можно воспользоваться соотношением (2.32) с константами междолинного взаимодействия $D_{ij} = 10^9$ эВ/см и плотностью вещества $\rho = 5,32$ г/см³.

Акустическое рассеяние в GaAs очень слабое и проявляется лишь при отсутствии конкурирующих видов рассеяния: полярного оптического и примесного. Это заметно на рис. 2.27, б, представляющем зависимость подвижности от температуры, при $N_d = 0$: в области температур ниже ~ 70 К в логарифмических координатах зависимость $\mu(T)$ имеет вид прямой с наклоном $-1,5$, характерным именно для акустического рассеяния. При более высоких температурах до $T \approx 600$ К наблюдается уменьшение

подвижности за счет возрастающей роли рассеяния на полярных оптических фононах. При еще более высоких температурах включаются механизмы рассеяния в L- и X-долины, что приводит к еще более крутому спаду кривой $\mu(T)$.

На фоне этих механизмов примесное рассеяние проявляется тем сильнее, чем выше концентрация примеси, и соответственно возрастает температура, при которой наблюдается максимум подвижности. Поэтому, как и для Si и Ge, зависимость подвижности от концентрации примеси (рис. 2.24) спадает тем быстрее, чем выше N_d . Приведенные результаты расчета хорошо совпадают с экспериментальными и расходятся с ними в области высоких концентраций, при которых начинается вырождение полупроводника.

Фосфид индия. Для расчетов времен релаксации были использованы такие значения параметров: плотность – 4,8 г/см³, скорость звука – $5,2 \cdot 10^5$ см/с, акустический деформационный потенциал, – 6,8 эВ, статическая и высокочастотная диэлектрическая проницаемость – 12,4 и 9,5 соответственно, энергия полярных оптических фононов – 43 меВ, энергия междолинных фононов – 30 мэВ, константа междолинной связи – $1,1 \cdot 10^9$ эВ/см. Значения параметров материала взяты преимущественно из книг [2.14, 2.16]. Механизмы полярного оптического рассеивания, как и в арсениде галлия, являются преобладающими. Как и в GaAs в InP возможны разные виды электронных переходов между долинами при участии оптических фононов с энергией 30 мэВ. Для анализа примесного рассеивания здесь также используется модель Конуелл-Вайскопфа. При самых низких температурах примесное рассеивание преобладает на ионизированных атомах примеси. Наличие примеси снижает подвижность тем больше, чем выше ее концентрация. Это сильнее сказывается при уменьшении температуры, когда фононные механизмы рассеивания ослабевают. Поэтому примесное рассеивание ослабевает с уменьшением примеси и увеличением температуры.

Результаты расчета обратных времен релаксации импульса, которые показывают вклад разных механизмов рассеивания, изображены на рис. 2.28, а.

С ростом температуры растет и роль фононных механизмов рассеивания, однако в сравнении с GaAs оптическое рассеивание в InP имеет больший вклад, о чем свидетельствует более крутой спад подвижности с максимальным наклоном кривой $\mu(T)$ в логарифмических координатах (рис. 2.28, б). Там же приведены результаты измерений подвижности, взятые из [2.9].

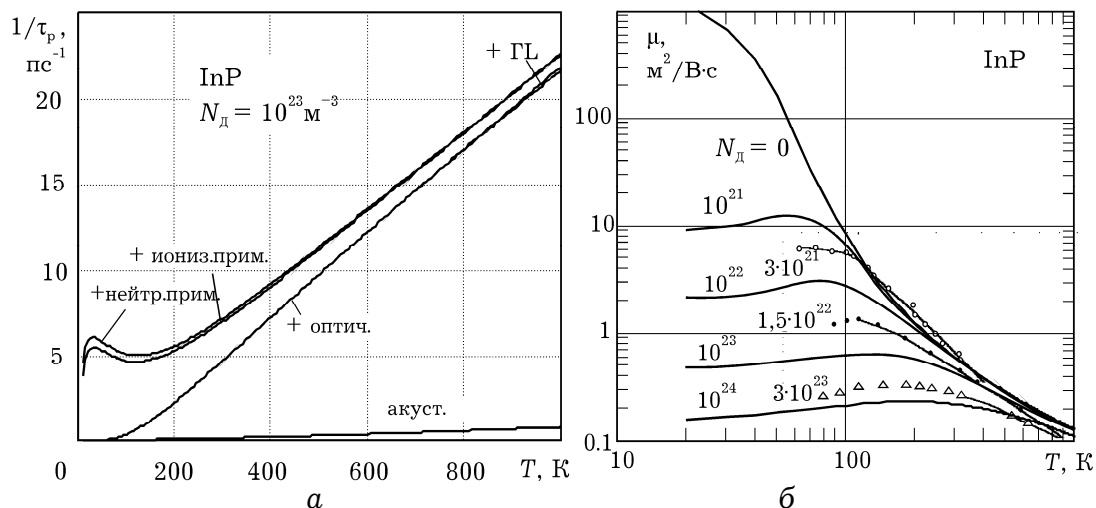


Рис. 2.28. Температурная зависимость обратных времен релаксации импульса при концентрации примеси $N_d = 10^{23} \text{ м}^{-3}$ (а) и подвижности при разных концентрациях примеси (б) для фосфида индия. Графики обратных времен релаксации для каждого вида суммируются с предыдущими

Нитрид галлия. Аналогичная картина наблюдается и для нитрида галлия. Из фононных механизмов для GaN характерны акустическое и полярное оптическое рассеивание. Акустическое рассеивание растет с повышением температуры, но его вклад проявляется только при отсутствии конкурирующих примесного и оптического рассеиваний. Для расчета времени релаксации акустического рассеивания были использованы значения акустического деформационного потенциала $D_a = 10,1$ эВ, скорости звука $v_{зв} = 4,57 \cdot 10^3$ м/с, плотности $\rho = 6,15$ г/см³, диэлектрической проницаемости $\epsilon = 7,5$ и эффективные массы плотности состояний в соответствии с каждой долиной.

В нитриде галлия так же, как и в других соединениях, преобладающим является полярное оптическое рассеивание [2.16], которое проявляется при повышении температуры до ≈ 120 К. При расчетах времен релаксации выбирались значения статической и высокочастотной диэлектрической проницаемости 9,6 и 5,5 соответственно, энергии полярных оптических фононов 65 мэВ.

Вклад междолинных механизмов с энергией междолинных фононов около 59 мэВ и константой междолинной связи $1,1 \cdot 10^9$ эВ/см незначителен даже при высоких температурах. При низких температурах растет роль примесного рассеивания, причем при самых низких температурах примесное рассеивание преобладает как на нейтральных, так и на ионизированных, атомах примеси.

Вклад оптических механизмов рассеивания в сравнении с InP в GaN имеет меньшее значение (но больше в сравнении с

GaAs), которое сказывается в более пологом спаде кривой $\mu(T)$. Итогом такого анализа являются зависимости на рис. 2.29.

Приведенные результаты расчетов близки к известным экспериментальным, но расходятся с ними в области высоких концентраций, при которых начинается вырождение полупроводника.

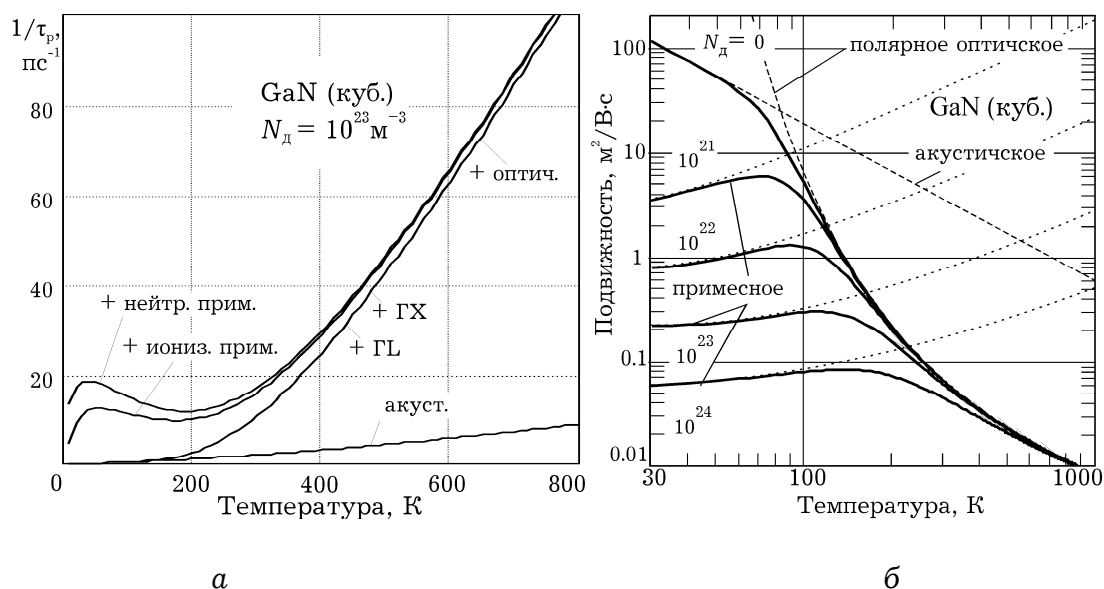


Рис. 2.29. Температурная зависимость обратных времен релаксации импульса при концентрации примеси $N_d = 10^{23} \text{ м}^{-3}$ (а) и подвижности при разных концентрациях примеси (б) для нитрида галлия. Кривые для каждого вида рассеивания суммируются с предыдущими

Значения параметров материалов выбирались в основном из книг [2.15, 2.16]. Необходимо заметить, что в справочной литературе имеется существенная неопределенность в значениях ряда констант. Это относится прежде всего к междолинным расстояниям, акустическому и оптическому деформационным потенциалам, константам междолинной связи, а также к значениям эффективных масс. Поэтому можно ожидать, что при других наборах используемых параметров совпадение результатов расчета с экспериментальными будут более удовлетворительными

Литература к главе 2

2.1. Москалюк В. А., Синекон Ю. С., Кассинг Р. Физика электронных процессов: Учеб. пособие. – Ч. I: Электронные состояния. – К.: УкрИНТЭИ, 2001. – 148 с.

2.2. Braunstein, R., Moore, A.R., Herman, F.: Phys. Rev. 109(1958) 695.

- 2.3. *Ортмонд Б.Ф.* Введение в физическую химию и кристаллохимию полупроводников./Пер. с англ. ./ Под ред. В.М. Глазова- М.: Высшая школа, 1982. – 323с.
- 2.4. *Пасынков В.В., Сорокин В.С.* Материалы электронной техники.- М.: Высшая школа, 1986. – 311с.
- 2.5. *Бонч-Бруевич В. Л., Калашников С. Г.* Физика полупроводников.- М.: Наука, 1990. – 688с.
- 2.6. *Vurgaftman I., Meyer J.R.* Band parameters for III-V compound semiconductors and their alloys// J. Appl. Phys. – 2001. – Vol. 89, №11. P. 5815–5862.
- 2.7. *Vurgaftman I., Meyer J.R.* Band parameters for nitrogen-containing semiconductors// J. Appl. Phys. – 2003. – Vol. 94, №4. P. 3675–3691.
- 2.8. *Chen S.* High-field properties of carrier transport in bulk wurtzite GaN:A Monte Carlo perspective / S. Chen, G. Wanga // J. Appl. Phys. – 2008. – Vol.103. – P.023703-1-6.
- 2.9. Физикотехнический институт им. А. Ф. Иоффе РАН, New Semiconductor Materials. Characteristics and Properties URL: <http://www.ioffe.ru/SVA/NSM/>
- 2.10. *Давыдов С.Ю.* Оценки параметров нитридов элементов третьей группы: BN, AlN, GaN и InN / С.Ю Давыдов // Физика и техника полупроводников. – 2002. – №36. – Вып. 1. – С. 45–47.
- 2.11. *Perssona C.* Effective electronic masses in wurtzite and zinc-blende GaN and AlN / C. Perssona, A Ferreira da Silvab., R. Ahujaa, B. Johanssona // J. Cryst. Grow. – 2001. – №231. – P. 397–406.
- 2.12. *S. Adachi*, "GaAs, AlAs, and AlGaAs: Material Parameters for Use in research and device Applications," J.App.Phys. 58(3), 1 Aug. 1985, R1.
- 2.13. *Зеегер К.* Физика полупроводников: Пер. с англ. – М.: Мир, 1977. – 616 с.
- 2.14. *Зи С.* Физика полупроводниковых приборов: В 2 кн.: Пер. с англ. – М.: Мир, 1984. – Кн. 1. – 456 с.
- 2.15. *Zi, J., X. Wan, G. Wei, K. Zhang, X. Xie*, J. Phys. Condens. Matter 8 (1996), 6323-6328.
- 2.16. *Иващенко В. М., Митин В. В.* Моделирование кинетических явлений в полупроводниках. Метод Монте-Карло. – К.: Наук. думка, 1990. – 192 с.
- 2.17. *Куликов К. В., Москалюк В. А.* Высокочастотные параметры нитрида галлия// Техника и приборы СВЧ, 2008, № 2, с. 48-52.

3. ГОРЯЧИЕ НОСИТЕЛИ В СИЛЬНОМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ

3.1. РАЗОГРЕВ ЭЛЕКТРОННОГО ГАЗА

Для большинства микроэлектронных приборов с характерными размерами порядка микрометров и напряжениях в несколько вольт значения напряженности электрического поля имеет порядок 10^6 В/м. Такие поля обычно считают сильными. Из-за большого разнообразия эффектов, которые они вызывают, наблюдается отклонение от омической зависимости между током и напряжением, или между дрейфовой скоростью и напряженностью поля. Однако разные проявления эффектов сильного электрического поля имеют в своей основе, как правило, одну первопричину – значительное разогревание электронного газа. Превышение температуры электронного коллектива над температурой кристаллической решетки наблюдается и в слабых полях, но оно не вызывает существенного изменения кинетических параметров движения электронов.

Анализ процессов дрейфового движения предполагают независимость от электрического поля времени свободного пробега в “одночастичной” модели или времени релаксации импульса в коллективной модели, что возможно только для слабых электрических полей. В этих условиях скорость дрейфового движения остается пропорциональной напряженности электрического поля, что соответствует омической зависимости между плотностью тока и напряженностью. В сильных электрических полях приближение времени релаксации становится, строго говоря, несправедливым, так как скорость изменения функции распределения перестает быть пропорциональной ее отклонению от равновесной. Но во многих случаях можно, формально сохраняя это приближение, считать время релаксации (т.е. и подвижность) зависящим от средней энергии электрона, а значит и от напряженности электрического поля. Это вызывает нелинейность вольтамперной характеристики однородных образцов материала.

Зависимость подвижности от напряженности электрического поля можно учесть, лишь рассматривая *разогрев* электронного коллектива, то есть анализируя степень влияния поля на среднюю энергию хаотического теплового движения E_T , мерой которой и является электронная температура, определяемая обычно из соотношения

$$E_T = 3k_B T_e / 2. \quad (3.1)$$

Основной физической причиной, объясняющей повышение электронной температуры в сильных полях, является неупругость соударений, что формально отражается в различии времен релаксации импульса и энергии. Электрон, ускоряясь электрическим полем на длине свободного пробега, приобретает импульс $p = m^* v$ и соответственно энергию $E_d = m^* v^2 / 2$ направленного дрейфового движения. Но направленность движения практически полностью теряется уже после первого соударения, эффективно хаотизирующего движение, так как время релаксации импульса одного порядка со средним временем свободного пробега. Но полного рассеяния приобретенной за это время энергии не происходит, так как все виды соударений обладают определенной степенью неупругости. Для этого потребовалось бы время порядка τ_E , обычно большее, чем τ_p . Оставшаяся энергия пополняет запас не направленного, а хаотического теплового движения, что и означает увеличение электронной температуры, определяемой из равенства (3.1).

Рассмотрим вначале механизм разогрева **в однодолинной модели**, пригодной для кремния или германия, при сохранении в зоне проводимости концентрации электронов ($dn/dt = 0$), используя уравнения (1.43) и (1.47) в форме:

$$d\vec{p}/dt - e\vec{\mathcal{E}} = -\vec{p}/\tau_p, \quad (3.2)$$

$$dE/dt - e\vec{v}\vec{\mathcal{E}} = -(E - E_0)/\tau_E. \quad (3.3)$$

Для количественного анализа разогрева электронов необходимо из уравнений баланса полной энергии и энергии дрейфового движения получить уравнение баланса для энергии только хаотического движения. В уравнении баланса *полной* энергии (3.3) слагаемое $e\vec{v}\vec{\mathcal{E}}$ характеризует энергию, приобретаемую от поля за единицу времени, а слагаемое $(E - E_0)/\tau_E$ — теряемую при столкновениях за это же время. Для получения уравнения баланса энергии *дрейфового* движения домножим все слагаемые уравнения (3.2) скалярно на \vec{v} и получим

$$dE_d/dt - e\vec{v}\vec{\mathcal{E}} = -2E_d/\tau_p. \quad (3.4)$$

Здесь правая часть представляет собой энергию дрейфа, теряемую за 1с, а $\tau_p/2$ – время ее релаксации. Вычитая последнее уравнение из (3.3) получим уравнение баланса энергии только хаотического *теплового* движения, которое и определяет динамику изменения электронной температуры:

$$\frac{dE_T}{dt} = -\frac{E_T - E_0}{\tau_E} - \frac{E_d}{\tau_E} + \frac{2E_d}{\tau_p} = -\frac{E_T - E_0}{\tau_E} + \frac{E_d}{\tau_E} \left(\frac{2\tau_E}{\tau_p} - 1 \right). \quad (3.5)$$

Слагаемое $2E_d/\tau_p$ равно энергии дрейфа, превращающейся за 1 с непосредственно в тепловую, но ее часть, равная E_d/τ_E , за это же время релаксирует, переходя к кристаллической решетке. В итоге увеличение тепловой энергии за счет дрейфовой за одну секунду составляет $E_d(2\tau_E/\tau_p - 1)/\tau_E$. Отношение $2\tau_E/\tau_p \equiv \gamma$ является параметром, характеризующим эффективность перехода энергии дрейфа в тепловую и определяется степенью *неупругости* соударений по отношению к направленной энергии. Поэтому разогрев электронного газа возможен только при условии, что скорость рассеяния импульса превышает скорость рассеяния энергии, точнее, при $\tau_E > \tau_p/2$. И только за счет превышения электронной температуры над равновесной ($E_0 = 3k_B T/2$), равной температуре кристаллической решетки, возможно *стационарное* состояние, при котором отдаваемая и получаемая электронами энергии уравниваются друг друга. Это наступает, когда электронная температура устанавливается на уровне, соответствующем равенству

$$E_T = E_0 + E_d(\gamma - 1). \quad (3.6)$$

Принимая во внимание, что в стационарном состоянии из (3.2) скорость дрейфа равна $v = e\tau_p\mathcal{E}/m^* = \mu\mathcal{E}$, где $\mu = e\tau_p/m^*$ – подвижность, получим для электронной температуры

$$T_e = T + \mu^2\mathcal{E}^2 m^*(\gamma - 1)/3k_B. \quad (3.7)$$

Здесь надо иметь в виду, что времена релаксации для различных механизмов рассеяния зависят от электронной

температуры весьма сложно (см. разд. 2). В литературе, например, [3.2] имеется зависимость $T_e(\mathcal{E})$, полученная в предположении преобладания только одного вида рассеяния (обычно акустического) и не учитывающая слагаемого E_d/τ_E в (3.5). Но удельный вес различных механизмов существенно изменяется в процессе разогрева, в сильных электрических полях преобладают механизмы с участием оптических фононов. Поэтому в общем случае получение зависимости $T_e(\mathcal{E})$ в явном виде практически невозможно и приходится ее находить как обратную к $\mathcal{E}(T_e)$:

$$\mathcal{E} = \sqrt{\frac{3k_B(T_e - T)}{m^* \mu^2 (\gamma - 1)}} = \frac{v_T}{\mu} \sqrt{\frac{T_e/T - 1}{\gamma - 1}}, \quad (3.8)$$

где $v_T = \sqrt{3k_B T / m^*}$ – равновесная тепловая скорость „холодных” электронов, соответствующая температуре решетки.

Кремний. На рис. 3.1 представлены зависимости обратных времен релаксации импульса и энергии, а также коэффициента неупругости рассеяния от электронной температуры, рассчитанные для кремния при температуре $T = 300$ К по формулам, приведенным в разд. 2.3, при тех же числовых значениях параметров, которые использовались при расчете слабополевой подвижности (разд. 2.4). Здесь следует обратить внимание на то, что преобладающим механизмом является оптическое рассеяние, акустическое рассеяние из-за большой упругости практически не влияет на время релаксации энергии, а примесное заметно только при очень слабом разогреве ($T \approx T_e$).

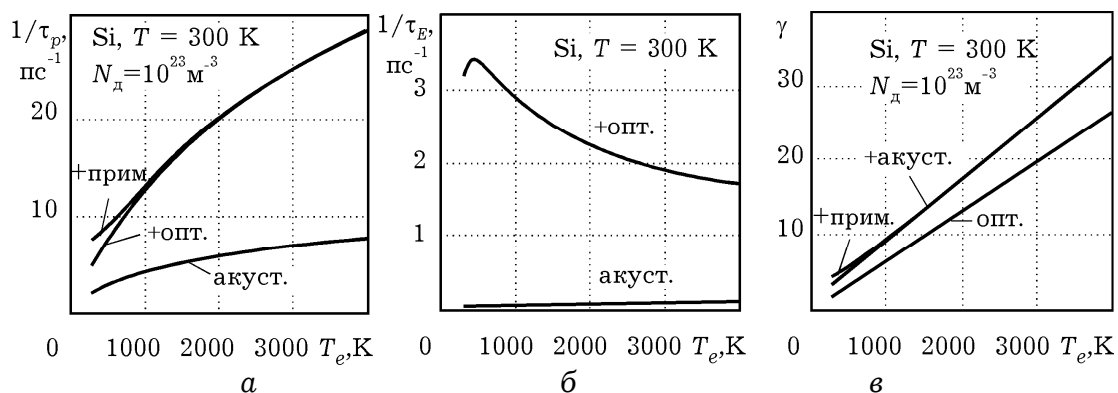


Рис. 3.1. Зависимости обратных времен релаксации импульса (а), энергии (б) и коэффициента неупругости (в) от электронной температуры в кремнии. Кривые для каждого вида рассеяния суммируются с предыдущими.

Примечательно также, что коэффициент неупругости γ с повышением электронной температуры возрастает практически линейно. Это характерно для оптических механизмов рассеяния в соответствии с формулами (2.38), где следует принять $E = 3k_B T_e / 2$.

Приведенные результаты позволяют рассчитать важные характеристики вещества (рис. 3.2) в виде зависимостей от напряженности электрического поля: электронную температуру (а), как обратную функцию к (3.8), времена релаксации импульса (б) и энергии (в).

Надо обратить внимание на то, что электронная температура и время релаксации энергии практически не зависят от концентрации примеси, а время релаксации импульса, то есть и подвижность, эту зависимость проявляют в относительно слабых полях до 10...20 кВ/см. При этом τ_p уменьшается с ростом напряженности электрического поля, причем в сильном поле $1/\tau_p \sim \mathcal{E}$, а τ_E — возрастает, так что неравенство $\tau_E > \tau_p$ усиливается, что означает увеличение степени неупругости рассеяния.

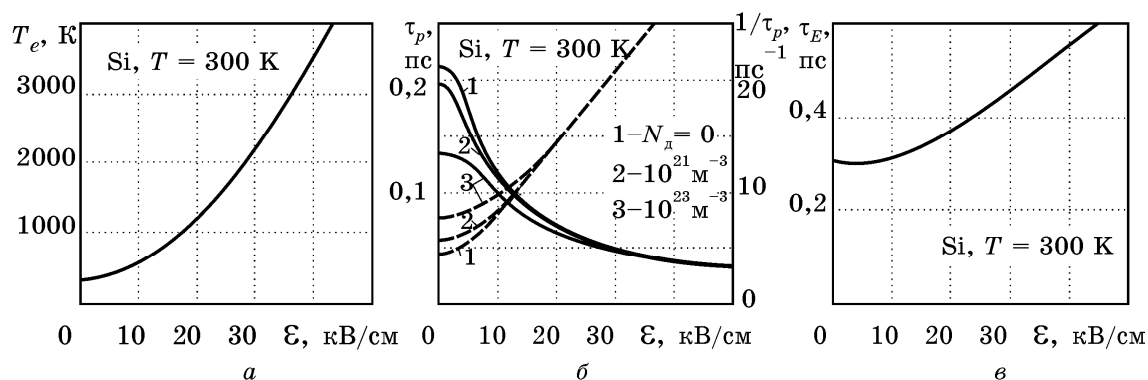


Рис. 3.2. Зависимость электронной температуры (а), времен релаксации импульса (б) и энергии (в) от напряженности электрического поля для кремния. Штриховые линии (б) характеризуют влияние концентрации примеси на обратное время релаксации импульса.

Германий. Подобные результаты расчетов для германия представлены на рис. 3.3. Основные отличия от кремния связаны прежде всего с преобладанием в германии при высоких температурах акустического вида рассеяния, что сказывается на величине времени релаксации импульса (рис. 3.3, а) и особенно заметно увеличивает время релаксации энергии (рис. 3.3, б, в).

Это объясняется прежде всего большей упругостью акустического рассеяния, что требует увеличения количества соударений для рассеяния энергии, приобретенной в электрическом поле. Характерно также, что коэффициент неупругости γ (рис. 3.3, в) нелинейно возрастает с повышением

электронной температуры, что на первый взгляд противоречит соотношению (2.38). Но по мере повышения T_e увеличивается и энергия акустических фононов (разд. 2.3.2), которые могут участвовать в рассеянии. Поэтому γ возрастает не так быстро, как это наблюдается при оптическом рассеянии, в котором участвуют фононы с почти одинаковой энергией, но сами значения γ почти на порядок выше, чем в кремнии. Кроме того большая упругость акустических соударений, а также более высокая подвижность вызывают большую степень разогрева: в тех же полях электронная температура в германии на порядок выше.

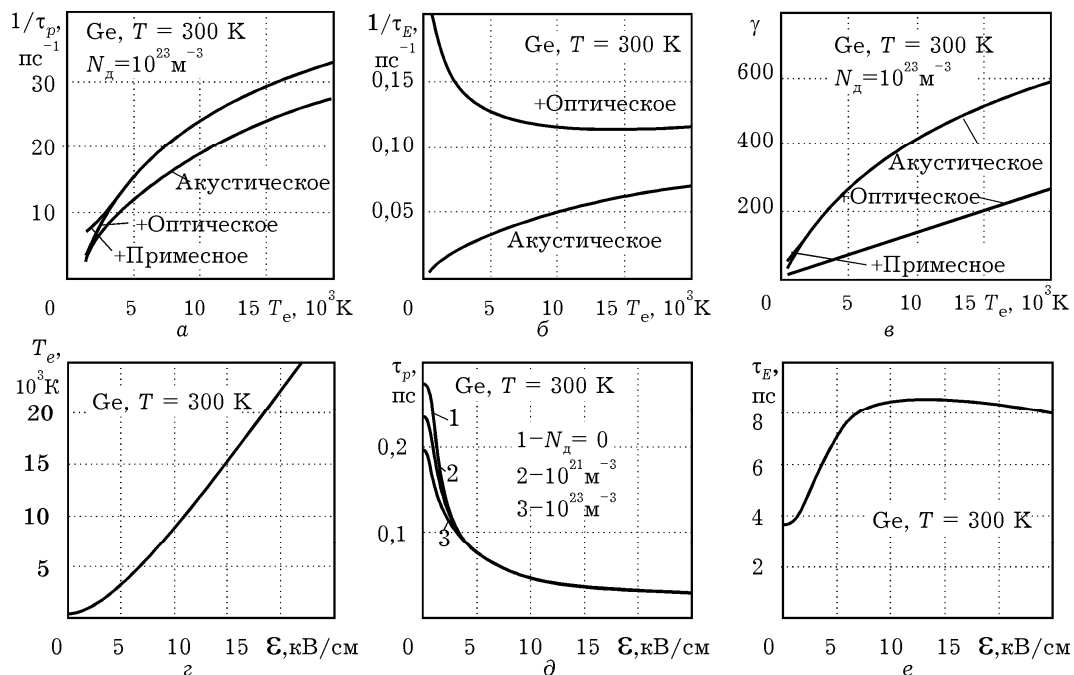


Рис. 3.3. Те же зависимости, что и на рис. 3.1 и 3.2, для германия. Следует обратить внимание на существенно больший вес акустического рассеяния.

В многодолинных полупроводниках типа GaAs, кроме Γ -долины у дна зоны проводимости, имеются и более высокие L- и X-долины. В каждой из них механизм разогрева в принципе тот же самый, но ситуация усложняется тем, что, во-первых, электроны имеют в разных долинах различные параметры и степени их разогрева будут отличаться. Во-вторых, по мере возрастания электронной температуры изменяется и концентрация электронов в долинах. Считая, что полная концентрация электронов $n(T)$ в зоне проводимости остается постоянной и определяется температурой решетки, а их перераспределение по долинам в результате междолинного рассеяния — электронной температурой, относительную

заселенность долин можно характеризовать коэффициентами ($i, j = \Gamma, L, X$):

$$b_i(T_e) = n_i/n = n_i/(n_i + n_j) = 1/(1 + n_j/n_i) = 1/(1 + \tau_{nji}/\tau_{nij}), \quad (3.9)$$

которые определяются через концентрационные времена релаксации в соответствии с (2.32)

$$\tau_{nij}/\tau_{nji} = n_i/n_j = (m_{di}^*/m_{dj}^*)^{3/2} \exp(-\Delta_{ij}/k_B T_e).$$

Выражение (3.9) нетрудно обобщить и на случай трех долин.

В качестве исходных в многодолинной модели необходимо использовать уравнения баланса импульса и энергии электронного коллектива в форме (3.2) и (3.3), но отдельно для каждой долины. Соответственно уравнение баланса только тепловой энергии получается суммированием уравнений типа (3.5) с весовыми коэффициентами, равными относительной заселенности долин b_i . Вводя понятие средней тепловой энергии электронов $E_T = b_\Gamma E_{T\Gamma} + b_L E_{TL}$ (для сокращения записи рассмотрим только двухдолинную ΓL -модель), получим для ее значения в стационарном состоянии

$$E_T = E_0 + b_\Gamma E_{д\Gamma} (\gamma_\Gamma - 1) + b_L E_{дL} (\gamma_L - 1). \quad (3.10)$$

Отсюда и устанавливается связь между средней для долин электронной температурой и напряженностью электрического поля:

$$T_e = T + \mathcal{E}^2 [b_\Gamma \mu_\Gamma^2 m_\Gamma^* (\gamma_\Gamma - 1) + b_L \mu_L^2 m_L^* (\gamma_L - 1)] / 3k_B. \quad (3.11)$$

Так же, как и в однодолинной модели, зависимость электронной температуры от напряженности электрического поля находится как обратная к $\mathcal{E}(T_e)$:

$$\mathcal{E} = \sqrt{\frac{3k_B(T_e - T)}{b_\Gamma \mu_\Gamma^2 m_\Gamma^* (\gamma_\Gamma - 1) + b_L \mu_L^2 m_L^* (\gamma_L - 1)}}, \quad (3.12)$$

поскольку параметры b_i, μ_i, γ_i являются функциями электронной температуры.

Для **арсенида галлия** результаты расчетов обратных времен релаксации импульса и энергии в Γ - и L -долинах в зависимости от электронной температуры при $T = 300$ К представлены на рис.

3.4 в виде диаграмм, где кривая для каждого вида рассеяния суммируется с предыдущей.

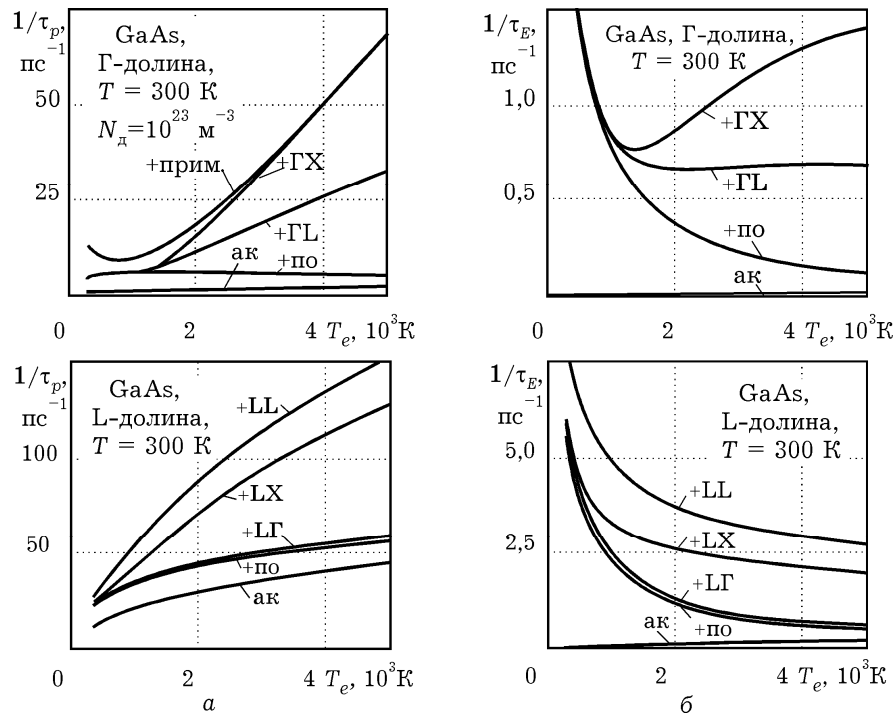


Рис. 3.4. Зависимости обратных времен релаксации импульса (а) и энергии (б) от электронной температуры в Γ - (верхний ряд) и L-долинах (нижний ряд). Кривые для каждого вида рассеяния суммируются с предыдущими.

Сравнив эти результаты с соответствующими скоростями рассеяния на рис. 2.19, следует обратить внимание, что примесное рассеяние, являясь упругим, не участвует в рассеянии энергии и заметно только для Γ -долины, причем его роль ослабевает при повышении электронной температуры, а акустическое влияет практически только на время релаксации импульса. Междолинное рассеяние становится заметным в Γ -долине лишь при температурах, выше ~ 1000 К, а в L-долине можно пренебречь рассеянием в Γ -долину из-за низкой плотности состояний в последней.

Знание времен релаксации в отдельных долинах и расчет заселенностей долин (3.9) с использованием (2.36) позволяет рассчитать зависимость (3.12) между электронной температурой и напряженностью электрического поля, а по ней определить поле-температурную характеристику $T_e(\mathcal{E})$. На этой зависимости (рис. 3.5, а), можно выделить три характерных участка: область слабых полей с напряженностью < 4 кВ/см, где разогрев определяется практически только параметрами Γ -долины; переходную область (до $20 - 40$ кВ/см), в которой существенно изменяется

заселенность долин, и область очень сильных полей, для которой характерно заполнение электронами преимущественно верхних долин (рис. 3.5, б) и медленное, практически линейное, возрастание электронной температуры.

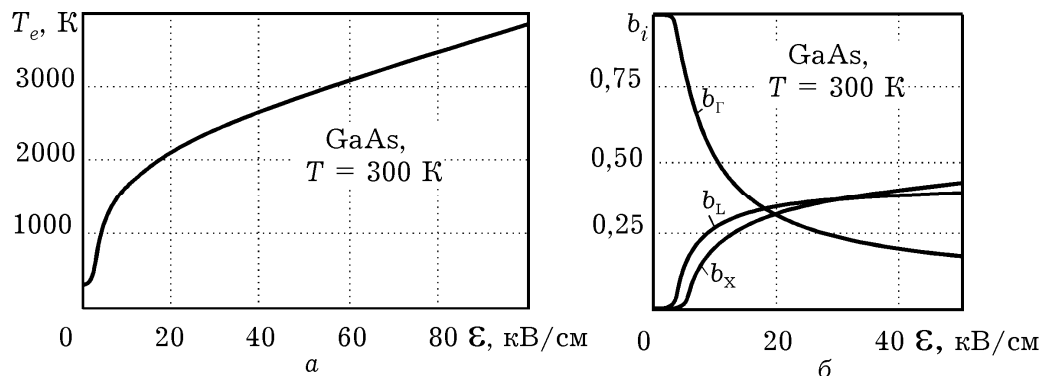


Рис. 3.5. Зависимости электронной температуры (а) и заселенности долин (б) от напряженности электрического поля в арсениде галлия для комнатной температуры.

Фосфид индия. Зависимости времен релаксации для фосфида индия рассчитаны при температуре $T = 300 \text{ K}$ для тех же значений параметров, которые были использованы для расчета слабопольовой подвижности. В InP оптическое рассеивание преобладает. Как и в GaAs, механизмы акустического рассеивания из-за большой упругости практически не влияют на время релаксации импульса и энергии в Γ -долинах, но существенны для L-долин. Междолинное рассеивание становится заметным лишь при температурах выше $\sim 1200 \text{ K}$. Так же как и для арсенида галлия рассеиванием из L-долины в Γ -долину можно пренебречь из-за низкой плотности состояний в ней. При этом достаточно весомым является вклад эквивалентного междолинного LL-рассеивания. Примесное рассеивание на ионизированных атомах заметно только в Γ -долине, причем оно имеет значение при низких температурах и с увеличением электронной температуры его вклад уменьшается. Результаты расчетов приведены на рис. 3.6.

Для поле-температурной характеристики фосфида индия характерна аналогичная картина (рис. 3.5). Здесь также можно выделить три участка. Область слабых полей с напряженностью менее 10 кВ/см , в которой, как и в арсениде галлия, разогрев электронного газа, определяется исключительно параметрами Γ -долин. Переходная область, где происходит значительное перераспределение электронов по долинам, занимает участок приблизительно до 20 кВ/см . Область очень сильных полей ($\sim 20 \text{ кВ/см}$) в основном характерна заполнением электронами преимущественно L-долин и сравнительно меньшим вкладом

электронов X-долин. В отличие от GaAs в InP происходит более резкий рост электронной температуры и значение $T_e = 3500$ К достигается уже при напряженности ~ 60 кВ/см (рис. 3.7).

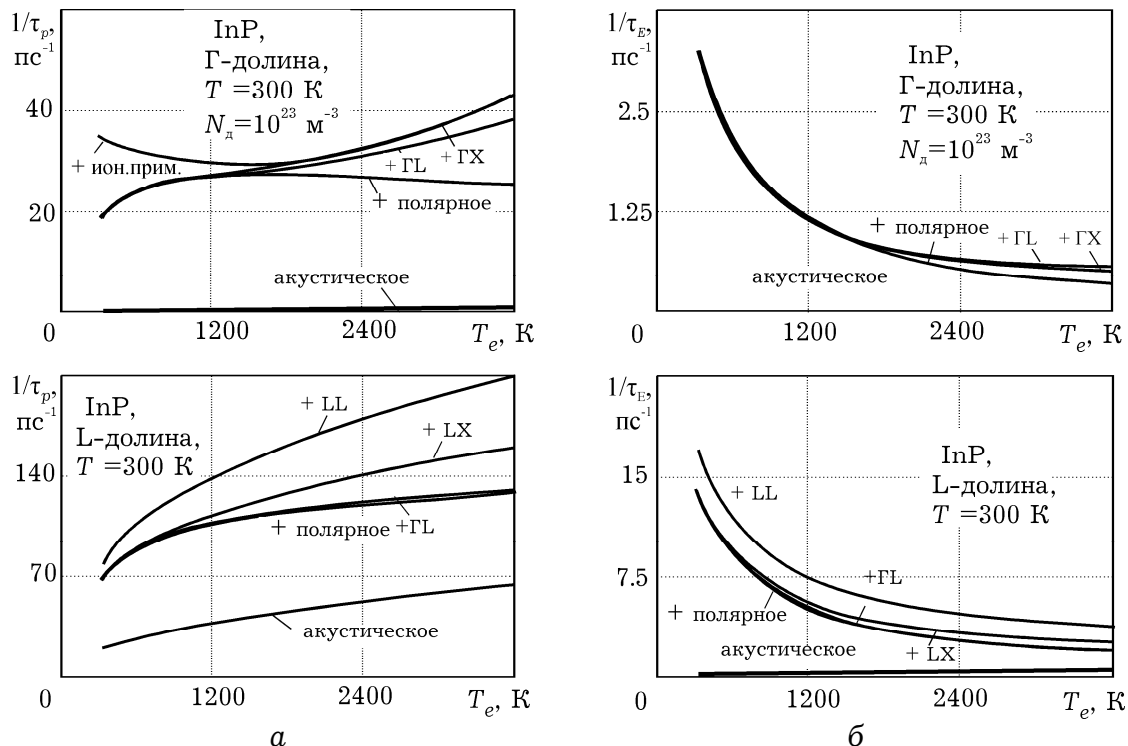


Рис. 3.6. Зависимость обратных времен релаксации импульса (а) и энергии (б) от электронной температуры в Γ - (верхний ряд) и L- (нижний ряд) долинах для фосфида индия. Кривые для каждого вида рассеивания суммируются с предыдущими

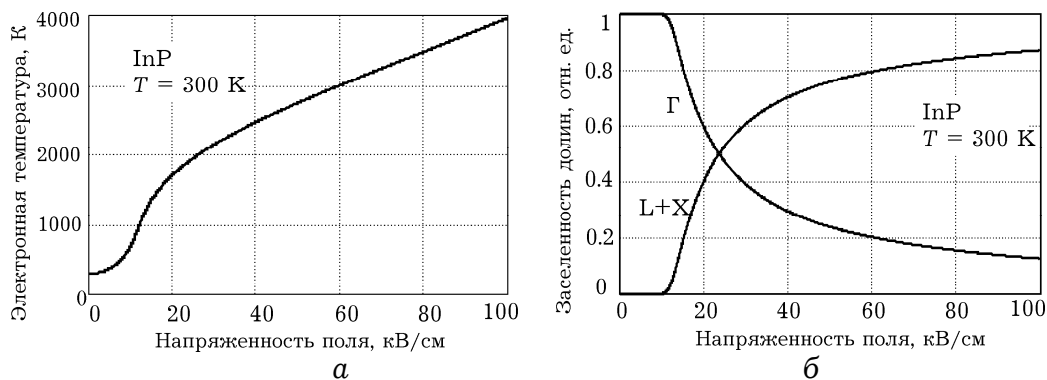


Рис. 3.7. Поле-температурная зависимость (а) и зависимость заселенности долин от напряженности электрического поля (б) в InP для комнатной температуры

Нитрид галлия. Аналогичная ситуация наблюдается и для кубической структуры нитрида галлия: преобладающими являются механизмы оптического рассеивания; акустическое, из-за упругости, практически не влияет на времена релаксации и является значимым лишь для X-долин. Междолинное рассеивание

достаточно весомо проявляется в Γ -долине при повышенные температуры до ~ 2500 К, что видно из диаграмм, представленных на рис. 3.8:

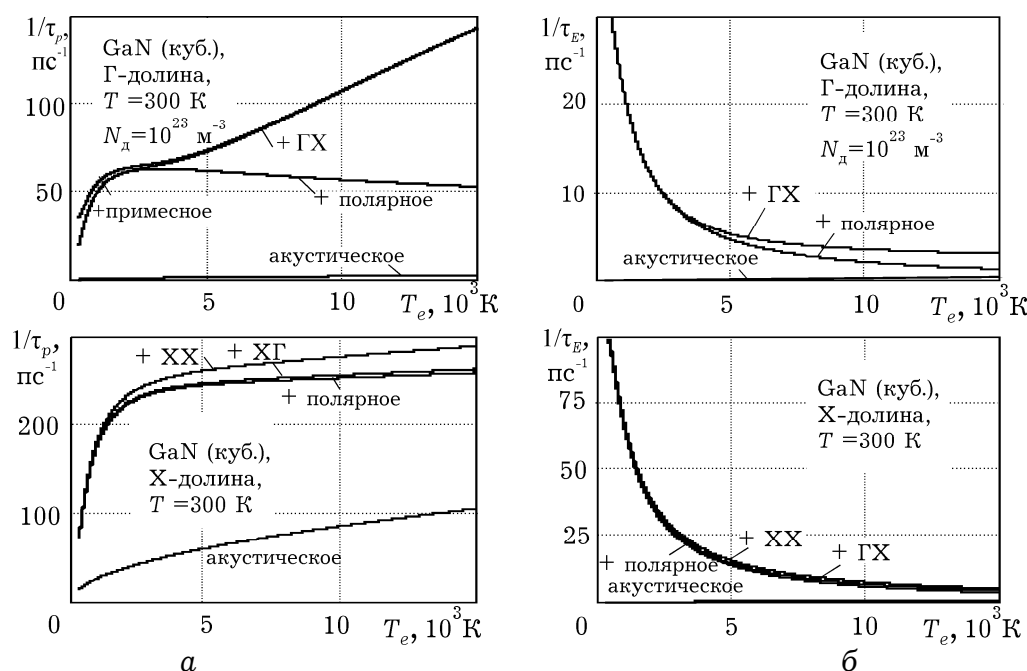


Рис. 3.8. Зависимость обратных времен релаксации импульса (а) и энергии (б) от электронной температуры в Γ - (верхний ряд) и X- (нижний ряд) долинах для нитрида галлия. Кривые для каждого вида рассеивания суммируются с предыдущими

Подобная, но несколько иная картина наблюдается для поле-температурной характеристики нитрида галлия, в котором заметный рост электронной температуры проявляется лишь со значений 120 – 125 кВ/см. А в области сильных полей (начиная с 200 кВ/см) происходит достаточно быстрый рост электронной температуры и достигаются значения порядка $\sim 10^4$ К (рис. 3.9).

Полученные результаты позволяют представить функциями напряженности поля, а не электронной температуры, и другие важные характеристики, усредненные по долинам (рис. 3.10):

– подвижность

$$\mu = b_{\Gamma}\mu_{\Gamma} + b_{L}\mu_{L}, \quad (3.13)$$

– омическую эффективную массу (рис.3.10, а)

$$1/m^* = b_{\Gamma}/m_{\Gamma}^* + b_{L}/m_{L}^*, \quad (3.14)$$

– время релаксации импульса (рис.3.10, б)

$$\tau_p = \mu m^*/e, \quad (3.15)$$

– время релаксации энергии (рис.3.10, в)

$$\tau_E = \tau_p/2 + 3k_B(T_e - T)/2e\mu\varepsilon^2. \quad (3.16)$$

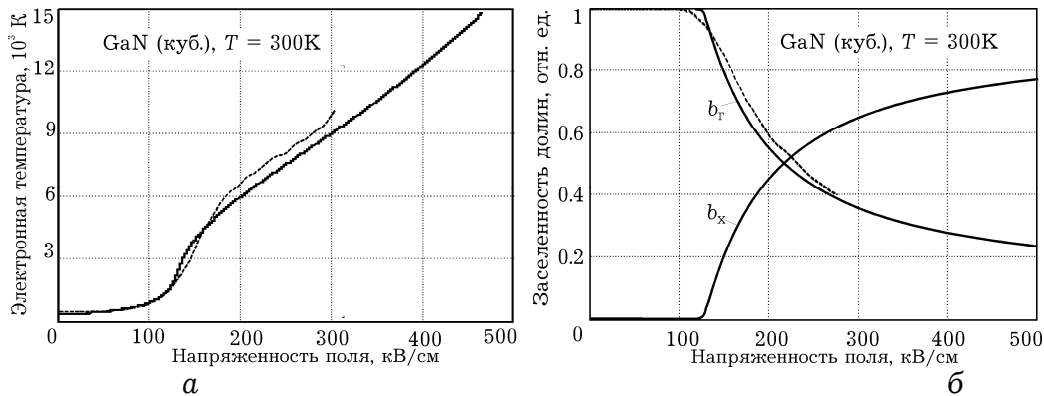


Рис. 3.9. Поле-температурная зависимость (а) и зависимость заселенности долин от напряженности электрического поля (б) в GaN для комнатной температуры. Штриховые линии взяты из [3.3, 3.4]

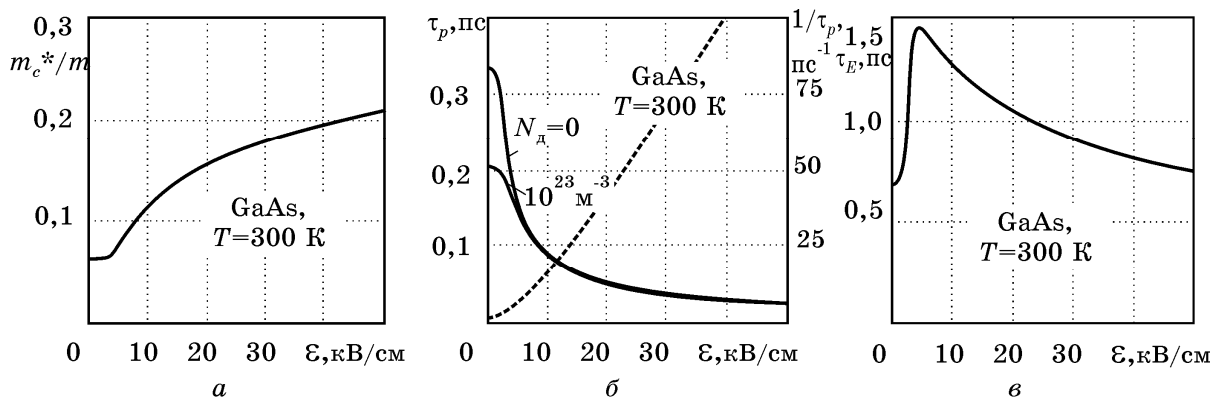


Рис. 3.10. Зависимости усредненных по долинам приведенной эффективной массы (а), времен релаксации импульса (б) и энергии (в) от напряженности электрического поля для арсенида галлия. Штриховая линия (б) – для обратного времени релаксации импульса при $N_d = 0$.

Последнее выражение получается из (3.6) с учетом выражений для коэффициента неупругости γ и энергии дрейфового движения E_d . Полученный таким образом набор усредненных по долинам параметров в виде функций от \mathcal{E} позволяет использовать уравнения однодолинной модели для многодолинных полупроводников в приближении заданного поля.

3.2. НАСЫЩЕНИЕ ДРЕЙФОВОЙ СКОРОСТИ

Разогрев электронного газа в сильных электрических полях сопровождается так называемым насыщением дрейфовой скорости, то есть возникновением ситуации, когда направленная

скорость становится практически не зависящей от напряженности поля, а на вольтамперной характеристике появляется горизонтальный участок.

Простейшие представления о причинах такого явления можно составить на примере одночастичной модели дрейфа [3.1]. Среднее время свободного пробега обратно пропорционально средней скорости, которая имеет как хаотическую (тепловую), так и направленную (дрейфовую) составляющую. В слабых электрических полях $v_{др} \ll v_{хаот}$ и средняя скорость определяется только хаотической составляющей. Поэтому подвижность остается постоянной, а вольтамперная характеристика – линейной. Но по мере того, как с ростом \mathcal{E} дрейфовая скорость становится соизмеримой с хаотической, средняя скорость на свободном пробеге возрастает как за счет дрейфовой составляющей, так и за счет хаотической, увеличивающейся при разогреве электронного газа. В результате сокращается время свободного пробега, уменьшается подвижность, и вольтамперная характеристика становится более пологой, приобретая тенденцию к насыщению. Отсюда ясен и возможный критерий, определяющий термин “сильное электрическое поле”, – соизмеримость дрейфовой и хаотической составляющих скорости электронов.

Более точное описание этого явления возможно с использованием коллективной модели дрейфа. Как видно на примере кремния (рис. 3.2, б), обратное время релаксации импульса возрастает с увеличением напряженности поля, причем практически линейно. Такое наблюдается при $\mathcal{E} > 10...20$ кВ/см, когда самым существенным становится рассеяние с участием оптических фононов, причем с преобладающим их возбуждением. В этих условиях дрейфовая скорость электронов не может возрастать выше некоторого значения v_n , так как чем больше напряженность поля, тем пропорционально выше скорость рассеяния. Потому как только ускоренный полем электрон достигнет определенной скорости (в среднем!), равной скорости насыщения v_n , с большой долей вероятности происходит рассеяние с возбуждением оптического фонона и уменьшением энергии электрона на величину $\hbar\omega_0$. Таким образом рост напряженности электрического поля приводит к возрастанию не дрейфовой, а хаотической скорости, т.е. электронной температуры.

Некоторым аргументом в пользу такого представления является тот факт, что у германия (см. далее), в котором преобладает акустическое рассеяние, переход к режиму насыщения заметно

“затянут” в область высоких электронных температур, а в режиме насыщения имеется заметный подъем скорости.

Для выяснения факторов, определяющих значение скорости насыщения, иногда используют следующую простую модель. Как следует из уравнения баланса импульса (3.2), рассеиваемый за время τ_p импульс $m^* v_n$ равен приобретенному от поля $-e\mathcal{E}\tau_p$. Соответственно уравнению баланса энергии (3.1) для дрейфового движения приобретаемая от поля за это же время энергия дрейфа $e v_n \mathcal{E} \tau_p$ в стационарном состоянии равна энергии излучаемого фонона $\hbar\omega_0$. Исключая из этих равенств напряженность электрического поля, получим

$$v_n = (\hbar\omega_0/m^*)^{1/2}. \quad (3.17)$$

Это соотношение, обычно используемое в учебной литературе [3.5], следует воспринимать лишь как оценочное. Для получения расчетных соотношений необходимо учитывать особенности механизмов рассеяния в конкретных материалах.

Кремний. Из соотношения (3.8) непосредственно после умножения на подвижность следует выражение для дрейфовой скорости

$$v = v_T \sqrt{\frac{T_e/T - 1}{\gamma - 1}}. \quad (3.18)$$

Это выражение позволяет, рассчитывая времена релаксации, с учетом ранее полученной связи между электронной температурой и напряженностью электрического поля получить так называемую поле-скоростную характеристику $v(\mathcal{E})$.

Для кремния она представлена на рис. 3.11, где в слабых полях скорость пропорциональна напряженности, а в сильных ($\mathcal{E} > 20$ кВ/см) наблюдается очень пологий участок, который и соответствует насыщению. В этом режиме практически не наблюдается зависимости скорости от концентрации примеси, что говорит о малой роли примесного рассеяния и соответственно о преобладании фононных механизмов в сильных электрических полях.

Для более детального анализа влияния различных факторов на скорость насыщения и получения соответствующих аналитических соотношений необходимо сделать некоторые допущения. Во-первых, доля акустического рассеяния даже в сильных полях

остаётся заметной по сравнению с оптическим. К тому же соотношение между ними остаётся практически постоянным, так как при больших электронных температурах ($k_B T_e \gg \hbar\omega_0$) параметр $t_e = \hbar\omega_0/2k_B T_e$ мал и $K_1(t_e) \approx 1/t_e$. В этих условиях для обратного времени релаксации импульса при оптическом рассеянии (2.26) можно использовать более простое выражение

$$\frac{1}{\tau_{po}} \approx \frac{m^{*3/2} D_o^2}{\pi^{3/2} \rho \hbar^2} \cdot \frac{\sqrt{2k_B T_e}}{\hbar\omega_0 \text{sh } t} \quad (3.19)$$

Так как время релаксации импульса для акустического рассеяния (2.24) также пропорционально $\sqrt{T_e}$, то их отношение, равное

$$\frac{\tau_{po}}{\tau_{pa}} = \sqrt{\frac{3\pi}{2}} \frac{D_a^2 \hbar\omega_0 \text{sh } t}{D_o^2 (\hbar v_{зв})^2} k_B T, \quad (3.20)$$

не зависит от электронной температуры. В частности, для кремния при комнатной температуре оно равно $\approx 0,38$. Это позволяет считать, что общее время релаксации импульса пропорционально времени оптической релаксации и его можно представить в виде

$$\tau_p = \tau_{po} / (1 + \tau_{po} / \tau_{pa}). \quad (3.21)$$

Во-вторых, коэффициент неупругости соударений $\gamma = 2\tau_E / \tau_p$ зависит от T_e практически линейно (рис. 3.1, в). Если пренебречь релаксацией энергии за счёт почти упругого акустического рассеяния и считать, что τ_E определяется только оптическим рассеянием ($\tau_E = \tau_{Eo}$), то с учетом (3.21) получим

$$\gamma = \gamma_o (1 + \tau_{po} / \tau_{pa}), \quad (3.22)$$

где γ_o — коэффициент неупругости оптического рассеяния, который определяется в соответствии с (2.40):

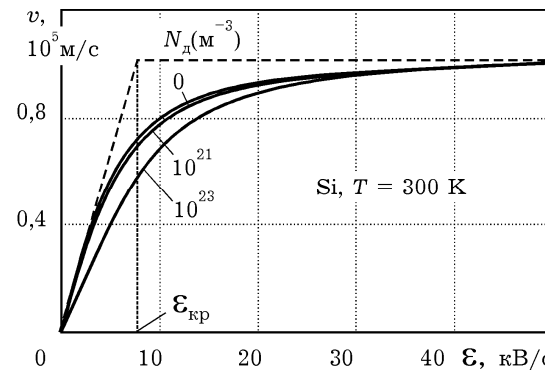


Рис. 3.11. Зависимость дрейфовой скорости от напряженности электрического поля в кремнии при различных концентрациях примеси и ее кусочно-линейная аппроксимация (штриховая линия).

$$\gamma_o = \frac{2\tau_{Eo}}{\tau_{po}} = 6 \frac{T_e}{T} \left(\frac{k_B T}{\hbar \omega_o} \right)^2 = \frac{3}{2tt_e} = \frac{3}{2t^2} \frac{T_e}{T}. \quad (3.23)$$

В результате полный коэффициент неупругости удобно представить в виде

$$\gamma = \frac{3(1 + \tau_{po}/\tau_{pa})}{2t^2} \frac{T_e}{T} \equiv \beta \frac{T_e}{T}, \quad (3.24)$$

откуда наглядно видно, что он пропорционален электронной температуре.

Коэффициент пропорциональности β , имеющий смысл коэффициента неупругости фононного рассеяния при $T = T_e$, во-первых, не зависит от T_e , а только от температуры решетки T , и, во-вторых, имеет обычно значения, большие 1. Например, для кремния $\beta = 2,6$ при $T = 300$ К. Это позволяет при больших электронных температурах пренебречь единицей в знаменателе и числителе (3.18) и получить для скорости насыщения простое выражение

$$v_H = \frac{v_T}{\sqrt{\beta}} = \frac{v_T}{\sqrt{6(1 + \tau_{po}/\tau_{pa})}} \frac{\hbar \omega_o}{k_B T}. \quad (3.25)$$

Значение v_T пропорционально $T^{1/2}$, поэтому, если отношение τ_{po}/τ_{pa} мало, можно приближенно считать скорость насыщения пропорциональной $T^{-1/2}$. Характерно также, что значение v_H не зависит от концентрации примеси, так как в области насыщения преобладают оптические механизмы рассеяния.

В итоге принятые допущения позволяют для поле-скоростной характеристики вместо (3.18) применять более простое выражение:

$$v = v_H \sqrt{\frac{T_e/T - 1}{T_e/T - 1/\beta}}. \quad (3.26)$$

При известных значениях v_H и слабополевой подвижности μ_0 иногда используют кусочно-линейную аппроксимацию (см. рис. 3.11) поле-скоростной характеристики $v(\mathcal{E})$

$$v = \mu_0 \mathcal{E} \text{ при } \mathcal{E} < \mathcal{E}_{кр}, \quad v = v_H \text{ при } \mathcal{E} > \mathcal{E}_{кр}, \quad (3.27)$$

где $\mathcal{E}_{\text{кр}} = v_{\text{н}}/\mu_0$ – так называемая *критическая напряженность* поля, соответствующая условной границе раздела между сильными и слабыми полями. Такая аппроксимация удобна, поскольку используемые здесь параметры $v_{\text{н}}$ и μ_0 относительно просто могут быть идентифицированы.

Еще более удобной и достаточно точной может оказаться явная зависимость $v(\mathcal{E}) = \mu(\mathcal{E}) \cdot \mathcal{E}$, где зависимость подвижности от поля задается приближенным выражением

$$\mu(\mathcal{E}) = \mu_0/[1 + (\mathcal{E}/\mathcal{E}_{\text{кр}})^2]^{1/2}. \quad (3.28)$$

Эта зависимость пригодна только для фононных механизмов рассеяния. Влияние примеси заметно в слабом поле и может быть приближенно учтено, если в числителе (3.28) вместо μ_0 использовать подвижность $\mu(N_{\text{д}})$ для данной концентрации примеси, но $\mathcal{E}_{\text{кр}}$ рассчитывать через μ_0 .

Германий. Преобладание акустического рассеяния вызывает, как уже отмечалось в предыдущем разделе, нелинейность зависимости $\gamma(T_e)$. Но тем не менее общий коэффициент неупругости удобно представить, как и для кремния, через коэффициент неупругости только для оптического рассеяния. При этом в отличие от (3.22) нельзя пренебречь вкладом акустического рассеяния в релаксацию энергии, и эта связь будет иметь вид

$$\gamma = \gamma_0 \frac{1 + \tau_{\text{po}} / \tau_{\text{pa}}}{1 + \tau_{\text{Eo}} / \tau_{\text{Ea}}}. \quad (3.29)$$

Коэффициент неупругости оптического рассеяния γ_0 , пропорционален T_e и при характерных для германия параметрах равен $\gamma_0 \approx 3,7 T_e/T$. Отношение времен релаксации импульса в числителе (3.29), как и в (3.20), не зависит от T_e и имеет значение $\approx 2,7$ при $T = 300$ К. Но отношение времен релаксации энергии при высоких электронных температурах зависит от T_e практически линейно $\tau_{\text{Eo}}/\tau_{\text{Ea}} \approx 0,014 T_e/T$. Поэтому полный коэффициент неупругости можно представить так же как и (3.24) в форме $\gamma = \beta T_e/T$, но коэффициент β уменьшается с ростом T_e по закону $\beta \approx \beta_0/(1 + 0,014 T_e/T)$, где $\beta_0 \approx 10,2$.

В итоге можно считать, что в соответствии с (3.18) при $T_e/T \gg 1$ наступает насыщение, но в режиме насыщения наблюдается сравнительно медленное возрастание скорости

дрейфа. Хорошим подтверждением подобного представления могут служить результаты, приведенные на рис. 3.12, где изображены поле-скоростная характеристика (а) и зависимость подвижности от напряженности электрического поля (б), полученные с учетом поле-температурной характеристики $T_e(\mathcal{E})$ (рис. 3.3, з).

На рисунках также приведены результаты экспериментов, взятые из книги [3.6], где описаны данные, полученные из различных первоисточников.

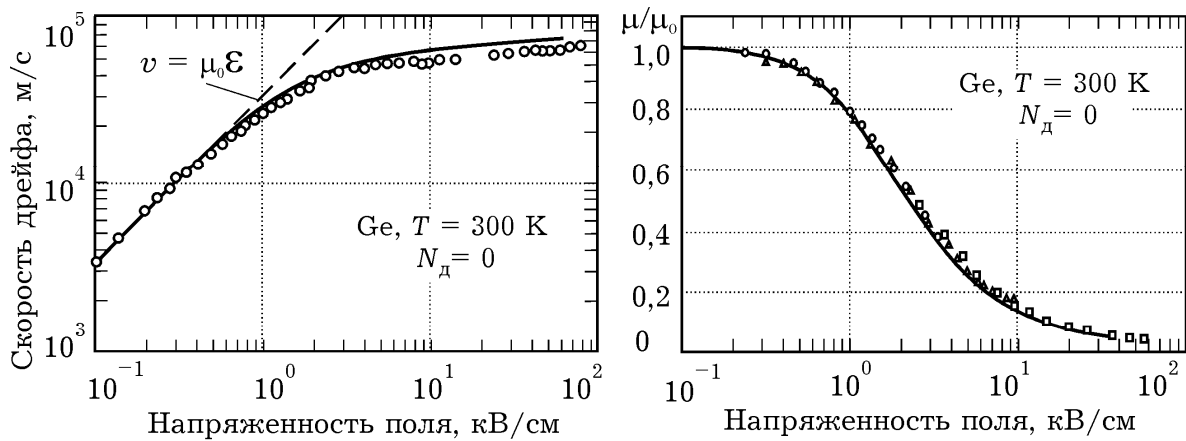


Рис. 3.12. Зависимости дрейфовой скорости (а) и подвижности (б), нормированной на ее слабополевое значение, от напряженности электрического поля в германии при комнатной температуре. Точками обозначены экспериментальные значения различных авторов

Арсенид галлия. Поле-скоростная характеристика GaAs имеет принципиально иной вид, что вызвано прежде всего многодолинной структурой закона дисперсии, а также особенностями механизмов рассеяния в разных долинах.

В центральной Γ -долине при относительно небольшом разогреве ($T_e \approx 1000$ К) из фоновых механизмов преобладает полярное оптическое рассеяние (рис. 3.4, а). Акустическое сказывается только на рассеянии импульса, его доля определяется отношением $\tau_{рпо}/\tau_{ра}$, которое в широкой области электронных температур не превышает 0,1. Оценка скорости насыщения, которую должны обеспечить полярные оптические фононы с энергией $\hbar\omega_0 = 0,034$ эВ, по формуле, аналогичной (3.25), дает значение $v_n' \approx 2,2 \cdot 10^5$ м/с. Но из-за особенностей полярного оптического рассеяния значение $1/\tau_{рпо}$ падает с ростом T_e и скорость уменьшается, что частично проявляется на верхней кривой рис. 3.13, а.

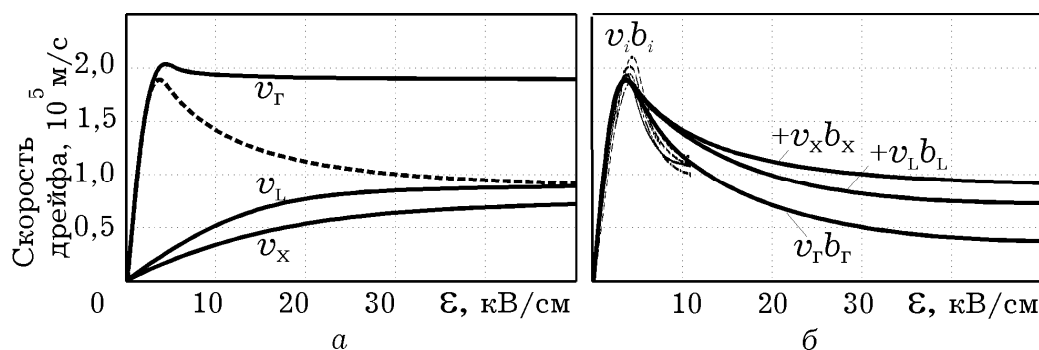


Рис. 3.13. Зависимости от напряженности электрического поля скоростей (а) дрейфа электронов для каждой долины и их усредненная зависимость (штриховая линия). Та же зависимость (б), построенная суммированием скоростей в отдельных долинах с весовыми коэффициентами, равными их заселенностям. Штриховые линии (б) от 0 до 10 кВ/см взяты из [3.7, 3.8]

Этот спад прекращается, поскольку при электронных температурах около 1500 К включаются механизмы междолинного рассеяния, в которых участвуют фононы с энергией $\hbar\omega_{ij} \approx 0,8\hbar\omega_0 = 0,26$ эВ. При больших электронных температурах эти механизмы становятся преобладающими и общий коэффициент неупругости в этих условиях удобно выразить через его значение для междолинного рассеяния

$$\gamma = \gamma_{ij} \frac{1 + \tau_{pij} / \tau_{рпо}}{1 + \tau_{Eij} / \tau_{Eпо}}. \quad (3.30)$$

Такое представление позволяет рассчитывать скорость насыщения по формуле (3.25), но используя параметры междолинных фононов и поправочный коэффициент, учитывающий вклад других фононов, в виде дроби из последнего выражения. Для Γ -долины, например, при $T_e = 6000$ К этот коэффициент равен 1,02, стремясь к 1 при дальнейшем повышении электронной температуры. Поэтому с хорошей точностью можно считать, что в этих условиях неупругость соударений определяется только неупругостью междолинного рассеяния, которая в соответствии с (2.38) практически линейно возрастает при увеличении T_e/T в сильных электрических полях с коэффициентом пропорциональности $\beta \approx 6,5$. А так как тепловая скорость электронов в Γ -долине равна $v_{T\Gamma} = \sqrt{3k_B T / m_\Gamma^*}$, что при комнатной температуре составляет $4,6 \cdot 10^5$ м/с, то скорость дрейфа достигает значения насыщения $v_{н\Gamma} \approx 1,8 \cdot 10^5$ м/с.

В итоге поле-скоростная характеристика для Γ -долины (рис. 3.13) достигнув максимального значения скорости $\approx 2,1 \cdot 10^5$ м/с, в

результате смены преобладающего вида рассеяния с полярного оптического на междолинное несколько снижается и насыщается при $v_{нГ}$. Пиковое значение скорости имеет при значении напряженности электрического поля, несколько меньшем 4 кВ/см, которое обычно и принимают в качестве критического поля.

В L-долинах при достаточно высоких электронных температурах также преобладают междолинные виды рассеяния, как в эквивалентные долины (LL-), так и в неэквивалентные (LG-, LX-). Но заметную роль играет также акустическое и в меньшей мере полярное оптическое. Поэтому коэффициент неупругости можно определить, как и в Г-долине, по (3.29) с поправочным коэффициентом, равным 1,2...1,4. Это дает значение $\beta \approx 9$ и соответственно значение скорости насыщения $v_{нL} \approx 1,1 \cdot 10^5$ м/с. Соответствующая поле-скоростная характеристика для электронов L-долины представлена на рис. 3.10, а.

В очень сильных электрических полях нельзя ограничиться двухдолинным приближением, так как заметная доля электронов находится в состояниях X-долин. Параметры электронов в них обеспечивают скорость насыщения $v_{нX} \approx 0,7 \cdot 10^5$ м/с, которая может быть рассчитана так же, как и в L-долинах.

Средняя скорость электронов находится суммированием скоростей в отдельных долинах с весовыми коэффициентами, равными соответствующим заселенностям (3.9), которые также можно представить функциями напряженности поля, используя предварительно найденную зависимость (3.12) между T_e и \mathcal{E} . На рис. 3.13, б представлена результирующая поле-скоростная характеристика и вклады в нее электронов каждой долины, а также результаты (штриховые линии) от 0 до 10 кВ/см, взятые из [3.7, 3.8].

Фосфид индия имеет подобную GaAs зонную структуру и соответственно поле-скоростную характеристику (рис. 3.14, а). Ее наклон в слабом поле, который определяется подвижностью, несколько меньше из-за большей эффективной массы в Г-долине. Но из-за большего энергетического расстояния до верхних долин (0,59 эВ) скорость достигает большего максимального значения $\approx 2,5 \cdot 10^5$ м/с соответственно при большем значении критической напряженности поля, близком к 10 кВ/см. Для сравнения на том же рисунке приведены результаты моделирования (сплошные) и эксперимента (штриховые и пунктирные) до напряженности 30 кВ/см, взятые из [3.4].

Нитрид галлия с кубической кристаллической структурой имеет еще меньшую подвижность в слабом поле прежде всего из-за большей эффективной массы. Самым существенным отличием

его поле-скоростной характеристики является то, что переход к режиму насыщения происходит при значительно более высокой напряженности электрического поля около 100 кВ/см. Кроме того, спад дрейфовой скорости, вызванный интенсификацией переходов в верхние долины, начинается при напряженности около 150 кВ/см, что соответствует междолинному расстоянию 1,3 эВ.

В чем-то подобные параметры поле-скоростной характеристики имеет нитрид галлия с гексагональной структурой кристаллической решетки. При расчетах принималось, что междолинные переходы происходят не только из нижней Γ_1 -долины в верхнюю Γ_2 -долину (1,9 эВ), но и в еще более высокие ЛМ-долины (2,1 эВ). Хотя эффективные массы в них почти одинаковы ($m^* \approx m$), за счет большего количества последних будет преобладать Γ_1 -ЛМ рассеивания. В итоге интенсивные междолинные переходы начинаются при еще больших (~200 кВ/см) напряженностях, чем для кубической структуры. На рис. 3.14 б, в для сравнения приведены результаты моделирования поле-скоростных характеристик методом Монте-Карло [3.9].

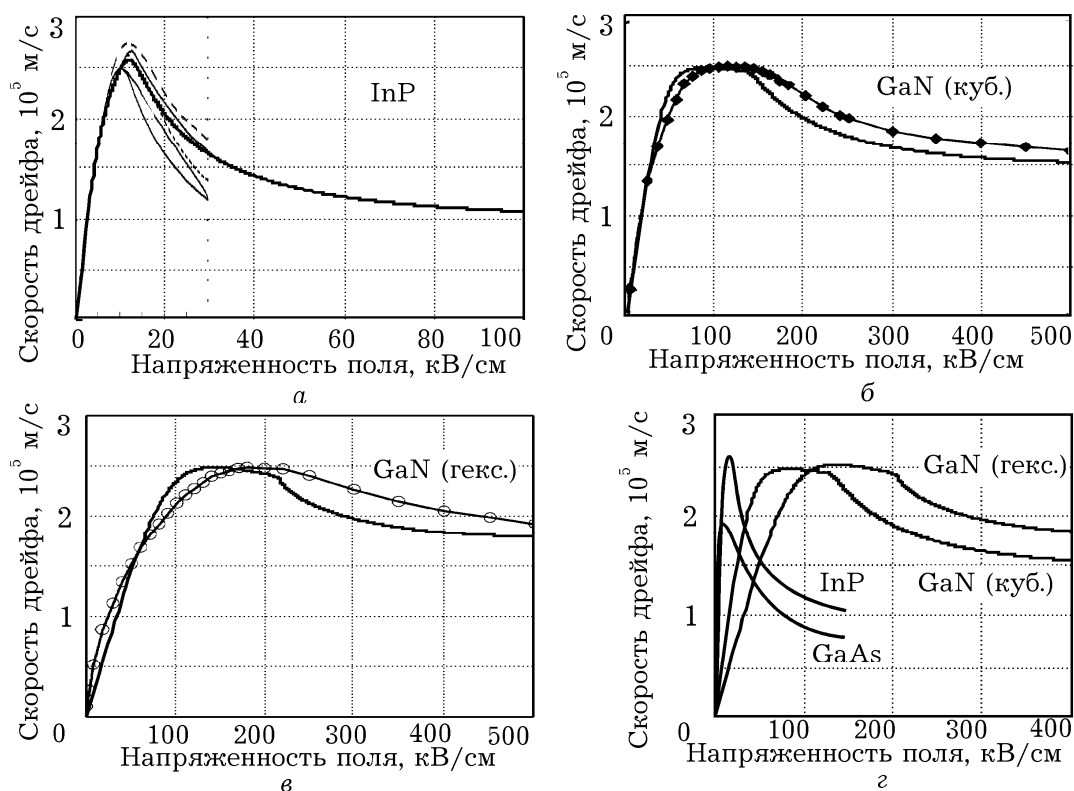


Рис. 3.14. Поле-скоростные характеристики фосфида индия (а), нитрида галлия, с кубической (б) и гексагональной (в) структурой. Для удобства сравнения на (г) эти характеристики приведены в одинаковом масштабе

Приведенные на рис. 3.14, г характеристики указывают на особенности этих материалов и на преимущество нитрида галлия для использования в приборах повышенной мощности.

3.3. ДИНАМИЧЕСКАЯ ПРОВОДИМОСТЬ

3.3.1. Дифференциальная проводимость

Статическая удельная электропроводность определяется как коэффициент пропорциональности между плотностью тока проводимости и напряженностью электрического поля: $j = \sigma \mathcal{E}$. Соответственно определим динамическую (дифференциальную) проводимость как коэффициент пропорциональности между изменениями dj и $d\mathcal{E}$: $\sigma_d = dj/d\mathcal{E}$. Так как плотность тока $j = env$, то при неизменной концентрации $dj = en \cdot dv$. В слабых электрических полях скорость дрейфа пропорциональна напряженности поля $v = \mu \mathcal{E} = (e\tau_p/m^*)\mathcal{E}$, а так как подвижность постоянна, то $dv = \mu d\mathcal{E}$ и дифференциальная проводимость совпадает со статической. Но в сильных полях из-за особенностей оптического рассеяния проявляется насыщение скорости, уменьшается время релаксации импульса и соответственно подвижность. В этих условиях из стационарного уравнения баланса импульса следует, что $(e/m^*)d\mathcal{E} = (\tau_p^{-1}) \cdot dv + v d(\tau_p^{-1})$. В режиме насыщения $dv = 0$ и $d(\tau_p^{-1}) = (e/m^* v_n) d\mathcal{E}$. Если эту обратную связь между подвижностью и напряженностью поля экстраполировать на переходную область, то можно найти, что $dv = \mu(1 - v/v_n)d\mathcal{E}$ и

$$\sigma_d = \sigma(1 - v/v_n). \quad (3.31)$$

В слабых полях $v \ll v_n$ и дифференциальная проводимость совпадает со статической, но по мере приближения к режиму насыщения уменьшается до 0. Полученное соотношение (3.31) пригодно для однодолинных полупроводников (Si, Ge), если концентрация электронов в зоне остается неизменной.

Результаты предыдущего параграфа показывают, что характерной особенностью арсенида галлия и других многодолинных полупроводников может быть наличие падающего участка на зависимости дрейфовой скорости от напряженности электрического поля. Объясняется это тем, что по мере разогрева полем увеличивается доля электронов, которые при междолинном рассеивании переходят в верхние долины, где их эффективная масса становится большей. В результате уменьшается средняя дрейфовая скорость и соответственно ток проводимости. Такая

ситуация приводит к тому, что положительные изменения поля вызывают отрицательные изменения тока проводимости в материале, что и соответствует *отрицательной динамической проводимости* (ОДП). На рис. 3.15 (штриховая линия) представлена ее зависимость от напряженности поля, полученная численным дифференцированием функции $v(\mathcal{E})$.

Для анализа основных факторов, определяющих ее величину, примем во внимание, что ток создается различными группами электронов, имеющих в разных долинах свои характерные параметры, то есть $j = \sum_i en_i v_i = en \sum_i b_i v_i$, и поэтому $\sigma_d = \sum_i \sigma_{di}$, где суммирование ведется по Γ -, L-, а при необходимости и по X-долинам. Для каждой долины

$$\sigma_{di} = end(b_i v_i)/d\mathcal{E} = en[b_i dv_i/d\mathcal{E} + v_i db_i/d\mathcal{E}], \quad (3.32)$$

где первое слагаемое может быть описано соотношением, подобным (3.31).

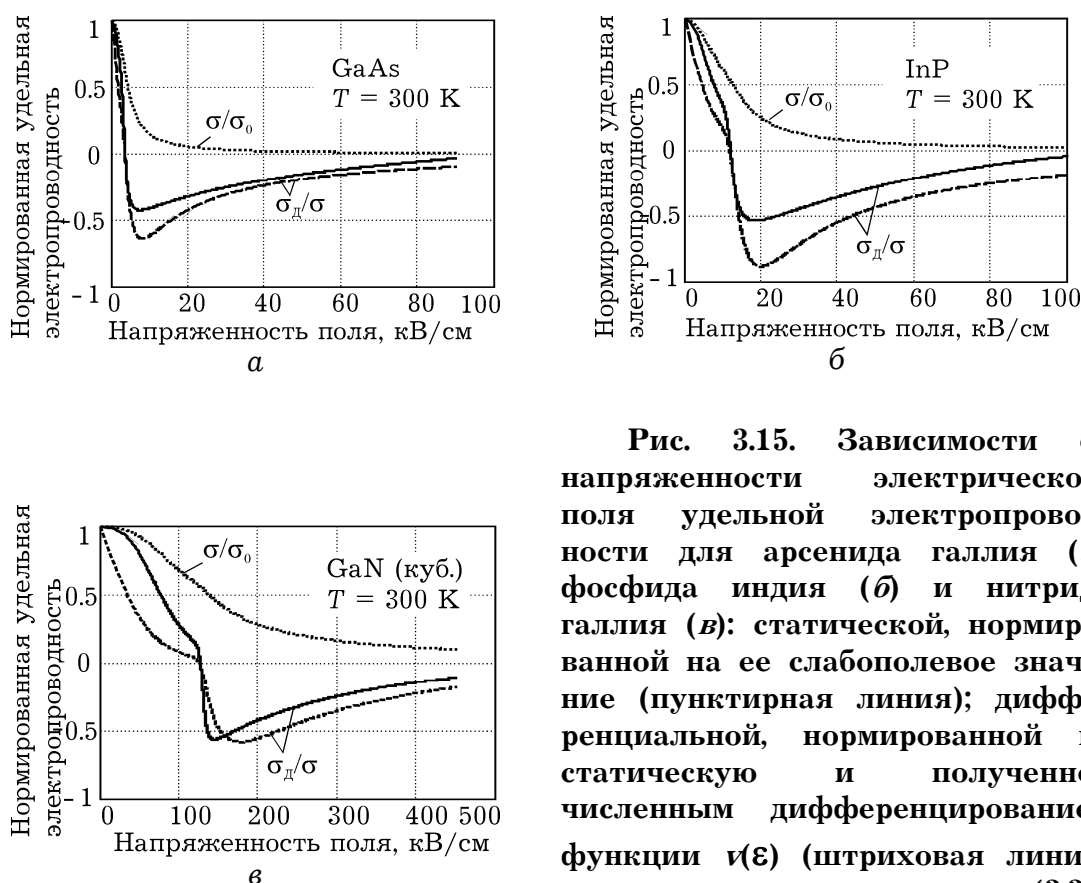


Рис. 3.15. Зависимости от напряженности электрического поля удельной электропроводности для арсенида галлия (а), фосфида индия (б) и нитрида галлия (в): статической, нормированной на ее слабополевое значение (пунктирная линия); дифференциальной, нормированной на статическую и полученной численным дифференцированием функции $v(\mathcal{E})$ (штриховая линия) и рассчитанную по (3.37) (сплошная линия)

Во втором слагаемом, связанном с перезаселением долин, надо иметь в виду, что заселенности долин $b_i = n/n_i$ зависят от напряженности поля, но не явно, а через электронную температуру. Поэтому

$$db_i = \frac{db_i/dT_e}{d\mathcal{E}/dT_e} d\mathcal{E}. \quad (3.33)$$

Для Γ -долины, используя соотношения (1.57) и (2.32), можно получить

$$\frac{db_\Gamma}{dT_e} = -b_\Gamma(1 - b_\Gamma) \frac{\Delta_{\Gamma L}}{k_B T_e^2}. \quad (3.34)$$

Производная $d\mathcal{E}/dT_e$ может быть легко найдена для режима насыщения, где и проявляются междолинные переходы, из условия $d\mathcal{E} = v_n d(\mu^{-1})$. Отсюда $d\mathcal{E}/dT_e = v_n d(\mu^{-1})/dT_e$, а так как в сильном поле при междолинном рассеянии можно считать, что $1/\tau_p \sim T_e$ (рис. 3.6, б), то

$$d\mathcal{E}/dT_e \approx v_n / \mu T_e. \quad (3.35)$$

В итоге для расчета дифференциальной проводимости, обусловленной электронами только Γ -долины, можно получить

$$\sigma_{д\Gamma} = en\mu_\Gamma b_\Gamma \left[1 - \frac{v_\Gamma}{v_{н\Gamma}} - \frac{v_\Gamma}{v_{н\Gamma}} (1 - b_\Gamma) \frac{\Delta_{\Gamma L}}{k_B T_e} \right]. \quad (3.36)$$

Полная проводимость включает подобное выражение для верхней (верхних) долины. Для упрощения можно принять, что в области наиболее резкого спада результирующей кривой $v(\mathcal{E})$ (рис. 3.13, а) скорость электронов в верхних долинах еще не достигает насыщения и их подвижность можно считать постоянной и к тому же почти на порядок меньшей. В итоге хорошим приближением может служить выражение

$$\sigma_d \approx \sigma(1 - v_\Gamma/v_{н\Gamma} - b_\Gamma b_L \Delta_{L\Gamma}/k_B T_e), \quad (3.37)$$

в котором вклад электронов верхних долин учтен в суммарной статической электропроводности σ , а в последнее слагаемое может быть введен корректирующий коэффициент, равный 0,6...0,8.

На рис. 3.15 представлены нормированные зависимости σ_d/σ от напряженности электрического поля: рассчитанные по (3.37) и полученные численным дифференцированием кривой $v(\mathcal{E})$. Наиболее характерной особенностью данных результатов является наличие области отрицательной дифференциальной проводимости, которая проявляется практически сразу с началом междолинных переходов ($\mathcal{E} \approx 4$ кВ/см) и имеет максимальное значение при $\mathcal{E} \approx 5...10$ кВ/см, когда концентрации электронов в верхней и нижней долинах приблизительно одинаковы ($b_U b_L \approx 0,25$). Стоит обратить внимание, что из рассмотренных материалов наибольшее относительное значение ОДП имеет фосфид индия.

Приведенный расчет дифференциальной проводимости применим лишь для медленных изменений электрического поля и соответственно тока. С повышением скорости изменений (частоты) необходимо учитывать инерционность различных процессов: запаздывание столкновений, которое характеризуется временем релаксации импульса (τ_p), релаксации энергии (τ_E), инерционность междолинных переходов (τ_n), а для этого использовать нестационарные уравнения сохранения.

3.3.2. Высокочастотная проводимость

Для высокочастотных изменений электрического поля ограничимся расчетом так называемой малосигнальной проводимости. Для этого все переменные величины $a(t)$ представляются в виде суммы постоянной составляющей a_0 и малой гармонически меняющейся во времени переменной с амплитудой a_{\sim} : $a(t) = a_0 + a_{\sim} e^{i\omega t}$. Пренебрегая затем членами второго порядка малости и учитывая, что $d/dt = i\omega$, используемые уравнения разделяем на не зависящие от времени для постоянных составляющих и линейные уравнения для переменных.

Использование такой процедуры для плотности тока $j = env$ дает для его высокочастотной составляющей соотношение

$$j_{\sim} = en_0 v_{\sim} + e v_0 n_{\sim}. \quad (3.38)$$

Отсюда видно, что переменная составляющая тока определяются двумя факторами: переменной скоростью и переменной концентрацией электронов за счет междолинных переходов. Соответственно высокочастотная проводимость для каждой долины включает два слагаемых. Вычисляя v_{\sim} и n_{\sim} как

пропорциональные переменной амплитуде электрического поля \mathcal{E} , получим два характерных слагаемых для проводимости.

Перенося детальные расчеты в приложение ПЗ, здесь приведем лишь их результат. Для первого слагаемого в соответствии с (3.38) результат можно представить в таком виде:

$$\sigma_{\omega}^{(1)} = \frac{\sigma_0}{1 + i\omega\tau_{p0}} \left[1 - \frac{v_0}{v_H} \cdot \frac{1}{1 + i\omega\tau_{E0}} \right]. \quad (3.39)$$

Если считать, что величина проводимости определяется в основном электронами одной долины, то (3.39) при $\omega \rightarrow 0$ совпадает с формулой (3.37), описывающей низкочастотную дифференциальную проводимость.

Второе слагаемое, которое определяется переменной составляющей концентрации электронов, как дополнение к (3.39) для высокочастотной электропроводности, можно получить в виде

$$\sigma_{\omega}^{(2)} = - \frac{e^2 n_0 \tau_{p0}}{m^*} \cdot \frac{v_0}{v_H} \cdot \frac{\tau_{E\omega}}{\tau_{E0}} \cdot \frac{b_{i0} b_{j0}}{1 + i\omega\tau_M} \cdot \frac{\Delta_{ji}}{k_B T_e}. \quad (3.40)$$

Здесь введено понятие общего времени междолинной релаксации через обратные времена релаксации между отдельными долинами

$$\tau_M = (1/\tau_{ij0} + 1/\tau_{ji0})^{-1}, \quad (3.41)$$

В итоге величина высокочастотной электропроводности, нормированная на ее статическое значение, может быть рассчитана из соотношения

$$\frac{\sigma_{\omega}}{\sigma} = \frac{\tau_{p\omega}}{\tau_{p0}} - \frac{v_{0\Gamma}}{v_{H\Gamma}} \cdot \frac{\tau_{E\omega}}{\tau_{E0}} \left(\frac{\tau_{p\omega}}{\tau_{p0}} + \frac{b_{i0} b_{j0}}{1 + i\omega\tau_M} \cdot \frac{\Delta_{ji}}{k_B T_e} \right). \quad (3.42)$$

Прежде всего необходимо отметить, что проводимость имеет комплексный характер, который проявляется на высоких частотах. На рис. 3.16 представлены частотные зависимости действительной (а) и мнимой (б) частей (верхний ряд) электропроводности, нормированной на ее статическое значение $\sigma = e^2 n \tau_{p0} / m^*$.

Частотная зависимость σ_ω определяется множителями типа $(1 + i\omega\tau)^{-1}$, содержащими три временные константы τ_p , τ_E и τ_m . Первое слагаемое зависит от частоты столкновений, то есть от времени релаксации импульса, и объясняется инерционностью изменений тока при изменении поля. Второе слагаемое связано с переходом дрейфовой скорости к режиму насыщения, что требует участия оптических фононов и возможно только при повышении температуры электронов. Инерционность процессов разогрева определяется временем релаксации энергии, поэтому второе слагаемое содержит еще и τ_E . Для процессов, связанных с перераспределением электронов по долинам, характерна инерционность как процессов разогрева, так и междолинных переходов. Поэтому в третье слагаемое входят времена τ_E и τ_m (3.41). Существенно также, что оно пропорционально энергетическому расстоянию Δ_{ij} между долинами и предопределяет отрицательные значения электропроводности.

Из этих трех времен наибольшее значение (около 1 пс для GaAs, несколько меньше для InP и около до 0,3 пс для GaN) имеет время τ_E , поэтому верхний частотный предел ОДП зависит прежде всего от инерционности процессов разогрева электронов и составляет десятки (для GaAs) и сотни (для GaN) гигагерц. В слабом поле (верхняя кривая на рис. 3.16) процессы разогрева и междолинных переходов несущественны, поэтому спад проводимости определяется величиной τ_p и происходит на частоте около 100 ГГц.

Реактивная проводимость проявляется на более низких частотах (порядка единиц или десятков ГГц), что объясняется особенностями множителей $(1 + i\omega\tau)^{-1} = (1 - i\omega\tau)/(1 + \omega^2\tau^2)$: реактивная проводимость в отличие от активной пропорциональна первой степени $\omega\tau$. В слабом поле она имеет индуктивный характер, но в области ОДП меняет знак вместе с активной составляющей и становится емкостной. В дальнейшем на очень высоких частотах за время периода поля межзонные переходы не успевают происходить (высокочастотная составляющая заселенности долин b_{\sim} уменьшается) и реактивная проводимость снова становится индуктивной. При этом, поскольку большая часть электронов находится в верхних долинах, величина проводимости определяется параметрами этих долин.

Учет тока смещения $j_{cm} = i\omega\epsilon\epsilon_0\mathcal{E}$ приводит к появлению в комплексной проводимости σ_ω только мнимой составляющей $i\omega\epsilon\epsilon_0$, а для нормированной величины σ_ω/σ ее вклад равен $\Delta\sigma_\omega/\sigma =$

$= i\omega\varepsilon\varepsilon_0/en\mu = i\omega\tau_{\text{мкс}}$. Здесь $\tau_{\text{мкс}} = \varepsilon\varepsilon_0/\sigma$ – максвеллово время релаксации, которое и определяет инерционность процессов, связанных с токами смещения. Его величина может изменяться в широких пределах в зависимости от концентрации примеси, но для типичных значений N_d порядка 10^{23} м^{-3} это время в сильных электрических полях составляет $10^{-1} \dots 10^{-2}$ пс.

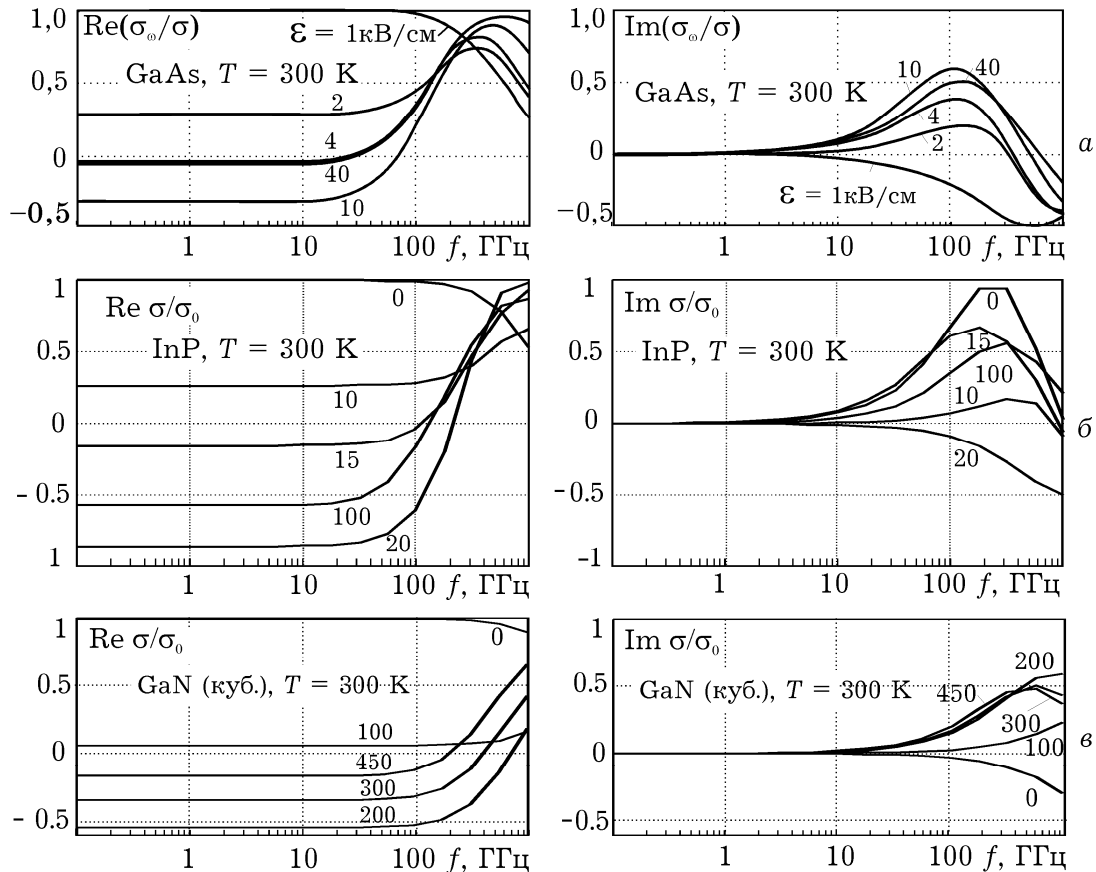


Рис. 3.16. Частотные зависимости действительной и мнимой частей динамической удельной электропроводности, нормированной на статическое значение для арсенида галлия (а), фосфида индия (б) и нитрида галлия (в)

Соответственно влияние тока смещения заметно лишь на частотах в сотни гигагерц. Но уже при концентрациях примеси 10^{21} м^{-3} максвеллово время релаксации становится большим, чем время релаксации энергии, и инерционность проводимости будет определяться в основном токами смещения.

Важным результатом проведенного анализа является возможность расчета граничной частоты существования отрицательной динамической проводимости. Его можно провести, пренебрегая инерционностью междолинных переходов ($\omega\tau_m \ll 1$), запаздыванием столкновений ($\omega\tau_p \ll 1$) и считая, что в режиме

ОДП скорость электронов достигает насыщения ($v_{0Г} = v_{нГ}$). Тогда формула (3.42) упрощается и приобретает вид:

$$\frac{\sigma_{\omega}}{\sigma} \approx 1 - \frac{1 - i\omega\tau_{E0}}{1 + \omega^2\tau_{E0}^2} \left(1 + b_{i0}b_{j0} \frac{\Delta_{ji}}{k_B T_e} \right). \quad (3.43)$$

Отсюда из условия $\text{Re } \sigma_{\omega} = 0$, которое и соответствует предельной ситуации, легко получить

$$\omega_m \tau_{E0} = \sqrt{b_{i0}b_{j0} \Delta_{ji} / k_B T_e}. \quad (3.44)$$

При напряженности поля, когда заселенности верхних и нижней долины одинаковы и значение ОДП максимально, последнее выражение упрощается и граничная частота равна

$$f_m = \frac{1}{4\pi\tau_{E0}} \sqrt{\frac{\Delta_{ji}}{k_B T_e}}. \quad (3.45)$$

Как видно из рис. 3.17, граничная частота определяется временем релаксации энергии и имеет значения порядка десятков гигагерц, уменьшаясь с ростом напряженности электрического поля.

Полученные результаты, приведенные на рис. 3.17, для предельной частоты для арсенида галлия и фосфида индия в целом соответствуют известным результатам [3.10], но для нитрида галлия существенно превышают результаты прогноза в [3.11].

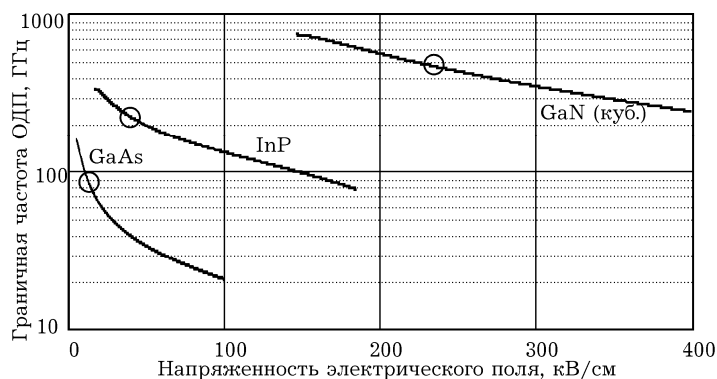


Рис. 3.17. Зависимость максимальной частоты существования отрицательной дифференциальной проводимости от напряженности электрического поля.

3.4. «ВСПЛЕСК» ДРЕЙФОВОЙ СКОРОСТИ

Дрейфовые процессы проявляют интересные особенности в импульсном электрическом поле. При резком увеличении напряженности поля значение скорости дрейфа может быть получено в несколько раз большим, чем стационарное. Это явление известно как «всплеск» (overshoot) дрейфовой скорости. Характерными представляются результаты численного эксперимента [3.7], в котором напряженность поля в GaAs скачком изменяется в момент времени t_0^+ от значения $\mathcal{E}_1 = 2$ кВ/см до $\mathcal{E}_2 = 40$ кВ/см. Решением (1.44) в данном случае должна быть функция, подобная (2.40), но с начальным условием $v(t_0^+) = \mu_1 \mathcal{E}_1$:

$$v(t) = v(t_0^+) + [\mu_2 \mathcal{E}_2 - v(t_0^+)] \{1 - \exp[-(t - t_0^+)/\tau_p]\}. \quad (3.46)$$

Так как этим значениям напряженности электрического поля соответствуют по статической поле-скоростной характеристике (рис. 3.13) приблизительно одинаковые значения скорости дрейфа около $1 \cdot 10^5$ м/с, то переходного процесса не должно наблюдаться. Однако численное интегрирование полной системы уравнений сохранения импульсов, энергий и концентраций в Г-, L- и X-долинах с расчетом времен релаксаций на каждом шаге интегрирования дает результаты, которые представлены на рис. 3.18. Наиболее ярким из них является кратковременное увеличение («всплеск») скорости до значений, в 6...7 раз превышающих ее стационарное значение.

Это явление связано с инерционностью процессов рассеяния и с зависимостью времен релаксации от напряженности поля. При начальном значении напряженности 2 кВ/см (которое еще можно считать слабым) электроны разогреваются до температуры ≈ 700 К, которой соответствуют времена релаксации импульса $\approx 0,2$ пс и энергии ≈ 1 пс, причем практически все электроны занимают состояния Г-долины. При включении поля с напряженностью 40 кВ/см параметры движения некоторое время сохраняются прежними и нарастание скорости можно описать (3.46), но со временем релаксации импульса 0,2 пс.

На начальном участке ускорение определяется как в вакууме отношением $e\mathcal{E}_2/m^*$, а скорость возрастает по закону $v(t) = v(t_0^+) + (e/m^*)\mathcal{E}_2(t - t_0^+)$ и за время $\tau_{p1} = 0,2$ пс она достигла бы значения

$\mu_1 \mathcal{E}_2 \approx 20 \cdot 10^5$ м/с. То есть в процессе нарастания скорость стремится к значению, которое определяется высоким значением подвижности в слабом поле \mathcal{E}_1 и большим значением напряженности \mathcal{E}_2 .

Однако этого значения скорость достичь не успевает, так как одновременно с возрастанием скорости происходит дальнейший разогрев электронного газа с преобладанием почти упругого акустического и частично полярного оптического рассеяния. Инерционность этого процесса определяется не столько процессами релаксации энергии, как это можно было бы ожидать из (2.42), сколько интенсивной хаотизацией направленного движения и разогревом электронов почти без релаксации энергии, что формально отображается в преобладающем весе третьего слагаемого в уравнении (3.5), так как $\tau_{E1} > \tau_{p1}$.

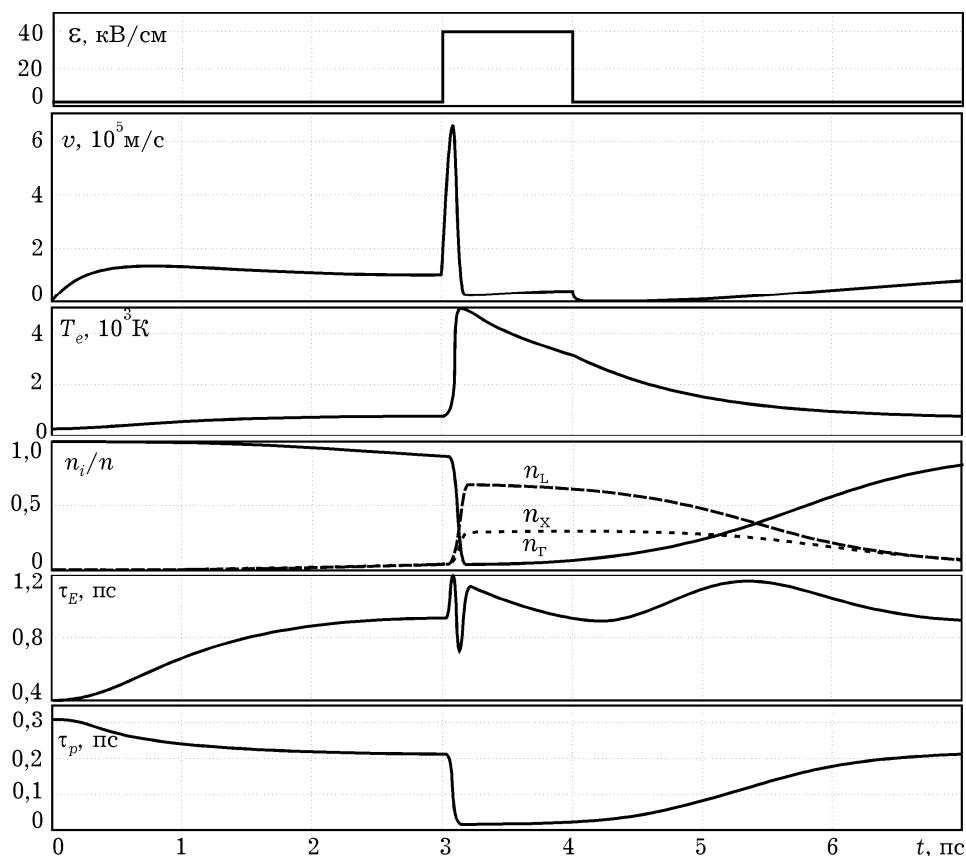


Рис. 3.18. Динамические (временные) зависимости при включении импульса электрического поля от 2 до 40 кВ/см для (сверху вниз): средней скорости дрейфа, электронной температуры, заселенности долин, усредненных по долинам времен релаксации энергии и импульса. Следует обратить внимание на экстремальные значения скорости дрейфа с учетом того, что при напряженностях электрического поля 2 и 40 кВ/см ее статические значения приблизительно одинаковы и равны $1 \cdot 10^5$ м/с.

Такая ситуация приводит к «всплеску» электронной температуры, которая на этом этапе достигает значений ≈ 5000 К, что почти вдвое превышает ее стационарное значение. Но уже при $T_e \approx 2000$ К включаются междолинные механизмы рассеяния и происходит быстрое перераспределение электронов по долинам, уменьшение τ_p до значений $\approx 0,02$ пс и возрастание τ_E до $\approx 1,4$ пс. Такое изменение τ_p вызывает быстрое уменьшение скорости дрейфа, причем после резкого спада скорость может падать ниже стационарного значения для $\mathcal{E}_2 = 40$ кВ/см, а при обратном фронте импульса поля она может быть даже отрицательной. Это явление известно как отрицательный «всплеск» (*undershoot*) и связано с различными временами релаксации электронов, которые движутся вдоль поля или навстречу ему. Анализ последнего явления не может быть проведен на основе релаксационных уравнений, использующих усредненные времена релаксации, а доступен только методу Монте-Карло.

Приведенное объяснение подтверждают также результаты численного моделирования переходных процессов (рис. 3.19) для разных амплитуд импульса электрического поля. С увеличением напряженности поля, естественно, растет максимальное значение дрейфовой скорости (рис. 3.19, а). Одновременно с этим наблюдается и «всплеск» электронной температуры (рис. 3.19, б), который влечет уменьшение времени релаксации импульса с увеличением напряженности поля и соответственно сокращает длительность «всплеска» скорости.

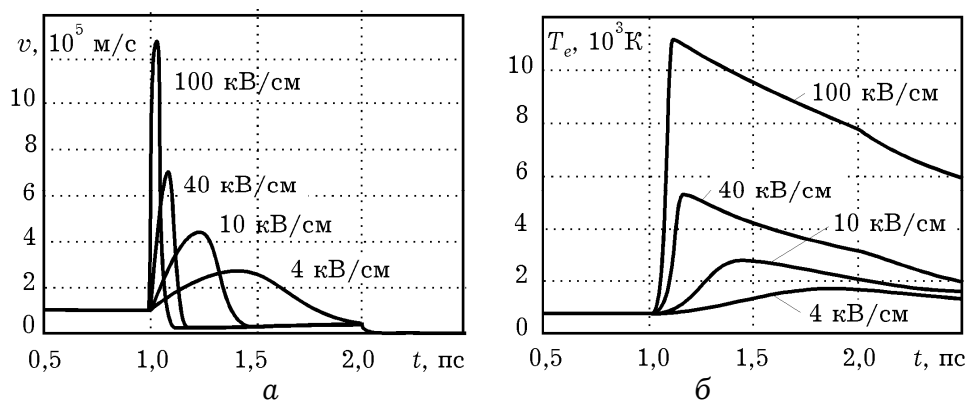


Рис. 3.19. Динамические зависимости дрейфовой скорости и электронной температуры для разных значений амплитуды напряженности поля. Длительность импульса поля 1 пс

Аналогичные процессы происходят и в статическом, но пространственно неоднородном электрическом поле. Это легко понять, приняв во внимание, что в общем случае уравнения

сохранения содержат полную производную по времени от функций координат и времени: $d/dt = \partial/\partial t + \vec{v}\partial/\partial\vec{r}$. В переменном во времени, но однородном в пространстве поле $\partial/\partial\vec{r} = 0$ и $d/dt = \partial/\partial t$, а в статическом, но пространственно изменяющемся поле $d/dt = \vec{v}\partial/\partial\vec{r}$. Поэтому уравнения и соответственно их решения будут иметь подобный вид при замене $x = \int_{t_0}^t v(t') dt'$.

Иллюстрацией может служить рис. 3.20, где представлена пространственная зависимость скорости электронов, дрейфовавших в статическом поле с напряженностью $\mathcal{E}_1 = 2$ кВ/см со скоростью $v_0 = \mu_1\mathcal{E}_1$ и влетающих в поле $\mathcal{E}_2 = 40$ кВ/см. Там же пунктиром представлена зависимость $v(x) = \sqrt{v_0^2 + 2(e/m^*)\mathcal{E}_2x}$, характеризующая движение "свободного" (в приближении эффективной массы) электрона в однородном поле с начальной скоростью v_0 без столкновений ($\tau_p = \infty$). Штриховая линия соответствует функции

$$\frac{x}{\tau_{p1}} = v_{ст} \ln \left| \frac{v_{ст} - v_0}{v_{ст} - v(x)} \right| + v_0 - v(x), \quad (3.47)$$

являющейся решением уравнения баланса импульса для $\tau_p = \tau_{p1} = \text{const}$. В данном расчете время релаксации импульса соответствует слабому полю, а $v_{ст} = \mu_1\mathcal{E}_2$ — скорость дрейфа в установившемся режиме в поле с напряженностью 40 кВ/см, но с подвижностью μ_1 , такой же как и перед влетом электрона в сильное поле.

Из графиков видно, что после влета в пространство сильного поля наблюдается так называемое баллистическое движение, т.е. движение без рассеяния, подобное движению в вакууме. Оно длится около 0,1 пс, за это время электрон успевает пролететь расстояние в несколько сотых долей мкм (порядка длины свободного пробега) и ускориться до $\sim 7 \cdot 10^5$ м/с. После этого, как и в динамическом случае, в результате начинающихся столкновений повышается электронная температура, сопровождающаяся переходами в верхние долины и уменьшением времени релаксации почти на порядок. В дальнейшем скорость релаксирует до значения, которое определяется низким значением подвижности в сильном электрическом поле.

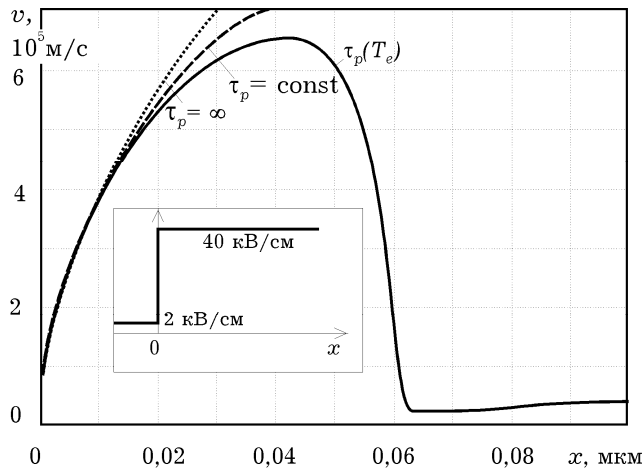


Рис. 3.20. Пространственное изменение скорости электронов, дрейфовавших в поле 2 кВ/см и влетающих в поле 40 кВ/см в точке $x = 0$. Пунктирная линия соответствует баллистическому движению без столкновений, штриховая – движению с теми же параметрами, что и в слабом поле.

Эффект "всплеска" может проявляться и в том случае, если напряженность поля меняется в пространстве не скачкообразно, а относительно плавно. Он заметен на расстояниях порядка $\sim 0,1$ мкм, поэтому даже в субмикронных приборах средняя скорость пролета может быть получена существенно выше, чем стационарная. Это наглядно видно из рис. 1.6 и 1.7, где приведены распределения скорости вдоль канала полевого транзистора с длиной затвора 0,5 мкм. Кроме того он может быть использован для получения сверхкоротких импульсов.

3.5. ДИФFUЗИЯ ГОРЯЧИХ ЭЛЕКТРОНОВ

С особенностями механизма разогрева электронов также связана характерная для многодолинных полупроводников зависимость коэффициента диффузии D от напряженности электрического поля (рис. 3.21).

Рассчитать ее можно, используя соотношение Эйнштейна: $D = \mu k_B T_e / e$. Учитывая зависимость подвижности от электронной температуры и зависимость электронной температуры от напряженности поля, можно объяснить ход функции $D(\mathcal{E})$: в слабых полях время релаксации импульса, а следовательно и подвижность, почти не зависят от напряженности, потому что рост коэффициента диффузии объясняется только повышением температуры электронного газа; в сильных полях, где преобладают оптическое и междолинное рассеивание, разогрев не такой значительный и коэффициент диффузии спадает из-за

уменьшения времени релаксации импульса и, соответственно, подвижности.

Максимальное значение коэффициент диффузии имеет при напряженности поля, близкой к критической. Влияние концентрации примеси наблюдается практически только в слабом поле из-за увеличения примесного рассеивания и соответствующего уменьшения подвижности при более высоком уровне легирования.

Объяснить ход зависимости коэффициента диффузии от напряженности электрического поля для фосфида индия (рис. 3.21, б) можно теми же механизмами, что и для арсенида галлия (рис. 3.21, а). Несколько меньшие значения коэффициента диффузии в слабом поле связаны с меньшей подвижностью электронов, а положение максимума отвечает значению критической напряженности поля, которое для фосфида индия составляет ~ 10 кВ/см

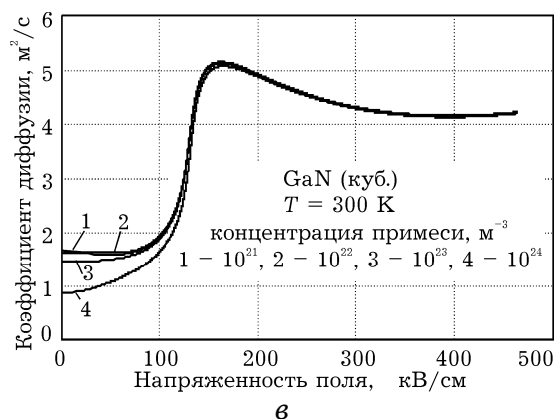
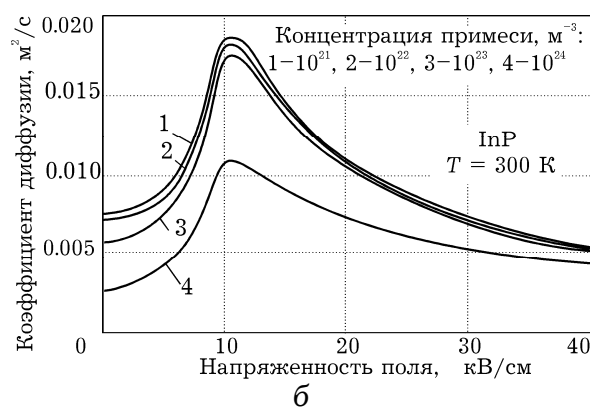
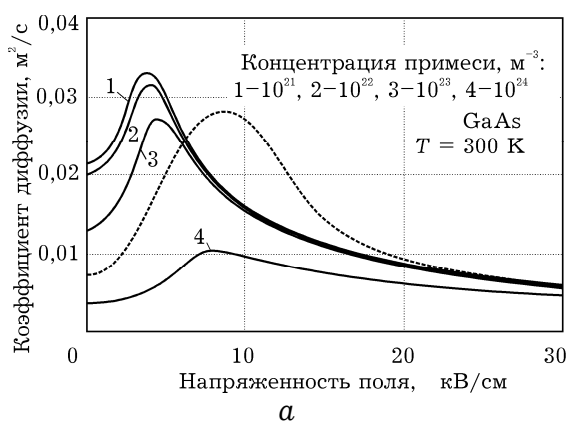


Рис. 3.21. Зависимости коэффициента диффузии от напряженности электрического поля для разных концентраций примеси для арсенида галлия (а), фосфида индия (б) и нитрида галлия (в). Штриховая линия построена по результатам работы [3.10]

Для нитрида галлия (рис. 3.21, в) зависимость коэффициента диффузии от напряженности электрического поля в целом имеет подобный вид. Почти на порядок меньшие значения в слабом поле в сравнении с арсенидом галлия вызваны меньшими значениями подвижности. Соответственно более низкие значения и при приближении к критическим полям, а в очень сильных полях

коэффициент диффузии имеет практически такие же значения, несмотря на то, что в нитриде галлия электронная температура может достигать существенно более высоких значений.

Следует отметить, что имеющиеся в научной литературе [3.3] экспериментальные и теоретические зависимости $D(\mathcal{E})$, в целом имея подобный вид, отличаются между собой как максимальным значением (до $0,1 \text{ м}^2/\text{с}$), так и положением максимума. Это связано, вероятнее всего, с разными методами измерений и некоторой неопределенностью условий эксперимента (концентрации примеси).

Литература к главе 3

- 3.1. Москалюк В. А., Синекон Ю. С., Кассинг Р. Физика электронных процессов: Учеб. пособие. – Ч. I: Электронные состояния. – К.: УкрИНТЭИ, 2001. – 148 с.
- 3.2. Зеегер К. Физика полупроводников: Пер. с англ. – М.: Мир, 1977. – 616 с.
- 3.3. Физикотехнический институт им. А.Ф.Иоффе Российской Академии Наук, New Semiconductor Materials. Characteristics and Properties URL: <http://www.ioffe.ru/SVA/NSM/>
- 3.4. U. V. Bhapkar and M. S. Shur, Monte Carlo calculation of velocity-field characteristics of wurtzite GaN, [J. Appl. Phys. 82, 4 \(1997\), 1649-1655.](#)
- 3.5. Бонч-Бруевич В.Л., Калашиников С.Г. Физика полупроводников.- М.: Наука,1977. – 367с.
- 3.6. Конуэлл Э. Кинетические свойства полупроводников в сильных электрических полях: Пер. с англ. – М.: Мир, 1970. – 384 с
- 3.7. Шур М. Современные приборы на основе арсенида галлия: Пер. с англ. – М.: Мир, 1991. – 632 с.
- 3.8. Blakemore, J. S., J. Appl. Phys. 53, 10 (1982) R123-R181.
- 3.9. Kolnik, J., Oguzman I.H., Brennan K.F., Wang R., Ruden P.P., Wang, Y. Electronic transport studies of bulk zincblende and wurtzite phases of GaN based on an ensemble Monte Carlo calculation including a full zone band structure, [J. Appl. Phys. 78\(2\) \(1995\), 1033-1038.](#)
- 3.10. Прохоров Э. Д., Белецкий Н. И. Полупроводниковые материалы для приборов с междолинным переносом. – Х.: Вища шк.; Изд. при Харьк. ун-те, 1982. – 144 с.
- 3.11. Данилин В., Жукова Т. Транзисторы на GaN пока самый крепкий орешек // Электроника: наука, технологии, бизнес. – 2005. – Т. 26, вып. 4. – С. 20–28.

4. СУБМИКРОННЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ

4.1. ОБЩИЕ СВОЙСТВА

Сверхвысокочастотные полевые транзисторы изготавливают обычно из арсенида галлия, реже из фосфида индия или нитрида галлия. Основанием структуры прибора служит полуизолирующая подложка (рис. 4.1, а), на которой выращен активный эпитаксиальный слой с высокой концентрацией донорной примеси, а также тонкий n^+ -контактный слой, нанесенный на активный эпитаксиальный слой с концентрацией примеси $\sim 1 \cdot 10^{19} \text{ см}^{-3}$. Такой слой необходим для формирования омических контактов стока и истока. Для уменьшения их омического сопротивления обычно используют сплав Au-Ge-Ni. Толщина активного эпитаксиального слоя составляет от 0,1 до 0,2 мкм при концентрации примеси примерно $\sim 3 \cdot 10^{17} \text{ см}^{-3}$.

Конструкции субмикронных ПТШ, применяемые в малошумящих усилителях (МШУ), как правило, представляют собой симметричные двухсекционные структуры. В настоящее время существует множество конструкций ПТШ, однако в качестве базовых структур могут рассматриваться структуры, приведенные на рис. 4.1. Применение той или иной конструкции связано с выбором согласующих элементов. Это могут быть цепи с сосредоточенными компонентами, а также отрезки линий передачи. Конкретный тип конструкции транзистора связан с типом усилителей – это могут быть, например, маломощные МШУ и усилители мощности. В мощных цепях используются многосекционные транзисторные структуры и согласующие цепи на отрезках линий передачи.

Затвор, расположенный между контактами стока и истока, образует с материалом полупроводника барьер Шоттки. Металл затвора должен иметь высокую проводимость, хорошую адгезию к GaAs, высокую надежность и не вступать в реакцию с GaAs вплоть до температур 400°C. В качестве материала затвора предпочтительна система тугоплавких металлов, таких, как Ti-Ta-Mo-Au. Барьер Шоттки на затворе может быть также выполнен из тугоплавкого материала Mo или W с нанесенным сверху слоем золота, который необходим для монтажа проволочных выводов. В поверхностном n -слое пластины вытравливается углубление до

нужного напряжения отсечки и тока стока насыщения, в котором затем формируется субмикронный затвор. На рис. 4.1 показано углубление затвора в структуре с n^+ -контактным эпитаксиальным слоем. Если даже на поверхности нет n^+ -контактного слоя, углубление затвора необходимо как для уменьшения сопротивления истока, так и для более эффективного управления током транзистора.

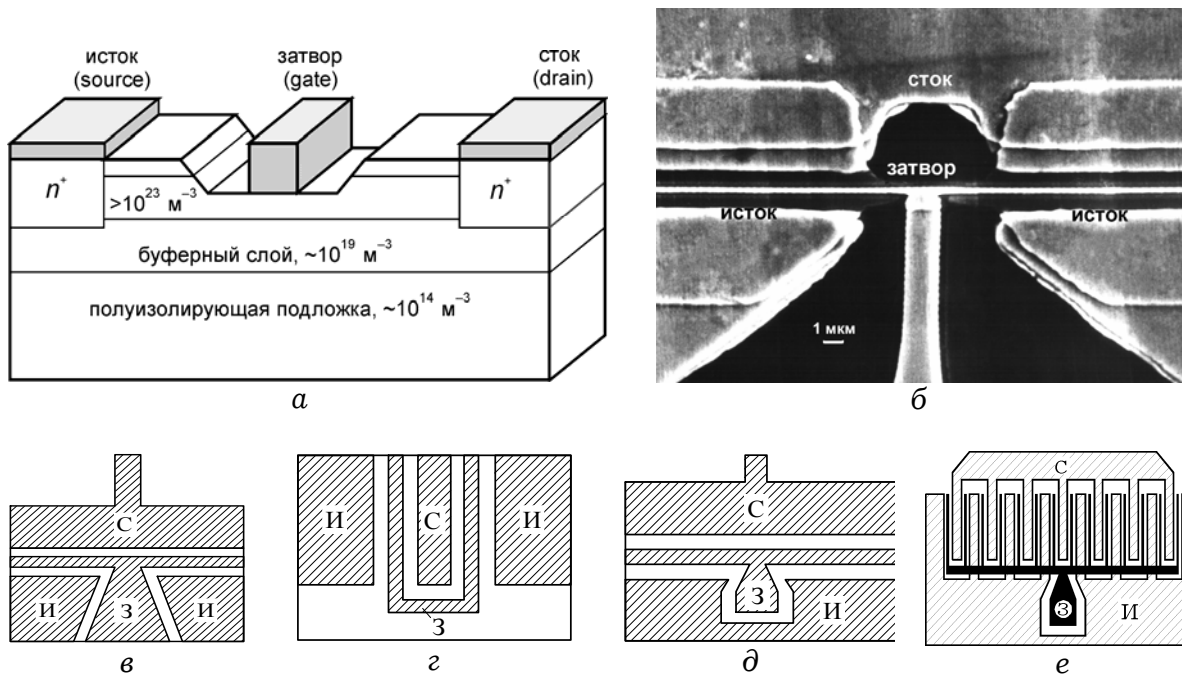


Рис. 4.1. Типичная структура (а) и фотография (б) полевого транзистора с затвором в виде барьера Шоттки и варианты конструкций малошумящих транзисторов: в – Hewlett Packard HP (США); г – Siemens (ФРГ); д – NEC (Япония), а также мощного (е) транзистора

Контактная разность потенциалов барьера совместно с приложенным внешним напряжением, отрицательным по отношению к истоку потенциалом смещения (транзистор обычно включается по схеме с общим истоком), создают под затвором обедненную область, формирующую проводящий канал для электронов между стоком и истоком. Толщина обедненной области в приближении ее резкой границы зависит от разности потенциалов затвора V_{gs} (с учетом контактной разности на барьере Шоттки V_b) и потенциалом канала $V(x)$, увеличивающимся от истока к стоку, и определяется известным соотношением

$$d(x) = \sqrt{\frac{2\epsilon\epsilon_0[V(x) + V_b - V_{gs}]}{eN_d}}. \quad (4.1)$$

Изменения потенциала затвора приводят к модуляции толщины канала и соответственно к управлению током стока. В определенном смысле канал можно рассматривать как нелинейный резистор, управляемый потенциалом затвора. Стоковые характеристики полевого транзистора при малых напряжениях стока близки к линейным и хорошо описываются такой «резистивной» моделью управления. При больших стоковых напряжениях при сужении канала напряженность электрического поля возрастает до таких значений, при которых наблюдается насыщение дрейфовой скорости электронов, что вызывает выполаживание стоковых характеристик или «насыщение» тока стока.

Простейшая аналитическая теория полевых транзисторов, созданная Шокли [4.1], основана на так называемом приближении «плавного канала», предполагающем медленное изменение толщины обедненной области и соответственно наличие в канале только продольного электрического поля. Модель Шокли приводит к следующей зависимости тока стока I_d от напряжения стока V_{ds} , напряжения затвора и физико-топологических параметров транзистора

$$I_d = G_0 \left\{ V_{ds} - \frac{2}{3} \frac{\left[(V_{ds} + V_b - V_{gs})^{3/2} - (V_b - V_{gs})^{3/2} \right]}{V_p^{1/2}} \right\}. \quad (4.2)$$

Здесь параметр $G_0 = e\mu N_d W a / L$ соответствует проводимости необедненного канала толщиной a , шириной W и длиной L с концентрацией электронов N_d , а V_p — так называемый потенциал перекрытия, равный потенциалу затвора, при котором толщина обедненной области у истока становится равной толщине эпитаксиального слоя a :

$$V_p = -\frac{eN_d a^2}{2\epsilon\epsilon_0} - V_b. \quad (4.3)$$

Функция (4.2) справедлива до значения напряжения стока, при котором толщина обедненной области в стоковой части становится равной a и канал «перекрывается», а напряженность поля достигает значений выше критического и скорость электронов насыщается. После этого значение тока остается неизменным (рис. 4.2) при увеличении напряжения стока и равным току насыщения:

$$I_{dH} = G_0 \left\{ \frac{1}{3} V_p - \frac{2}{3} \frac{(V_b - V_{gs})^{3/2}}{V_p^{1/2}} - V_b + V_{gs} \right\}. \quad (4.4)$$

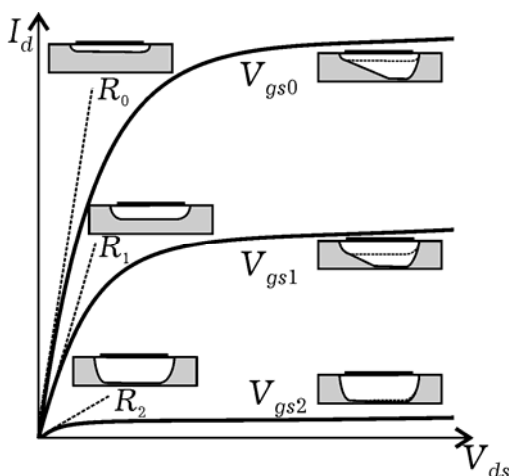


Рис. 4.2. Изменение формы канала при разных напряжениях затвора и стока

Модель Шокли пригодна для длинноканальных транзисторов, у которых длина затвора составляет микрометры, что гораздо больше толщины эпитаксиального слоя. Кроме того, она предполагает, что скорость дрейфа определяется *локальным значением* напряженности электрического поля. Но уменьшение длины затвора, необходимое для повышения рабочих частот, делает допущения этой модели несостоятельными. Это можно проиллюстрировать на примере стоковых характеристик транзистора – укорочение затвора до субмикронных размеров качественно меняет их вид (рис. 4.3)

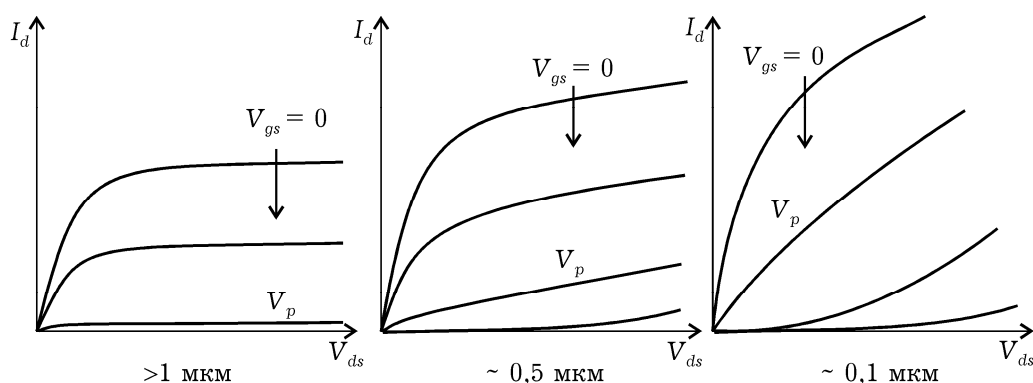


Рис. 4.3. Трансформация формы стоковых характеристик при уменьшении длины затвора

Для субмикронных длин затвора на вольтамперных характеристиках не наблюдается насыщения – выходная проводимость становится значительной и тем выше, чем меньше длина затвора. Примечательно также, что при потенциале затвора, близком к потенциалу перекрытия, стоковая характеристика близка к линейной, а ее наклон, определяющий выходную проводимость $G_d = \partial I_d / \partial V_{ds}$, приближенно равен наклону остальных ВАХ в режиме «насыщения». Кроме того, ее вид близок к омическому. Такую ситуацию можно представить как параллельное соединение проводимости канала обычного транзистора и некоторой шунтирующей проводимости. Количественные оценки показывают, что такую проводимость могут обеспечить электроны обогащенного слоя, образующегося у границы эпитаксиального слоя и подложки (буферного слоя).

Подтверждением такой модели могут служить результаты моделирования транзистора методом Монте-Карло [4.2]. На рис. 4.4 представлен потенциальный рельеф и пространственное распределение электронов в канале открытого транзистора. Видно, что значительная часть электронов находится вне пределов эпитаксиального слоя, создавая ток, параллельный току канала.

Здесь надо обратить внимание на то, что условия дрейфового движения электронов в канале и подложке существенно различны. Переходя в подложку в истоковой части за счет хаотической скорости, они образуют контактный потенциальный барьер высотой чуть более 0,1 эВ, определяемой температурой и отношением между концентрациями примеси в эпитаксиальном слое и подложке. Из распределения электронов на рис. 4.4 видно также, что в стоковой части под затвором концентрация электронов в подложке значительно выше. Это происходит за счет разогрева электронов в канале (распределение электронов по долинам см. на рис. 1.3) и повышения их электронной температуры. В результате контактный потенциальный барьер существенно повышается и может достигать значений более 1 эВ. В итоге из распределения потенциального рельефа наглядно видно, что продольное электрическое поле в подложке более слабое, чем в канале, что создает различные условия при протекании тока через области канала и подложки.

Еще одна специфическая особенность приборов с субмикронными размерами связана с релаксационными процессами. Время релаксации некоторых процессов соизмеримо со временем пролета электронов в канале. Соответственно длины релаксации могут составлять сотые и десятые доли микрометров, что делает непригодной локальную модель дрейфа, при которой скорость дрейфа определяется текущим значением напряженности

электрического поля $-v = \mu\mathcal{E}$. В этих условиях необходимо использовать кинетическое уравнение Больцмана, либо релаксационные уравнения сохранения импульса, энергии и концентрации, описанные в первом разделе.

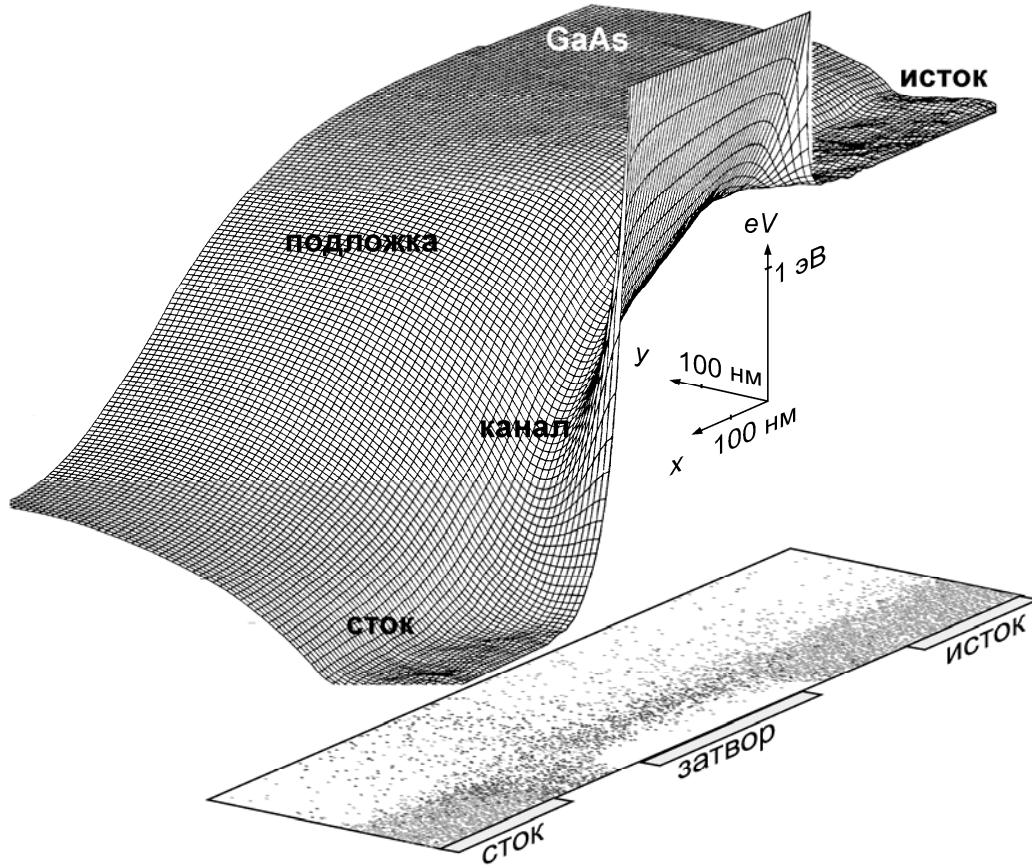


Рис. 4.4. Потенциальный рельеф и распределение электронов в канале открытого ПТШ. Соответствующее распределение электронов по долинам см. на рис. 1.3

4.2. КВАЗИДВУМЕРНАЯ МОДЕЛЬ ПТШ

В рассматриваемой модели используются обычные для моделей такого уровня допущения о резкой границе обедненной области и об отсутствии поперечной составляющей электрического поля в канале. Такие допущения позволяют использовать в качестве исходных одномерные уравнение Пуассона и уравнение непрерывности тока в следующей форме (ось x направлена вдоль канала от истока к стоку):

$$d\mathcal{E} / dx = e[n(x) - N_d] / \epsilon\epsilon_0, \quad (4.5)$$

$$I = en(x)v(x)h(x)W, \quad (4.6)$$

а толщину обедненной области $d(x)$ рассчитывать в зависимости от потенциала канала $V(x)$ относительно истока:

$$d(x) = \{2\varepsilon\varepsilon_0[V_b - V_{gs} + V(x)] / eN_d\}^{1/2}. \quad (4.7)$$

В этих соотношениях e , $n(x)$, $v(x)$ – заряд, концентрация и скорость электронов, $h(x)$ – толщина проводящего канала, W – ширина затвора, N_d – концентрация донорной примеси в эпитаксиальном слое, $\varepsilon\varepsilon_0$ – диэлектрическая проницаемость GaAs, V_{gs} – напряжение смещения на затворе относительно истока, V_b – высота барьера Шоттки, $V(x)$ – переменный потенциал канала, определяющий продольную напряженность электрического поля:

$$\mathcal{E}(x) = dV(x) / dx. \quad (4.8)$$

Нестационарная динамика носителей может быть проанализирована с помощью релаксационных уравнений для импульса $p = m^*v$ и энергии E :

$$dp / dt = e\mathcal{E} - p / \tau_p, \quad (4.9)$$

$$dE / dt = e\mathcal{E}v - (E - E_0) / \tau_E. \quad (4.10)$$

Зависящие от энергии времена релаксации $\tau_p(E)$ и $\tau_E(E)$ могут быть найдены методами, описанными в гл. 2. В ранних работах они определялись из поле-скоростных и поле-температурных характеристик из условия стационарности ($d/dt = 0$): $\tau_p = p_c / e\mathcal{E}_c$, $\tau_E = (E - E_0) / e\mathcal{E}_c v_c$. Здесь индексом «с» обозначены стационарные значения переменных, а E_0 соответствует равновесной энергии электронов. Экспериментальные и теоретические стационарные зависимости $\tau_p(E)$ и $\tau_E(E)$ приведены в работах [4.3, 4.4]. Необходимо отметить, что эти расчеты с различной степенью точности описывают экспериментальные характеристики.

Подставляя в последние уравнения времена релаксации и заменяя в них $dt = dx/v$, получим:

$$d(m^*v) / dx = e(\mathcal{E} / v - \mathcal{E}_c / v_c), \quad (4.11)$$

$$dE / dx = e(\mathcal{E} - \mathcal{E}_c v_c / v), \quad (4.12)$$

которые и являются исходными для анализа нестационарных эффектов в рассматриваемой модели. Для полноты в систему

уравнений (4.5)-(4.12) необходимо включить выражение для расчета толщины проводящего канала.

Основным отличием рассматриваемой модели от описанных в [4.5, 4.6], кроме процедуры расчета, является учет шунтирующего влияния подложки, которое, как показывает анализ моделей, использующих метод частиц (Монте-Карло), может быть весьма существенным. Это объясняется тем, что на границе эпитаксиального слоя с подложкой возникает контактное электрическое поле, вызывающее в ней режим обогащения, что приводит к увеличению эффективной толщины проводящего канала.

Анализируя влияние подложки на вольтамперные характеристики транзистора, необходимо учитывать два фактора.

Первый из них состоит в том, что на границе эпитаксиальный слой-подложка из-за различных концентраций носителей возникает контактная разность потенциалов, которая определяется соотношением концентраций в разных слоях

$$V_k = (kT / e) \ln(N_d / N_{дп}). \quad (4.13)$$

Контактное поле образовано перемещением (диффузией) подвижных носителей из активного слоя в подложку. Переходный слой из-за обычно большого различия в концентрациях примеси располагается в основном в подложке и является обогащенным слоем. Соответствующие графики распределения потенциала, напряженности поперечного электрического поля и концентрации носителей по глубине представлены на рис. 4.5. Этот обогащенный слой непосредственно примыкает к проводящему каналу и увеличивает его эффективную толщину.

Второй фактор, который во многих случаях необходимо учитывать, состоит в том, что более глубокие слои подложки, не находящиеся в области сильного тянущего поля, обладают чисто омической проводимостью из-за конечной концентрации носителей. В результате подложка шунтирует нелинейную проводимость канала и может заметно влиять на вольтамперные характеристики транзистора.

Количественный анализ этих явлений возможен лишь на основе решения двумерного уравнения Пуассона. Однако, приближенные оценки параметров обогащенного слоя можно провести с помощью одномерного анализа. Выбирая направление оси y от поверхности вглубь подложки и связывая распределение потенциала $V(y)$ с распределением равновесной концентрации $n(y)$ электронов в подложке соотношением $n(y) = N_{дп} \exp[eV(y) / kT]$, запишем уравнение Пуассона

$$d^2V / dy^2 = -\rho(y) / \varepsilon\varepsilon_0 = eN_{\text{дп}} / \varepsilon\varepsilon_0 [\exp(eV(y) / kT) - 1]. \quad (4.14)$$

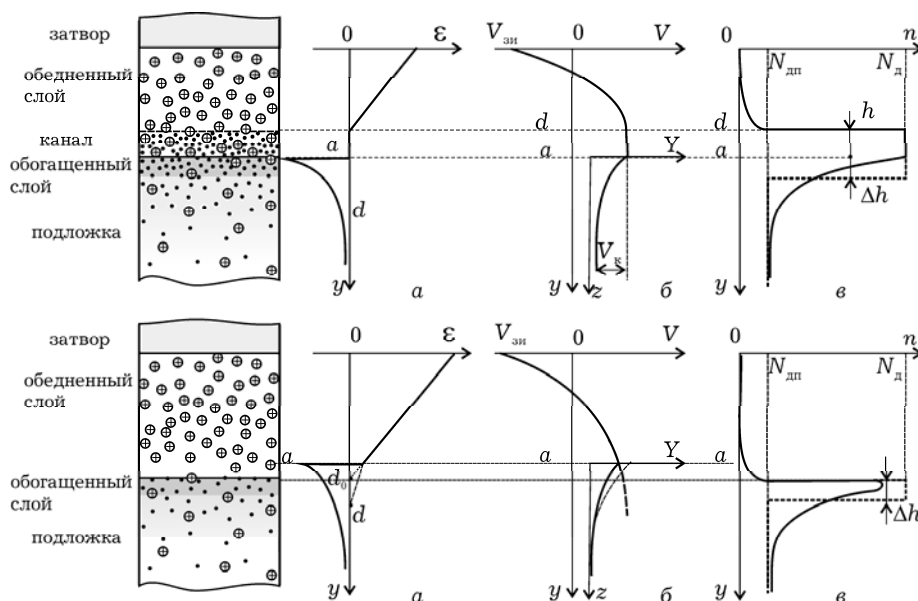


Рис. 4.5. Распределение напряженности поперечного электрического поля (а), потенциала (б) и концентрации электронов (в) при непокрытом канале (верхний ряд) и перекрытом (нижний)

Вводя локальные (рис. 4.5) нормированные координаты $Y=eV/kT$ и $z=y/l_{\text{эп}}$, где $l_{\text{эп}} = \sqrt{\varepsilon\varepsilon_0 kT / e^2 N_{\text{дп}}}$ - дебаева длина экранирования для подложки, получим

$$d^2Y / dz^2 = e^Y - 1. \quad (4.15)$$

Интегрирование этого уравнения при условии, что в глубине подложки $Y \rightarrow 0$ и $dY/dz \rightarrow 0$, приводит к выражению:

$$dY / dz = -\sqrt{2(e^Y - Y - 1)}. \quad (4.16)$$

Выбранный перед корнем знак соответствует направлению электрического поля. Дальнейшее аналитическое интегрирование этого уравнения весьма затруднительно, однако, для нахождения проводимости подложки можно обойтись без знания распределения потенциала и концентрации носителей. Так, например, весь избыточный электронный заряд, сконцентрированный в слое от y_1 до y_2 под единичной площадкой на границе раздела

эпитаксиальный слой-подложка, может быть найден интегрированием:

$$\begin{aligned} \Delta Q_1(\Delta y) &= -\int_{y_1}^{y_2} \rho(y) dy = \varepsilon \varepsilon_0 \int_{y_1}^{y_2} (d^2 V / dy^2) dy = \varepsilon \varepsilon_0 \int_{y_1}^{y_2} d(dV / dy) = \\ &= (\varepsilon \varepsilon_0 k T / e l_{\text{эп}}) (dY / dz \big|_{y_2} - dY / dz \big|_{y_1}). \end{aligned} \quad (4.17)$$

Полный заряд обогащенного слоя Q_1 может быть найден при $y_1 = a$ и при $y_2 \rightarrow \infty$. Так как в глубине подложки $dY / dz \rightarrow 0$, а на поверхности $dY(0) / dz = \sqrt{2(e^{Y_k} - Y_k - 1)}$ и $Y(0) = Y_k = \ln(N_d / N_{\text{дп}})$, то

$$Q_1 = e N_{\text{дп}} l_{\text{эп}} \sqrt{2(e^{Y_k} - Y_k - 1)}. \quad (4.18)$$

Если упрощенно считать, что весь заряд равномерно с концентрацией N_d распределен в некотором слое толщиной Δh , то

$$\Delta h_1 = Q_1 / e N_d = l_{\text{эп}} \sqrt{2(e^{Y_k} - Y_k - 1)} (N_{\text{дп}} / N_d). \quad (4.19)$$

Эти оценки позволяют принять, что эффективная толщина канала увеличивается на величину Δh_1 .

Однако эта оценка пригодна лишь для случая неперекрытого канала, когда $d < a$. Если же поперечное поле затвора, создающее режим обеднения, проникает в подложку, то толщина обедненного слоя в этом случае может быть рассчитана по формуле:

$$d = \sqrt{a^2 + (d_0^2 - a^2) N_d / N_{\text{дп}}}, \quad (4.20)$$

где величина d_0 находится по формуле (4.7) и представляет собой толщину обедненного слоя, которая была бы в однородной среде с концентрацией N_d . Последнее выражение несложно может быть получено при решении одномерного уравнения Пуассона для режима обеднения. Из него же следует и выражение для распределения поперечного поля затвора в подложке:

$$dV / dy = e N_{\text{дп}} (d - y) / \varepsilon \varepsilon_0. \quad (4.21)$$

Так как результирующее поле можно представить в виде суммы поля затвора (4.21) и контактного поля (4.16), то проделав процедуру, подобную предыдущей, можно считать, что весь заряд

в режиме перекрытия канала распределен равномерно с концентрацией N_d в эквивалентном слое

$$\Delta h_2 = \Delta h_1 - (d - a)N_{дп} / N_d. \quad (4.22)$$

Расчет толщины проводящей части канала можно провести с учетом схемы (рис. 4.6), поясняющей определение формы обедненной области под затвором и толщины канала для различных участков активной области транзистора.

Из этой схемы ясно, что в основной части под затвором, когда $d < a$, толщина канала является переменной величиной и равна

$$h(x) = a - d(x) + \Delta h_1. \quad (4.23)$$

Если под стоковой частью затвора возникает перекрытие ($d > a$), то

$$h(x) = \Delta h_2(x). \quad (4.24)$$

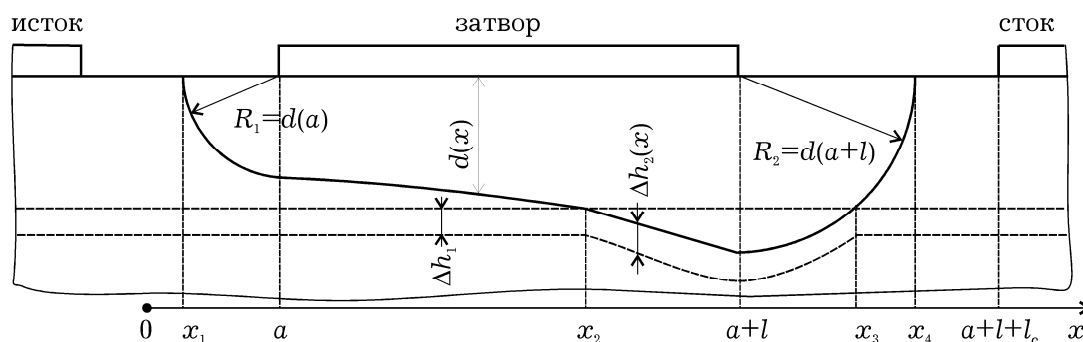


Рис. 4.6. Схема расчета формы обедненной области под затвором и толщины проводящего канала

Здесь надо заметить, что и в этом случае толщина канала — переменна, так как входящая в (4.22) величина d зависит от меняющегося потенциала канала.

Краевые эффекты, возникающие у концов затвора, могут быть учтены аппроксимирующими дугами окружностей для обедненной области. У истокового края затвора радиус дуги равен толщине обедненной области у начала затвора и для его нахождения надо предварительно рассчитать потенциал канала $V(a) = IR_{gs}$. Радиус дуги у стокового конца затвора равен толщине обедненной области у края затвора и в рассматриваемой модели он находится в ходе вычислений. Надо иметь в виду, что в этой области (точка x_3 на рис. 4.6) надо произвести замену Δh_2 на Δh_1 .

За пределами обедненной области ($x < x_1$ и $x > x_4$) канал полностью открыт и

$$h = a + \Delta h_1. \quad (4.25)$$

Следует также заметить, что выбор эквивалентных слоев $\Delta h_{1,2}$, увеличивающих эффективную толщину канала не совсем корректен. Это объясняется тем, что верхняя часть обогащенного слоя находится в области сильного тянущего поля стока и только ее необходимо учитывать при расчете эффективной толщины канала. Нижняя часть обогащенного слоя находится, как правило, в области слабого поля и эти электроны создают практически омическую шунтирующую проводимость подложки. Однако, четкие критерии для такого разделения в рамках рассматриваемой одномерной модели отсутствуют.

Для численного решения системы уравнений (4.5)...(4.12) канал разбивается на участки длиной Δx , и производные вычисляются через конечные приращения. Исходные уравнения в этом случае могут быть записаны для i -ой точки в виде:

$$\begin{aligned} \mathcal{E}_i &= \mathcal{E}_{i-1} - e\Delta x(n_i - N_d) / \varepsilon\varepsilon_0, \\ I &= en_i v_i h_i W, \\ V_i &= V_{i-1} + \mathcal{E}_i \Delta x, \\ m_i^* v_i &= m_{i-1}^* v_{i-1} + e\Delta x(\mathcal{E}_i / v_i - \mathcal{E}_{ci} / v_{ci}), \\ E_i &= E_{i-1} + e\Delta x(\mathcal{E}_i - \mathcal{E}_{ci} v_{ci} / v_i), \\ h_i &= a - d_i + \Delta h_i. \end{aligned}$$

Стационарные значения величин v_{ci} , \mathcal{E}_{ci} и значение эффективной массы $m_i^*(E)$ могут быть найдены по значению энергий E_{i-1} с использованием функций $\mathcal{E}_c(E)$, $v_c(E)$ и $m^*(E)$, рассчитанных по методу Монте-Карло и приведенных, например, в [4.7].

На рис. 4.7 представлены результаты вычисления формы обедненной области под затвором, распределение скорости электронов, потенциала и энергии, концентрации носителей и напряженности электрического поля вдоль канала транзистора при длине затвора $L = 0,5$ мкм, ширине затвора $W = 200$ мкм, толщине эпитаксиального слоя $a = 0,15$ мкм, концентрации доноров в канале $N_d = 10^{17}$ см⁻³ и подложке $N_{дп} = 5 \cdot 10^{14}$ см⁻³, напряжении на затворе $V_{gs} = 0$ В, токе стока 30 мА.

Рис. 4.7 показывает, что скорость не „следует” за изменениями напряженности поля в соответствии со статической поле-скоростной характеристикой. Пространственное проявление эффекта «всплеска» приводит к заметному превышению

мгновенной дрейфовой скорости под затвором над ее стационарными значениями, что уменьшает время пролета электронов в канале и повышает быстродействие прибора. Уменьшение скорости у стокового края затвора связано с интенсивным увеличением эффективной массы электронов, вызванной их разогревом в сильном электрическом поле и переходом в верхнюю долину. Это же приводит к образованию статического домена с существенным перепадом концентрации носителей. Область существования домена выходит из стоковой части канала достаточно далеко в омическую область. Необходимо отметить, что энергия электронов остается заметно выше равновесной и вне области домена. Полное охлаждение электронного потока происходит на расстояниях около 0,3-0,5 мкм. Эта особенность важна как с точки зрения моделирования шумов в канале, так и в связи с выбором расстояния затвор-сток при проектировании прибора.

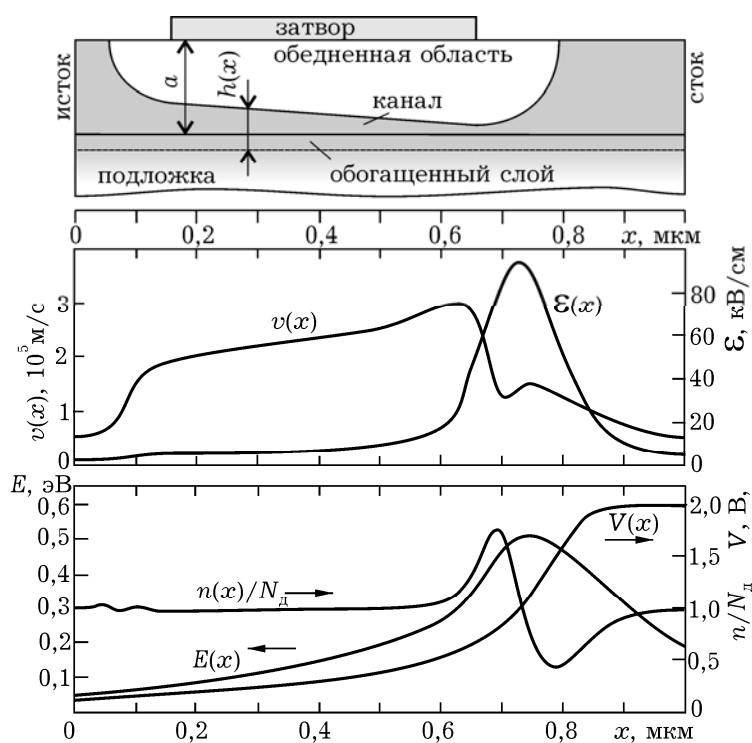


Рис. 4.7. Результаты вычисления формы обедненной области под затвором, распределения скорости электронов, потенциала и энергии, концентрации носителей и напряженности поля вдоль канала транзистора

Данная модель хорошо описывает увеличение выходной проводимости, т. е. наклон стоковой характеристики в транзисторах с длиной затвора около 0,5 мкм, но не адекватно отражает некоторые более тонкие явления, в частности, зависимость этой проводимости от напряжения затвора. Для этого необходимо учитывать, во-первых, разные условия движения электронов в подложке и канале транзистора и, во-вторых, принять во внимание, что приближение плавного канала теряет силу для длин затвора 0,2...0,3 мкм.

Описание нестационарного дрейфа в таких структурах также возможно на основе систем уравнений (4.5)-(4.12), но записанных для канала и подложки. Кроме уравнений непрерывности, Пуассона, уравнений сохранения импульса и энергии, они должны содержать уравнения связи, описывающие процессы перераспределения носителей между каналом и подложкой в широком диапазоне питающих напряжений, включая субпороговый режим [4.8] и режим открытого канала [4.9].

Необходимо принять во внимание, что температура электронов под затвором может существенно отличаться от равновесной вследствие разогрева и междолинных переходов. При этом «медленные» электроны верхних долин уменьшают среднюю дрейфовую скорость коллективного движения, что приводит к процессам накопления носителей в области сильного электрического поля под затвором. В работах [4.10, 4.11] в рамках квазидвумерной физико-топологической модели рассматривается диффузионный механизм инжекции горячих носителей из канала в подложку. На границе эпитаксиального слоя с подложкой имеется потенциальный барьер, обусловленный контактом полупроводников с различной концентрацией донорной примеси. Высота потенциального барьера на границе может изменяться из-за нарушения равновесия и инжекции горячих носителей и их перераспределения между каналом и подложкой. Таким образом, потенциальный барьер в подложке связан как с пространственным зарядом электронов, обогащающих подложку, так и носителями, инжектированными из канала в подложку.

По мере перекрытия канала увеличивается доля носителей, перешедших через потенциальный барьер в подложку и дрейфующих по подложке. Проникновение носителей в подложку связано с торможением в поперечном поле потенциального барьера. Вероятность междолинного рассеяния у этой части носителей уменьшается из-за потери энергии при инжекции через потенциальный барьер в подложку. Это приводит к уменьшению температуры электронного газа и увеличению средней дрейфовой скорости потока носителей. Кроме того, увеличение дрейфовой скорости связано с ослаблением примесного рассеяния из-за низкой концентрации донорной примеси подобно тому, как это происходит в гетероструктурных транзисторах типа *HEMT*. Как показано в работе [4.11] пиковые значения дрейфовой скорости примерно в 1,5...2 раза могут превышать значения скорости в канале, где преобладают эффекты разогрева и накопления носителей. Несмотря на то, что обратный переход электронов из подложки в канал за затвором связан с ростом энергии, когда приобретенная потенциальная энергия превращается в

кинетическую, выполнение условий для перехода электронов в верхнюю долину будет несколько задерживаться из-за конечного времени релаксации. В целом это приводит к ослаблению междолинного рассеяния и разогрева электронного газа.

Выражение, описывающее изменение средней энергии электронов канала и подложки, получено в работах [4.12, 4.13] в виде:

$$dE = (I_K dE_K + I_{\Pi} dE_{\Pi}) / (I_K + I_{\Pi}) = (E_{\Pi} - E_K + eV_K) d(I_{\Pi K} - I_{\Pi\Pi}) / (I_K + I_{\Pi}),$$

где $dI_{\Pi\Pi}$, $dI_{\Pi K}$ – приращения токов на участке dx за счет инжекции носителей из канала в подложку и их возврата в канал. При инжекции электронов из канала в подложку ($dI_{\Pi K} > dI_{\Pi\Pi}$) происходит уменьшение энергии носителей, что приводит к охлаждению электронного потока.

Рассмотренные физические процессы учитываются в физико-топологической модели [4.10, 4.11], включающей следующие уравнения для канала:

$$\begin{aligned} d(\mathcal{E}h) &= (e/\varepsilon\varepsilon_0)h[n(x) - N_D]dx, \\ d(envhW) &= dI_{\Pi\Pi} - dI_{\Pi K}, \\ dE &= e\mathcal{E}dx - [(E - E_0)/v\tau_E]dx + [(E - E_{\Pi} + eV_K)/I_K]dI_{\Pi K}, \\ d(m^*v) &= (e\mathcal{E}/v)dx - (m^*/\tau_p)dx + [(m_{\Pi}^*v_{\Pi} - m^*v)/I_K]dI_{\Pi K}, \\ h(x) &= a - \sqrt{2\varepsilon\varepsilon_0[V_B - V_{gs} + V(x)]/eN_D}, \\ dV &= \mathcal{E}dx, \end{aligned} \quad (4.26)$$

и аналогичные уравнения для подложки:

$$\begin{aligned} d(\mathcal{E}_{\Pi}h_{\Pi}) &= (e/\varepsilon\varepsilon_0)h_{\Pi}[n_{\Pi}(x) - n_{cp}]dx, \\ d(en_{\Pi}v_{\Pi}h_{\Pi}W) &= -dI_{\Pi\Pi} + dI_{\Pi K}, \\ dE_{\Pi} &= e\mathcal{E}_{\Pi}dx - [(E_{\Pi} - E_0)/v_{\Pi}\tau_{E_{\Pi}}]dx + [(E - E_{\Pi} - eV_K)/I_K]dI_{\Pi\Pi}, \\ d(m_{\Pi}^*v_{\Pi}) &= (e\mathcal{E}_{\Pi}/v_{\Pi})dx - (m_{\Pi}^*/\tau_{p\Pi})dx - [(m_{\Pi}^*v_{\Pi} - m^*v)/I_K]dI_{\Pi\Pi}, \\ h_{\Pi}(x) &= a + \sqrt{2}l_{\varepsilon\Pi} - \sqrt{a^2 + 2\varepsilon\varepsilon_0[V_B - V_{gs} + V(x) - V_p]/eN_D}, \\ dV_{\Pi} &= \mathcal{E}_{\Pi}dx, \end{aligned} \quad (4.27)$$

а также уравнения, описывающие токообмен между каналом и подложкой и связывающие две предыдущие системы в единую систему уравнений:

$$dI_{\Pi\Pi} = 0,901eD_{\Pi}W[n_{\Pi} - 0,637\sqrt{N_D N_{D\Pi}}]dx/l_{\varepsilon\Pi}, \quad (4.28,a)$$

$$dI_{\Pi K} = 0,366eD_KW[n(x) - N_D]dx/l_{\varepsilon}. \quad (4.28,b)$$

Здесь \mathcal{E} , \mathcal{E}_Π — напряженность продольного электрического поля в канале и в подложке, h , h_Π — соответственно толщина канала и эффективная толщина подложки, N_d , $N_{дп}$ — уровни легирования эпитаксиального слоя и подложки, n , n_Π — текущие значения концентрации электронов в канале и подложке, v , v_Π — средние значения дрейфовой скорости носителей в канале и в подложке, $dI_{кп}$, $dI_{пк}$ — составляющие полного тока соответственно из канала в подложку и из подложки в канал, E , E_Π — средние значения энергии носителей канала и подложки, E_0 — равновесное значение энергии, τ_E , τ_p — времена релаксации энергии и импульса носителей в канале и, с соответствующим индексом, в подложке, m^* , m_Π^* — средние по долинам значения эффективной массы носителей в канале и в подложке, V_b , V_{gs} — потенциал барьера Шоттки и разность потенциалов затвор-исток, $V(x)$, $V_\Pi(x)$ — текущее значение потенциала канала и подложки (потенциал подложки имеет смысл только при полностью перекрытом канале), $n_{ср}$ — среднее (равновесное) значение концентрации в подложке, V_p — потенциал «отсечки» тока стока, l_ε , $l_{\varepsilon\Pi}$ — длины экранирования в канале и в подложке, D_k , D_Π — коэффициенты диффузии для канала и для подложки, V_k — контактный потенциал (высота потенциального барьера канал-подложка), a — толщина эпитаксиального слоя, W — ширина транзистора, T — температура кристалла. Остальные обозначения ясны из рис. 4.8.

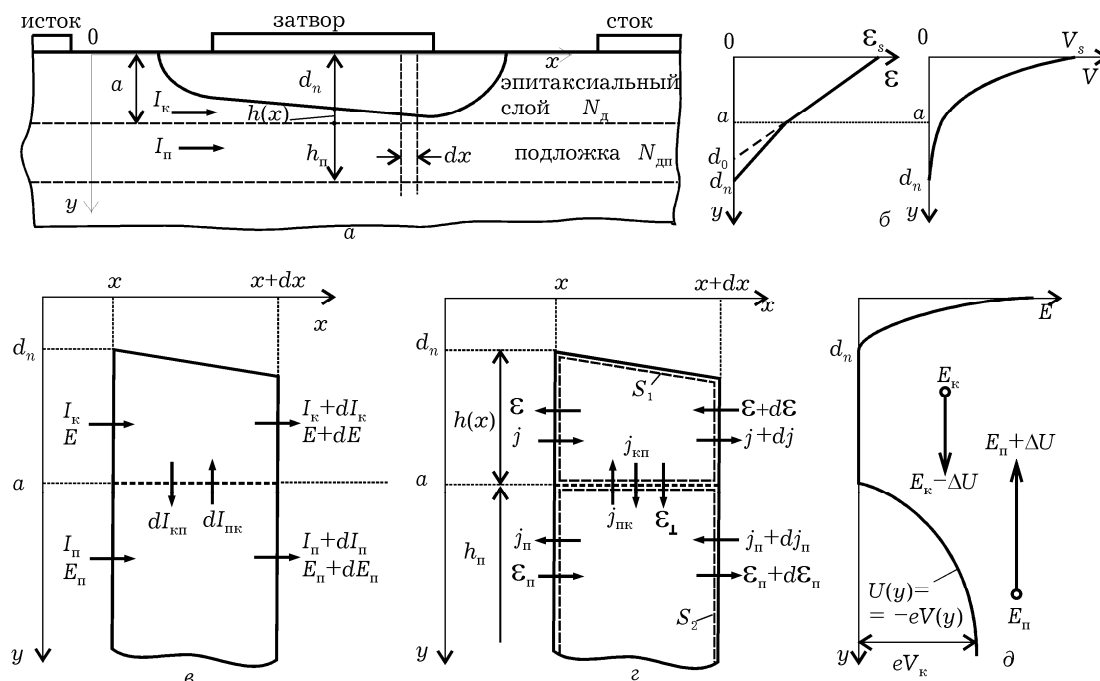


Рис. 4.8. Схемы, поясняющие вывод основных уравнений для канала и подложки (см. приложения П4.1 – П4.4)

Вид уравнений непрерывности, Пуассона, баланса импульса и энергии, входящих в системы (4.26) – (4.28), приведен в приложениях П4.1 – П4.4. Обратим внимание, что эти уравнения содержат производные от толщин канала и подложки, что оправдывает использование приближения плавного канала.

Для численного решения системы уравнений (4.26) – (4.28) активная область транзистора разбивается на участки длиной Δx и дифференциалы заменяются конечными приращениями. Системы для канала и подложки с учетом уравнений связи (4.28) решаются совместно на каждом шаге интегрирования. Область интегрирования начинается на расстоянии от истокового конца затвора, равном толщине эпитаксиального слоя, где эффекты разогрева электронов практически отсутствуют. Оставшаяся область до истока принимается омической. В ней полный ток стока делится на ток канала и ток подложки в соответствии с их проводимостями с учетом обогащенного слоя в подложке. Эти значения токов определяют (по закону Ома) напряженность электрического поля и соответственно слабополевым подвижностям – дрейфовую скорость и энергию электронов. Начальное значение концентрации выбирается равным N_d , а для подложки – $n_{cp} = (N_d N_{дп})^{1/2}$. Из-за отсутствия эффектов всплеска стационарные значения скорости и напряженности поля, необходимые для нахождения времен релаксации, выбираются равными текущим значениям. Начальное значение потенциала в канале равно падению напряжения на омическом сопротивлении истока.

При перекрытии канала интегрируется только система уравнений для подложки. На выходе из подложки и расширении канала у стокового конца затвора для обеспечения непрерывности искомых функций используются граничные условия, при которых переменным канала присваиваются значения переменных подложки. В дальнейшем процессы релаксации описываются полной системой уравнений модели.

Процедура интегрирования продолжается до тех пор, пока переменные величины не достигнут равновесных значений. Входящие в уравнения связи (4.28) коэффициенты диффузии в канале и подложке в общем случае зависят от напряженности электрического поля и могут описываться некоторой аппроксимирующей функцией в зависимости от напряженности поля, либо по соотношению Эйнштейна, пропорционально подвижностям и электронным температурам в канале и подложке (разд. 3.5). Важной особенностью квазидвумерной модели является возможность моделирования текущих значений основных

параметров переноса электронов, как в канале, так и в подложке, что дает возможность находить схемные параметры полевых транзисторов на основе анализа физических процессов. Так, например, по мере интегрирования системы уравнений модели рассчитывается температура электронного газа в канале и подложке, а также усредненная температура. Эти величины используются для расчета меры разогрева электронов в шумовой модели для описания разогревного шума. Кроме того, в процессе интегрирования рассчитываются приращения потенциалов и зарядов под затвором для малосигнального анализа схемных параметров, а также время пролета электронов в активной области транзистора, необходимое для расчета высокочастотной крутизны транзистора.

Результаты моделирования распределений характерных величин вдоль канала представлены на рис. 4.9 для транзисторов с длиной затвора 0,25 мкм (нижний ряд) и 0,5 мкм (верхний ряд) при двух значениях напряжения смещения 0 В (1) и -1 В (2). Толщина эпитаксиального слоя моделируемых транзисторов 0,12 мкм, концентрация примеси $2 \times 10^{23} \text{ м}^{-3}$, ширина транзисторов 300 мкм. Точками обозначены результаты моделирования с помощью кинетической модели, непрерывными линиями – квазидвумерной.

В основе *кинетической модели* лежит метод частиц, в котором рассчитывается движение модельных частиц в самосогласованном электрическом поле, с учетом механизмов рассеяния и граничных условий [1.6, 4.2]. При моделировании использовалась трехдолинная Γ -L-X модель зонной структуры GaAs с учетом непараболичности. Время свободного пробега каждой частицы, вид рассеяния и новое состояние волнового вектора после рассеяния определялось методом Монте-Карло. Параметры материала, необходимые для учета механизмов рассеяния, принимались как, например, в работе [4.3]. При этом учитывались процессы рассеяния на акустических, полярных оптических, междолинных фононах и на ионизированных примесях.

Двумерное уравнение Пуассона решалось методом циклической редукции с использованием метода матрицы емкости для учета электродов. Граничные условия для уравнения Пуассона $-\varphi = \text{const}$ на электродах и $d\varphi/dn = 0$ на остальной границе полупроводника. При моделировании транзистора использовалась сетка размерностью 180×65 с ячейкой 10×5 нм. Шаг времени пересчета поля был принят 25 фс. Количество моделируемых частиц выбиралось до 10000. Поскольку для каждой частицы на каждом временном шаге известны ее координаты x , y и компоненты волнового вектора k_x , k_y , k_z то,

производя статистическую обработку, можно получить интересные средние зависимости $n_{\text{cp}}(x)$, $v_{\text{cp}}(x)$, $\mathcal{E}(x)$ и т.д.

Средняя дрейфовая скорость определялась, например, следующим образом. Весь прибор вдоль координаты x разбивался на M интервалов. На каждом временном шаге подсчитывалось количество частиц, попавших в i -ый интервал, N_i и определялась их средняя дрейфовая скорость в данном интервале $v_{\text{ср}xi}$ по формуле

$$v_{\text{ср}xi} = (\hbar / N_i) \sum_{j=1}^{N_i} k_{xj} / (m_j^* (1 + 2\alpha_j E_j)),$$

где k_{xj} – x -я составляющая волнового вектора, m_j^* – эффективная масса, α_j – коэффициент непараболичности, E_j – энергия j -й частицы.

Для уменьшения статистических флуктуаций проводилось дополнительное усреднение по временному интервалу 2...4 пс. Средний ток подложки $I_{\text{п}}(x)$ определялся по формуле $I_{\text{п}}(x) = qv_{\text{п}x}(x)N_{\text{п}x}(x)/\Delta x$, где q – заряд частицы; $v_{\text{п}x}(x)$ – средняя дрейфовая скорость в подложке; $N_{\text{п}x}(x)$ – среднее число частиц в подложке в интервале Δx , в котором производится усреднение.

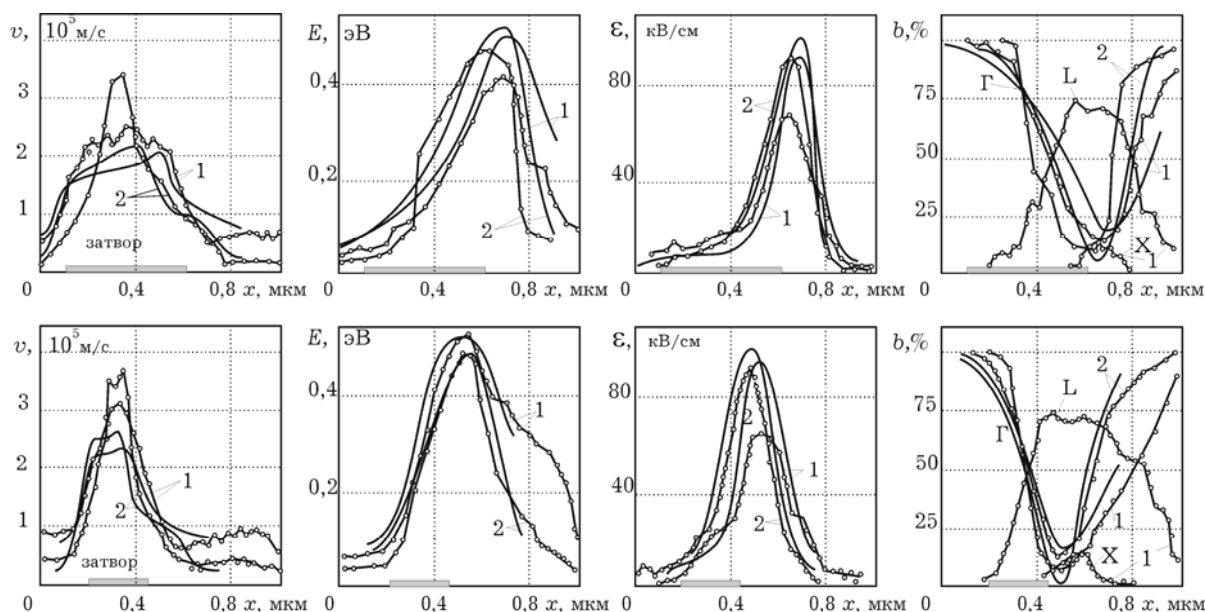


Рис. 4.9. Распределения дрейфовой скорости, энергии, напряженности электрического поля, заселенности долин для потенциалов на затворе 0 В (кривые 1) и –1,0 В (кривые 2). Сплошными линиями и линиями с кружками обозначены соответственно результаты расчетов по квазидвумерной и кинетической моделям. Верхний ряд – для длины затвора 0,5 мкм, нижний – 0,25 мкм

Результаты моделирования наглядно иллюстрируют эффект всплеска дрейфовой скорости выше значения $\sim 1,8 \cdot 10^5$ м/с, которое характерно для стационарного дрейфа в однородном поле в GaAs. Этот всплеск более сильный для транзистора с затвором 0,25 мкм, так как в нем электрическое поле достигает характерных значений (около 20 кВ/см) на меньших расстояниях, пока разогрев электронов еще незначительный. Пик дрейфовой скорости наблюдается в обоих транзисторах, когда приблизительно 30% электронов переходит в верхние долины. Кинетическая модель дает также отрицательный всплеск, наблюдающийся при спаде электрического поля. Пиковое значение скорости заметно меньше в квазидвумерной модели, что объясняется более медленным нарастанием электрического поля на начальном участке подзатворной области и менее интенсивным переходом электронов в верхние долины. Это расхождение может быть уменьшено коррекцией при расчете времен релаксации для различных механизмов рассеяния в квазидвумерной модели.

Надо отметить также, что более сильный всплеск наблюдается при увеличении отрицательного напряжения на затворе, так как при этом в канале существует более сильное электрическое поле. Кроме того, увеличение скорости связано и с вытеснением тока в подложку, где подвижность электронов выше из-за существенного ослабления примесного рассеяния, а также с влиянием потенциального барьера на границе эпитаксиальный слой – подложка.

Из сравнения заметно характерное для субмикронных транзисторов несовпадение максимумов дрейфовой скорости и средней энергии электронов. Энергия продолжает расти даже при уменьшении скорости, что связано с интенсивным переходом электронов в верхние долины и возрастанием их эффективных масс. Процессы релаксации наблюдаются при уменьшении электрического поля, пик которого находится левее стокового конца затвора и совпадает с экстремумами функций $E(x)$, $b(x)$. Характерно, что обе модели, несмотря на различные значения электрического поля, дают близкие значения максимальной энергии, не превышающие $\sim 0,5$ эВ.

Отчетливо видно, что процесс охлаждения при открытом канале происходит значительно дольше, чем при закрытом. Однако это не значит, что время релаксации в этих режимах существенно различно. Общее количество электронов, составляющих ток канала при смещении -1 В, приблизительно втрое меньше, чем в открытом канале. Поэтому естественно, что усреднение энергии и количества горячих электронов при их попадании в омическую область за затвором, где электроны

преимущественно холодные, будет приводить к кажущемуся тем более быстрому их охлаждению, чем меньше ток канала.

Наиболее заметное различие между используемыми моделями наблюдается при моделировании тока, вытесняемого из канала в подложку (рис. 4.10). Видно, что максимальные значения тока подложки приблизительно совпадают, а при смещении -1 В канал полностью перекрывается и весь ток вытесняется в подложку. Однако заметно различие в функциях распределения тока подложки вдоль затвора, что говорит о необходимости более точного моделирования процессов токораспределения в квазидвумерной модели. Этого можно достичь, учитывая изменяющуюся вдоль канала контактную разность потенциалов канал-подложка, принимая различные коэффициенты диффузии для материалов подложки и эпитаксиального слоя.

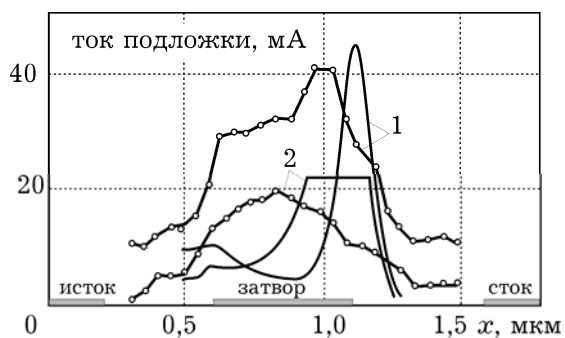


Рис. 4.10. Распределения тока подложки в кинетической (точки) и квазидвумерной (сплошная) модели при разных смещениях на затворе

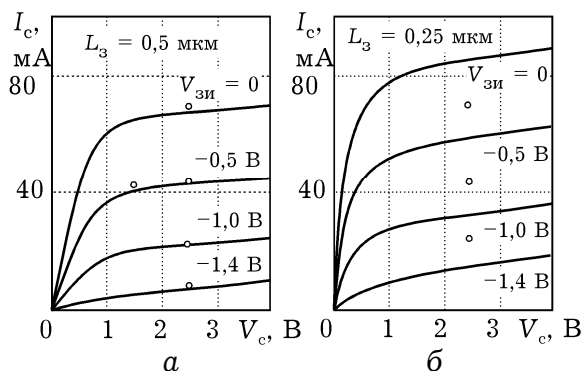


Рис. 4.11. Вольтамперные характеристики ПТШ с длиной затвора $0,5$ мкм (а) и $0,25$ мкм (б)

Несмотря на указанное различие результаты моделирования вольтамперных характеристик (рис. 4.11.) для $0,5$ мкм и для $0,25$ мкм можно признать удовлетворительными. Квазидвумерная модель для транзистора с длиной затвора $0,25$ мкм дает заниженные значения тока стока и соответственно крутизны сток-затворной характеристики. Это связано не только с расхождением значений дрейфовой скорости, но и, по-видимому, с тем, что для данных размеров активной части канала ($0,25 \times 0,12$ мкм) приближения квазидвумерной модели о форме обедненной области становятся уже достаточно грубыми. Тем не менее эта модель правильно учитывает увеличение наклона ВАХ при укорочении затвора, режим запираения транзистора.

Несмотря на некоторое различие результатов кинетического и квазидвумерного моделирования (рис. 4.10), распределения тока подложки в целом подтверждают правильность физических предпосылок квазидвумерной модели о роли подложки [4.10, 4.11, 4.14-4.16].

Это же подтверждает и хорошее совпадение расчетных и экспериментальных статических вольтамперных характеристик, сток-затворной и стоковой, приведенных на рис. 4.12, для транзистора со следующими физико-топологическими параметрами: длина затвора 0,5 мкм, ширина 300 мкм, толщина эпитаксиального слоя 0,12 мкм, концентрация примеси в эпитаксиальном слое $1,9 \cdot 10^{17} \text{ см}^{-3}$, в подложке $1 \cdot 10^{14} \text{ см}^{-3}$, межэлектродные расстояния исток-затвор и затвор-сток по 0,6 мкм, высота металлизации полосы затвора 0,6 мкм. К особенностям статических вольтамперных характеристик можно отнести отсутствие участков насыщения тока стока, а также близкую к омической зависимость тока стока при потенциале затвора ниже $-1,4 \text{ В}$, который с учетом барьерной разности потенциалов соответствует потенциалу перекрытия канала.

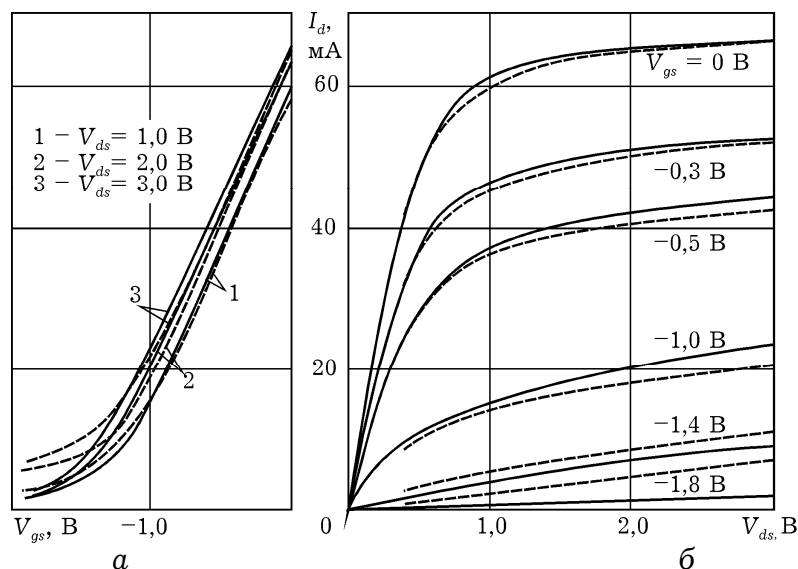


Рис. 4.12. Сток-затворные (а) и стоковые (б) характеристики ПТШ: сплошные линии – экспериментальные, штриховые – расчетные

Заметное отличие характеристик наблюдается в области напряжений затвора, близких к потенциалу перекрытия ($\cong 1,4 \text{ В}$). Особенности моделирования этого режима будут подробно рассмотрены в разд 4.4.

4.3. СХЕМНЫЕ МОДЕЛИ ПТШ

Схемная модель позволяет, при известных ее параметрах, рассчитать выходные сигнальные и шумовые характеристики. При составлении схемных моделей активных твердотельных компонентов надо иметь ввиду, что схемная модель с сосредоточенными параметрами не является однозначной, так как

твердотельные компоненты по своей природе являются распределенными структурами.

Чаще всего используются схемные модели с сосредоточенными параметрами, которые определяются по физико-топологической модели, содержащей уравнения Пуассона, уравнение непрерывности тока, уравнения движения, уравнения сохранения импульса и энергии.

В режиме малого сигнала по физико-топологической модели рассчитываются приращения токов, напряжений и зарядов для расчета параметров схемной модели.

4.3.1. Малосигнальная модель

Простейшая идеальная схемная модель транзистора, в котором ток стока I_d управляется напряжениями затвора V_{gs} и стока V_{ds} , на низких частотах может быть представлена цепью, содержащей идеальный зависимый источник тока SV_{gs} , управляемый входным напряжением, и параллельно соединенной выходной проводимостью G_d . Здесь S – крутизна сток-затворной вольтамперной характеристики ПТШ.

Типичный вид схемной модели транзистора представлен на рис. 4.13. Кроме источника тока, управляемого напряжением, и выходной проводимости она включает компоненты, моделирующие некоторые характерные области транзистора. Это, прежде всего, компоненты активной части, которые моделируются в рамках квазидвумерной модели: емкость затвор-исток C_{gs} , емкость сток-затвор C_{gd} , сопротивление канала R_i , емкость C_f и сопротивление R_f домена в стоковой части канала, сопротивления омических областей эпитаксиального слоя у истока и стока, которые дополняют переходные сопротивления омических контактов до полных сопротивлений R_s и R_d .

Малосигнальные параметры схемной модели определяются по задаваемым малым приращениям напряжения затвора и тока стока. Методы их нахождения описаны в приложении П4.

В области крайне высоких частот в схемной модели субмикронного ПТШ необходимо учитывать также паразитные параметры корпуса и выводов. Как правило, синтез и параметрическая идентификация паразитных цепей связана с измерениями и применением процедур оптимизации, однако в первом приближении их можно рассчитать, как указано в приложении П4.5.

Из перечисленных выше параметров наибольший вес, как с точки зрения отражения специфики субмикронных транзисторов,

так и оптимизации характеристик, имеют крутизна S , выходная проводимость G_d , входная (C_{gs}) и проходная (C_{gd}) емкости.

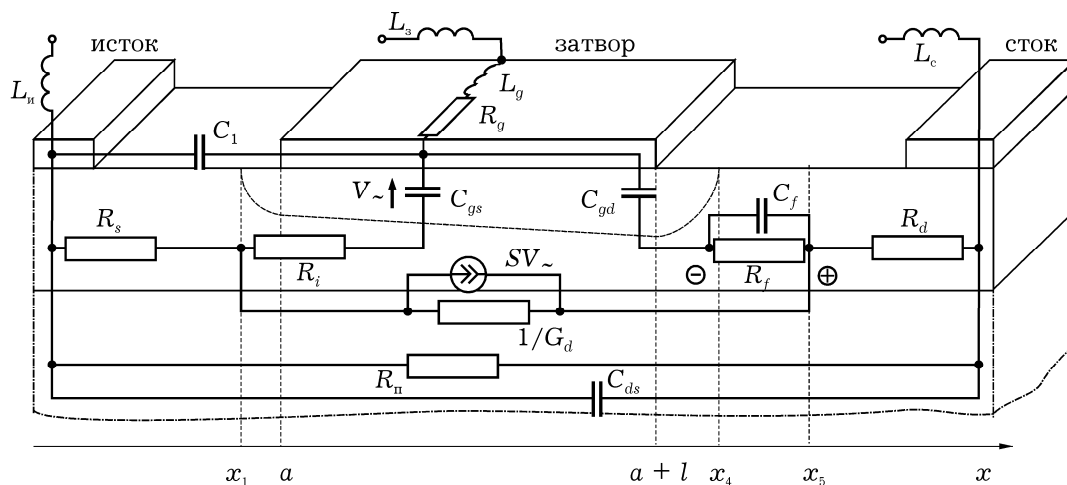


Рис. 4.13. Схемное представление областей субмикронного ПТШ на сверхвысоких частотах

Наибольший интерес представляет анализ зависимостей выходной проводимости и крутизны от электрического режима (рис. 4.14). Уменьшение выходной проводимости при увеличении потенциала стока связано, прежде всего, с расширением обедненной области и уменьшением толщины канала и, соответственно, разогревом электронов в сильном поле, их переходом в верхние долины и увеличением эффективной массы. Эти эффекты ослабевают по мере вытеснения электронов в подложку и некоторого их охлаждения за счет преодоления потенциального барьера на границе эпитаксиального слоя с подложкой. Преимущественный дрейф в подложке при потенциалах затвора, близких к потенциалу отсечки, способствует росту скорости, а значит, приводит к увеличению выходной проводимости. Этому потенциалу соответствует максимальный наклон стоковой характеристики и максимальная выходная проводимость.

Зависимости входной и проходной емкостей (рис. 4.15) от электрического режима на электродах достаточно монотонны при потенциалах стока выше 1 В. Различие между расчетными и экспериментальными значениями входной емкости на некоторую постоянную величину, равную приблизительно 0,2 пФ, можно объяснить наличием некоторой паразитной емкости между истоком и затвором, не связанной с изменением заряда обедненной области и поэтому не зависящей от потенциала на затворе. К особенностям зависимости проходной емкости можно

относительно ее некоторый рост при запираании транзистора, заметный на экспериментальных характеристиках.

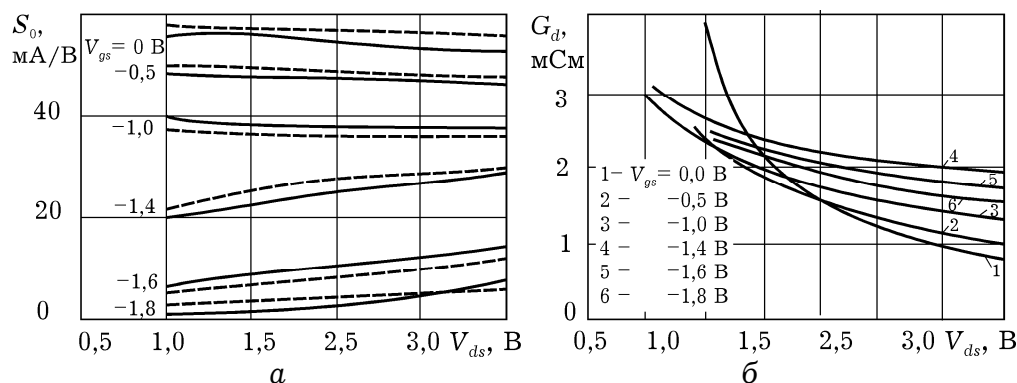


Рис. 4.14. Зависимости крутизны (а) и выходной проводимости (б) от напряжения стока

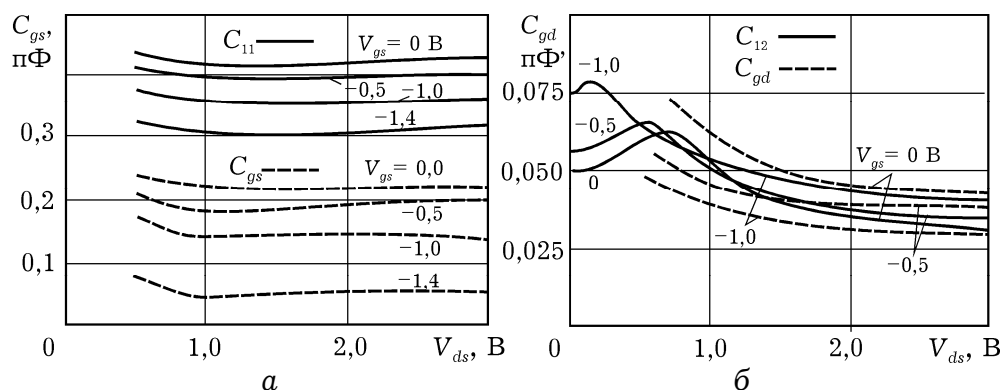


Рис. 4.15. Сравнение экспериментальных (сплошные линии) зависимостей для входной емкости и рассчитанных (штриховые) для емкости C_{gs} (а), а также проходной емкости и емкости C_{gd} (б) от напряжения стока

Представляет интерес анализ влияния геометрических размеров структуры транзистора, в частности, длины затвора на параметры схемной модели транзистора в рабочем электрическом режиме. С укорочением длины затвора проявляется ряд эффектов, присущих субмикронным структурам, влияющих на схемные параметры, особенно на передаточные характеристики. Кроме того, в связи с широким диапазоном изменения потенциалов на электродах требует адекватного моделирования субпороговый режим, когда значения тока стока на несколько порядков меньше, чем ток насыщения открытого транзистора.

На рис. 4.16 представлены параметры схемной модели: крутизны S_o , выходной проводимости G_d , входной емкости затвор-исток C_{gs} и проходной емкости затвор-сток C_{gd} в зависимости от потенциала затвора V_{gs} . Уменьшение длины затвора приводит к

улучшению параметров транзистора – увеличению крутизны, уменьшению входной и проходной емкостей практически до напряжений отсечки. Наблюдаемый при укорочении затвора рост крутизны и выходной проводимости транзистора связан с влиянием уже описанных короткоканальных эффектов. Необходимо отметить появление экстремумов на зависимости выходной проводимости от потенциала на затворе при укорочении затвора, в основном связанных с токораспределением между каналом и подложкой и эффектом «всплеска» дрейфовой скорости. При потенциалах, близких к потенциалу «отсечки» входная емкость практически не зависит от длины затвора и определяется краевой емкостью истокового конца затвора.

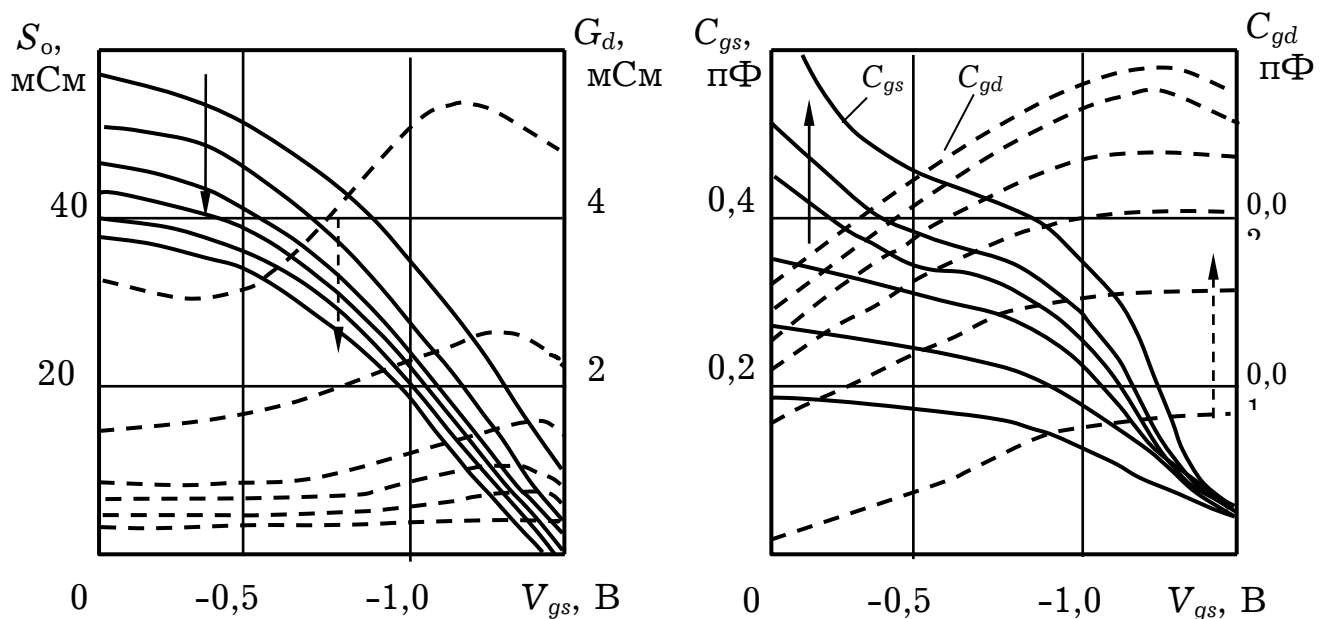


Рис. 4.16. Зависимости малосигнальных крутизны, выходной проводимости, входной и проходной емкостей от напряжения на затворе. Направление стрелок соответствует увеличению длины затвора от 0,25 мкм до 1,5 мкм с шагом 0,25 мкм.

Рост проходной емкости при увеличении длины затвора может быть связан с усиливающимися эффектами накопления носителей у стокового конца затвора и образованием статического домена.

Приведенные зависимости получены для структуры с характерными физико-топологическими параметрами: толщиной эпитаксиального слоя 0,12 мкм, концентрацией доноров в этом слое $1,9 \cdot 10^{17} \text{ см}^{-3}$, межэлектродными расстояниями 0,8 мкм.

Переход от физико-топологической модели к малосигнальной схемной модели позволяет, используя хорошо отработанный аппарат теории цепей с сосредоточенными параметрами [4.18],

провести расчет Y -матрицы транзистора, которая затем по известному соотношению пересчитывается в матрицу рассеяния:

$$[S] = ([1] - [Y])([1] + [Y])^{-1}.$$

Типичные частотные зависимости S -параметров транзистора приведены на рис. 4.17. С использованием аппарата СВЧ цепей они позволяют рассчитывать входные устройства, где обычно используются ПТШ. Впрочем, квазидвумерная модель позволяет, анализируя значения малосигнальных параметров в относительно широком диапазоне питающих напряжений, использовать ее для расчетов нелинейных устройств.

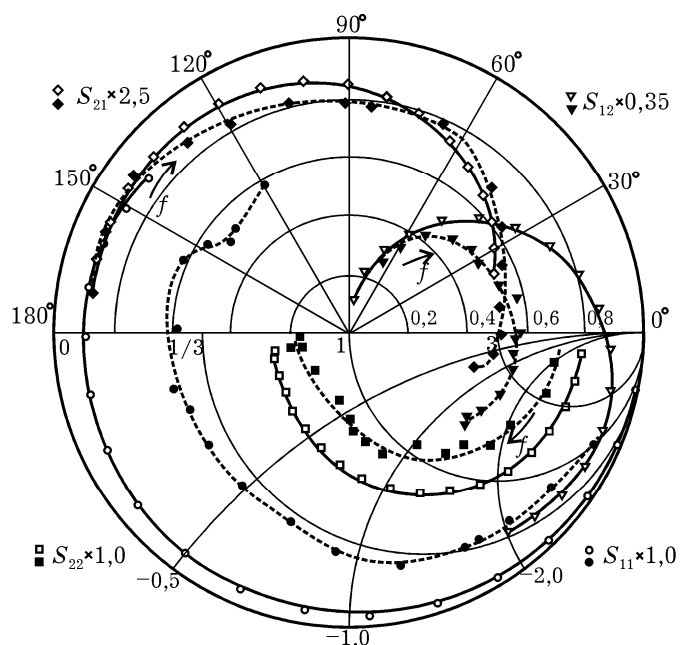


Рис. 4.17. Представление частотных характеристик элементов матрицы рассеяния в полярных координатах (вверху) и на сетке полных проводимостей (внизу). Частотные метки приведены, начиная с частоты 1 ГГц, в направлении стрелок с шагом 1 ГГц. Штриховые линии – экспериментальные. Следует обратить внимание на масштабные множители

4.3.2. Шумовая модель

Известно множество моделей ПТШ (аналитических, квазидвумерных, двумерных, кинетических), отличающихся размерностью решаемых уравнений, уровнем строгости и объемом вычислительных затрат. Однако, в этих работах отсутствует подробный анализ составляющих шума и не дается физическая интерпретация малозумящего режима работы ПТШ.

Один из способов схемотехнического анализа минимального коэффициента шума в субмикронных ПТШ основывается на применении квазидвумерной физико-топологической модели [4.17], позволяющей рассчитывать параметры схемной модели (рис. 4.13) и шумовой модели с различного рода локальными источниками шума (рис. 4.18).

В термодинамическом равновесии спектральная плотность шума описывается соотношением Найквиста. При этом справедливо соотношение Эйнштейна $D = \mu k_B T / e$, которое, строго говоря, выполняется, когда электроны имеют распределение Максвелла.

В сильных электрических полях проявляются эффекты разогрева и значение приобретает так называемый диффузионный шум, т.е. шум, обусловленный рассеянием носителей [4.19].

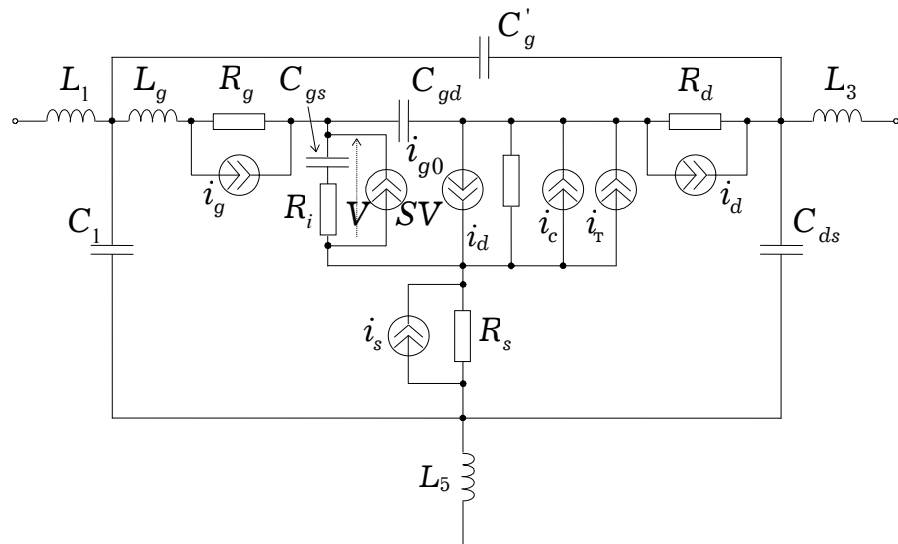


Рис. 4.18. Шумовая схемная модель субмикронного ПТШ.

Спектральная плотность теплового шума, связанная с флуктуациями скорости, описывается соотношением:

$$S_i(\omega) = 4k_B T_e n(x) \mu \Delta y \Delta z / \Delta x = 4kT_B / \Delta R,$$

где $\Delta R = \Delta x / e \mu n(x) \Delta y \Delta z$ – сопротивление ячейки $\Delta x \Delta y \Delta z$.

Наиболее характерными для СВЧ диапазона и субмикронных полевых структур являются следующие источники шума.

Шум канала полевого транзистора [4.20] определяется соотношением $\overline{|i_c|^2} = P 4k_B T_e \Delta f S_0$, где P – коэффициент, учитывающий нетепловую природу шума канала (в сильных электрических полях увеличивается вклад шума канала за счет уменьшения подвижности электронов), T_e – температура электронного газа.

Шумы областей, находящихся в термодинамическом равновесии – омических контактов стока и истока, металлизации затвора – описываются соотношением Найквиста:

$$\overline{|i_s|^2} = 4k_B T_e \Delta f / R_s, \quad \overline{|i_g|^2} = 4k_B T_e \Delta f / R_g, \quad \overline{|i_d|^2} = 4k_B T_e \Delta f / R_d.$$

Шум, наведенный на затворе, связан с тем, что на высоких частотах флуктуации тока приводят за счет емкостной связи с затвором к флуктуациям наведенного тока затвора (частотно-зависимый шум):

$$\overline{|i_{g0}|^2} = 4k_B T_e \Delta f R \omega^2 C_{gs}^2 / S_0,$$

где R – коэффициент, близкий к единице, и зависящий от потенциала на электродах транзистора. Этот источник коррелирован с источником, описывающим шум канала $\overline{|i_c|^2}$. Степень корреляции определяется по формуле:

$$C = \overline{j i_{g0} \cdot i_c^*} / \sqrt{i_{g0}^2 \cdot i_c^2},$$

откуда $\overline{i_c \cdot i_{g0}^*} = -jC \sqrt{\overline{|i_c|^2} \overline{|i_{g0}|^2}} = -j4k_B T_e \Delta f \cdot \omega C_{gs} \cdot C \sqrt{P \cdot R}$.

В субмикронных транзисторах необходимо учитывать шум токораспределения между каналом и подложкой. Он связан со случайным характером распределения тока между двумя цепями. По аналогии с вакуумным тетродом можно записать:

$$\overline{|i_T|^2} = 2e(I_K I_{II} / I_C) \Delta f,$$

где I_d , I_K , I_{II} – ток стока, составляющие тока канала и подложки. Как видно из последнего соотношения, максимальный шум токораспределения имеет место при $I_K \approx I_{II}$.

Из анализа составляющих шума субмикронного ПТШ можно сделать следующие выводы.

Малозумящий режим связан с напряжением на затворе, близком к потенциалу отсечки тока стока. В этом режиме носители под затвором движутся по подложке, преодолев потенциальный барьер канал-подложка и потеряв часть энергии. Это приводит к уменьшению вероятности междолинного рассеяния и уменьшению шума, связанного с разогревом носителей, который, как будет показано, является преобладающим в субмикронных ПТШ. Шум токораспределения при напряжениях на затворе, близких к напряжению отсечки $|V_{gs}| \approx |V_p|$,

минимален, так как полный ток через транзистор близок или равен по величине току подложки. Вклад этой составляющей шума в рассматриваемом режиме минимален.

Шум, наведенный на затворе, в малошумящем режиме, также минимален, так как входная емкость C_{gs} при увеличении напряжения на затворе $|V_{gs}|$ уменьшается.

Шум канала близок к минимальному, так как зависимость малосигнальной крутизны $S_0 = f(V_{gs})$ обратно пропорциональна и зависимость температуры шума от напряжения на затворе имеет минимум.

На основе приведенной методики и анализа составляющих шума для типичной транзисторной структуры с размерами затвора $0,5 \times 300$ мкм в оптимальном электрическом режиме получены следующие частотные зависимости шума (рис. 4.19)

Минимальный коэффициент шума обеспечивается как выбором электрического режима, так и условиями согласования [4.19]. Методика расчета минимального коэффициента шума может быть сведена к следующим этапам.

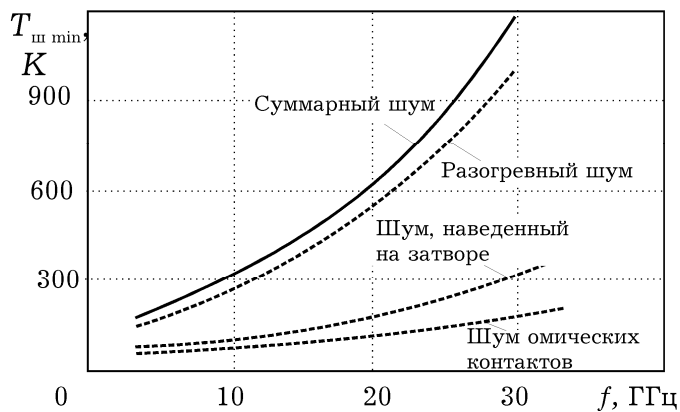


Рис. 4.19. Составляющие шумовой температуры субмикронного ПТШ

На первом этапе используя принцип суперпозиции для линейных цепей: локальные шумовые источники шума пересчитываются ко входу и выходу ПТШ, представленного в виде четырехполюсника, при этом суммарные токи на входе и на выходе:

$$I_{ш \text{ ВХ}} = i'_g + i'_{g0} + i'_T + i'_s + i'_d,$$

$$I_{ш \text{ ВЫХ}} = i''_g + i''_{g0} + i''_T + i''_s + i''_d.$$

На втором этапе рассчитываются волновые шумовые τ -параметры.

$$\tau_{11} = \frac{I_{ш \text{ ВХ}} \cdot I_{ш \text{ ВХ}}^*}{4k_B T \Delta f R_0 Y_\Gamma^2}, \quad \tau_{22} = \frac{I_{ш \text{ ВЫХ}} \cdot I_{ш \text{ ВЫХ}}^*}{4k_B T \Delta f R_0 Y_H^2},$$

$$\tau_{12} = \frac{I_{ш \text{ ВХ}} \cdot I_{ш \text{ ВЫХ}}^*}{4k_B T \Delta f R_0 Y_\Gamma Y_H}, \quad \tau_{12} = \tau_{21}^*,$$

где R_r – стандартное нормировочное сопротивление (волновое сопротивление линий), Y_r и Y_n – проводимости генератора и нагрузки, соответственно.

В предположении большого усиления ($S_{12} \ll S_{21}$), когда шумами нагрузки можно пренебречь, минимальный коэффициент шума запишется [4.20]:

$$F_{\min} = 1 + \frac{\tau_{22} - b + \sqrt{(\tau_{22} + b)^2 - 4|c|^2}}{2|S_{21}|^2},$$

где $b = \tau_{11} \cdot |S_{21}|^2 - 2 \operatorname{Re}(\tau_{12} \cdot S_{21} \cdot S_{11}^*)$, $c = \tau_{22} \cdot S_{11} - \tau_{12} \cdot S_{21}$,

Оптимальный коэффициент отражения по входу с учетом согласования входа четырехполюсника определяется:

$$\Gamma_{1\text{opt}} = \frac{2c^*}{\tau_{22} + b + \sqrt{(\tau_{22} + b)^2 - 4|c|^2}}.$$

При допущении отсутствия влияния нагрузки на шумы на входе, что практически всегда выполняется, и при условии согласования на выходе четырехполюсника на максимальную передачу мощности (комплексно-сопряженное согласование) для коэффициента отражения на выходе можно записать:

$$\Gamma_2 = \left[S_{22} + \frac{\Gamma_1 \cdot S_{12} \cdot S_{21}}{1 - \Gamma_1 \cdot S_{11}} \right]^*.$$

Далее по известным частотным характеристикам шума и матрице рассеяния ПТШ рассчитываются частотные зависимости «сопутствующего» коэффициента усиления по мощности, оптимальных нагрузок на входе и на выходе, коэффициент устойчивости и другие характеристики.

На основе рассмотренных источников шума из схемной шумовой модели можно рассчитать мощность шума $\Delta P_{\text{ш вых}}$, создаваемого ими на выходе транзистора, и найти эквивалентную шумовую температуру $T_{\text{ш}} = \Delta P_{\text{ш}} / K_p k_B \Delta f$. Некоторые результаты моделирования шумовых свойств транзистора представлены на рис. 4.20.

Из этих результатов видна особенность шумовых свойств субмикронных полевых транзисторов – преимущественное влияние разогривных эффектов, которые могут ослабляться при

наличии полуизолирующей подложки. Это становится очевидным при совместном анализе распределений дрейфовой скорости, составляющей тока через подложку (рис. 4.10), а также шумовой температуры (рис. 4.20). Максимальное значение дрейфовой скорости в почти закрытом транзисторе ($V_{gs} = -1,0$ В) примерно в 1,5 раза превышает значение скорости в открытом ($V_{gs} = -1,0$ В). В открытом канале максимальный ток подложки у стокового края затвора составляет около 70% от полного тока, тогда как при $V_{gs} = -1,0$ В – около 90% на всем протяжении затвора. Это приводит к заметному проявлению эффекта «всплеска» скорости и ослаблению разогревных эффектов. Сравнение распределений скорости и тока подложки с полученными при моделировании методом частиц (рис. 4.10), а также совпадение минимума шума и максимума выходной проводимости при $V_{gs} = -1,4$ В (рис. 4.21), когда весь ток вытеснен в подложку, подтверждает вывод об охлаждающем влиянии подложки, ограничивающем разогрев электронов.

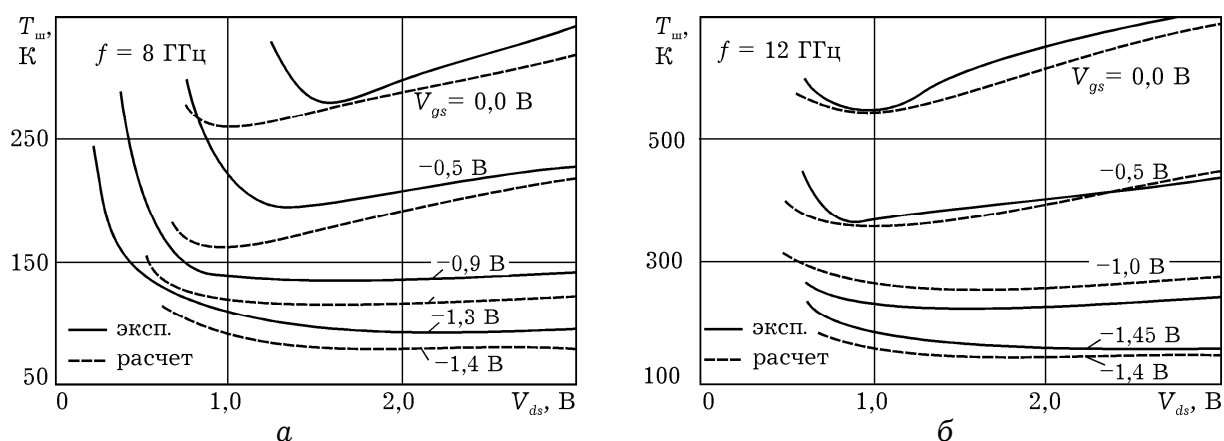


Рис. 4.20. Экспериментальные (сплошные линии) и расчетные (штриховые) зависимости шумовой температуры от напряжения стока на частоте 8 ГГц (а) и 12 ГГц (б)

Следует отметить, что наличие максимума на зависимости выходной проводимости от напряжения на затворе (рис. 4.21) является характерным для малозумящих транзисторов. На большой выборке малозумящих транзисторов экспериментально установлено, что максимум выходной проводимости достигается вблизи напряжений на затворе, соответствующих малозумящему режиму транзистора. При этом, чем более заметным является максимум проводимости, тем более ярко выражен минимум на зависимости минимальной шумовой температуры (минимального коэффициента шума) от напряжения на затворе.

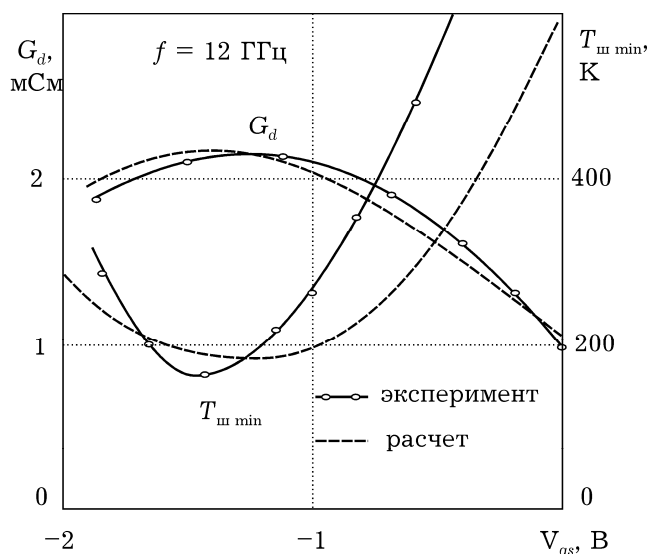


Рис. 4.21. Экспериментальная и расчетная зависимости шумовой температуры и выходной проводимости субмикронного ПТШ от напряжения на затворе

При проектировании субмикронных транзисторов для сверхбыстродействующих схем [4.21-4.23] предъявляется условие хорошей «отсечки» тока стока – на сток-затворной характеристике отсутствует характерный экспоненциальный «хвост» при напряжениях, близких к отсечке. Это достигается за счет уменьшения проводимости подложки и требует создания более высоких и резких потенциальных барьеров на границе эпитаксиальный слой – подложка. С одной стороны, при этом существенно уменьшается экстремум на зависимости выходной проводимости от напряжения на затворе, что уменьшает степень нелинейности динамических характеристик транзистора. С другой стороны, отсутствие «охлаждающего» потенциального барьера со стороны подложки ослабляет эффект «всплеска» дрейфовой скорости и уменьшается средняя дрейфовая скорость носителей в канале, что приводит к ухудшению скоростных характеристик транзистора. Таким образом, разработка субмикронных транзисторов для сверхбыстродействующих интегральных схем требует оптимального выбора конструкции, топологии и электрофизических параметров (уровней легирования канала, концентрации примесей в подложке, геометрических размеров и др.). Для оптимизации характеристик сверхбыстродействующих устройств также могут использоваться квазидвумерные модели [4.10, 4.11], сочетающие высокую физическую информативность и экономичность при расчетах.

4.4. МОДЕЛИ ПТШ ДЛЯ РЕЖИМА БОЛЬШОГО СИГНАЛА

В данном подразделе предлагаются модифицированные модели, в которых в отличие от работы [4.9] учтены особенности токопереноса вблизи режима отсечки тока (субпороговый режим) [4.3], при отрицательном потенциале стока (инверсный режим), а также при положительных потенциалах затвора [4.9], когда существенную роль играют токи прямосмещенного барьера Шоттки. Одним из результатов моделирования являются стоковые вольтамперные характеристики (рис. 4.22), рассчитанные в широком диапазоне питающих напряжений.

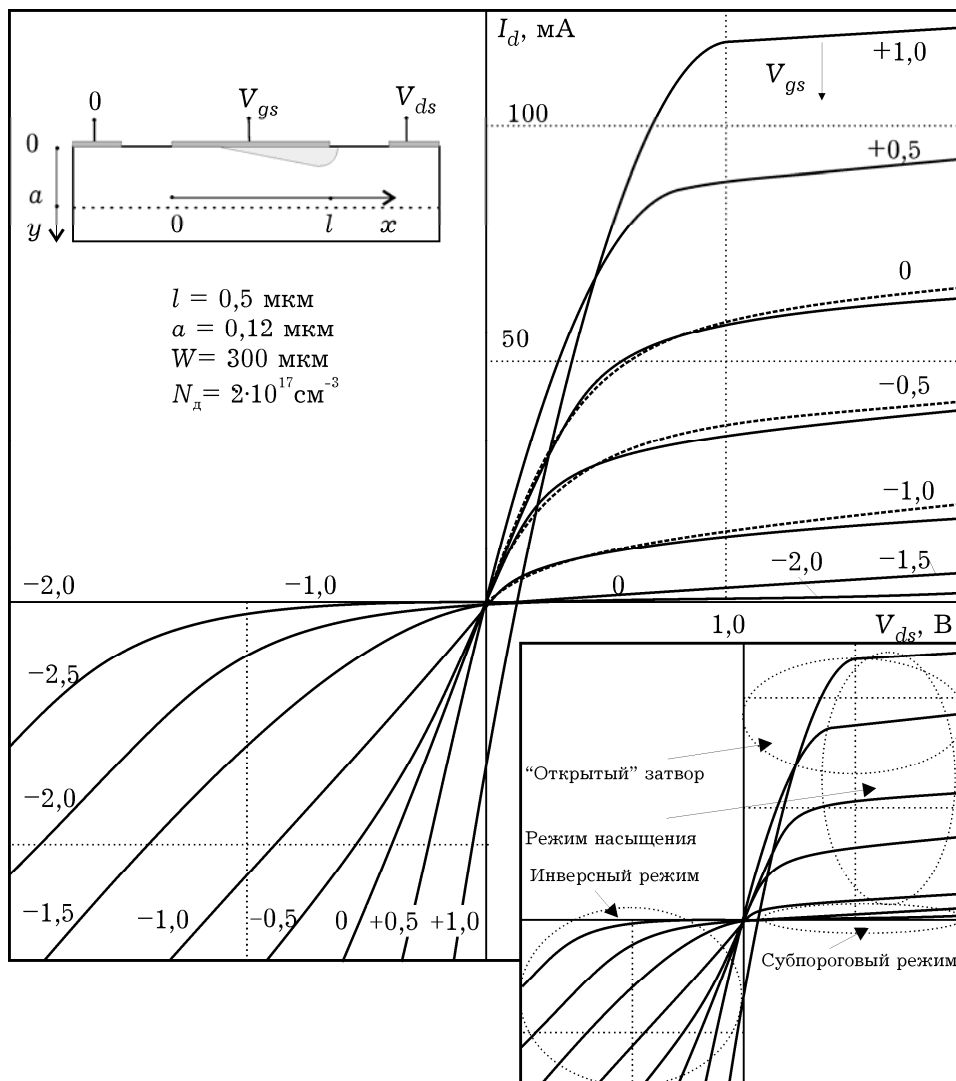


Рис. 4.22. Семейство выходных характеристик, полученное из нелинейной модели ПТШ. Штриховые линии – экспериментальные характеристики. На вставке обозначены характерные нелинейные режимы.

4.4.1. Субпороговый режим

Рассмотрим особенность работы субмикронного арсенид-галлиевого полевого транзистора в субпороговом режиме (вблизи режима отсечки тока стока) и покажем возможность управления током стока не только с помощью изменения толщины проводящего канала, но и за счет надбарьерного прохождения электронов в истоковой части канала при его перекрытии.

Арсенид-галлиевые полевые транзисторы с барьером Шоттки используются при создании различных СВЧ устройств и сверхбыстродействующих цифровых интегральных схем [4.11]. В транзисторах с длиной затвора менее 1 мкм время переключения может составлять несколько пикосекунд (без учета паразитных реактивных элементов). Расчет времени переключения требует правильного моделирования субпорогового режима. В субмикронных транзисторах существенную роль играет проводимость подложки, которая становится особенно заметной вблизи режима отсечки, что проявляется, в частности, в существовании тока стока при напряжении на затворе выше порогового (рис. 4.23,а).

Экспериментальная сток-затворная характеристика для субпорогового режима хорошо линейаризуется (рис. 4.23,б) в полулогарифмических координатах [4.8], что указывает на ее экспоненциальный характер.

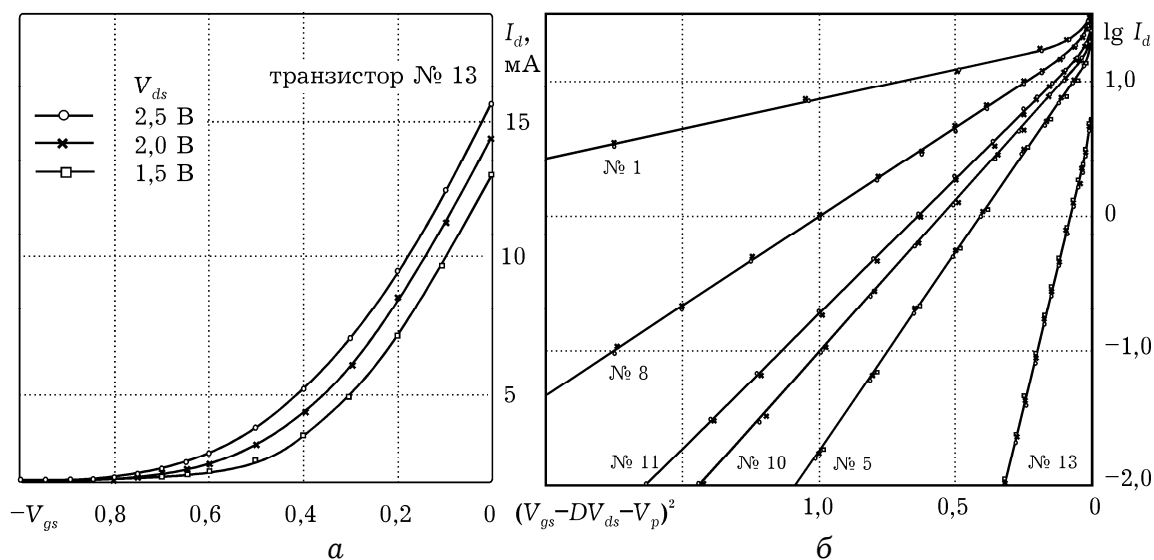


Рис. 4.23. Экспериментальные сток-затворные характеристики ПТШ при разных стокowych напряжениях в естественных и спрямляющих координатах (возле кривых указаны условные номера образцов транзисторов).

Экспоненциальный характер убывания тока стока при запираании транзистора может быть объяснен с помощью кинетической модели Монте-Карло. При пороговом напряжении канал перекрывается в истоковой части канала, где образуется потенциальный барьер, хорошо заметный на рис. 4.24. Существование тока стока при этом объясняется надбарьерным прохождением электронов за счет начальных тепловых скоростей от истока в область тянущего поля канала.

В квазидвумерной модели этот механизм можно учесть, задавая в начальных условиях концентрацию электронов, значение которой находится при максвелловском распределении при известной высоте потенциального барьера. Однако расчет высоты барьера невозможен в рамках традиционного допущения о резкой границе обедненной области и требует учета более реальной формы поперечного сечения канала, которая образуется суперпозицией параболического распределения потенциала обедненной области под затвором и логарифмического распределения в обогащенном слое подложки [4.8].

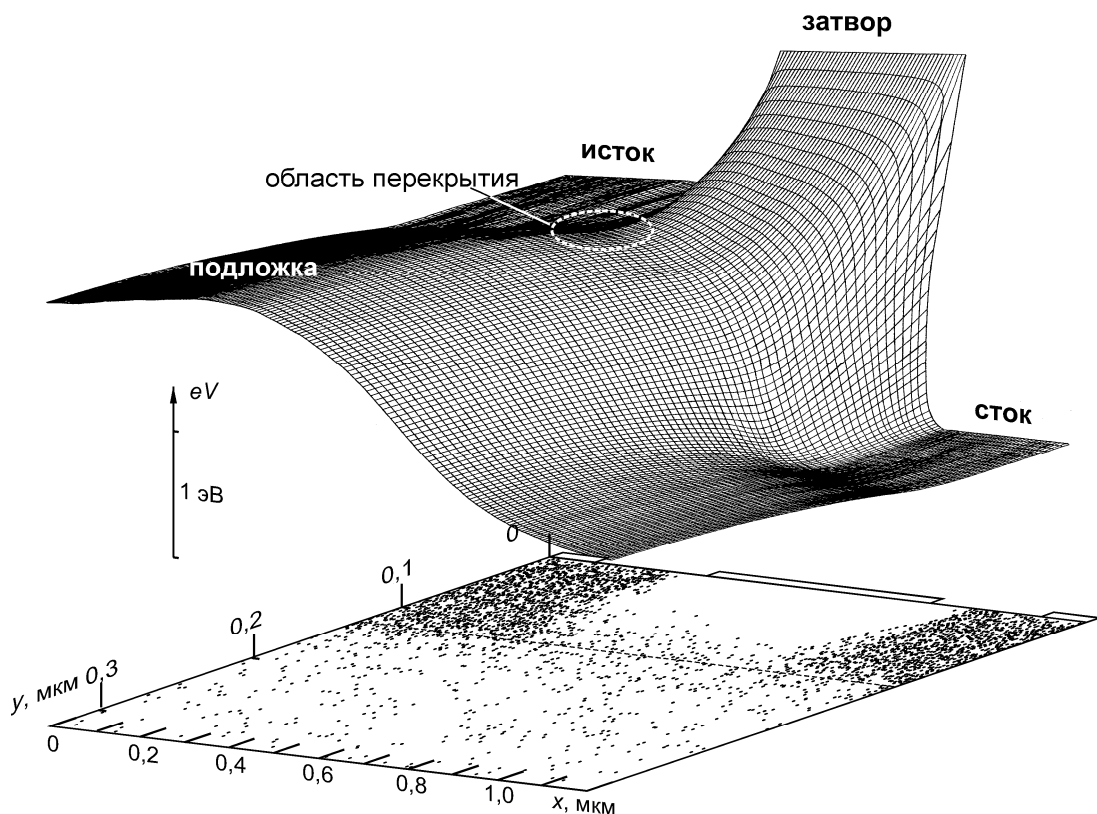


Рис.4.24. Потенциальный рельеф и распределение электронов в канале ПТШ вблизи режима перекрытия

Наиболее полно это можно смоделировать с помощью кинетических моделей (Монте-Карло) но их использование

нерационально из-за большого времени вычислений, необходимого в этом режиме для проведения статистического усреднения. Для моделирования субпорогового режима транзистора при нелинейном схемотехническом анализе приближение резкой границы обедненной области не позволяет правильно описать физические процессы протекания тока, и необходимо более точное описание формы проводящего канала. В режиме перекрытия потенциальный рельеф в поперечном сечении (рис. 4.25) создается суперпозицией поля обедненной области под затвором и поля пространственного заряда электронов, инжектированных в подложку. Трудность адекватного моделирования связана с тем, что распределение потенциала у затвора известно лишь в предположении резкой границы обедненной области. В действительности это распределение заканчивается экспоненциальными «хвостами» с пространственной константой, равной длине экранирования l_3 . Вблизи режима перекрытия распределение потенциала в поперечном сечении канала можно представить (вставка на рис. 4.25) как сумму:

$$\begin{aligned} V_{\text{подл}}(y) &= V_0 \exp [-(y + h/2)/l_3] \quad \text{и} \\ V_{\text{обедн}}(y) &= V_0 \exp [(y - h/2)/l_3]. \end{aligned}$$

Здесь h – ширина канала, не равная обычно рассматриваемой, а отличающаяся от нее на некоторую постоянную величину. Также неопределенным является потенциал V_0 .

Суммарное распределение потенциальной энергии электрона может быть представлено функцией

$$\begin{aligned} U(y) &= -eV(y) = 2U_0 e^{-h/2l_3} \text{ch}(y/l_3) = \\ &= U_{\text{min}} \text{ch}(y/l_3) \approx U_{\text{min}} [1 + (y/l_3)^2]. \end{aligned} \quad (4.29)$$

Распределение электронов в поперечном сечении канала можно представить с помощью модели искривленных зон [$E_c(y) = E_{c0} - eV(y)$]:

$$\begin{aligned} n(y) &= N_c \exp[(E_{\text{Ф}} - E_c)/k_B T] = N_d \exp[eV(y)/k_B T] = \\ &= N_d \exp(-U_{\text{min}}/k_B T) \exp(-U_{\text{min}} y^2 / k_B T l_3^2). \end{aligned} \quad (4.30)$$

Использование максвелловского распределения здесь оправдано тем, что в точке максимального перекрытия канала у истока продольное тянущее поле практически равно нулю и электроны преодолевают потенциальный барьер высотой U_{min} за счет тепловой скорости.

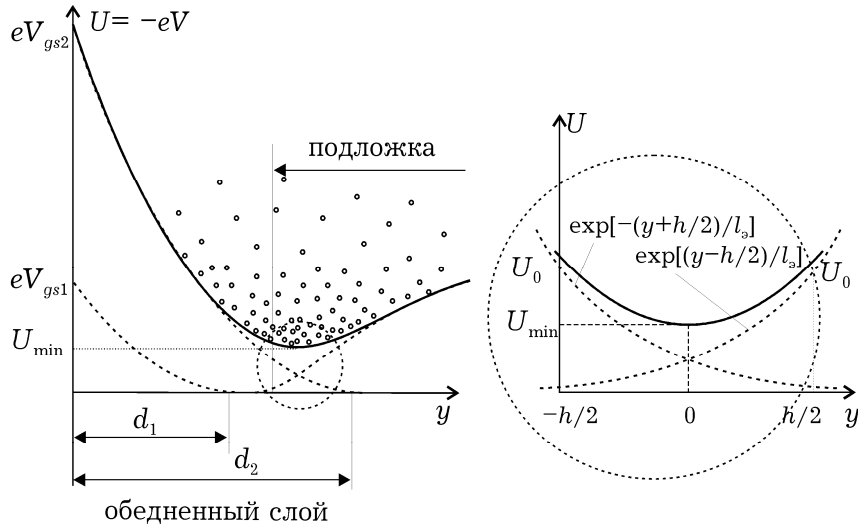


Рис. 4.25. Форма потенциального рельефа в поперечном сечении канала у истока (а), иллюстрирующая надбарьерное прохождение электронов, и его параболическая аппроксимация (б).

Для моделирования процессов в транзисторе с использованием квазидвумерной модели [4.10, 4.12] удобно ввести среднюю концентрацию электронов, обеспечивающую протекание тока стока в субпороговом режиме:

$$n_{\text{cp}} = (1/h) \int_{-\infty}^{\infty} n(y) dy = (2N_{\text{д}}/h) \int_0^{\infty} \exp[eV(y)/k_{\text{B}}T] dy. \quad (4.31)$$

После интегрирования с использованием (4.29) и табличного интеграла получим функцию

$$n_{\text{cp}} = N_{\text{д}} \sqrt{\pi} (l_3/h) \exp(-U_{\text{min}}/k_{\text{B}}T) / \sqrt{U_{\text{min}}/k_{\text{B}}T}. \quad (4.32)$$

С учетом обозначения $U_{\text{min}} = 2U_0 e^{-h/2 l_3}$ получим

$$n_{\text{cp}} = N_{\text{д}} \frac{\sqrt{\pi} l_3}{h} \frac{\exp[-(2U_0/k_{\text{B}}T) \exp(-h/2 l_3)]}{\sqrt{(2U_0/k_{\text{B}}T) \exp(-h/2 l_3)}}. \quad (4.33)$$

Электроны с такой концентрацией создают надбарьерный ток за счет тепловой скорости $v_{\text{T}} = \sqrt{3k_{\text{B}}T/m^*}$ с плотностью $j = en_{\text{cp}}v_{\text{T}}/4$. При ширине транзистора W и толщине канала порядка длины экранирования l_3 они обеспечивают при потенциале

перекрытия ток порядка 10 мА, что хорошо согласуется с экспериментом (рис. 4.23).

Полученное выражение может быть использовано в моделях, основанных на численном интегрировании системы одномерных уравнений, однако оно содержит трудно идентифицируемые параметры (U_0 и h) и неудобно для простых моделей.

В аналитических моделях можно применять соотношение

$$I_d = I_{d0} \exp[\alpha(V_{gs} - V_p - DV_{ds})^2], \quad (4.34)$$

которое следует непосредственно из экспериментальных результатов. На рис. 4.23 видно, что сток-затворные характеристики линеаризируются в координатах $\ln I_d$ и $(V_{gs} - V_p - DV_{ds})^2$, совпадая для разных стоковых напряжений, и их «хвосты» могут быть описаны выражением (4.34).

Параметры I_{d0} , α , D легко определяются из результатов эксперимента: коэффициент α – по наклону линеаризованных характеристик, величина D , равная отношению $\Delta V_{си} / \Delta V_{gs}$ при $I_d = \text{const}$, показывает относительную степень влияния стокового напряжения на ток стока, а значение I_{d0} находится как экстраполированное значение тока в точку $V_{gs} - V_p - DV_{ds} = 0$. Следует обратить внимание, что для исследованной партии транзисторов значения I_{d0} практически одинаковы и составляют $I_{d0} \cong 25$ мА. Эту величину можно интерпретировать как максимальный ток надбарьерной «эмиссии» (из «виртуального катода»), обеспечиваемой омическим контактом при отсутствии тянущего поля. Значения D имеют порядок 0,01 и по физическому смыслу аналогичны проницаемости сетки в вакуумном триоде. Наибольший разброс наблюдается для коэффициента α , который имеет значения порядка 1...10 В⁻². Он наиболее критичен к технологическим погрешностям при производстве, поэтому при соответствующей его идентификации может служить хорошим критерием при тестировании транзисторов.

Полученные результаты свидетельствуют о том, что в субпороговом режиме в отличие от обычного для полевых транзисторов управления током за счет толщины канала («резистивный» механизм) возможен и *другой механизм управления*. Он связан с проникновением поля стока в область истока и подобен влиянию поля анода на пространственный заряд у катода в вакуумном триоде. Это подтверждается изменением вида сток-затворных характеристик: от «пентодного» вида при микронных длинах затвора до «триодного» при длинах около 0,1 мкм (рис. 4.3).

4.4.2. Открытый затвор

Проектирование СВЧ устройств, работающих в большом динамическом диапазоне, требует разработки нелинейных моделей активных компонентов [4.24 - 4.27].

Рассмотрим модель транзистора, которая учитывает один из характерных нелинейных эффектов, возникающих при открывании затворного барьера при положительном смещении.

В рамках квазидвумерной модели эффект открывания затворного барьера можно учесть, уменьшая на каждом шаге интегрирования ток канала на величину, соответствующую вольтамперной характеристике барьера Шоттки при известной разности потенциала затвора и текущего потенциала канала. Особенностью такой процедуры является то, что потенциал затвора зависит от падения напряжения на сопротивлении металлизации за счет протекания полного тока затвора, значение которого определяется только в конце процедуры интегрирования.

Картину физических процессов для выработки приемлемых допущений можно представить на рис. 4.26, где изображена стоковая характеристика транзистора. В режиме насыщения сравнительно большой потенциал стока обеспечивает существование обедненной области практически вдоль всего канала даже при положительном потенциале затвора. С уменьшением потенциала стока снижается и потенциал вдоль канала. Барьер Шоттки открывается, прежде всего, со стороны истока, и часть электронов уходит из канала, создавая затворный ток. Дальнейшее уменьшение потенциала стока приводит к расширению открытой части затвора и открыванию барьера со стороны стока. Более того, при потенциалах стока, меньших затворного, ток стока становится отрицательным.

Режим работы транзистора при отсутствии разогревных явлений можно промоделировать, пренебрегая при малых стоковых напряжениях обедненной областью. Для канала толщиной a и шириной W с концентрацией электронов n , обладающих подвижностью μ , очевидно, что на участке длиной dx ток канала I вызывает падение напряжения

$$dV = I(x)dx/en\mu Wa. \quad (4.35)$$

Ток канала является переменным, так как на каждом участке dx он уменьшается на величину dI за счет ухода электронов из канала в затворную цепь. Принимая вольтамперную

характеристику открытого барьера Шоттки линейной для положительных смещений, можно записать:

$$dI = g[V_{gb} - V(x)]Wdx. \quad (4.36)$$

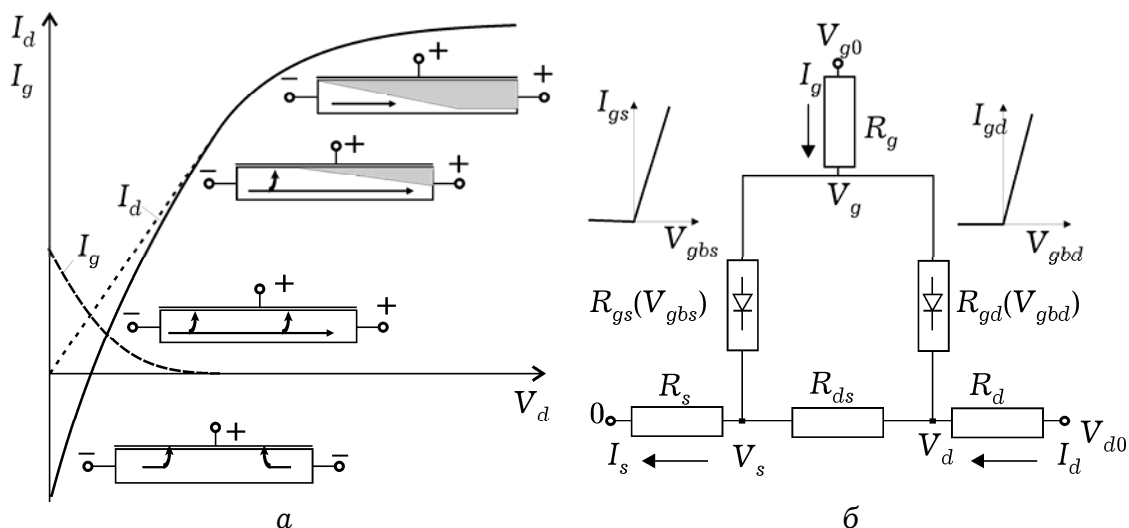


Рис. 4.26. К расчету начального участка сток-затворной характеристики ПТШ при положительных потенциалах затвора: *а* – физическая модель, показывающая трансформацию обедненной области и направления токов стока и затвора в разных режимах; *б* – схемная модель подзатворной области при положительных потенциалах затвора;

Здесь $V_{gb} = V_{gs} - V_b$ – потенциал затвора с учетом контактной разности потенциалов на барьере (типичное значение $V_b \approx 0.8$ В); $V(x)$ – текущий потенциал канала; g – проводимость перехода затвор-канал единичной площади. Этот параметр модели необходимо идентифицировать из дополнительных измерений. В самом простом случае, пренебрегая переходным сопротивлением контакта, можно считать, что g определяется поперечным сопротивлением канала половинной толщины: $g = 2en\mu/a$.

Систему уравнений (4.35) и (4.36) легко разделить повторным дифференцированием и перекрестной подстановкой:

$$d^2I/dx^2 = \alpha^2I; \quad (4.37)$$

$$d^2V/dx^2 = -\alpha^2[V_{gb} - V(x)]; \quad (4.38)$$

Здесь обозначено $\alpha^2 \equiv g/en\mu a$.

Решение этих уравнений, удовлетворяющее краевым условиям $V(0) = V_s$ и $V(L) = V_d$, можно представить в виде

$$V_{gb} - V(x) = [V_{gbs}\text{sh}\alpha(L-x) + V_{gbd}\text{Sh}\alpha x]/\text{sh}\alpha L, \quad (4.39)$$

$$I(x) = G[V_{gbs}\text{ch}\alpha(L-x) + V_{gbd}\text{ch}\alpha x]/\text{sh}\alpha L, \quad (4.40)$$

где $G = W\sqrt{gen\mu a}$.

Надо иметь в виду, что полученные решения справедливы для $V_{gb} - V(x) > 0$. В противном случае $I(x) = I(x_0) = \text{const} = I_d$ и

$$V(x) = V(x_0) + I(x_0)(x - x_0)/en\mu W a, \quad (4.41)$$

где x_0 — точка, в которой $V(x_0) = V_{gb}$.

Из полученных решений непосредственно следуют выражения для тока каждого из электродов

$$I_s = I(0) = G[V_{gbs}(\text{ch}\alpha L - 1) + V_{сн}]/\text{sh}\alpha L, \quad (4.42)$$

$$I_d = I(L) = G[V_{gs} - V_{gbd}(\text{ch}\alpha L - 1)]/\text{sh}\alpha L, \quad (4.43)$$

$$I_g = I_s - I_d = G(V_{gbs} + V_{gbd})(\text{ch}\alpha L - 1)/\text{sh}\alpha L. \quad (4.44)$$

Данные выражения определяют вид вольтамперных характеристик непосредственно активной части канала транзистора и соответственно вид схемной модели. Последняя может быть представлена в виде треугольника из сопротивлений открытого затвора $R_{gs} = R_{gd} = \text{sh}\alpha L/G(\text{ch}\alpha L - 1)$ и сопротивления сток-исток $R_{ds} = \text{sh}\alpha L/G$ (рис. 4.26). Надо иметь в виду, что сопротивления R_{gs} и R_{sd} зависят от электрического режима и управляются напряжениями V_{gbs} и V_{gbd} в соответствии с вольтамперными характеристиками, приведенными на том же рисунке. В общем случае ВАХ затвора может быть описана и другой приемлемой функцией, что требует задания соответствующего вида нелинейности для величины G .

В отличие от встречающегося представления подзатворной области в виде двух диодов ее схемная модель должна быть представлена тремя нелинейными резисторами, два из которых эквивалентны диодам и управляются напряжениями затвор-исток и затвор-сток, а третий учитывает их взаимодействие. Для моделирования этого режима необходимо дополнительно идентифицировать три параметра (V_e , i_{s0} , m) при использовании аппроксимации ВАХ затвора в виде $i = i_s [\exp(eV/mk_B T) - 1]$, $i_s = i_{s0} \exp(-eV_b/k_B T)$, или два параметра (G , V_e) при кусочно-линейной аппроксимации $i = 0$ при $V \leq V_e$, $i = G(V - V_e)$ при $V > V_e$.

На рис. 4.27 представлены результаты расчета по полной схемной модели транзистора, включающей паразитные сопротивления стока, истока и металлизации затвора. В этом

случае аналитическая модель (4.42 – 4.44) должна быть дополнена соотношениями:

$$V_{d0} = V_d + I_d R_d, \quad V_{g0} = V_g + I_g R_g, \quad V_s = I_s R_s,$$

позволяющими перейти от «внутренних» потенциалов к потенциалам на выводах транзистора в схеме с общим истоком. На рассчитанных в соответствии с этой моделью начальных участках стоковой характеристики транзистора для различных значений потенциала затвора отчетливо видны два «излома», соответствующие открыванию затвора у истокового и стокового края.

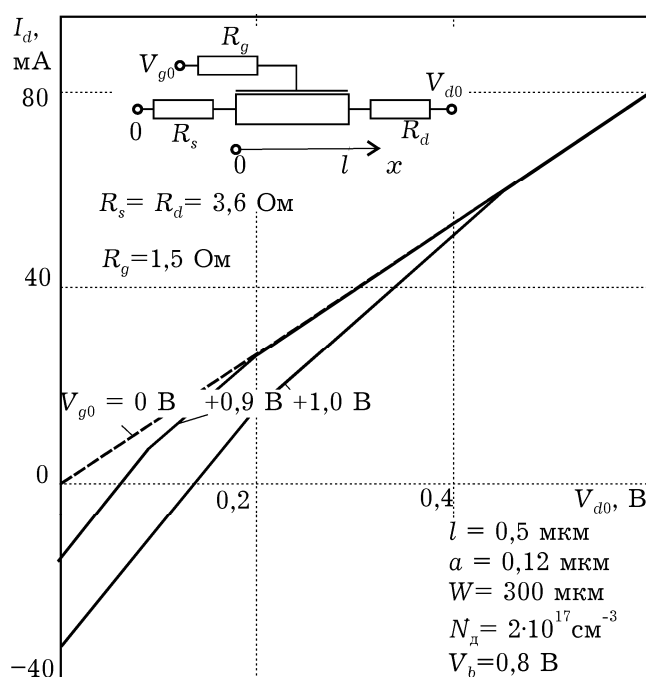


Рис. 4.27. Результаты моделирования стоковой характеристики ПТШ при положительных потенциалах затвора при кусочно-линейной аппроксимации ВАХ барьера Шоттки.

Данная модель пригодна для небольших стоковых напряжений, и вопрос о поведении транзистора в режиме насыщения при больших положительных смещениях на затворе остается за пределами принятых допущений.

4.4.3. Инверсный режим

Моделирование инверсного режима ПТШ связано с необходимостью формирования процедуры интегрирования основных уравнений модели [4.9] в направлении навстречу электронному потоку. Несмотря на принципиальную возможность

создания такой процедуры, более удобно, мысленно поменяв сток и исток местами, использовать отработанную модель для анализа физических процессов. Но при расчете параметров схемной модели надо иметь в виду, что способ их определения в используемой физико-топологической модели соответствует схеме с общим стоком. Характерным для инверсного режима является то, что при потенциалах стока, более отрицательных, чем потенциал затвора (относительно истока), на величину барьерного потенциала, транзистор работает в режиме открытого затвора.

Пример моделирования статических ВАХ для ПТШ с длиной затвора 0,5 мкм в сравнении с экспериментальными данными приведен на рис. 4.22. Наклон огибающей кривых на крутом участке соответствует омическому сопротивлению стока и истока. Лишь для потенциала затвора +0,8 В заметно уменьшение тока стока из-за открывания затворного барьера Шоттки. При повышении стокового напряжения часть открытого барьера уменьшается и ВАХ на пологом участке имеет обычный вид. При отрицательных стоковых напряжениях открывание барьера заметно для потенциала +0,5 В; для остальных значений – вне пределов представленных графиков. Из вида ВАХ для инверсного режима следует, что транзистор работает в неактивном режиме, так как степень влияния потенциалов затвора и стока на ток стока одинакова.

Физико-топологический уровень описания ПТШ позволяет после этапа идентификации провести расчет схемных параметров для любой рабочей точки или в заданном диапазоне значений питающих напряжений. Эти массивы параметров затем могут быть использованы непосредственно в системах проектирования с применением, например, сплайн-аппроксимации.

При моделировании высоковольтного режима необходимо учитывать особенность работы ПТШ, состоящую в том, что генерируемые в процессе лавинного размножения дырки вытягиваются из канала полем затвора и не создают «дырочной» лавины. Это дает возможность пользоваться простой теорией лавинных процессов и считать, что ток канала на каждом шаге интегрирования увеличивается на $\Delta i_k = \alpha i_k \Delta x$ только за счет генерируемых электронов. На такую же величину возрастает ток затвора, что на схемной модели отображается включением между стоком и затвором нелинейного резистора. Здесь требуется дополнительная идентификация параметров зависимости коэффициента ионизации α от напряженности электрического поля, которая может быть выбрана в виде: $\alpha = \alpha_0 \exp(\mathcal{E}_0/\mathcal{E})^n$.

Описанные нелинейные модели субмикронных ПТШ, позволяют учесть эффекты при уровнях входных воздействий, сравнимых с напряжениями питания транзистора. Модели пригодны для нелинейного моделирования устройств средствами САПР и сочетают в себе возможности описания различных нелинейных эффектов в субмикронных ПТШ и не требуют значительных дополнительных вычислительных затрат.

4.5. ДВУМЕРНОЕ МОДЕЛИРОВАНИЕ СУБМИКРОННЫХ ПТШ

Необходимость двумерного моделирования физических процессов, когда длина затвора и размеры активной области не превышает 0.5 мкм, обусловлена их сравнимостью с характерными размерами высоколегированных слоев и наличием областей с высокими градиентами функций, описывающих электрические поля и физические характеристики электронного газа: концентрацию, скорость, энергию.

При этом моделирование сравнительно «больших» транзисторных структур (с размерами около 1 мкм и больше) может осуществляться по упрощенной («локальной») двумерной модели, содержащей уравнения Пуассона и уравнение непрерывности тока, а полевая зависимость подвижности электронов связана с аппроксимацией «стационарных» значений дрейфовой скорости. При этом совместно решаются двумерные уравнение Пуассона и уравнение непрерывности, которые в стационарном случае представляют собой уравнения в частных производных эллиптического типа. Уравнение Пуассона для моделируемого ПТШ запишется:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{e}{\epsilon \epsilon_0} [n(x, y) - N_d(x, y)] \quad (4.45)$$

где $N_d(x, y)$ - функция, задающая профиль легирующей примеси в активной области транзистора.

Уравнение непрерывности тока транзистора записывается как:

$$\operatorname{div}[D \operatorname{grad} n(x) - \mu \operatorname{grad} V(x)] = 0,$$

где D - коэффициент диффузии, μ - подвижность электронов.

В слабых полях коэффициент диффузии описывается соотношением Эйнштейна: $D = \mu k_B T / e$. В сильных полях

зависимости подвижностей электронов и дырок от напряженности электрического поля, температуры и уровней легирования донорной и акцепторной примеси аппроксимируются функциями, предложенными в работе [4.28].

Недостатком использования аппроксимаций является необходимость проведения экспериментальных исследований полевой зависимости подвижности для различных материалов, а также «интегральный» характер этих зависимостей, не позволяющий идентифицировать вклад различных механизмов рассеяния в релаксационные процессы. Наиболее физичный способ определения времен релаксации для различных механизмов рассеяния связан с подходом, приведенным в гл. 2. При этом времена релаксации рассчитываются по аналитическим соотношениям для отдельных механизмов рассеяния и можно учитывать наиболее существенные из них.

Система уравнений локальной модели более экономична при численном анализе с точки зрения вычислительных затрат, поскольку не включает релаксационные уравнения и исходная система дифференциальных уравнений в частных производных имеет меньшую размерность, но она не пригодна для анализа нестационарных эффектов.

С позиций адаптивного моделирования и построения системы моделирования субмикронных приборов использование локальной двумерной модели оправдано для сравнительно «больших» структур и может также использоваться для получения начальных приближений, получения оценок относительно шага дискретизации уравнений и количества итераций вычислительного процесса, определения размеров активных и пассивных областей структуры и др. [4.29].

Большой интерес для анализа представляет нелокальная двумерная модель. Рассмотрим уравнения нестационарной двумерной модели субмикронного ПТШ (нелокальной модели), основанную на решении уравнений Пуассона и уравнения непрерывности, уравнений сохранения импульса и энергии, которая в сочетании с аппроксимациями времен релаксации энергии и импульса в сильных электрических полях, позволяет рассчитать как физические характеристики транзистора, так и статические (вольтамперные) характеристики в широком диапазоне питающих напряжений.

Исходная система уравнений нелокальной двумерной модели имеет следующий вид:

$$v_x \frac{\partial n}{\partial x} + v_y \frac{\partial n}{\partial y} = 0, \quad (4.46)$$

$$v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} = \frac{e\mathcal{E}_x}{m^*} - \frac{v_x}{\tau_p}, \quad (4.47)$$

$$v_x \frac{\partial v_y}{\partial x} + v_y \frac{\partial v_y}{\partial y} = \frac{e\mathcal{E}_y}{m^*} - \frac{v_y}{\tau_p}, \quad (4.48)$$

$$v_x \frac{\partial T_e}{\partial x} + v_y \frac{\partial T_e}{\partial y} = -\frac{T_e - T_0}{\tau_E} + \frac{2E_d}{3k_B}(\gamma - 1), \quad (4.49)$$

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{e}{\epsilon\epsilon_0} [n(x, y) - N_d(x, y)] \quad (4.50)$$

где дрейфовая энергия определяется как

$$E_d = m^* v^2 / 2 = \frac{m^*}{2} (v_x^2 + v_y^2).$$

Для решения уравнений область транзистора представляется в виде конечных элементов. Выбор метода разбиения обусловлен необходимостью аппроксимации области со сложной конфигурацией, а также тем, что сильные поля локализуются в области с размерами 0,2-0,3 мкм у стокового края затвора и требуется более мелкая сетка. Триангуляция в методе конечных элементов связана с представлением моделируемой области в виде совокупности треугольников. По известным распределениям потенциала в узлах сетки рассчитываются напряженности электрического поля :

$$\vec{\mathcal{E}} = -\text{grad } V(x). \quad (4.51)$$

Для решения системы уравнений (4.46)-(4.50) на электродах транзистора задаются граничные условия Дирихле для потенциала ($V = \text{const}$). На боковых границах задается однородное условие Неймана ($\partial V / \partial \mathbf{n} = 0$), где \mathbf{n} - нормаль к этим границам. Аналогичные условия задаются на свободной поверхности между электродами транзистора.

Алгоритм решения системы уравнений (4.46-4.50) и с учетом аппроксимаций времен релаксации и уравнения для напряженности электрического поля (4.51) представляет собой несколько вложенных циклов. Каждое из уравнений, в свою очередь, является системой конечно-элементных уравнений большой размерности. Разреженные матрицы систем обладают диагональным преобладанием и имеют ленточную структуру. Для решения систем используется метод прогонки в сочетании с итерационным многошаговым методом Ньютона. После получения решения с заданной точностью для потенциалов (уравнение

Пуассона) совместно решается система уравнений (4.46-4.50) с расчетом на каждом шаге времен релаксации и эффективной массы с учетом заселенности долин по соответствующим аппроксимациям. При этом рассматривается двухдолинная модель энергетических зон.

Выходной ток ПТШ рассчитывался как нормальная составляющая тока к контакту стока по известным двумерным распределениям концентрации электронов, напряженности электрического поля и подвижности по результатам решения исходной системы уравнений при достижении заданной точности решений в узлах сетки. При невозможности достижения заданной точности в пределах итерационных процессов (внутреннего и внешнего циклов) осуществлялась процедура дробления шага триангулярной сетки.

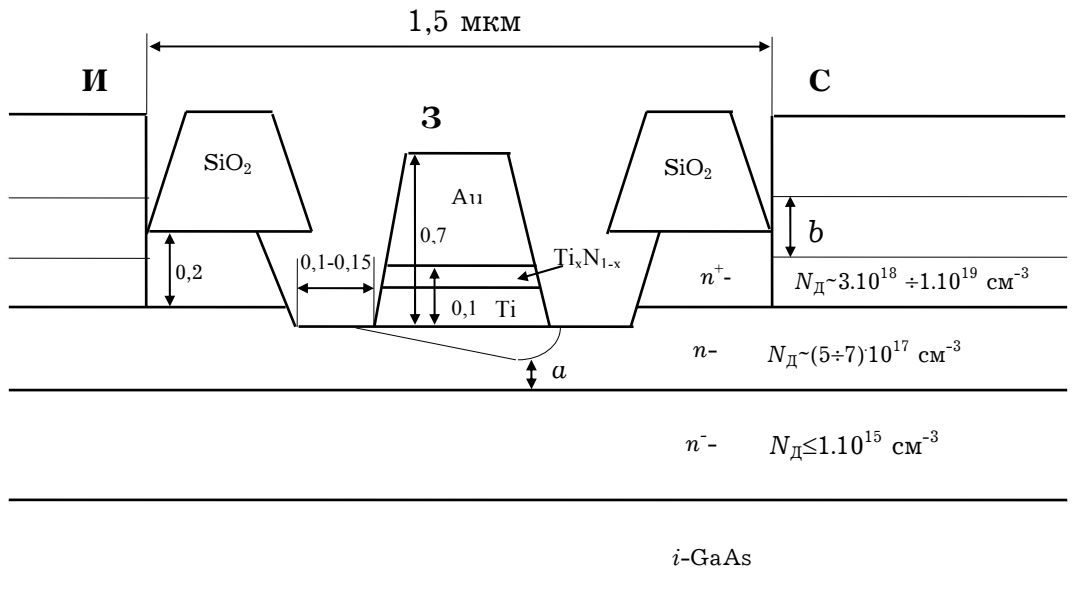


Рис. 4.28. Топология моделируемого субмикронного ПТШ

В качестве тестовой использовалась структура ПТШ с размерами затвора $0,25 \times 100$ мкм, толщиной эпитаксиального слоя $0,1$ мкм со ступенчатым профилем легирующей примеси, с концентрацией доноров в канале и полуизолирующей подложке соответственно $5 \times 10^{17} \text{ см}^{-3}$ и $5 \times 10^{15} \text{ см}^{-3}$.

Другие размеры, параметры слоев и материалов ПТШ представлены на рис. 4.28. Результаты моделирования физических процессов в канале исследуемого транзистора представлены на рисунках 4.29, 4.30. На всех ниже приведенных рисунках, где представлены двумерные распределения физических величин

субмикронного ПТШ с длиной затвора 0,25 мкм, затвор располагается симметрично относительно координаты $x = 0$.

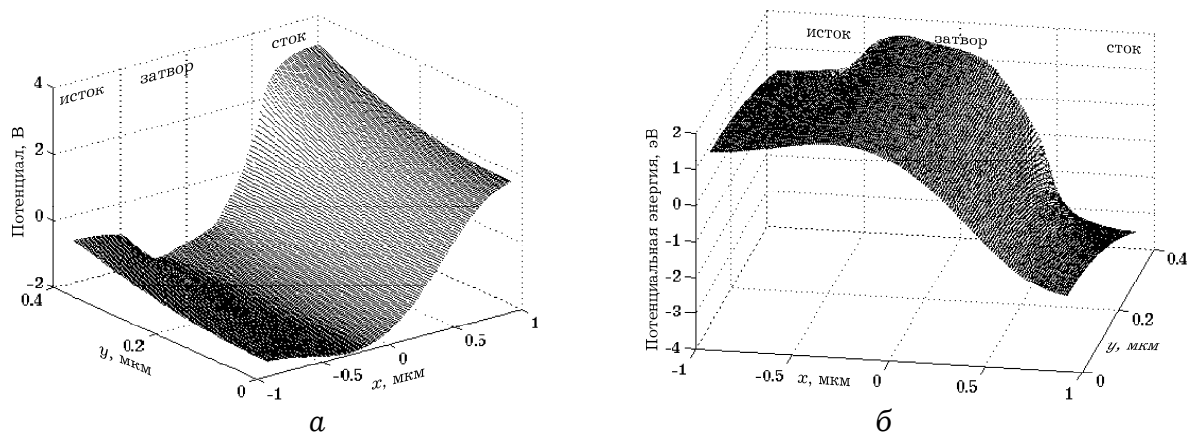


Рис. 4.29. Распределение потенциала (а) и потенциальной энергии (б) в структуре субмикронного ПТШ

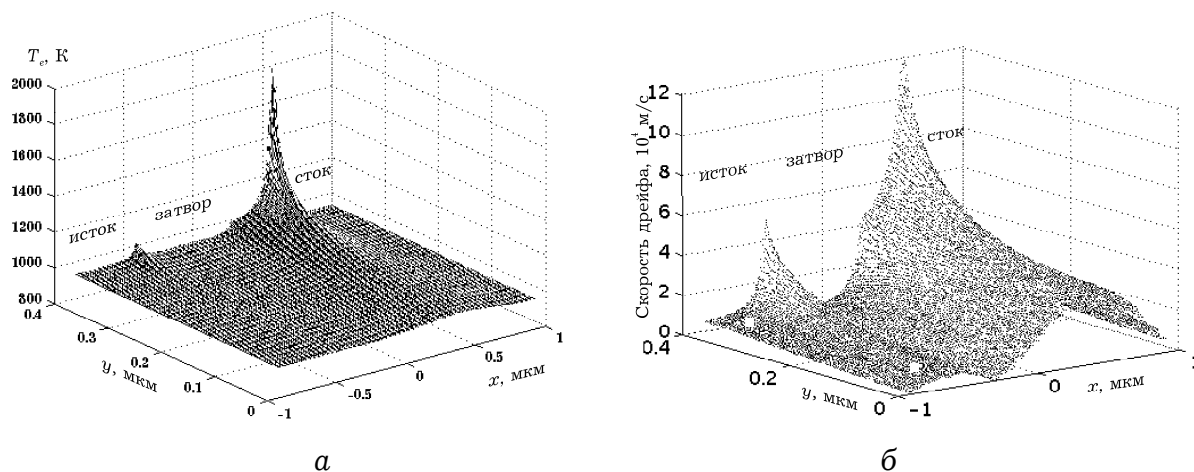


Рис. 4.30. Распределение электронной температуры (а) и продольной скорости (б) для субмикронного ПТШ с длиной затвора 0,25 мкм

Полученные результаты соответствуют следующему электрическому режиму на электродах транзистора: потенциал истока равен нулю, потенциалы на затворе и стоке равны соответственно $-1,0$ В (с учетом высоты барьера Шоттки $0,8$ В) и $4,0$ В.

Как показывают результаты сравнения расчетов и измерений [4.10] для подобной транзисторной структуры с размерами затвора $0,3 \times 300$ мкм в короткоканальных структурах (длина затвора менее $0,5$ мкм) на вольтамперных характеристиках области насыщения тока стока выражены не так ярко, как в относительно более «длинных» структурах (длина затвора более $0,5$ мкм), имеющих более пологие характеристики.

С точки зрения анализа интегральных характеристик транзистора это связано с большими значениями динамической

крутизны «коротких» структур. Сравнение проведенных расчетов с экспериментом указывает на необходимость учета в исходной системе релаксационных уравнений, так как значения рассчитанных токов стока в пологой области выходных характеристик ПТШ меньше для всех значений питающих напряжений. Приведенные данные подтверждают результаты, полученные при моделировании по кинетической модели и в квазидвумерном приближении [4.11], о том, что эффект «всплеска» дрейфовой скорости приводит к увеличению средней скорости электронного газа, что, в свою очередь, приводит к росту динамической крутизны и выходной проводимости ПТШ.

4.6. МОДЕЛИРОВАНИЕ ЭФФЕКТОВ ТРЕТЬЕГО ИЗМЕРЕНИЯ

4.6.1. Распределенные эффекты в ПТШ

При увеличении степени интеграции СВЧ ИС геометрические размеры их компонентов могут быть сравнимыми с длиной электромагнитной волны в материале полупроводника. При этом распределенные эффекты могут влиять на усилительные и шумовые характеристики активных компонентов [4.30].

Для анализа распределенных эффектов в затворной линии субмикронного ПТШ рассмотрим его структуру (рис. 4.31).

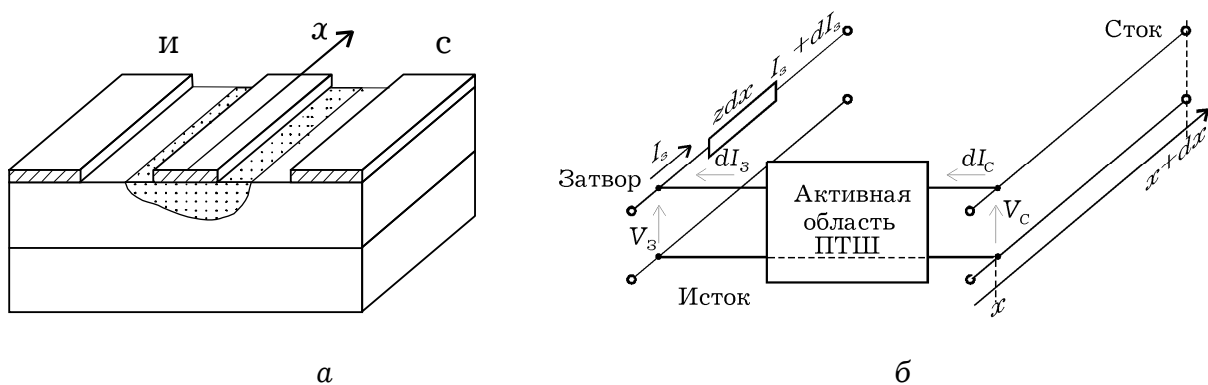


Рис. 4.31. Затворная линия транзистора: а – структура ПТШ; б – схемная модель дифференциального отрезка линии

Для количественного анализа процессов распространения волн тока и напряжения вдоль затворной линии передачи рассмотрим дифференциальные уравнения (ось x направлена от точки входа сигнала вдоль ширины транзистора):

$$dV_c(x) = -I_s(x)zdx;$$

$$\begin{aligned} dI_s(x) &= y_{11}dxV_s(x) + y_{12}dxV_c; \\ dI_c(x) &= y_{21}dxV_s(x) + y_{22}dxV_c. \end{aligned} \quad (4.52)$$

Здесь z – погонное сопротивление затворной линии передачи; а y_{ik} ($i, k = 1, 2$) – параметры погонной матрицы проводимости активной области транзистора; т. е. матрицы проводимости ПТШ (без учета сопротивления металлизации затвора) единичной ширины.

Первое уравнение характеризует падение напряжения на участке затворной линии dx . Два других уравнения устанавливают связь между токами и напряжениями на четырехполюснике дифференциальной секции транзистора. Уравнения записаны в предположении эквипотенциальности стока и истока (их сопротивления растекания пренебрежимо малы) для трех переменных: тока стока $I_c(x)$ и затвора $I_s(x)$, а также напряжение затвор-исток $V_s(x)$ (рис. 4.31). Выходное напряжение сток-исток – величина постоянная и не зависит от x .

Первые два уравнения можно рассматривать как систему уравнений для определения $V_s(x)$ и $I_s(x)$. Разделение переменных осуществляется повторным дифференцированием и последующей перекрестной подстановкой. Обозначая $\sqrt{zy_{11}} \equiv \dot{\gamma} = \alpha + j\beta$, получим

$$\begin{cases} \frac{d^2 I_s(x)}{dx^2} = \gamma^2 I_s(x) \\ \frac{d^2 V_s(x)}{dx^2} = \gamma^2 V_s(x) + \gamma^2 (y_{12}/y_{11}) V_c \end{cases}.$$

Решение уравнений можно представить в форме

$$\begin{cases} I_s(x) = I_s^+ e^{-\gamma x} - I_s^- e^{\gamma x} = \frac{V_s^+}{Z_B} e^{-\gamma x} - \frac{V_s^-}{Z_B} e^{\gamma x} \\ V_s(x) = V_s^+ e^{-\gamma x} + V_s^- e^{\gamma x} - (y_{12}/y_{11}) V_c \end{cases}.$$

Подставим эти решения в одно из исходных уравнений первого порядка. Приравнявая слагаемые при одинаковых функциях, получим связь между двумя парами постоянных:

$$\frac{V_3^+}{I_3^+} = \frac{V_3^-}{I_3^-} = \frac{z}{\gamma} = \sqrt{\frac{z}{y_{11}}} \equiv Z_0.$$

Здесь величина Z_0 по аналогии с длинными линиями соответствует комплексному значению волнового сопротивления.

Еще две постоянные интегрирования можно найти из условий $V_3(0) = V_{\text{вх}}$ и $I_3(W) = 0$. Подставляя в уравнения предыдущей системы, получим

$$V_3^\pm = \frac{[V_{\text{вх}} + (y_{12}/y_{11})V_c]e^{\pm\gamma W}}{\text{ch } \gamma W}.$$

В итоге выражения, характеризующие распределение амплитуд напряжения и тока, можно представить в форме:

$$I_3(x) = \frac{[V_{\text{вх}} + (y_{12}/y_{11})V_c] \text{sh } \gamma(W-x)}{Z_0 \text{ch } \gamma W};$$

$$V_3(x) = \frac{[V_{\text{вх}} + (y_{12}/y_{11})V_c] \text{ch } \gamma(W-x)}{\text{ch } \gamma W} - \left(\frac{y_{12}}{y_{11}}\right)V_c.$$

Полученные функции позволяют установить связь амплитуд входного и выходного токов для транзистора в целом с амплитудами напряжений.

Из последнего выражения для тока затвора $I_3(x)$ при $x = 0$ получим

$$I_{\text{вх}} = I_3(0) = [V_{\text{вх}} + (y_{12}/y_{11})V_c] \text{th } \gamma W / Z_0.$$

Подставляя $V_3(x)$ в третье уравнение исходной системы (4.52), интегрированием по всей ширине затвора получим выражение для полного тока стока

$$I_{\text{вых}} = \int_0^W dI_c(x) = y_{21}W (\text{th } \gamma W / \gamma W) V_{\text{вх}} +$$

$$+ \left[y_{22}W + \frac{y_{21} \cdot y_{12}}{y_{11}} W \left(\frac{\text{th } \gamma W}{\gamma W} - 1 \right) \right] V_c.$$

Два последних уравнения связи токов $I_{вх}$ и $I_{вых}$ с напряжениями $U_{вх}$ и $U_{вых}$ позволяют перейти к параметрам сосредоточенной модели транзистора без учета сопротивления металлизации затвора. Если вместо погонных y -параметров перейти к параметрам сосредоточенной модели транзистора $Y_{ik}^c = y_{ik}W$, то результат можно представить в следующей форме:

$$Y_{11} = Y_{11}^c F; \quad Y_{12} = Y_{12}^c F; \quad Y_{21} = Y_{21}^c F;$$

$$Y_{22} = Y_{22}^c F + \frac{Y_{12}^c \cdot Y_{21}^c - Y_{11}^c \cdot Y_{22}^c}{Y_{11}^c} (F - 1).$$

Здесь $F \equiv \text{th } \gamma W / \gamma W$ - функция, которая определяет степень влияния распределенных эффектов на характеристики транзистора. Ее аргумент удобно представить в виде $\gamma W = \sqrt{Z_g Y_{11}^c}$, где Z_g - полное сопротивление металлизации затвора.

Если упрощенно считать это сопротивление активным R_g , а проводимость Y_{11}^c смоделировать проводимостью последовательного соединения омического сопротивления R_s , сопротивления канала R_k и емкости затвор-исток C_{gs} (что обычно достаточно хорошо выполняется в широком диапазоне частот), то

$$\gamma W = \sqrt{R_g / (R_k + R_s - i / \omega C_{gs})}.$$

На рис. 4.32 представлены частотные зависимости F и Z_0 для значений $R_g = 2$ Ом, $R_k = 25$ Ом, $R_s = 8$ Ом, $C_{gs} = 0,1$ пФ (на 100 мкм ширины транзистора).

Кривые 1...5 соответствуют значениям ширины затвора 100...500 мкм. Уменьшение модуля F наблюдается на частотах выше 10 ГГц, при этом происходит и ухудшение параметров передачи транзистора. Следует отметить, что наибольший фазовый сдвиг, вызываемый распределенными эффектами, соответствует области частот, где становятся равными активная и реактивная составляющие проводимости Y_{11}^c .

В качестве критерия отсутствия распределенных эффектов можно рассматривать неравенство $\gamma W \ll 1$. При слабо выраженных распределенных эффектах можно использовать приближенное соотношение $\text{th } \gamma W / \gamma W \approx 1 - (\gamma W)^2 / 3$. Полученные выражения для расчета Y -параметров транзистора достаточно

просты и компактны, могут быть использованы с моделью практически любого уровня.

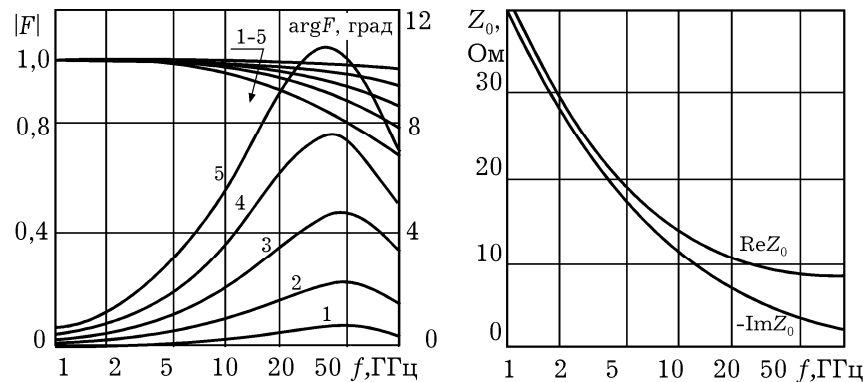


Рис. 4.32. Влияние распределенных эффектов на параметры транзистора

Рассмотренные распределённые эффекты должны влиять не только на усилительные, но и шумовые свойства. Активная область ПТШ описывается схемной моделью с сосредоточенными параметрами, учитывающей свойства линии на полупроводниковой подложке, в которой происходит дрейф горячих носителей. Распространяющаяся электромагнитная волна локализуется в области пространственного заряда под затвором, представляющую собой низкопроводящую область, близкую по своим свойствам к диэлектрику. Проникновение поля в обедненный слой подложки ограничивается высокопроводящим слоем канала. В затворной линии передачи, обладающей заметным активным сопротивлением, происходит затухание сигнала и генерация тепловых шумов, что повышает коэффициент шума всего транзистора.

4.6.2. Анализ ПТШ с распределенными параметрами во временной области

Расширение рабочего диапазона субмикронных транзисторов в миллиметровую область и быстродействия до пикосекундного диапазона приводит к необходимости учета распределенных эффектов вдоль третьей координаты – вдоль ширины затвора [4.31–4.33] и анализа их влияния на временные характеристики.

Для моделирования переходных процессов в микроволновых полевых структурах во временной области, состоящих из линейных и нелинейных компонентов с распределенными параметрами можно использовать метод, описанный в [4.34], где компоненты представлены характеристическими моделями во временной области, а для уменьшения вычислительных затрат,

заменяются схемными моделями. Затем модели компонентов доопределяются граничными значениями напряжений и токов, связанных с распространением сигнала вдоль выбранного направления, тем самым учитывается влияние геометрических размеров полевой структуры на величину проходящего сигнала. Распределенная аналитическая схемная модель, позволяющая рассчитывать Y -матрицу транзистора, импедансные и передаточные характеристики с учетом затухания и изменения фазы сигнала в затворной линии в частотной области, разработана в работе [4.30]. Для цепей миллиметрового диапазона, работающих в импульсном режиме, представляет также интерес информация о форме сигнала, что связано с возможностью рассчитывать распределения токов и напряжений вдоль затворной линии во временной области. Схематически такая линия представлена на рис. 4.33.

Анализ транзисторной структуры вдоль третьей координаты во временной области связан с разработкой моделей, параметры которых в поперечных и продольной плоскостях зависят от времени. Один из подходов связан с дискретизацией структуры вдоль ширины транзистора на отрезки ΔW (на рис. 4.31 вдоль координаты x) и анализом параметров схемной модели транзистора в поперечной к направлению распространения сигнала плоскости во временной области, используя схемную модель с сосредоточенными параметрами. По известным значениям токов и напряжений в начале отрезка линии ΔW решаются волновые уравнения для тока и напряжения в линии. В качестве начальных условий задаются токи и напряжения на электродах, полученные после решения системы узловых дифференциальных уравнений во временной области, т.е. используется декомпозиция системы, когда поочередно решается система уравнений транзистора в поперечной плоскости (без затворной линии) и волновые уравнения для затворной линии во временной области. Постоянная распространения и волновое сопротивление линии определяются ее параметрами в поперечной плоскости.

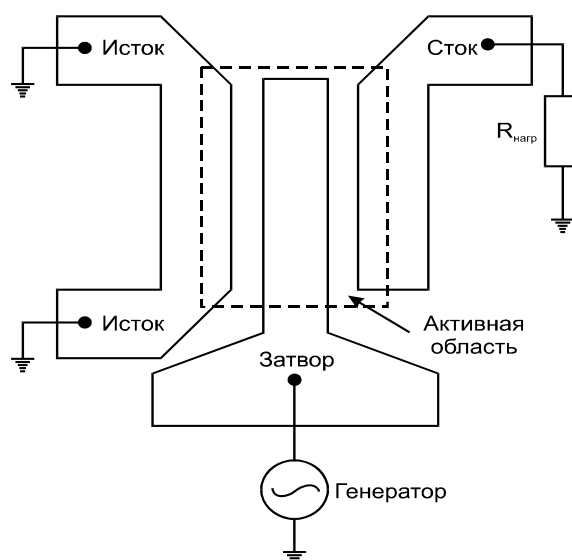


Рис. 4.33. Затворная линия передачи транзистора

На рис. 4.34 C_{gs} , C_{gd} , C_{ds} – входная, проходная и выходная емкости соответственно; I_{gd} , I_{gs} – токи через диоды затвор-сток и затвор-исток соответственно; I_{drain} – зависимый источник тока, описывающий нелинейную зависимость тока стока от напряжения на затворе V_{gs} и стоке V_{ds} ; R_s – сопротивление контакта истока; R_{go} , L_{go} и R_{do} – погонные параметры затворной и стоковой линии соответственно.

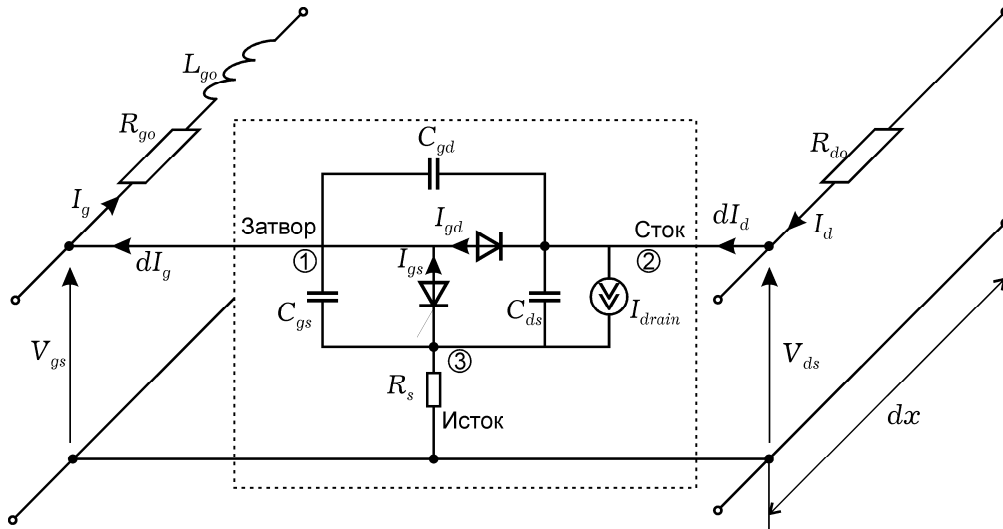


Рис. 4.34. Нелинейная модель дифференциального отрезка активной линии передачи

Используя закон Кирхгофа для токов, получим систему уравнений, описывающих поведение схемы во временной области:

$$\begin{aligned}
 \text{узел 1:} \quad & I_{C_{gd}} - I_{gd} - I_{gs} + I_{C_{gs}} + I_{R_g} = 0; \\
 \text{узел 2:} \quad & I_{R_d} - I_{C_{ds}} - I_{C_{gd}} + I_{drain} - I_{R_n} + I_{gd} = 0; \\
 \text{узел 3:} \quad & I_{C_{ds}} + I_{gs} - I_{C_{gs}} - I_{drain} + I_{R_s} = 0.
 \end{aligned} \tag{4.53}$$

Используя зависимость $I = CdU/dt$, запишем токи через емкости:

$$\begin{aligned}
 C_{gd} \frac{dU_{gd}}{dt} - I_{gd} - I_{gs} + C_{gs} \frac{dU_{gs}}{dt} + I_{R_g} &= 0; \\
 I_{R_d} - C_{ds} \frac{dU_{ds}}{dt} - C_{gd} \frac{dU_{gd}}{dt} + I_{drain} - I_{R_n} + I_{gd} &= 0; \\
 C_{ds} \frac{dU_{ds}}{dt} + I_{gs} - C_{gs} \frac{dU_{gs}}{dt} - I_{drain} + I_{R_s} &= 0.
 \end{aligned} \tag{4.54}$$

Заменим напряжения на элементах разностями соответствующих узловых потенциалов:

$$U_{gs} = V_2 - V_4; \quad U_{gd} = V_3 - V_2; \quad U_{ds} = V_4 - V_3; \quad (4.55)$$

Подставив полученные значения (4.55) в (4.54), получим следующую систему уравнений:

$$\begin{aligned} & C_{gd} \frac{d(V_3 - V_2)}{dt} + C_{gs} \frac{d(V_2 - V_4)}{dt} = \\ & = I_{gd} \{V_3 - V_2\} + I_{gs} \{V_2 - V_4\} - I_{Rg}; \\ & -C_{ds} \frac{d(V_4 - V_3)}{dt} - C_{gd} \frac{d(V_3 - V_2)}{dt} = \\ & = I_{Rn} - I_{drain} \{V_2 - V_4, V_4 - V_2\} - I_{gd} \{V_3 - V_2\} - I_{Rd}; \\ & C_{ds} \frac{d(V_4 - V_3)}{dt} - C_{gs} \frac{d(V_2 - V_4)}{dt} = \\ & = I_{drain} \{V_2 - V_4, V_4 - V_2\} - I_{Rs} - I_{gs} \{V_2 - V_4\}. \end{aligned} \quad (4.56)$$

Поскольку узловые потенциалы V_1 и V_5 являются задающими источниками напряжения $U_{in}(t)$ и E_c , то система может быть представлена в виде:

$$\begin{aligned} & C_{gd} \frac{d(V_3 - V_2)}{dt} + C_{gs} \frac{d(V_2 - V_4)}{dt} = I_{gd} \{V_3 - V_2\} + I_{gs} \{V_2 - V_4\} - I_{Rg}; \\ & -C_{ds} \frac{d(V_4 - V_3)}{dt} - C_{gd} \frac{d(V_3 - V_2)}{dt} = I_{Rn} - I_{drain} \{V_2 - V_4, V_4 - V_2\} - I_{gd} \{V_3 - V_2\} - I_{Rd}; \\ & C_{ds} \frac{d(V_4 - V_3)}{dt} - C_{gs} \frac{d(V_2 - V_4)}{dt} = I_{drain} \{V_2 - V_4, V_4 - V_2\} - I_{Rs} - I_{gs} \{V_2 - V_4\}. \end{aligned}$$

В матричной форме:

$$\mathbf{M} \frac{d\mathbf{V}}{dt} = \mathbf{B}, \quad \text{где}$$

$$\mathbf{M} = \begin{bmatrix} C_{gs} - C_{gd} & C_{gd} & -C_{gs} \\ C_{gd} & C_{ds} - C_{gd} & -C_{ds} \\ -C_{gs} & -C_{ds} & C_{gs} + C_{ds} \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} I_{gd} \{V_3 - V_2\} + I_{gs} \{V_2 - V_4\} - \frac{V_2 - U_{in}(t)}{R_g} \\ \frac{V_3}{R_n} - I_{drain} \{V_2 - V_4, V_4 - V_2\} - I_{gd} \{V_3 - V_2\} - \frac{V_5 - V_3}{R_d} \\ I_{drain} \{V_2 - V_4, V_4 - V_2\} - \frac{V_4}{R_s} - I_{gs} \{V_2 - V_4\} \end{bmatrix}.$$

Решение данной системы связано с необходимостью учета плохо обусловленной матрицы коэффициентов \mathbf{M} . В этом случае необходимо использовать адаптивный подход, связанный с анализом собственных частот системы [4.35].

Для анализа временных характеристик вдоль затворной линии передачи (вдоль третьей координаты транзистора) и определения распределения токов и напряжений вдоль координаты x необходимо рассмотренную систему уравнений, позволяющую найти токи и напряжения в поперечной к направлению распространению сигнала плоскости, дополнить системой волновых уравнений. Для отрезка линии можно записать систему телеграфных уравнений во временной области с погонными параметрами затворной линии в виде:

$$\begin{aligned} -\frac{\partial V}{\partial x} &= Ri + L \frac{\partial i}{\partial t}, \\ -\frac{\partial i}{\partial x} &= GV + C \frac{\partial V}{\partial t}. \end{aligned}$$

Здесь параметры R и L определяются сопротивлением и индуктивностью металлизации затвора, а G и C соответственно выходной проводимостью и входной емкостью транзистора.

Разделение переменных в телеграфных уравнениях достигается повторным дифференцированием с перекрестной подстановкой и переходом к волновым уравнениям для тока и напряжения:

$$\nabla^2 V - \frac{1}{v^2} \frac{\partial^2 V}{\partial t^2} = 0; \quad \nabla^2 i - \frac{1}{v^2} \frac{\partial^2 i}{\partial t^2} = 0.$$

Если, в простейшем случае, сигнал на входе гармонический или если он представим в виде конечного числа гармоник, то задачу можно свести к решению волновых уравнений в комплексной области для каждой из гармоник.

4.6.3. Моделирование размерных эффектов и технологических неоднородностей

В реальных структурах (активных и пассивных) имеют место неоднородности, возникающие вследствие технологических погрешностей их изготовления. Подобный анализ актуален, так как при субмикронных размерах технологические неоднородности могут быть сравнимы с размерами структуры (рис. 4.35). Учесть их можно, например, представляя транзистор в виде некоторой дискретной структуры вдоль третьей координаты (ширины затвора).

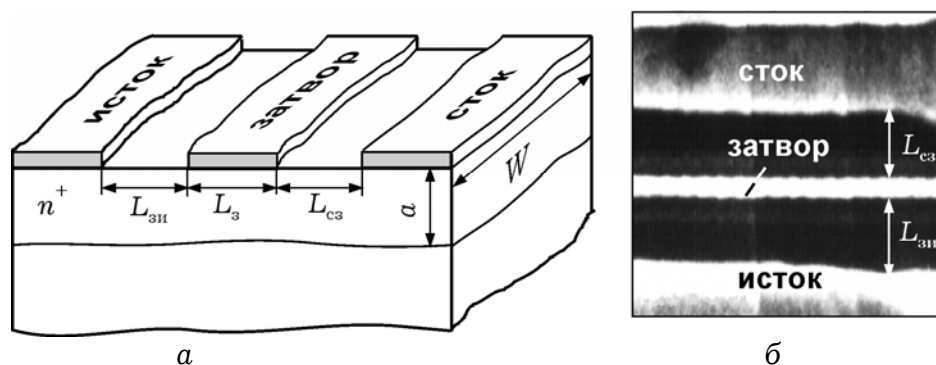


Рис. 4.35. Топология транзистора с технологическими неоднородностями (а) и фрагмент (б) фотографии на рис. 4.1б.

Тогда каждую из секций можно будет описывать своим набором физико-топологических параметров и соответственно, Y -матрицей, т. е. структуру можно представить в виде параллельно соединенных четырехполюсников:

$$Y = \sum_{i=1}^n [Y_i].$$

Параметры (геометрические, физические) в каждой секции (рис. 4.36) можно задавать либо детерминированной функцией, либо случайным образом (по соответствующему закону распределения – нормальному, равномерному и др., используя метод Монте-Карло). После получения результирующей Y -матрицы высокочастотные параметры рассчитываются известными методами теории электронных цепей.

Актуальным также является анализ результатов теоретических и экспериментальных исследований влияния морфологии омических контактов на СВЧ параметры GaAs полевых транзисторов с барьером Шоттки [4.36, 4.37].

С повышением степени интеграции монокристаллических интегральных схем, а также расширением рабочего частотного диапазона полевых транзисторов с барьером Шоттки в

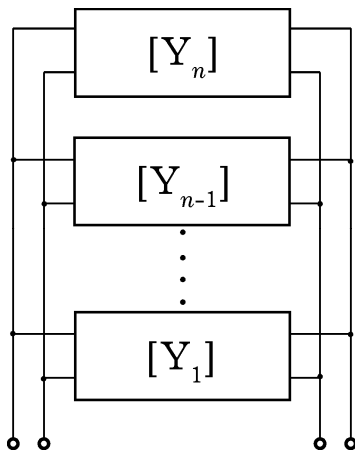


Рис. 4.36. Представление транзистора в виде параллельных четырехполюсников

миллиметровую область элементы их топологии приближаются к субмикронным размерам. При этом к пассивным областям транзистора предъявляются все более жесткие требования. Это относится к величине удельного контактного сопротивления ρ_k , тепловой стабильности, радиационной стойкости, а также повторяемости геометрических размеров структуры в процессе изготовления. Растут требования к морфологии поверхности контактов, однородности распределения электрических параметров контактов по его плоскости. Этим обусловлена необходимость усовершенствования технологии, включая омические контакты, а также моделирование размерных эффектов и физико-топологических неоднородностей в ПТШ, обусловленных структурой самих контактов.

Для исследования влияния технологических неоднородностей использовалась тестовая структура ПТШ с топологией и геометрическими параметрами, приведенными на (рис. 4.28).

Транзисторы изготавливались на эпитаксиальных структурах GaAs, выращенных газовой эпитаксией металлоорганических соединений на полуизолирующей подложке с буферным слоем. Топология стока и истока в виде омических контактов и затвора в виде барьерного контакта формировались в процессе литографии со средним значением длины затвора в пределах 0,3 мкм и шириной 200 мкм. Такие размеры обеспечивались применением электронно-лучевой литографии для получения рабочей частоты транзистора около 12 ГГц.

Для исследования влияния параметров контактов, в частности, с различными размерами зерен – 1,0–1,5 мкм и около 3,0 мкм, использовались две тестовые структуры транзистора [4.36].

Для моделирования неоднородностей использовалась секционированная модель вдоль третьей координаты ПТШ. В основе такой модели лежит дискретизация структуры ПТШ вдоль ширины затвора W . Каждый из n элементарных транзисторов

их топологии приближаются к субмикронным размерам. При этом к пассивным областям транзистора предъявляются все более жесткие требования. Это относится к величине удельного контактного сопротивления ρ_k , тепловой стабильности, радиационной стойкости, а также повторяемости геометрических размеров структуры в процессе изготовления. Растут требования к морфологии поверхности контактов, однородности распределения электрических параметров контактов по его плоскости. Этим обусловлена

шириной ΔW имеет индивидуальный набор геометрических параметров: толщину эпитаксиального слоя, длину затвора, межэлектродные расстояния исток-затвор и затвор-сток, толщину слоя металлизации затвору. Подобные модели использовались для исследования влияния технологических неоднородностей структур с размерами затвора 0,5 мкм [4.37–4.39].

С уменьшением длины затвора влияние неоднородностей становится более заметным, однако, в силу малости размеров это предполагает совершенствование методов исследования. В частности, исходная система уравнений для анализа нестационарных эффектов должна быть дополнена релаксационными уравнениями для канала и подложки в приближении времени релаксации.

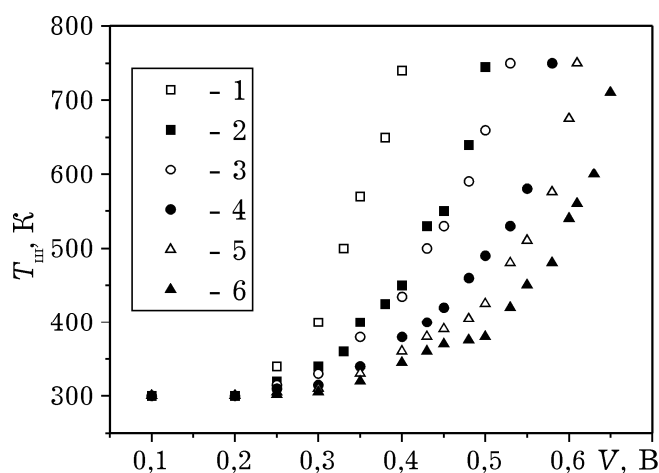


Рис. 4.37. Зависимость шумовой температуры канала от напряжения на активной области образца на частоте 10 ГГц при температуре 300 К: 1, 2 – $L_з = 0,2$ мкм, 3, 4 – $L_з = 0,6$ мкм, 5, 6 – $L_з = 1,0$ мкм; 1, 3, 5 – стандартная технология, 2, 4, 6 – улучшенная.

Экспериментальные зависимости, приведенные на рис. 4.37 и 4.38 для транзисторов с различной длиной затвора для комнатной температуры кристаллической решетки и при охлаждении до 80 К, подтверждают как зависимость шумовой температуры ПТШ от электрического режима, так и от неоднородностей контактов, связанных с их морфологией.

Это подтверждает необходимость учета неоднородностей субмикронных структур при моделировании для поиска путей оптимизации параметров и достижения минимального коэффициента шума. Кроме адекватного моделирования технологических неоднородностей, необходимо определять оптимальные геометрические и электрофизические параметры (межэлектродные расстояния, толщины слоев и их концентрации и др.) и электрический режим.

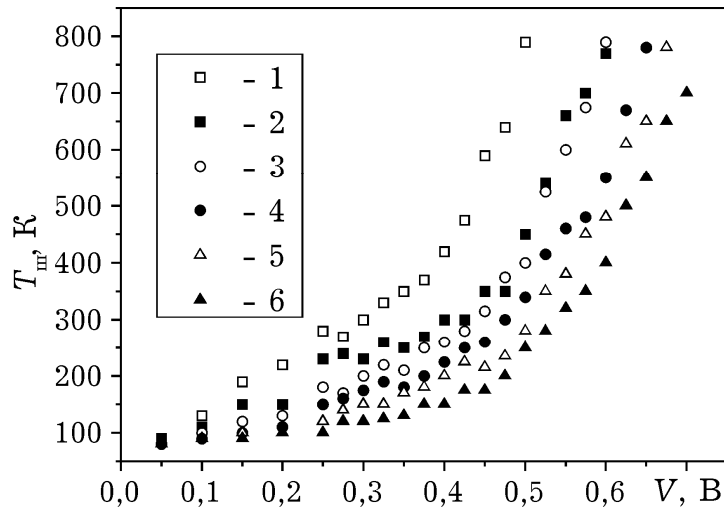


Рис. 4.38. Зависимость шумовой температуры канала от напряжения на активной области образца на частоте 10 ГГц при температуре 80 К: 1, 2 – $L_3 = 0,2$ мкм, 3, 4 – $L_3 = 0,6$ мкм, 5, 6 – $L_3 = 1,0$ мкм; 1, 3, 5 – стандартная технология, 2,4,6 – улучшенная.

Методика моделирования опробована на исследовании влияния на параметры ПТШ неоднородностей сопротивления истока и стока, обусловленных разным удельным контактным сопротивлением и глубиной проплавления омических контактов, что определяется морфологией. При этом параметры активной области транзистора выбирались фиксированными, т.е. задавались как некоторые средние эффективные значения.

Транзистор с шириной затвора W представляется в виде эквивалентного четырехполюсника, который описывается мало сиг-нальной схемой моделью с сосредоточенными параметрами, рассчитанными на основе физико-топологической модели, учитывающей нестационарные эффекты. Параметры четырехполюсника в виде матрицы проводимости и матрицы рассеяния рассчитываются известными методами анализа электронных схем. Очевидно, что полная структура транзистора представляет собой цепочку параллельно соединенных четырехполюсников. Шумовые параметры отдельных секций и транзистора в целом моделировались на основе матрицы спектральных плотностей мощности шума (τ -матрицы) при условии аддитивности шумов структуры и предположении отсутствия корреляции между шумами отдельных секций. Параметры, которые характеризуют неоднородности каждой секции, задаются детерминированной функцией, если известен закон их изменения вдоль ширины затвора, или же как случайные величины методом Монте-Карло.

Длина выборки случайных чисел по заданному закону распределения определяется количеством разбиений структуры транзистора, т. е. количеством секций. Результирующая матрица проводимости цепочки представляет собой сумму матриц проводимости элементарных секций. Суммарная Y -матрица по известному соотношению пересчитывается в матрицу рассеяния:

$$[S] = ([1] - [Y])([1] + [Y])^{-1}.$$

Такой подход реализован в программе моделирования ПТШ с учетом топологических неоднородностей. В качестве примера для расчетов исследовались структуры серийных ПТШ с параметрами: $L_{из} = 0,8 \div 3,8$ мкм – расстояние исток-затвор; $L_з = 0,3$ мкм – среднее значение длины затвора; $L_{зс} = 0,8 \div 3,8$ мкм – расстояние затвор-сток; $h = 0,6$ мкм – высота металлизации затвора; $a = 0,11$ мкм – среднее значение толщины канала; $\rho_k = (10^4 \div 10^7)$ Ом·см² – удельное сопротивление омических контактов; $b = 0,1 \div 0,3$ мкм – глубина проплавления контактов.

Следует отметить, что значительный разброс (более 50%) величин $L_{из}$ и $L_{зс}$ объясняется тем, что в условиях неоптимизированной технологии формирования омических контактов размер отдельных зерен, рекристаллизованных из расплава, может достигать 3 мкм. В результате сопротивление как со стороны истока, так и со стороны стока, может возрасти больше чем на порядок. Количество секций выбиралось таким образом, чтобы ширина каждой равнялась не более 2 мкм, т.е. цепочка состояла не меньше чем из 100 элементарных транзисторов.

На рис. 4.39 приведены результаты расчетов коэффициента усиления по мощности K_{yP} и температуры шума $T_{ш}$ ПТШ на частоте 12 ГГц без учета и с учетом неоднородностей с максимальной дисперсией 50%, а также результаты измерений среднестатистических значений этих параметров на транзисторах, в которых сток и исток в виде омических контактов сформированы по типичной серийной технологии и по улучшенной технологии, где размеры рекристаллизованных зерен не превышают 0,1 мкм. Как видно из рисунка 4.39, K_{yP} уменьшается на 0,8 дБ, а $T_{ш}$ возрастает на 23 К, если использовать предложенную модель и учесть флуктуации в пассивных областях ПТШ с приведенной выше дисперсией параметров.

Эти расчеты в значительной мере совпадают с результатами СВЧ измерений транзисторов на рабочей частоте 12 ГГц. Так, для транзисторов с улучшенной морфологией наблюдается увеличение K_{yP} на 0,6 дБ и уменьшение $T_{ш}$ на 13 К при $T_0 = 100$ К и на 10 К

при $T_0 = 340$ К соответственно, где T_0 – температура, при которой проводились измерения СВЧ параметров.

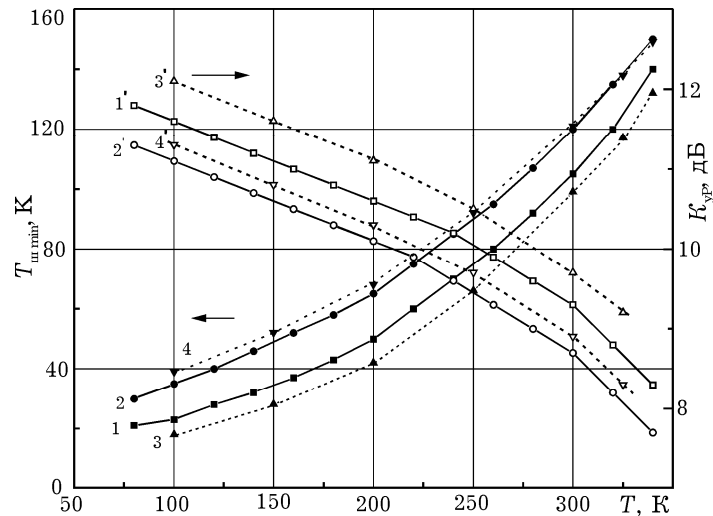


Рис. 4.39. Температурные зависимости $K_{ур}$ и $T_{ш}$ на частоте 12 ГГц: 1, 1' – ПТШ с улучшенной морфологией омических контактов; 2, 2' – серийный ПТШ, изготовленный по серийной технологии; 3, 3' – рассчитанные параметры модели ПТШ без учета неоднородностей; 4, 4' – с учетом неоднородностей.

Следует заметить, что рассчитанные параметры ПТШ по предложенной модели без учета влияния неоднородностей лучше, чем у реального транзистора. Это легко объясняется тем, что даже при максимально усовершенствованной технологии в реальных приборах влияние неоднородностей будет иметь место. Кроме этого, при усовершенствованной технологии пассивных областей субмикронных структур всегда имеет место влияние физико-топологических неоднородностей в активной области ПТШ.

Проведенные исследования показывают также, что температура шума $T_{ш}$ со снижением температуры от 300 до 80 К уменьшается примерно в 4 раза для типичных транзисторов и в 5 раз для транзисторов с улучшенной морфологией стока и истока и достигает меньше 20 К. Величина $K_{ур}$ при этом увеличивается примерно на 2,5 дБ. Аналогичный характер температурных зависимостей $K_{ур}$ и $T_{ш}$ подтверждается результатами моделирования.

В значительной мере это можно объяснить тем, что в ПТШ с улучшенной морфологией омических контактов при охлаждении за счет уменьшения влияния термополевой эмиссии ρ_k будет иметь меньшее влияние на параметры прибора.

Литература к главе 4

- 4.1. *Shockley W.*, A unipolar field-effect transistor, Proc. IRE, 40, 1365 (1952)
- 4.2. *Минаков В.В., Москалюк В.А., Касиян В.Т.* Программа моделирования полупроводниковых приборов субмикронных размеров многочастичным методом Монте-Карло, Электронная техника, сер. 1, электроника СВЧ, 1988, вып.8(412) с.71-73.
- 4.3. *M.Littlejohn, J.Rausser, T.Glisson.* Velocity field characteristics of GaAs with $\Gamma_6^c - L_6^c - X_6^c$ conduction-band ordering // J.Appl.Phys.-1977.-V.48.-№11.-P.4587-4590.
- 4.4. *K.Brennan, K.Hess.* High field transport in GaAs, InP and InAs // Solid State Electronics.-1984.-V.27.-№27.-P.347-350.
- 4.5. *Carnez B., Cappy A., Kaszynski A., Constant E., Salmer G.* Modeling of a Submicrometer Gate Field-Effect Transistor Including Effects of Nonstationary Electron Dynamics // J. Appl. Phys.-1980.-Vol.51-№1.-P.784-790.
- 4.6. *Кальфа А.А., Пашковский А.Б., Тагер А.С.* Исследование вольтамперных характеристик полевых транзисторов с субмикронным затвором // Электронная техника, сер. «Электроника СВЧ». -Вып.4. -1984. -С.27-30.
- 4.7. *Cook R.K., Frey J.* An Efficient Technique for Two-Dimensional Simulation of Velocity Overshoot Effects in Si and GaAs Devices. COMPEL-The International Journal for Computation and Mathematics in Electrical and Electronic Engineering.-1982.-Vol.1.-№2.-P.65-87.
- 4.8. *Максименко А.Г., Москалюк В.А., Тимофеев В.И.* О моделировании субпорогового режима ПТШ // Вестник КПИ, сер. Радиоэлектроника. -1992.- N29.-С.15-18.
- 4.9. *Москалюк В.А., Тимофеев В.И.* Модель полевого транзистора с барьером Шоттки при открытом затворе // Вестник КПИ, сер. Радиоэлектроника. -1994.- N30.-С.7-10.
- 4.10. *Москалюк В.А., Тимофеев В.И.* Моделирование процессов переноса электронов в канале и подложке субмикронного полевого транзистора // Электронная техника. Сер. Электроника СВЧ. - 1990.-Вып. 2 (426).- С. 22-27.
- 4.11. *Москалюк В. А., Минаков В. В., Тимофеев В.И.* Анализ моделей субмикронных полевых транзисторов с позиций адаптивного проектирования // Изв. ВУЗов. Сер. Радиоэлектроника. - 1988. -Т.31.-N6.-С.64-72.
- 4.12. *Москалюк В.А., Тимофеев В.И.* Моделирование малощумящего режима субмикронного ПТШ // Изв. ВУЗов. Сер. Радиоэлектроника. - 1993.- Т.36. - N3. - С.57-63.

- 4.13. *Москалюк В.А., Тимофеев В.И.* Моделирование малошумящего режима субмикронного полевого транзистора // Труды XII Всесоюз. конф. твердотельной электронике СВЧ.– Киев.–1990.–С.63–64.
- 4.14. *Москалюк В.А., Минаков В.В., Тимофеев В.И.* О роли подложки в субмикронных полевых транзисторах // Труды III Всесоюз.конф. "Математическое моделирование физических процессов в полупроводниках и приборах".–Вильнюс.–1989.– С.130–133.
- 4.15. *Москалюк В.А., Тимофеев В.И.* Моделирование субпорогового режима арсенидгаллиевых полевых транзисторов с барьером Шоттки // Труды II республиканского совещания "Численные методы и средства проектирования и испытания элементов твердотельной электроники".–Таллин.–1989.–С.89–91.
- 4.16. *Москалюк В.А., Тимофеев В.И.* Особенности работы и моделирования субмикронных полевых транзисторов в малошумящем режиме // Труды междунар. семинара "Актуальные вопросы разработки и производства средств приема спутникового телевидения"– Севастополь.–1990.–С.18.
- 4.17. *Москалюк В.А., Тимофеев В.И.:* Моделирование схемных параметров полевого СВЧ транзистора // Автоматизация проектирования в электронике.–1987.–Т.36.–С. 3–11.
- 4.18. *Сигорский В. П., Петренко А. И.,* Алгоритмы анализа электронных схем. изд. 2-е, М., «Сов. радио»,. 1976, 608 с.
- 4.19. *Жалуд В., Кулешов В.Н.* Шумы в полупроводниковых устройствах. / Под редакцией А.К. Нарышкина. –М.: Сов. Радио,1977. – 416 с.
- 4.20. *Щепеткин Ф.В., Данич Ю.С.* Шумовые свойства транзисторного усилителя дециметрового диапазона // Электросвязь.–1973.–№2.–С.60–64.
- 4.21. *Москалюк В.А., Тимофеев В.И.* Особенности моделирования субмикронных арсенидгаллиевых полевых транзисторов для цифровых ИС // Труды II республиканского совещания "Численные методы и средства проектирования и испытания элементов твердотельной электроники".–Таллин.–1989.–С.86–88.
- 4.22. *Москалюк В.А., Тимофеев В.И.* Моделирование субмикронных арсенид–галлиевых полевых транзисторов для сверхскоростных интегральных схем // Труды I Всесоюз. конф. "Компьютерные методы исследования проблем теории и техники передачи дискретных сигналов по радиоканалам."–Москва.–"Радио и связь"–1990.–С.59–60.
- 4.23. *Москалюк В.А., Тимофеев В.И.* Модели полевых транзисторов для сверхскоростных и сверхвысокочастотных интегральных

схем // Труды Всесоюз. конф. "Математическое моделирование в энергетике" – Киев. – 1990. – С.161–162.

4.24 *Москалюк В.А., Тимофеев В.И.* Физико-топологическое моделирование ПТШ для режима большого сигнала // Труды Междунар. конф. "Проблемы автоматизированного моделирования в электронике." – Киев. – 1994. – С.195–199.

4.25. *Москалюк В.А., Тимофеев В.И.* О возможностях квазидвумерной модели для нелинейного моделирования субмикронных арсенидгаллиевых ПТШ // Труды XII Всесоюз. конф. твердотельной электронике СВЧ. – Киев. – 1990. – С.18–19.

4.26. *Тимофеев В.И.* Идентификация параметров субмикронных ПТШ для сверхбыстродействующих интегральных схем // Труды Междунар. семинара "Теория идентификации нелинейных динамических объектов". – г. Торговиште (Болгария). – 1990. – С.30–31.

4.27. *Москалюк В.А., Тимофеев В.И.* Физико-топологические модели ПТШ для режима большого сигнала // Труды Всеросс. конф. с междунар. участием "Разработка и применение САПР ВЧ и СВЧ электронной аппаратуры." – Владимир. – 1994. – С.105–106.

4.28. *Liou L.L., Ebel J.L., Huang C.I.* Thermal effects on the characteristics of AlGaAs/GaAs heterojunction bipolar transistors using two-dimensional numerical simulation // IEEE Transactions on electron devices. – 1993. – Vol.40, №1. – P.35–43.

4.29. *Тимофеев В.И.* Двумерное численное моделирование субмикронных полевых структур // Электроника и связь». – 2004, №22. – С.137–140.

4.30. *Москалюк В.А., Тимофеев В.И., Шовкун И.Д.* Распределенная модель СВЧ полевого транзистора // Изв. вузов. Сер. Радиоэлектроника. – 1986. – Т. 29. – № 6. – С. 87–89.

4.31. *Григорук Н.В., Приходько А.А., Тимофеев В.И.* Анализ переходных процессов в арсенид-галлиевых структурах с распределенными параметрами // Электроника и связь. – 1998. – №4. – С.359–361.

4.32. *Григорук А.А., Приходько Н.В., Тимофеев В.И.* Сверхширокополосные преобразователи миллиметрового диапазона для систем оптоволоконной связи // Электроника и связь. – 1999. – №.6. – Т.2 – С.256–264.

4.33. *Приходько Н.В., Тимофеев В.И.* Численный анализ полевых структур с распределенными параметрами во временной области // Электроника и связь. – 2000. – №8. – С.173–174.

4.34. *Chu Q.-X., Lau Y.-Pat, Chang F.-Y.* Transient analysis of microwave circuits based on time-domain characteristic models // IEEE Trans. on Microwave Theory and Techniques. – 1998. – Vol. 46, – №. 8, P.1097–1104.

- 4.35. Григорук А.А., Тимофеев В.И. Верификация нелинейных моделей электронных цепей СВЧ на классе жестко-устойчивых численных методов // Изв. ВУЗов. Радиоэлектроника.-1998.- № 1.- С.41-51.
- 4.36. Иващук А.В., Тимофеев В.И. Вплив морфології омичних контактів на надвисокочастотні параметри польових транзисторів // Оптоэлектроника и полупроводниковая техника.- 2000.-№35.- С.157-161.
- 4.37. Иващук А.В., Матуленіс А., Тимофеев В.И. Вплив морфології омичних контактів на шуми гарячих електронів в транзисторних структурах // Электроника и связь.- 1999.-№7.- С.96-99.
- 4.38. Москалюк В.А., Тимофеев В.И. Секционированная модель полевого СВЧ транзистора, учитывающая распределенные эффекты // Автоматизация проектирования в электронике.- 1986.- Вып.33.- С.93-97.
- 4.39. Тимофеев В.И., Андриец В.В. Шумы полевого СВЧ транзистора с топологическими неоднородностями // Вестник Киевского политехнического института. Радиоэлектроника.- 1987.- Вып.23.- С.1671-1708.

5. ГЕТЕРОСТРУКТУРНЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ

5.1. ОБЩИЕ СВОЙСТВА

Улучшение параметров ПТШ может быть достигнуто увеличением концентрации носителей в канале за счет повышения уровня легирования. Но при этом из-за примесного рассеяния падает подвижность и соответственно скорость носителей. Выход из этой ситуации может дать использование гетеропереходов. Основная идея при этом состоит в том, чтобы пространственно разделить области легирования с ионизированными донорными атомами и порожденных ими электронов, что должно существенно снизить примесное рассеяние и повысить подвижность. Одним из способов реализации этой идеи является использование гетероконтакта широкозонного высоколегированного полупроводника с нелегированным узкозонным. Различие энергий электронного сродства приводит к появлению разрыва дна зоны проводимости и, как следствие, к переходу электронов из широкозонного материала в нелегированный узкозонный. Наличие пространственного заряда электронов вызывает изгиб зоны проводимости и повышение концентрации электронов в образовавшейся потенциальной яме. Наиболее изученными и технологически отработанными являются гетероструктуры GaAs/AlGaAs. Подвижность электронов, перешедших в слой узкозонного материала и образующих там двумерный электронный газ, достигает при комнатной температуре значений, характерных для чистого материала (рис. 5.1).

Большая концентрация электронов в тонком слое, высокая подвижность, а также возможность управления концентрацией носителей в канале способствовали быстрой разработке нового типа полевых транзисторов на двумерном электронном газе. Первые полевые транзисторы, использующие гетеропереход GaAs/AlGaAs, с крутизной почти в три раза большей, чем крутизна аналогичного GaAs ПТШ были описаны в 1980 г. [5.2]. Эти транзисторы получили название *HEMT* (*High Electron Mobility Transistor*). В литературе встречаются также названия: *TEGFET* (*Two Dimensional Electron FET*) или *SDHFET* (*Selectively Doped*

Heterostructure FET). Последнее время чаще используется более общее название полевых транзисторов на основе гетероструктур ГСПТ (гетероструктурный полевой транзистор) или *HFET* (*Heterostructure Field Effect Transistor*). Схематическое изображение такого транзистора дано на рис. 5.2.

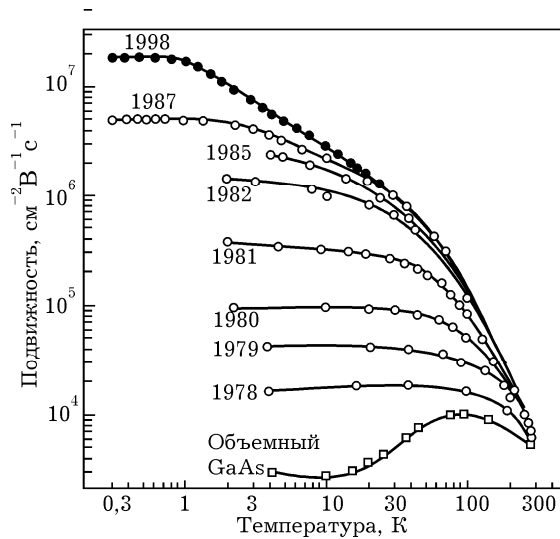


Рис. 5.1. Прогресс, достигнутый за 20 лет в повышении подвижности электронов в двумерной электронной системе модулированно-легированных гетероструктур GaAs/AlGaAs [5.1].

Основным структурным «элементом» в ГСПТ на основе GaAs является область двумерного электронного газа (2ДЭГ). Она формируется в зоне гетероперехода между барьерным $\text{Al}_x\text{Ga}_{1-x}\text{As}$ -слоем и нелегированным канальным GaAs-слоем. Ширина запрещенной зоны у $\text{Al}_x\text{Ga}_{1-x}\text{As}$ выше, чем у GaAs (рис. 2.14), и в области гетероперехода граница зоны проводимости GaAs может оказаться ниже уровня Ферми — образуется потенциальная яма (рис. 5.2, г).

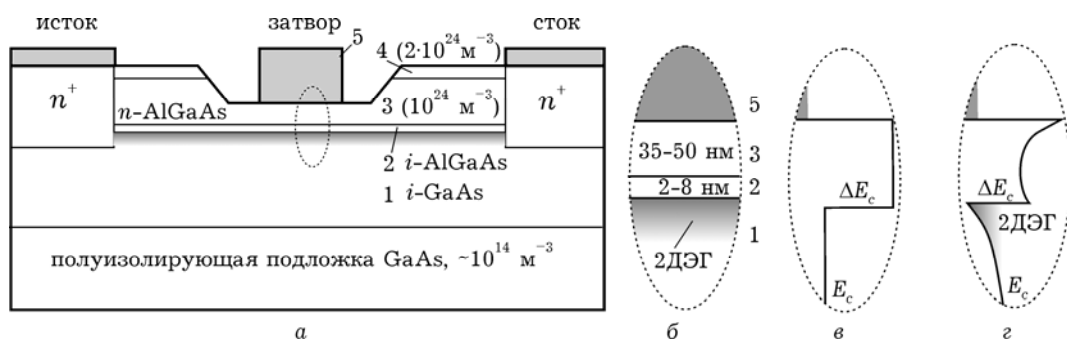


Рис. 5.2. Поперечное сечение ГСПТ (а) с характерными толщинами слоев (б): 1 — нелегированный GaAs, 2 — спейсер — нелегированный слой AlGaAs, 3 — барьерный слой — n AlGaAs, 4 — контактный слой — n^+ GaAs. Потенциальный профиль дна зоны проводимости исходных материалов (в) и в равновесном состоянии (г).

В ГСПТ на основе GaAs барьерный слой легируют кремнием (т.е. получают высоколегированный n - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ -слой), чтобы обеспечить источник свободных электронов. Переход электронов в потенциальную яму из области барьерного слоя вызван тем, что из-за различной энергии электронного сродства GaAs и $\text{Al}_x\text{Ga}_{1-x}\text{As}$ в области их контакта образуется достаточно сильное электрическое поле, моделируемое обычно разрывом запрещенной зоны ΔE_c , инициирующее переход свободных электронов из $\text{Al}_x\text{Ga}_{1-x}\text{As}$ в потенциальную яму. Зачастую выбирают долю Al $x \approx 0,3$, при которой $\Delta E_c \approx 0,3$ эВ. Таким образом, в канальном GaAs-слое непосредственно под гетеропереходом образуется тонкий слой пространственного заряда с высокой плотностью и подвижностью свободных электронов, который и называют двумерным электронным газом.

Между барьерным и канальным слоем формируют так называемый спейсер (*spacer*) - слой высокоомного нелегированного AlAs или $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Основное назначение этого тонкого (1–3 нм) слоя – предотвратить рассеяние носителей в канале на границе с гетеропереходом. Верхняя часть барьерного слоя образует с металлом затвора барьер Шоттки, который, как и в обычном ПТШ, управляет проводимостью канала. Зачастую в барьерном слое формируют область с повышенной концентрацией Si (так называемое δ -легирование), n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$ -слой для улучшения омических контактов стока и истока и увеличения концентрации носителей в канале.

Структуры современных ГСПТ весьма разнообразны, они могут существенно различаться по числу слоев, типу их материала, степени легирования и другим параметрам. На рис. 5.3, а изображена структура двухпереходного транзистора, в котором канал образуется в слое нелегированного арсенида галлия толщиной 30...50 нм между двумя слоями более широкозонного n^+ - $\text{Al}_x\text{Ga}_{1-x}\text{As}$. В электронном газе, ограниченном потенциальными барьерами с двух сторон, квантовые свойства 2ДЭГ проявляются еще в большей степени, чем в предыдущей конструкции.

Перспективными для использования в последнее время считаются нитриды галлия и алюминия. Наиболее технологически освоены структуры на гетеропереходах GaN/ $\text{Al}_x\text{Ga}_{1-x}\text{N}$. В монокристаллическом нелегированном GaN подвижность электронов проводимости оказывается весьма высокой – около $0,2 \text{ м}^2/(\text{В}\cdot\text{с})$ при комнатной температуре. Разрыв дна зоны проводимости пары GaN и $\text{Al}_x\text{Ga}_{1-x}\text{N}$ существенно больше, чем в гетеропереходе GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$, и составляет около 0,8 эВ. Поэтому электростатическое поле в области контакта еще более

сильное, а концентрация свободных электронов, перешедших из $\text{Al}_x\text{Ga}_{1-x}\text{N}$ в потенциальную яму GaN, может быть существенно выше. Кроме того, в области контакта GaN и $\text{Al}_x\text{Ga}_{1-x}\text{N}$ (или AlN) из-за несовпадения периода их кристаллических решеток образуется механические напряжения. А поскольку кристаллы AlN и GaN обладают сильной степенью поляризации, возникает пьезоэффект, стимулирующий дополнительный переход носителей в канал. В результате концентрация электронов проводимости в области 2ДЭГ может превышать 10^{13} см^{-2} . Отметим, что пьезоэффект столь велик, что его одного зачастую достаточно для обеспечения необходимой концентрации носителей в канале. Так, если в ГСПТ на основе GaAs барьерный слой обязательно легируют, чтобы обеспечить источник свободных электронов, то в случае нитрида галлия возможно использовать и нелегированный барьерный слой $\text{Al}_x\text{Ga}_{1-x}\text{N}$. Однако обычно барьерный слой в большинстве GaN-структур легирован. Сочетание высоких подвижности и концентрации носителей в канале при широкой запрещенной зоне (около 3,4 эВ при комнатной температуре) обуславливает уникальные частотные и мощностные свойства GaN ГСПТ.

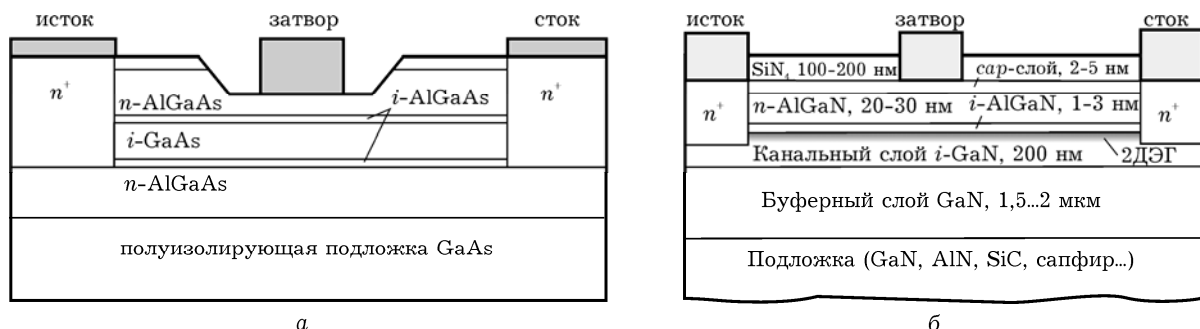


Рис. 5.3. Варианты структур двухпереходного арсенидгаллиевого ГСПТ (а) и гетероструктурного GaN/ $\text{Al}_x\text{Ga}_{1-x}\text{N}$ (б) транзистора.

Различных вариантов действующих транзисторов достаточно много. Обобщенная структура GaN ГСПТ представлена на рис. 5.3, б. Для минимизации влияния поверхностных состояний на ГСПТ поверх барьерного слоя формируют так называемый *cap*-слой (легированный либо нелегированный GaN, AlN, AlGaN), а также пассивирующий слой, например Si_3N_4 (последний характерен для любых полевых транзисторов). Канальный слой выращивают на толстом нелегированном или компенсированном буферном слое – высокоомном GaN-слое, в котором концентрация свободных носителей составляет менее 10^{13} см^{-3} , практически, как и в собственном полупроводнике. С технологической точки зрения, буферный слой обеспечивает при эпитаксии плавный переход от подложки к совершенной структуре канала. Зачастую структура

буферного слоя оказывается достаточно сложной – в ней формируют несколько различных слоев с обратными переходами, сверхрешетками и т.п. Основное требование к подложке, особенно для мощных транзисторов, – низкая электропроводность и высокая теплопроводность при приемлемых механических и структурных свойствах. Наиболее распространенные материалы подложек – сапфир, карбид кремния, кристаллический кремний с ориентацией решетки (111), а также GaN и AlN.

Возможности нитридгаллиевых транзисторов иллюстрирует рис. 5.4, на котором представлены стоковые характеристики [5.3], отличающиеся более высоким напряжением и плотностью тока стока.

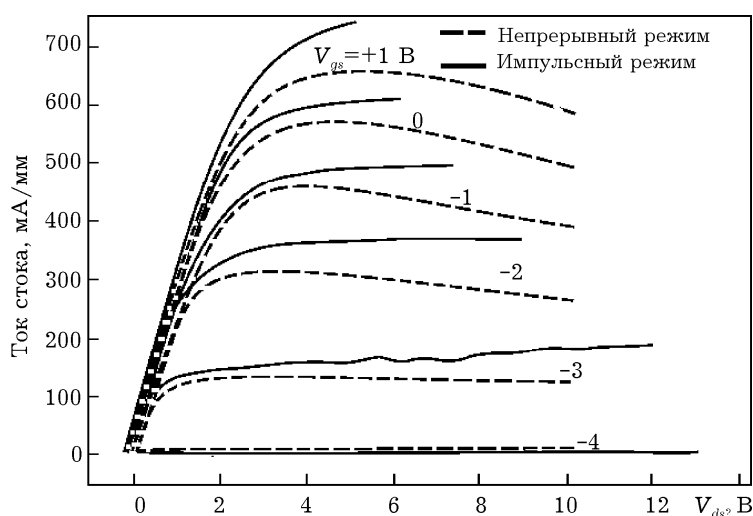


Рис. 5.4. Вольтамперные характеристики нитридгаллиевого ГСПТ, представленного на рис. 5.3,б, измеренные в импульсном (сплошные линии) и непрерывном (штриховые) режиме

5.2. КИНЕТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ГСПТ

Моделированию ГСПТ посвящено большое количество работ. Кинетические модели, позволяющие оценить характеристики прибора, рассматривались в [5.4, 5.5], в [5.6] приведена двумерная численная модель без учета квантовых эффектов. В работе [5.7] описана квазидвумерная гидродинамическая модель с учетом квантовых эффектов. Однако для моделирования нестационарных явлений переноса, происходящих на малых расстояниях в области сильного поля, необходим кинетический подход, основанный на самосогласованном решении кинетического уравнения Больцмана (1.12) и уравнения Пуассона.

Здесь приведены некоторые результаты моделирования с помощью программы, описанной в [5.8]. Для моделирования выбрана структура, аналогичная рис. 5.2, с длиной затвора 0,35

мкм. Слой $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ толщиной 50 нм равномерно легирован донорами с концентрацией 10^{14} см^{-3} . Вдоль границы гетероперехода расположен нелегированный спейсер толщиной 5 нм, под электродами стока и истока имеются n^+ -области, с помощью которых осуществляется контакт с двумерным электронным газом. При моделировании методом Монте-Карло использовалась трехдолинная Γ -L-X модель зонной структуры с параметрами материалов, взятыми из [5.9]. Из видов рассеяния учитывались фононные механизмы и рассеяние на ионизированных примесях в легированных областях. Граница гетероперехода GaAs/AlGaAs принималась резкой с разрывом зоны проводимости 0,3 эВ. Протекающий через электроды ток складывался из тока частиц, поглощаемых или инжектируемых данным электродом, и тока смещения. Ток каждого электрода определялся для всех временных шагов, что дает возможность отслеживать переходные характеристики прибора.

Примеры распределения потенциальной энергии и плотности электронного заряда в моделируемой области представлены на рис. 5.5. На верхнем рисунке следует обратить внимание на изменение вдоль канала формы потенциальной ямы, образованной «стенкой» разрыва дна зоны проводимости и пространственным зарядом электронов, перешедших в слой GaAs из легированного AlGaAs (рис. 5.5, б).

Распределение электронов в истоковой части канала локализовано у границы гетероперехода и, соответственно, в этой части потенциальная яма более крутая. По мере движения электронов к истоку под действием продольного греющего электрического поля их температура увеличивается. Поэтому, начиная с середины затвора, они глубже проникают в слой GaAs, и форма потенциальной ямы становится более полой.

Подобная информация содержится в распределении электронов по поперечному сечению (рис. 5.5, в). Здесь хорошо видна область обеднения под затвором, практически полностью лишенная подвижных носителей, а также обедненные области у границы гетероперехода со стороны AlGaAs. На этом рисунке также можно заметить начало образования параллельного паразитного транзистора в слое AlGaAs, однако протекающий через него ток мал, поскольку в модели нет омического контакта к слою AlGaAs.

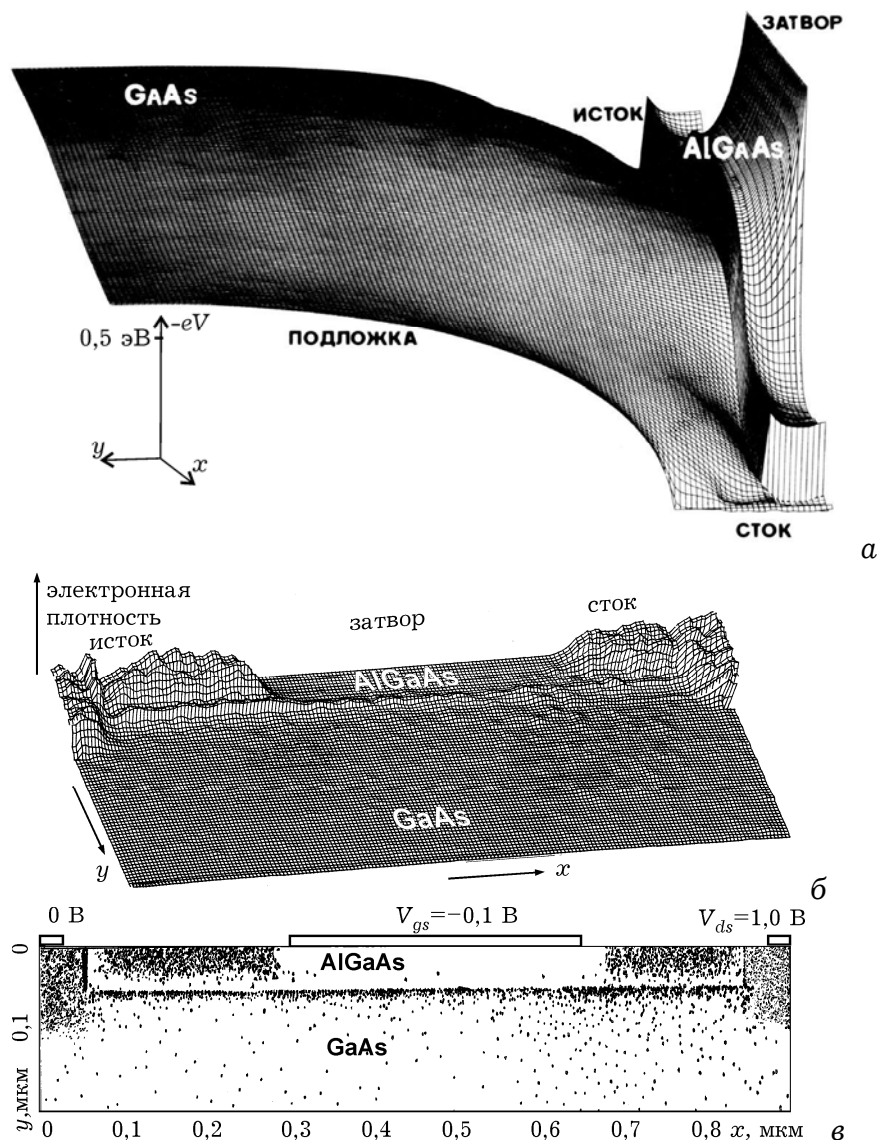


Рис. 5.5. Потенциальный рельеф дна зоны проводимости (а), распределение плотности заряда (б) и электронов (в) в поперечном сечении ГСПТ.

С точки зрения понимания особенностей физических процессов, протекающих в канале гетероструктурных транзисторов, интересны результаты моделирования, представленные на рис. 5.6. Напряженность электрического поля, что также следует из наклона потенциального рельефа (рис. 5.5,а), наибольшая в пространстве между затвором и стоком и превышает значение 80 кВ/см. При этом следует обратить внимание на график дрейфовой скорости. Во-первых, ее максимальное значение в полтора раза превышает статическое значение для GaAs, что является проявлением эффекта «всплеска», более сильного, чем в обычном ПТШ с такой же длиной затвора. Во-вторых, изменения скорости не следуют за изменением напряженности поля. Вблизи

стокового конца затвора скорость начинает падать, несмотря на резко возрастающую напряженность поля. Это объясняется тем, что энергия электронов здесь достигает значений около 0,3 эВ, достаточных для перехода электронов в L-, а затем и в X-долину. Большая эффективная масса электронов в верхних долинах приводит к уменьшению скорости дрейфа при возрастании их энергии. При этом характерно, что из-за большей плотности электронных состояний заселенность верхних долин может существенно превышать заселенность Γ -долины (рис. 5.6,б).

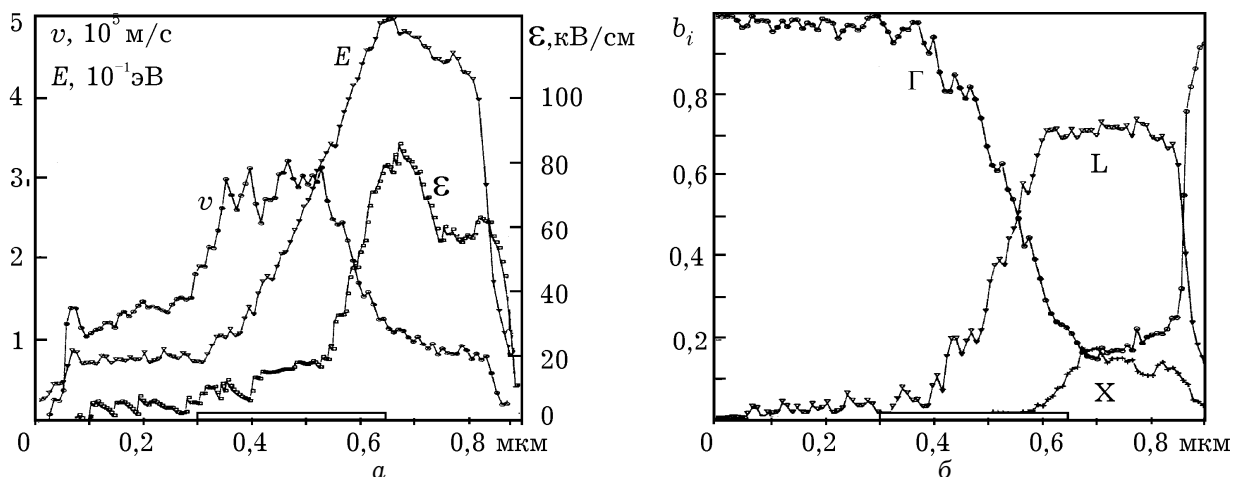


Рис. 5.6. Распределения (а) напряженности тянущего электрического поля (\mathcal{E}), дрейфовой скорости электронов (v) и их энергии (E), а также (б) распределения электронов по долинам вдоль канала транзистора.

У края затвора и напряженность поля и энергия начинают убывать, но заметной релаксации не наблюдается, поскольку электрическое поле, оставаясь сильным, удерживает большую часть электронов в верхних долинах. И лишь в n^+ -области стока происходит интенсивный возврат электронов в Γ -долину, а энергия и скорость уменьшаются до значений, близких к равновесным.

На рис. 5.7 представлены результаты моделирования и эксперимента стоковых характеристик транзистора с длиной затвора 0,35 μm и шириной 300 μm .

Малосигнальное значение крутизны прибора при стоковом напряжении 2 В и токе стока 60 мА составляет 137 мСм. На характеристиках незаметно значительного уменьшения крутизны при положительных смещениях на затворе, то есть образования параллельного паразитного транзистора в слое AlGaAs. Это связано, как уже отмечалось, с отсутствием в модели омического контакта к этому слою.

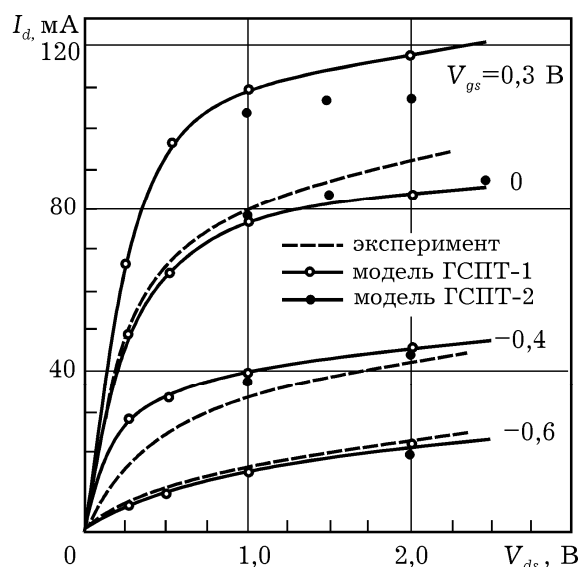


Рис. 5.7. Стоковые вольтамперные характеристики гетероструктурных транзисторов с одним гетеропереходом (ГСПТ-1) и двухпереходного (ГСПТ-2). Штриховая линия – результаты эксперимента для ГСПТ-1

О быстродействии транзистора можно судить по переходным характеристикам тока стока и тока затвора, представленным на рис. 5.8 и полученным при подаче на затвор ступеньки напряжения 0,3 В (а) и ступеньки напряжения на стоке $-1,0$ В (б).

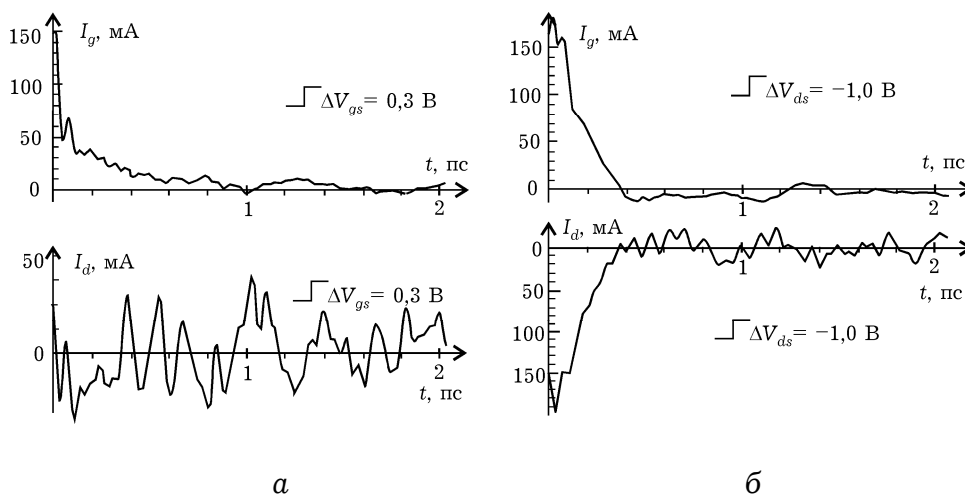


Рис. 5.8. Переходные характеристики токов электродов ГСПТ при ступенчатом изменении напряжения затвора (а) и напряжения стока (б)

Сам транзистор переключается довольно быстро, переходные процессы заканчиваются за время менее 1 пс. Однако существенное влияние на частотные свойства оказывают паразитные элементы. Для анализа их влияния по переходным характеристикам, используя преобразование Фурье, можно рассчитать параметры матрицы проводимости \mathbf{Y} , а затем и параметры матрицы рассеяния \mathbf{S} . На рис. 5.9 показаны

рассчитанные частотные зависимости S -параметров транзистора в рабочей точке $V_{ds} = 1,5$ В, $I_d = 30$ мА для внутренней активной части транзистора. Частотные метки приведены для частот 1 ГГц, 5 ГГц и далее с шагом 5 ГГц. Там же приведены зависимости S -параметров транзистора с учетом паразитных параметров, значения которых приведены на схемной модели (рис. 5.10).

Частотные зависимости параметра $|S_{21}|^2$, который определяет максимальный коэффициент усиления по мощности, позволяют оценить максимальную частоту прибора. Для внутреннего транзистора она составляет 62 ГГц, а для прибора с паразитными элементами 48 ГГц.

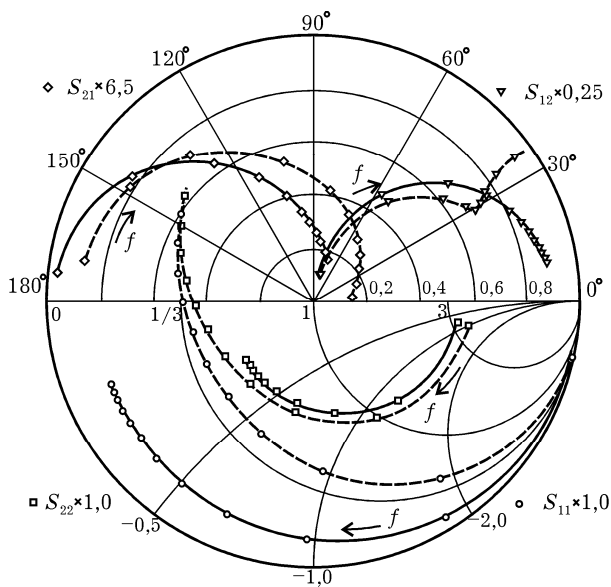


Рис. 5.9. Частотные зависимости S -параметров активной части транзистора (сплошные) и с учетом паразитных элементов (штриховые)

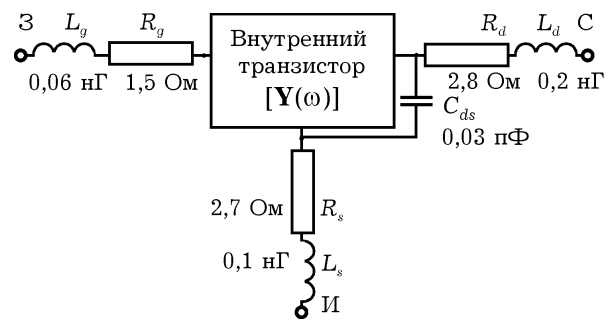


Рис. 5.10. Схемная модель ГСПТ с учетом паразитных элементов

На рис. 5.7 также представлены результаты моделирования ВАХ с двумя гетеропереходами (рис. 5.3, а) и аналогичными размерами затвора [5.10]. Потенциальный рельеф для такого двухпереходного транзистора представлен на рис. 5.11.

Сравнение вольтамперных характеристик транзисторов показывает, что наличие второго гетероперехода несколько снижает выходную проводимость, что особенно заметно при положительных потенциалах затвора. Это можно объяснить тем, что второй барьер локализует электроны в области узкозонного слоя GaAs, препятствуя их переходу в подложку из широкозонного AlGaAs и, тем самым, ослабляя ее «охлаждающую» и шунтирующую роль, как это имеет место в обычном ПТШ.

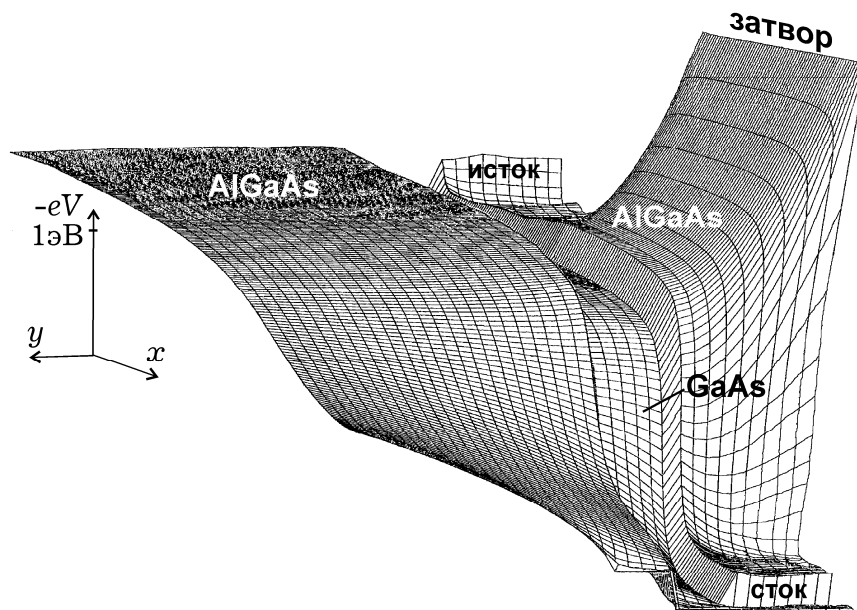


Рис. 5.11. Потенциальный рельеф для транзистора с двумя гетеропереходами, канал которого образован слоем узкозонного GaAs между AlGaAs

Квантовые эффекты. Из-за разрывов зоны проводимости и стимулированного этим накопления электронов и изгиба зоны у границы барьера образуется потенциальная яма (рис. 5.5). В двухпереходном транзисторе такая яма образуется между «стенками», образованными двумя переходами (рис. 5.11). В результате энергия электронов в таких ямах квантуется в направлении, нормальном к гетеропереходу (рис. 5.12), и образуются энергетические подзоны для продольного движения.

Поскольку математический аппарат квантовой теории переноса пока не позволяет моделировать реальные устройства, то наиболее часто для решения подобного рода задач используется приближение огибающей волновой функции. В соответствии с ним волновая функция двумерных электронов для m -ого состояния представляется в виде произведения плоской волны, параллельной гетеропереходу, и огибающей волновой функции $F_m(y)$ для поперечного направления:

$$\psi(\vec{r}) = F_m(y) A^{-1/2} \exp i(k_x x + k_z z). \quad (5.1)$$

Здесь A – нормировочная константа для плоскости xz . На рис. 5.12 представлены результаты решения уравнения Шредингера для огибающей волновой функции для потенциальных ям, характерных для подзатворной части в рассматриваемых транзисторах. Расстояния между рассчитанными уровнями

существенно превышают $k_B T$, что говорит о необходимости анализа квантовых эффектов в ГСПТ.

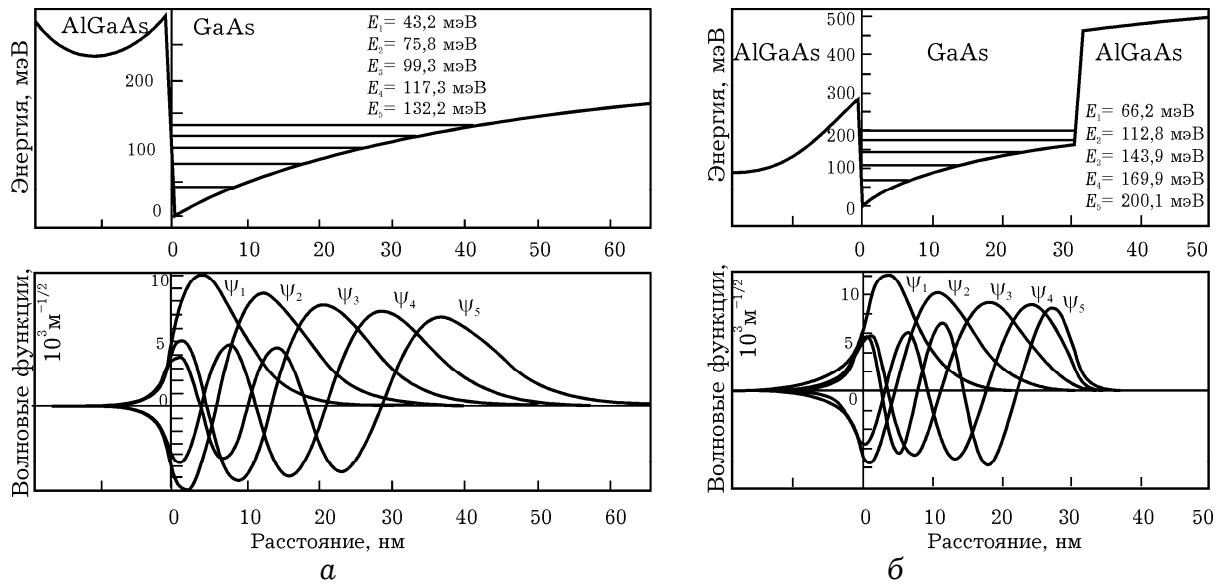


Рис. 5.12. Квантование энергетических состояний (верхний ряд) и соответствующие огибающие волновых функций (нижний ряд) для транзистора с одним (а) и двумя (б) гетеропереходами.

В рамках приближения огибающей волновой функции система уравнений для кинетического моделирования выглядит следующим образом:

кинетическое уравнение

$$\frac{\partial N(\vec{r}, \vec{k}, t)}{\partial t} + \vec{v} \cdot \frac{\partial N(\vec{r}, \vec{k}, t)}{\partial \vec{r}} - \frac{e\vec{\mathcal{E}}}{\hbar} \cdot \frac{\partial N(\vec{r}, \vec{k}, t)}{\partial \vec{k}} = \left[\frac{\partial N}{\partial t} \right]_c, \quad (5.2)$$

уравнение Пуассона

$$\nabla^2 V(\vec{r}) = -\rho(\vec{r}) / \epsilon \epsilon_0 \quad (5.3)$$

и уравнение Шредингера для огибающей волновой функции

$$\frac{\hbar^2}{2m^*} \frac{d^2 F_m(x, y)}{dy^2} + [U(x, y) - E_m(x)] F_m(x, y) = 0, \quad (5.4)$$

Здесь

$$\rho(\vec{r}) = e \left[N_d(x, y) - n(x, y) - \sum_{i=1}^k N_i(x, y) F_i^2(x, y) \right] \quad (5.5)$$

– объемная плотность заряда, включающая распределение концентрации ионизированных доноров $N_d(x, y)$, концентрацию «трехмерных» электронов $n(x, y)$ и в третьем слагаемом – заселенность i -го уровня $N_i(x, y)$, k – число учитываемых двумерных подзон; $U(x, y) = \Delta E_c(x, y) - eV(x, y)$ – потенциальная энергия электронов, определяемая распределением потенциала $V(x, y)$ и разрывом дна зоны проводимости $\Delta E_c(x, y)$.

В выражении (5.5) заряд «двумерных» электронов распределяется пропорционально пространственной плотности вероятности, т.е. квадрату амплитуд огибающих волновых функций, и заселенности уровней.

При решении такой системы либо на каждом шаге, либо через небольшое количество временных шагов самосогласованно решается система уравнений Шредингера и Пуассона, а затем кинетическое уравнение методом Монте-Карло, то есть интегрируются уравнения движения частиц в найденном электрическом поле. «Двумерные» электроны могут двигаться только в плоскости, параллельной гетеропереходу. Кроме системы уравнений (3.2) – (3.5) необходимо учитывать граничные условия при прохождении частицей гетероперехода. Классические (без учета туннелирования) граничные условия следуют из закона сохранения импульса, параллельного гетеропереходу, и энергии:

$$k'_{\parallel} = k''_{\parallel}, \quad (5.6)$$

$$\frac{(\hbar k'_{\parallel})^2}{2m^{*'}} + \frac{(\hbar k'_{\perp})^2}{2m^{*'}} = \frac{(\hbar k''_{\parallel})^2}{2m^{*''}} + \frac{(\hbar k''_{\perp})^2}{2m^{*''}} + \Delta E_c, \quad (5.7)$$

где $m^{*'}$ – эффективная масса электрона, k'_{\parallel} и k'_{\perp} – нормальная к гетеропереходу и ему параллельная составляющие волнового вектора в широкозонном материале; $m^{*''}$, k''_{\parallel} , k''_{\perp} – соответствующие величины в узкозонном материале.

При реализации полностью самосогласованной модели полевого гетеротранзистора возникают две основные трудности.

а) Форма потенциальной ямы $V(y)|_{x = \text{const}}$ меняется вдоль канала и, следовательно, волновые функции будут зависеть от координаты x , что усложняет решение системы уравнений Пуассона и Шредингера.

б) В двумерных ямах необходимо использовать соотношения для скоростей двумерного рассеяния. При этом имеется вероятность как внутривозонного, так и межвозонного

рассеяния, что значительно увеличивает размер матрицы скоростей рассеяния. Хотя для большинства механизмов рассеяния вероятность межподзонных переходов мала, для преобладающего при средних электрических полях полярного оптического рассеяния ей пренебречь нельзя. Еще одна трудность при учете двумерного рассеяния связана с тем, что матричные элементы зависят от формы огибающей волновой функции и определяются численным интегрированием. Это приводит к дополнительным вычислительным затратам, так как форма огибающей меняется в процессе решения задачи.

Однако при комнатных температурах и в сильных электрических полях, возникающих в реальных структурах в области затвор-сток, при напряжениях на стоке более 1 В, двумерные квантовые эффекты проявляются слабо [5.10]. К тому же при больших отрицательных напряжениях на затворе квантовая яма становится настолько широкой, что расстояние между уровнями квантования может стать меньше величины $k_B T$. Тем более, что разогрев электронов в сильном поле усиливает это неравенство и ослабляет проявление квантовых эффектов.

5.3. ДВУМЕРНОЕ МОДЕЛИРОВАНИЕ ГСПТ

Известны различные подходы к моделированию физических процессов в субмикронных гетеротранзисторах [5.11-5.14], составляющие часть иерархического ряда моделей – упрощенные аналитические модели [5.11] (учитывающие, например, бесстолкновительный дрейф носителей заряда при длинах затвора 0,15 мкм и температуре кристаллической решетки 77 К), одномерные для схемотехнических расчетов [5.12] и модели [5.13, 5.14], содержащие релаксационные уравнения сохранения энергии и импульса в квазидвумерном приближении. В последнем случае все уравнения интегрируются вдоль одной координаты в направлении тянущего поля, а размер области обеднения в приближении резкой границы на каждом шаге рассчитывается в зависимости от текущего значения потенциала в канале.

В упрощенных моделях гетероструктур используют аппроксимации зависимостей подвижности электронов от напряженности электрического поля. Эти уравнения представляют собой локальную модель токопереноса, в которой скорость (подвижность) носителей «локально» следует за изменениями напряженности электрического поля. В такой системе невозможно описывать нелокальные эффекты «всплеска» дрейфовой скорости электронов и дырок, эффекты накопления носителей вследствие изменения их энергии и эффективной массы при переходе в

верхние долины. В гетероструктурах, изготовленных на основе многодолинных полупроводников, несмотря на преимущественный дрейф электронов в двумерной потенциальной яме необходимо учитывать существенное влияние в сильных электрических полях междолинного рассеяния и различные условия разогрева и релаксации носителей в долинах [5.15, 5.16].

Структура моделируемого гетеротранзистора представлена на рис. 5.13. и характеризуется следующими физико-топологическими параметрами: длиной затвора 0,2 мкм, слоем широкозонного легированного $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ($x = 0,3$) под затвором, эффективной толщиной 0,1 мкм, слоем нелегированного $\text{Al}_x\text{Ga}_{1-x}\text{As}$ толщиной 50 Å (спейсер), слоем нелегированного GaAs на полуизолирующей подложке, толщиной в несколько десятков микрометров, в поверхностном слое которого в результате разрыва зон образуется потенциальная яма. Поскольку все электронные процессы связаны с поверхностным слоем и двумерным электронным газом (ДЭГ) в потенциальной яме, то для моделирования можно ограничиться эффективной толщиной транзистора 1-2 мкм.

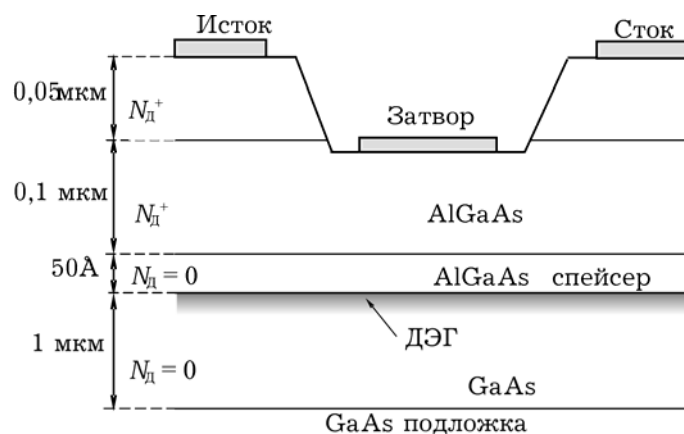


Рис. 5.13. Топологическая структура гетеротранзистора с двумерным электронным газом (ДЭГ)

Алгоритм моделирования и особенности его реализации приведены в приложении к разделу 5. На рисунках 5.14-5.17 приведены результаты двумерного моделирования основных физических характеристик в канале субмикронного гетеротранзистора. Затвор располагается симметрично относительно координаты $x = 0$.

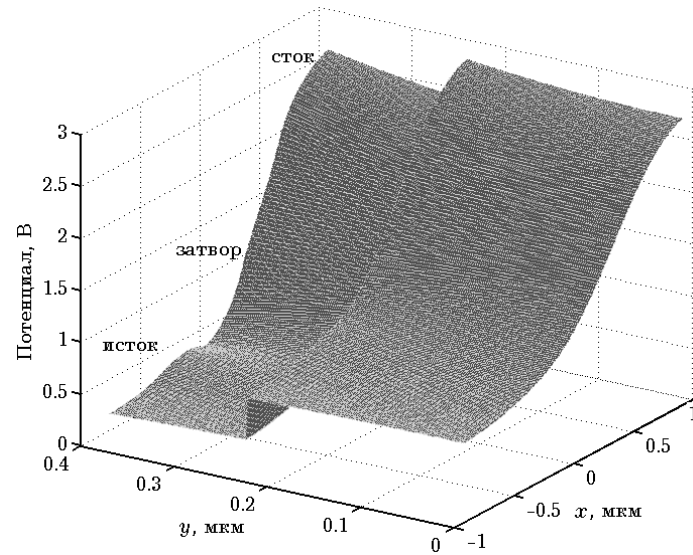


Рис. 5.14. Распределение потенциала в структуре гетеротранзистора

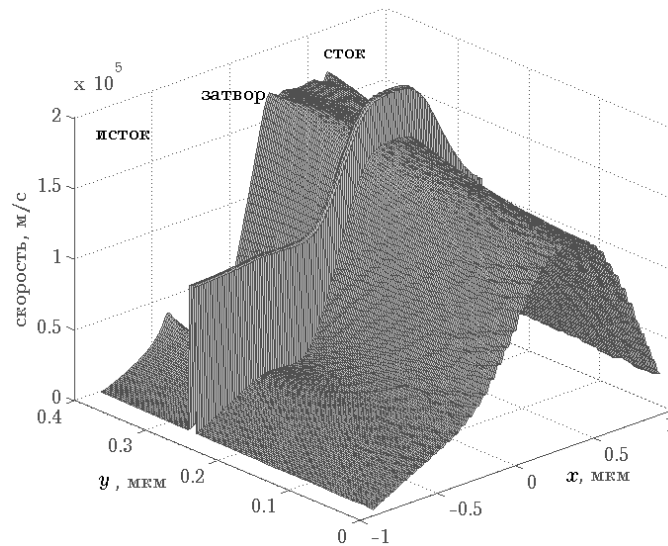


Рис. 5.15. Распределение дрейфовой скорости

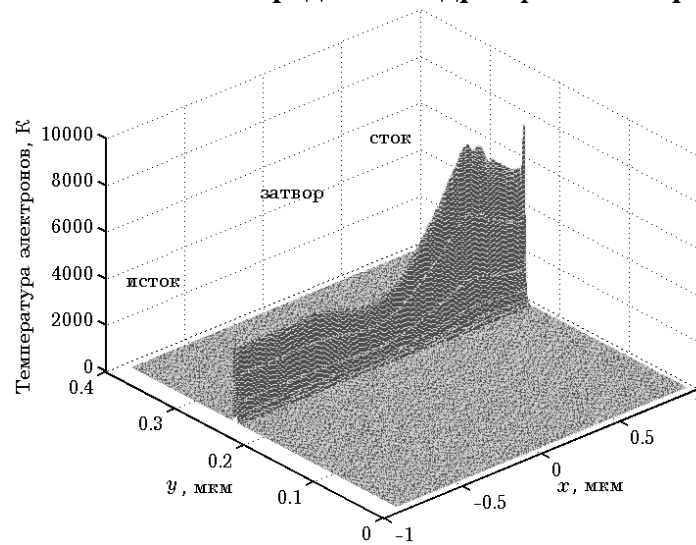


Рис. 5.16. Распределение температуры электронного газа

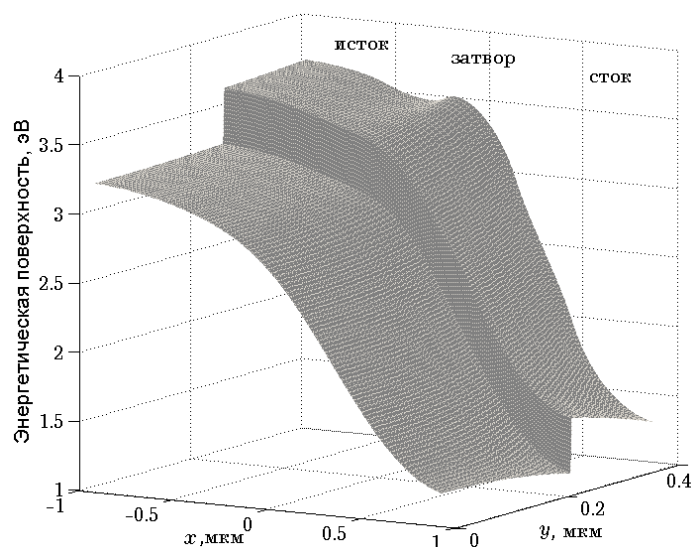


Рис. 5.17. Потенциальная энергетическая поверхность

Приведенные характеристики рассчитаны при электрическом смещении на электродах истока, затвора и стока, соответственно, 0 В, -0,5 В, 3,0 В. Разрыв зон на границе гетероперехода задавался при решении уравнения Пуассона как встроенный потенциал высотой около 0,4 В в предположении отсутствия квантовых эффектов для электронов в потенциальной яме.

Как показывают результаты моделирования, разогрев электронного газа и рост температуры электронов в большей степени связан с электронами, дрейфующими в потенциальной яме на границе гетероперехода. Эффект «всплеска» дрейфовой скорости заметен как для электронов гетероперехода, так и электронов, дрейфующих в высоколегированном слое $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Однако, поскольку область за затвором обеднена носителями заряда, лишь незначительное их количество имеет высокие значения скорости.

На рис. 5.18 показаны результаты сравнения моделируемых характеристик субмикронных гетеротранзисторов с длиной затвора 0,25 и 0,5 мкм, рассчитанных по двумерной модели с учетом релаксационных уравнений и кинетической модели Монте-Карло [5.17].

Наиболее существенно результаты моделирования отличаются на пологом участке характеристик. Отличие в наклоне характеристики, а значит и в выходной малосигнальной проводимости, может быть в большей степени связано с погрешностью определения времен релаксации при двумерном моделировании, в то время как в кинетической модели процессы рассеяния моделируются вероятностными методами на более высоком физическом уровне.

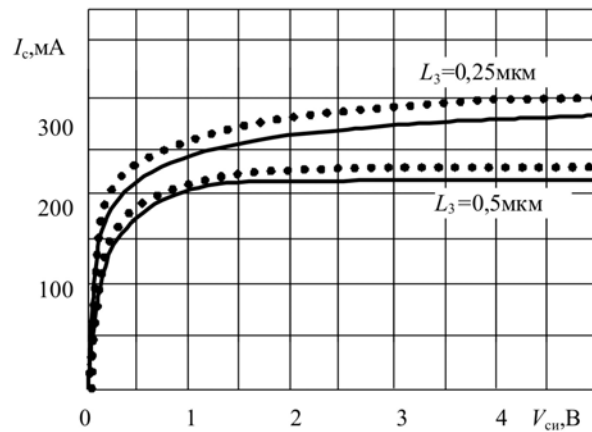


Рис. 5.18. Вольтамперные характеристики транзисторов с гетероселективным легированием с длиной затвора 0,25 и 0,5 мкм, рассчитанные по двумерной (сплошные линии) и кинетической модели (Монте-Карло).

Таким образом, модели процессов дрейфа в субмикронных транзисторах с селективным легированием на основе системы релаксационных уравнений в двумерном приближении позволяют сделать вывод о том, что эффектами, которые определяют работу субмикронных гетероструктур, является дрейф и разогрев электронов в условиях преимущественно междолинного и оптического рассеяния носителей заряда. Важной особенностью является то, что области сильного поля и разогрева электронного газа связаны с большими градиентами характеристик в подзатворной области и локализуются в ограниченной области за затвором.

Подобные модели пригодны для расчета субмикронных гетероструктур на основе широкого спектра соединений полупроводников $A^{III}B^V$ с междолинным переносом электронов, а также для определения параметров активных и пассивных областей транзистора, аппроксимаций характеристик для упрощенных моделей транзисторов и их схемотехнического анализа.

5.4. ГЕТЕРОТРАНЗИСТОРЫ С ДВУМЯ ПОТЕНЦИАЛЬНЫМИ ЯМАМИ

Рассмотрим модели и результаты моделирования субмикронных гетеротранзисторов с двумя потенциальными ямами (с двумя каналами с двумерным электронным газом). В подобных структурах более высокие значения дрейфовой скорости электронов достигаются за счет дрейфа носителей в двух

потенциальных ямах в условиях перераспределения их скоростей и энергий.

Повышение быстродействия транзисторов может быть также связано с созданием структур с субмикронными продольными и нанометровыми поперечными размерами, в том числе многослойными наноструктурами низкой размерности [5.18-5.20].

Характерные размеры затвора таких гетероструктурных транзисторов составляют 0,1-0,2 мкм при размерах высоколегированных слоев в активной области транзистора менее 0,1 мкм. Моделирование гетеротранзистора с двумя потенциальными ямами представляет интерес в связи с перераспределением носителей между двумя «каналами» и, несмотря на увеличение эффективной толщины транзистора (что может приводить к уменьшению крутизны транзистора), приводит к уменьшению разогрева электронов, что связано с уменьшением вероятности их междолинного рассеяния и приводит к повышению средней дрейфовой скорости электронного газа.

5.4.1. Численное моделирование

Как и в традиционных гетеротранзисторах, в субмикронных структурах с несколькими гетеропереходами преобладают эффекты сильного поля и субмикронные эффекты: квазибаллистический перенос, разогрев в условиях преобладания междолинного и оптического рассеяния носителей заряда, эффект «всплеска» дрейфовой скорости электронов и др. Область сильного поля и нестационарные процессы связаны с двумерной областью под затвором и двумерными распределениями характеристик электронного газа – электронной температуры (энергии), импульса (скорости), эффективной массы и др. Особенностью процессов в таких структурах является перераспределение носителей между потенциальными ямами, когда расстояние между ними не превышает длину свободного пробега электронов.

Моделирование субмикронных многослойных структур со сложной топологией связано с применением численных методов с заданным шагом дискретизации и адаптивной процедурой выбора шага, а также с неоднородными граничными условиями между слоями структуры.

Тестовая структура моделируемого гетеротранзистора представлена на рис. 5.19. Структура характеризуется следующими физико-топологическими параметрами: длиной затвора 0,2 мкм, слоем широкозонного легированного $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ($x = 0,3$) под затвором с эффективной толщиной 0,1 мкм, слоем нелегированного $\text{Al}_x\text{Ga}_{1-x}\text{As}$ толщиной 50Å (спейсер), слоем

нелегированного GaAs на полуизолирующей подложке, толщиной несколько десятков микрометров, в поверхностном слое которого в результате разрыва зон образуется потенциальная яма. Параметры второго канала ГСПТ аналогичны и показаны на рис. 5.19 (ось x направлена вдоль канала транзистора, ось y – перпендикулярно каналу от поверхности затвора).

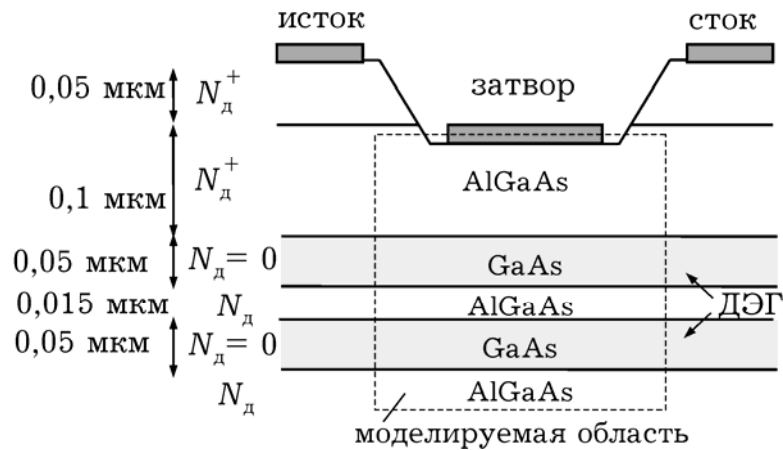


Рис. 5.19. Структура гетеротранзистора с двумя гетероканалами

Поскольку все электронные процессы связаны с поверхностным слоем и двумерным электронным газом (ДЭГ) в потенциальных ямах, то область моделирования ограничивается эффективной толщиной транзистора до 1 мкм.

На рис. 5.20–5.22 приведены двумерные распределения потенциальной энергии, электронной температуры и дрейфовой скорости в канале ГСПТ.

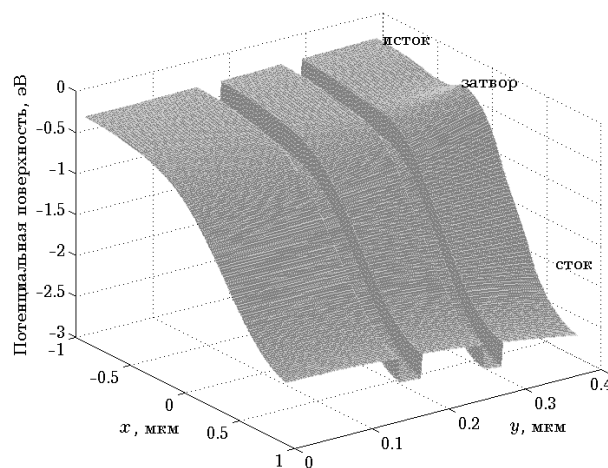


Рис. 5.20. Потенциальная рельеф в структуре гетеротранзистора с двумя гетероканалами

Выходной ток транзистора рассчитывается как нормальная составляющая тока к контакту стока по известным двумерным

распределениям концентрации электронов, напряженности электрического поля и скорости по результатам решения исходной системы уравнений при достижении заданной точности решений в узлах сетки.

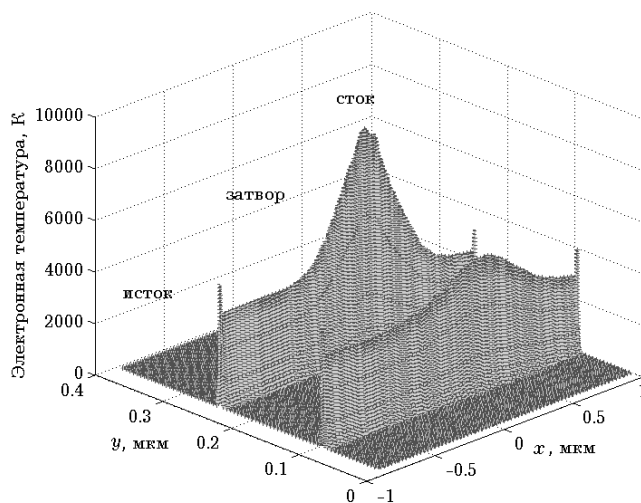


Рис. 5.21. Распределение электронной температуры

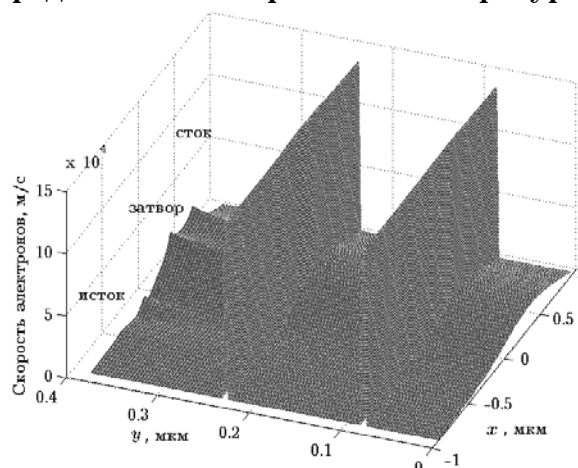


Рис. 5.22. Распределение дрейфовой скорости

Приведенные характеристики рассчитаны при электрическом смещении на электродах истока, затвора и стока, соответственно, 0, -0,5, 3,0 В. Разрыв зон на границе гетеропереходов задавался при решении уравнения Пуассона как встроенный потенциал высотой около 0,4 В в предположении отсутствия квантовых эффектов для электронов в обеих потенциальных ямах.

Как показывают результаты моделирования, разогрев электронного газа и рост температуры электронов в большей степени связан с электронами, дрейфующими в потенциальной яме на границе гетеропереходов. Пиковые значения температуры электронов (рис.5.21) в первом гетеропереходе (примыкающем к затвору) несколько выше, чем во втором, что свидетельствует о перераспределении электронов по энергиям. Эффект «всплеска» дрейфовой скорости заметен для электронов обоих гетеропере-

ходов. Результаты моделирования электронной температуры (рис. 5.21) и дрейфовой скорости (рис. 5.22) [5.21.-5.24] свидетельствуют о меньшей степени разогрева носителей электрическим полем и более высоких значениях дрейфовой скорости, чем в аналогичном гетеротранзисторе с одной потенциальной ямой примерно на 30-40 %. Это может быть связано с запаздыванием процессов разогрева носителей электрическим полем, так как дрейф носителей связан, в основном, с электронами в потенциальных ямах, а также «охлаждением» носителей при их вытеснении полем затвора из одной потенциальной ямы в другую и преодолением второго потенциального барьера с уменьшением энергии. При этом расстояние между гетеропереходами, как показывают результаты численных экспериментов, должно быть меньше длины свободного пробега электронов. Максимальные значения дрейфовой скорости для приведенной структуры достигаются при отрицательных смещениях на затворе 0,5–0,6 В.

Показано, что средние значения дрейфовой скорости в приведенной структуре выше, чем в ГСПТ с одной потенциальной ямой, при этом транзисторы с двумя гетеропереходами обладают более высокими значениями тока и мощности.

Приведенные модели пригодны (при соответствующей аппроксимации времен релаксации) для расчета субмикронных гетероструктур на основе широкого спектра соединений полупроводников $A^{III}B^V$ с междолинным переносом электронов, включая перспективные соединения на основе InP, GaN и др.

5.4.2. Аналитическая модель статических характеристик

Для получения аналитических соотношений структура моделируемого транзистора представляется в предположении резких границ гетеропереходов с конечной толщиной Δd (рис. 5.23).

Для анализа статических характеристик необходимо описать зависимость концентрации электронов в двумерном канале от напряжений на затворе и на стоке транзистора [5.25].

С учетом наличия двух слоев с двумерным электронным газом можно записать:

$$n_s = \frac{\varepsilon\varepsilon_0}{e(2d_i + d_d + 2\Delta d)} [V_g - V_{t_0} - V(x)], \quad (5.11)$$

где $(2d_i + d_d)$ – общая толщина слоя AlGaN, как показано на рис. 5.23, где V_{t_0} – пороговое напряжение, $2\Delta d \approx 100 \text{ \AA}$. Координата x отсчитывается от истока к стоку вдоль канала транзистора.

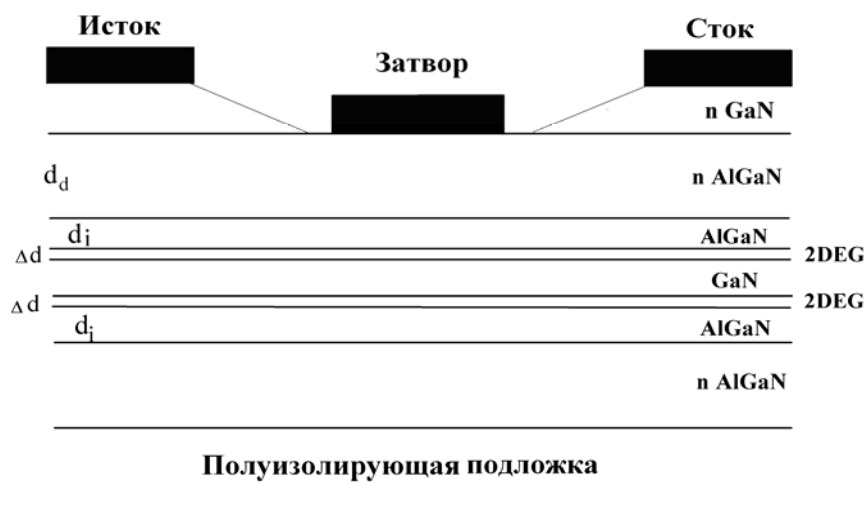


Рис. 5.23. Структура AlGaN/GaN транзистора с высокой подвижностью электронов с двумя гетероканалами

Для расчёта зависимости дрейфовой скорости электронов от напряжённости воспользуемся аппроксимацией работы [5.26]. Данная зависимость учитывает многодолинную структуру энергетических зон нитрида галлия, поэтому такая зависимость имеет максимум и, соответственно, участок с отрицательным сопротивлением. При подборе коэффициентов была использована экспериментальная зависимость дрейфовой скорости от электрического поля [5.27]. Ее аппроксимация может быть представлена соотношением:

$$v(x) = \frac{v_0 \varepsilon(x)/\varepsilon_0 + v_s [\varepsilon(x)/\varepsilon_0]^2}{1 + [\varepsilon(x)/\varepsilon_0]^2}, \quad (5.12)$$

где $v_0 = \mu_0 \varepsilon_0$, ε_0 - пороговое значение электрического поля (для GaAs равно 5 кВ/см, а для GaN – 100 кВ/см), μ_0 - низкополевое значение подвижности, v_s - скорость насыщения.

Для анализа температурных зависимостей и учета уровней легирования может использоваться более универсальная формула. При моделировании двухканального гетеротранзистора на основе нитрида галлия необходимо учитывать особенности его поле-скоростной характеристики. Следует отметить, значение запрещенной зоны нитрида галлия существенно больше (3,39 эВ) по сравнению с арсенидом галлия (1,42 эВ). Ток стока описывается выражением:

$$I_{ds} = en_s(x)v(x)W, \quad (5.13)$$

где W - ширина транзистора.

Совместно решая уравнения (5.11)-(5.13), а также уравнение для потенциала, можно получить следующее аналитическое выражение для вольт-амперной характеристики [5.25]:

$$I_{ds} = \beta \frac{(V_g - V_0)V_{ds} - V_{ds}^2/2}{1 + \alpha V_{ds}}, \quad (5.7)$$

где $\beta = \varepsilon\varepsilon_0\mu W/(2d_i + d_d + 2\Delta d)L_g$; μ – подвижность электронов, L_g – длина затвора; $\alpha = \frac{v(x)/v_0}{\varepsilon_0 L_g} + \beta R_s$; R_s - сопротивление истока.

В данной модели для расчета поверхностной концентрации и дрейфовой скорости электронов могут использоваться усредненные значения двумерного распределения концентрации и электрического поля, полученные при двумерном моделировании при заданных значениях потенциалов затвора и стока:

$$n_{cp} = \frac{\iint n(x, y) dx dy}{a_x a_y}, \quad \varepsilon_{cp} = \frac{\iint \varepsilon(x, y) dx dy}{a_x a_y},$$

где a_x и a_y размеры области двумерного моделирования.

Возможно также использование приведенной ниже аналитической модели гетеротранзистора при известном из двумерного моделирования значении дрейфовой скорости насыщения и расстоянии в канале, при котором происходит насыщение скорости.

Для получения тока и напряжения насыщения, вычислим производную $g_d = \partial I_d / \partial V_d$ и примем $g_d = 0$ получим:

$$V_{dSat} = \frac{\sqrt{1 + 2\alpha(V_g - V_{t_0})} - 1}{\alpha}.$$

Когда ток в канале достигает насыщения, уравнение Пуассона можно записать в виде [5.22]:

$$\frac{\partial^2 V}{\partial x^2} = \frac{1}{\varepsilon_0 \varepsilon} \cdot \frac{J}{v_s},$$

где плотность тока $J = I_{dSat}/W\Delta d$; а v_s - дрейфовая скорость носителей в точке насыщения.

Используя граничные условия:

$$V(0) = V_{dSat} \text{ и } \partial V(0)/\partial x = \epsilon_{dSat} \approx V_{dSat}/L_g,$$

можно получить аналитическое выражение для длины перекрытого канала в следующем виде [5.26]:

$$L_s = \gamma \left[\sqrt{1 + 2(V_{ds} - V_{dSat})/\gamma\epsilon_{dSat}} - 1 \right].$$

При напряжениях затвора, превышающих напряжение насыщения, ток в канале может рассчитываться по формуле (5.7), но вместо V_{ds} следует подставлять V_{dSat} , а вместо L_g - $(L_g - L_s)$.

Известны также и другие подходы к получению аналитических соотношений для расчета статических характеристик гетеротранзистора, в частности на соединениях AlGaAs [5.27, 5.28].

5.4.3. Схемные модели и шумовые параметры

Для расчёта малосигнальных крутизны, выходной проводимости, входной и проходной ёмкостей может использоваться подход, приведенный для расчета аналогичных параметров для субмикронного ПТШ, при известных статических выходных и проходных характеристиках, с учётом приращений заряда под затвором, как реакции на изменение входного и выходного напряжений транзистора.

При определении паразитных СВЧ элементов субмикронного полевого транзистора, как правило, используются малосигнальные S -параметры или параметры, измеренные при напряжении сток-исток $V_{ds} = 0$. Основной причиной этого является то, что при данных условиях схемную модель можно значительно упростить. Кроме того, одним из факторов, позволяющих осуществить такое упрощение, является относительная симметрия активного канала, определяемая симметрией обедненной области под затвором. Геометрическая асимметрия, обычно существующая между истоковым и стоковым контактами, связана с доменом сильного поля. В случае субмикронных длин затвора (порядка 0,1-0,2 мкм), когда длина затвора сравнима с толщиной канала, распределения

физических величин под затвором – электрического поля, концентрации носителей, их энергии и других имеют двумерный характер и асимметрия области пространственного заряда становится менее ярко выраженной. Это относится также к субмикронным гетероструктурам. Идентификация схемных малосигнальных и паразитных параметров, как правило, основывается на вычислительных процедурах оптимизации импедансов по экспериментальным S -параметрам в рабочем диапазоне частот. Элементы активной области определяются с помощью малосигнальной схемной модели транзистора.

Расчет минимального коэффициента шума в субмикронных гетеротранзисторах с двумя каналами также как и для субмикронных ПТШ основывается на применении схемной модели с различного рода локальными источниками шума. При этом методика расчета минимального коэффициента шума в субмикронных ПТШ может быть адаптирована для расчета гетероструктур после получения усредненных значений энергии, скорости, напряженности электрического поля и других характеристик активной области транзистора.

В гетеротранзисторах с двумя гетеропереходами присутствуют практически все составляющие шума, приведенные для короткоканальных транзисторов с затвором Шоттки. Главное отличие заключается в учете шума токораспределения.

Для двухканального гетеротранзистора шум токораспределения, кроме составляющей тока через подложку, связан также со случайным характером перераспределения электронов между гетеропереходами. Это должно приводить к дополнительной составляющей шума. Минимизация этой составляющей шума также связана с оптимальным выбором электрического смещения на затворе.

Для гетеротранзистора с двумя гетеропереходами при допущении об отсутствии тока через подложку шум токораспределения между двумя каналами (потенциальными ямами с двумерным электронным газом) запишется как:

$$\overline{|i_{\text{ТНЕМТ}}|^2} = 2e(I_{\text{к1}}I_{\text{к2}} / I_{\text{с}}) \Delta f,$$

где $I_{\text{к1}}$, $I_{\text{к2}}$, $I_{\text{с}}$ – составляющие тока соответственно первого и второго каналов и полный ток стока. Как видно из последнего соотношения, максимальный шум токораспределения имеет место при $I_{\text{к1}} \approx I_{\text{к2}}$.

Учитывая тот факт, что полного перекрытия гетероперехода не происходит, шум токораспределения в гетеротранзисторах с

двумя каналами, в отличие от традиционного субмикронного ПТШ, вносит больший вклад в суммарный шум транзистора.

Из анализа составляющих шума можно сделать следующие выводы. Малошумящий режим в гетеротранзисторах с двумя каналами связан с напряжением на затворе, существенно меньшим значений потенциала отсечки тока стока (в традиционных субмикронных ПТШ этот режим достигается при малых токах стока при потенциалах, близких к отсечке тока), когда носители первого канала за счет разогрева электрическим полем преодолевают потенциальный барьер и потеряв часть энергии попадают во второй гетеропереход. Это приводит к уменьшению вероятности междолинного рассеяния и уменьшению шума, связанного с разогревом носителей, который является преобладающим в субмикронных транзисторах. В случае гетеротранзистора с двумя гетеропереходами процессы, связанные с междолинными переходами, как уже было показано, ослабляются, что уменьшает среднюю температуру электронного газа и, соответственно, разогревный шум канала. Таким образом, в гетеротранзисторах с двумя каналами малошумящий режим достигается при больших значениях тока стока, что позволяет обеспечить более высокие уровни мощности транзистора.

Минимальный коэффициент шума обеспечивается как выбором электрического режима, так и условиями согласования. Представляет также интерес исследование гетеротранзистора с двумя потенциальными ямами и встроенными одномерными системами – квантовыми точками (КТ) (рис. 5.24).

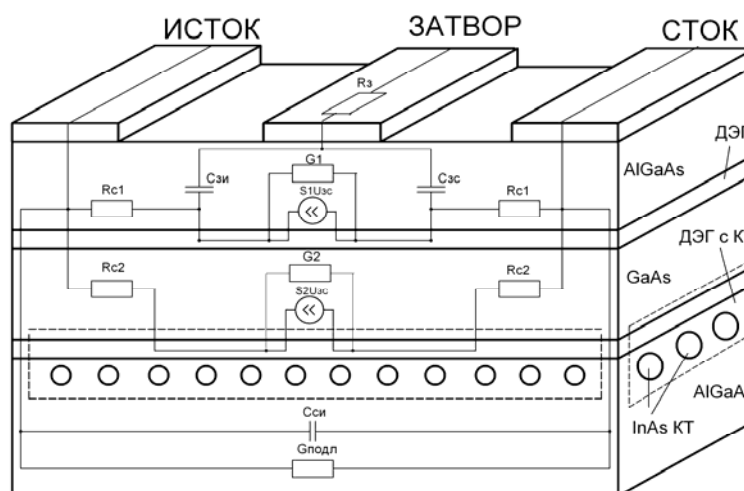


Рис. 5.24. Топология и схемная модель двухканального гетеротранзистора с квантовыми точками (КТ)

Представленные модели субмикронного гетеротранзистора с двумя гетеропереходами, учитывающие дополнительные

источники шума за счет токораспределения между потенциальными ямами, пригодны для расчета сигнальных и шумовых характеристик.

5.5. ГЕТЕРОСТРУКТУРНЫЕ ТРАНЗИСТОРЫ С КВАНТОВЫМИ ТОЧКАМИ

В качестве перспективных гетероструктур с КТ представляют интерес структуры на основе нитрида галлия. Как показывают результаты физико-топологического моделирования субмикронного полевого гетеротранзистора с квантовыми точками (КТ), в них скорость носителей выше, чем в обычном полевом транзисторе Шоттки.

Способ получения КТ основан на способности самоорганизации наноструктур. Важнейшими факторами, определяющими направление и конечный результат самоорганизации являются рассогласования постоянной решетки (создающие упругие напряжения в системе и повышающие её внутреннюю энергию), толщина слоя, границы раздела и т. д. Но на сегодняшний день возможности технологии еще не позволяют сделать выращивание КТ полностью контролируемым и воспроизводимым. Поэтому неоднородность островков квантовых точек и их параметров приводит к необходимости учета случайного характера расположения, размеров, энергетических и других характеристик КТ.

КТ из арсенида индия формируются вблизи гетероперехода в слое нитрида галлия [5.29]. Они захватывают часть электронов проводимости канала, которые с увеличением напряжённости электрического поля в канале инжектируются из КТ и увеличивают ток в канале транзистора. За счёт малой начальной энергии вероятность рассеяния таких электронов на оптических фононах практически равна нулю [5.30]. В итоге, общая энергия коллектива электронов в канале понижается, и скорость насыщения становится выше. Причем достигается это не предельными возможностями субмикрометровой литографии, а за счет преобладания двух физических факторов: безынерционной ударной ионизации КТ в сильном электрическом поле и существенного повышения дрейфовой скорости электронов в канале транзистора.

В работе [5.31] показано, что максимальные значения дрейфовой скорости в транзисторе с КТ в два раза выше, чем в традиционных ГСПТ, при этом необходимо учитывать трехмерный характер физических процессов в гетеротранзисторах с КТ, что связано как с особенностями топологии (ширина транзистора сравнима с размерами вдоль координаты x , связанной с тянущим

электрическим полем), так и случайным характером параметров КТ. На рис. 5.25, а приведена структура моделируемого транзистора. На рис. 5.25,б приведено зафиксированное на атомно-силовом микроскопе изображение слоя с самоорганизованными КТ [5.32].

При моделировании субмикронных приборов, имеющих многослойную структуру и содержащих искусственные неоднородности и дефекты, возникает ряд трудностей из-за наличия больших градиентов электрического поля, поэтому значительно повышаются вычислительные затраты за счет дополнительного дробления шага при использовании итерационных методов, сгущения сетки и т.д. Еще более актуальной становится задача адаптации численных методов к решению системы двумерных дифференциальных уравнений в частных производных для моделирования физических эффектов в канале ГСПТ с КТ.

В структурах со встроенными КТ, в которых движение частиц квантовано по всем трем координатам, спектр плотностей состояний представляет собой набор δ -функций [5.33].

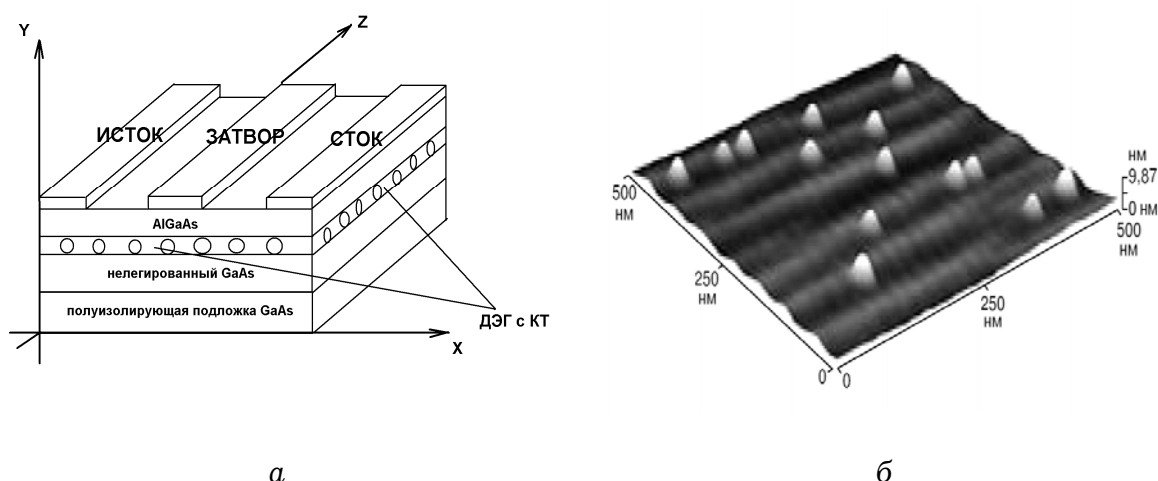


Рис. 5.25. Схематическое изображение а – гетероструктурного полевого транзистора с КТ, где ДЭГ – двумерный электронный газ; б – структура с самоорганизованными квантовыми точками.

Субмикронный гетеротранзистор с квантовыми точками моделируется на основе системы релаксационных уравнений с учетом преобладающих механизмов рассеяния. Для получения распределения потенциала с учетом квантовой ямы гетероперехода и квантовых точек в потенциальный рельеф соответственно встраиваются «ступеньки» потенциала 0,3 В и 0,7 В в отдельные узлы сетки без учета в них квантовых эффектов

Моделировалась структура GaN гетеротранзистора с КТ со следующими параметрами: длина затвора 0.4 мкм, напряжение на

затворе -0.2 В, напряжение на стоке 4 В. Канал и КТ получены путем встраивания в потенциальный рельеф ступенек потенциала: 0.3 В для квантовой ямы гетероперехода и $0,7$ В для КТ с учетом того, что плотность квантовых точек в реальной структуре составляет $\sim 10^{10}$ см $^{-2}$ [5.34].

Из распределения потенциала в моделируемой структуре (рис. 5.26) видно, что нахождение носителей в квантовой точке является энергетически выгодным. Это объясняет факт перераспределения носителей в канале между квантовой ямой и точками. При встраивании потенциальной КТ в гетеропереход меняется и картина распределения напряженности поля в транзисторе – в потенциальной яме с КТ напряженность максимальна, что приводит к росту дрейфовой скорости. Как видно из рис. 5.27, дрейфовая скорость носителей в гетеротранзисторе с КТ почти на порядок выше, чем в ПТШ без встроенных неоднородностей. Во-первых, это связано с тем, что носители движутся в квантовой яме практически без столкновений, во-вторых, рост средней скорости электронов связан с ионизацией КТ и инжекцией электронов в канал. Инжектированные из КТ электроны имеют меньшую начальную энергию, что приводит к понижению энергии всего коллектива носителей. Кроме того, наличие КТ приводит к квантованию оптических фононов, понижая вероятность оптического рассеяния. Описанные факторы приводят к росту дрейфовой скорости. Это связано с тем, что при малом количестве КТ не все оптические фононы квантуются, поэтому насыщение скорости наступает быстрее.

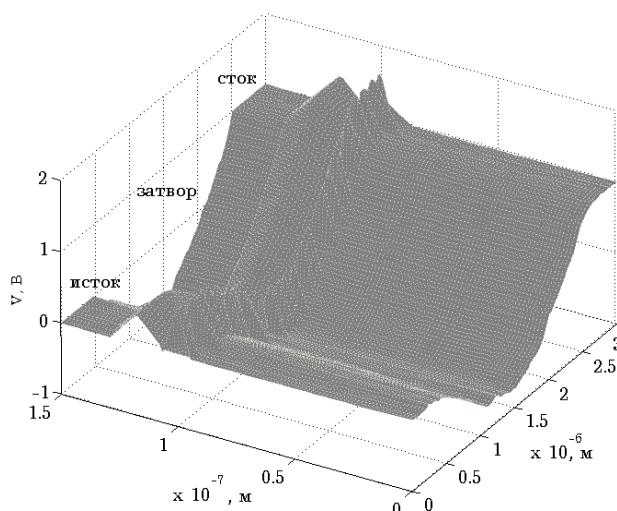


Рис. 5.26. Распределение потенциала в гетероструктурном транзисторе с КТ

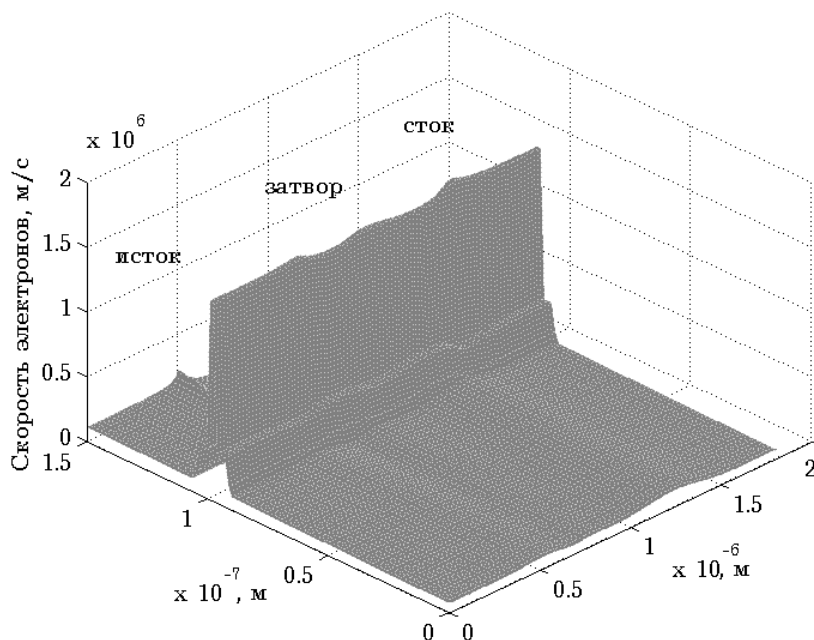


Рис. 5.27. Дрейфовая скорость носителей в гетеротранзисторе с КТ

Приведенные результаты позволяют описать механизм повышения дрейфовой скорости электронов. Полученные характеристики для поля, скорости и концентрации носителей позволяют судить как о распределении этих характеристик в структуре транзистора, так и об их количественных показателях. Анализ результатов моделирования показывает, что встраивание в гетеротранзистор даже небольшого количества КТ приводит к заметному росту скорости дрейфа носителей в канале транзистора.

Решение системы уравнений позволяет также определить распределения электронной температуры (рис. 5.28) и концентрации носителей заряда (рис. 5.29). При расчетах времен релаксаций учитывались следующие механизмы рассеяния: для времени релаксации импульса τ_p – акустическое, примесное, оптическое, полярное оптическое и междолинное рассеяния, а для времени релаксации энергии τ_E – те же, что и для времени релаксации импульса, но без учета примесного рассеяния. Известно [5.15], что из всех механизмов рассеяния в многодолинных полупроводниках наибольшую роль играют оптическое полярное и междолинное рассеяния. В области повышения электронной температуры время релаксации импульса при полярном оптическом рассеянии существенно превышает время релаксации импульса при междолинном. Обратное время релаксации энергии при полярном оптическом рассеянии уменьшается с ростом T_e , так как рассеяние энергии связано преимущественно с междолинными переходами.

Анализ распределений температуры электронного газа (рис. 5.28) и обратного времени релаксации импульса и энергии при междолинном рассеянии позволяют сделать вывод о том, что области сильного разогрева носителей (максимумы T_e) и области максимальных значений $1/\tau_p$ и $1/\tau_E$ совпадают.

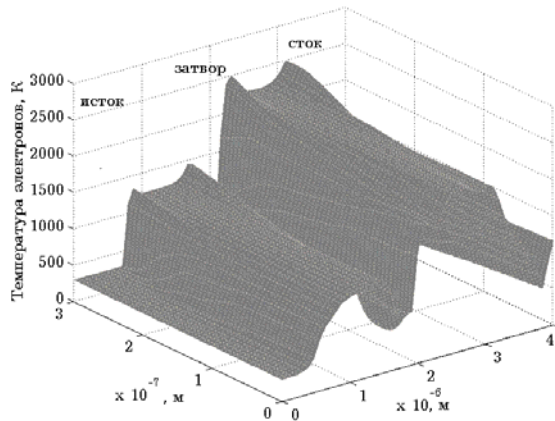


Рис. 5.28. Распределение электронной температуры

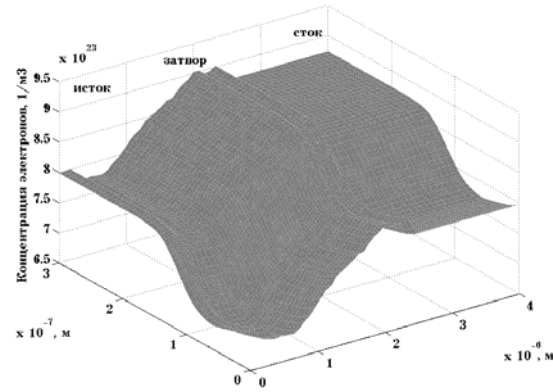


Рис. 5.29. Распределение концентрации носителей заряда

При встраивании в канал гетеротранзистора квантовых точек часть носителей вытесняется из канала (рис. 5.29) и пространственная локализация носителей в гетеротранзисторе с КТ происходит в области, примыкающей к стоку и в подложке. Область максимальных значений дрейфовой скорости у стока совпадает с областью максимальных значений концентрации подвижных носителей, что может объяснить эффект роста дрейфовой скорости носителей в гетеротранзисторах с КТ по сравнению с традиционными структурами.

В структурах с тонкими функциональными слоями также следует учитывать дискретизацию энергетического спектра и фермиевский характер заполнения соответствующих двумерных энергетических подзон квантования. Это возможно с помощью самосогласованного решения уравнений Пуассона и Шредингера:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = -\rho / \epsilon \epsilon_0$$

$$-\frac{\hbar^2}{2m^*} \Delta \psi(x, y) + [U(x, y) - E_0] \psi(x, y) = 0,$$

где E_0 – энергия основного уровня квантования.

Для расчета концентрации носителей заряда в области квантовой ямы необходимо принимать во внимание, что в двух

других направлениях электроны двигаются свободно, формируя подзоны размерного квантования. Концентрация подвижных носителей заряда без учета изгиба зон находится интегрированием по состояниям подзоны с учетом распределения Ферми-Дирака:

$$n(x, y) = \frac{m^* k_B T}{\pi \hbar^2} \ln \left[1 + \exp \left(\frac{E_\Phi - E_i}{k_B T} \right) \right] |\psi(x, y)|^2.$$

Энергии подзон размерного квантования в приближении прямоугольной симметричной квантовой ямы можно определить по формуле [5.35]:

$$\arcsin \frac{\hbar k}{\sqrt{2mU_0}} = \frac{n\pi - ka}{2}.$$

При использовании замены $\cos(\xi) = \pm \gamma \xi$, при n – нечетном, а $\sin(\xi) = \pm \xi \gamma$, когда n – четное, получим для равных значений энергий соотношение:

$$E = 2\xi^2 \hbar^2 / ma^2.$$

Для достижения большей точности нужно решать уравнение Шредингера численными методами. Но для оценки положения основного энергетического уровня, можно воспользоваться предположениям о «мелкой» квантовой яме, для которой положение основного уровня оценивается по формуле:

$$E_0 \approx U_0 - \frac{ma^2}{2\hbar^2} U_0^2.$$

Для произвольной доли алюминия в соединении $\text{Al}_x\text{Ga}_{1-x}\text{As}$ энергия сродства к электрону будет равняться: $4,07 - 1,1x$ эВ (при $x < 0,45$) и $3,64 - 0,14x$ эВ (при $x > 0,45$), и разрыв зон проводимости равняется $\Delta E_c = \chi_{\text{GaAs}} - \chi_{\text{AlGaAs}} \approx 0,4$ эВ при $x = 0,35$.

Если гетероструктура включает неоднородность в виде квантовых точек, то квантование спектра электронов в них должно учитываться отдельно путем вычисления положений энергетических уровней в соответствующей КТ.

Для определения энергетических уровней в сферически симметричной квантовой точке из арсенида индия воспользуемся методом, предложенным в [5.36].

Глубина потенциальной ямы в зоне проводимости $\Delta E_c = \chi_{\text{InAs}} - \chi_{\text{GaAs}} = 0,83$ эВ. Рассмотрим уравнение:

$$\sqrt{X - Y} = -\sqrt{Y} \operatorname{ctg} \left(\frac{\pi}{2} \sqrt{Y} \right),$$

где $X = E / E^*$, $E^* = \frac{\hbar^2}{2m^*} \left(\frac{\pi}{2a} \right)^2 n^2$; a – ширина потенциальной ямы;
 $n = 1, 2, 3 \dots$; $X = U_0 / E^*$, U_0 – высота стенок ямы.

Анализ данного уравнения показывает, что оно может иметь решение при $(n - 0,5) \leq \sqrt{Y}$. Тогда в КТ с такими параметрами будет один уровень энергии, который равняется $E_{\text{КТ}} = 0,036$ эВ.

На рис. 5.31 приведена энергетическая диаграмма для гетеротранзистора с квантовыми точками для проекции, поперечной к плоскости канала.

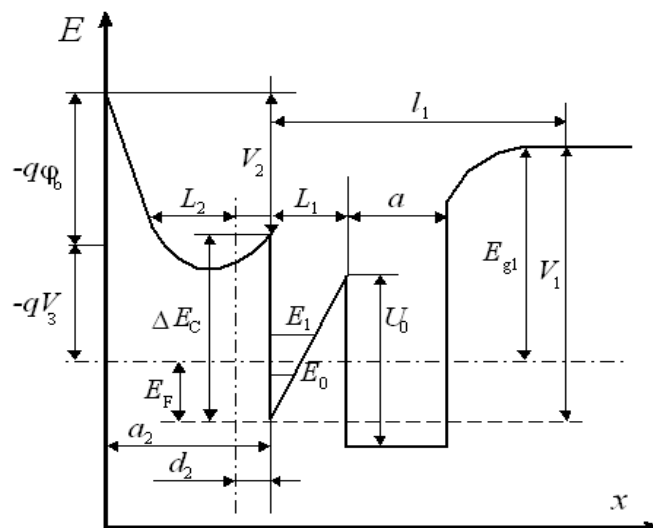


Рис. 5.30. Диаграмма зоны проводимости структуры гетеропереходного транзистора с КТ (U_0 – высота потенциального барьера квантовой точки, a – ее ширина)

В продольной проекции квантовые точки в канале могут представляться в виде чередующихся квантовых ям, т.е. структурой, аналогичной сверхрешетке. Отличие заключается в степени ограничений носителей заряда. Здесь может использоваться приближение о прямоугольной форме квантовых точек, хотя более вероятно, что в процессе роста, в том числе и по механизму Странски-Крастанова, на границе гетероперехода образуются КТ сферической формы. В условиях сильных полей, как показывают расчеты, отличия положений энергетических уровней в КТ в прямоугольном и сферическом приближении не являются существенными.

Эмиссия электронов из квантовых точек возможна при ударной ионизации, туннелировании носителей с КТ в квантовую яму (КЯ) и в результате тепловых выбросов носителей. Вероятность тепловых выбросов носителей из-за взаимодействия с колебаниями кристаллической решетки оценивалась с помощью больцмановского фактора [5.37]: $\delta = \exp(-\Delta E_a / k_B T)$, где $E_a = U_0 - E_i$.

Ускорение электронов до порога ударной ионизации квантовых точек зависит от соотношения двух факторов – ускорения во внешнем электрическом поле и рассеянии при столкновении с фононами, преимущественно с полярными оптическими и междолинными. Полученный с помощью решения системы релаксационных уравнений распределение электронной температуры учитывает эти факторы, и событие ионизации происходит, если значение электронной температуры достаточно в месте расположения КТ. При расчетах ионизация КТ учитывается в предположении, что энергия, которая передана локализованному в КТ носителю заряда, тратится на преодоление потенциального барьера и выбросы этого носителя из КТ, а в канал транзистора электрон инжектируется с начальной энергией, которая равняется энергии основного уровня КТ, с которого состоялась эмиссия.

С увеличением размеров КТ расстояние между энергетическими уровнями становится соизмеримым или меньше $k_B T$ и энергетический спектр квантовой ямы становится квазинепрерывным. Квантовую точку с площадью 1 мкм^2 уже можно рассматривать с точки зрения классической теории. Тогда на движение носителей влияет форма изменения потенциала в области КТ. В этом случае увеличение дрейфовой скорости связано с тем, что носители на время все равно захватываются квантовой точкой и теряя энергию эмиттируются «охлажденными».

Для учета влияния туннелирования необходима оценка коэффициента туннелирования, например, для треугольного барьера, при переходе электронов из КТ в КЯ (рис. 5.31). Также, необходимым условием (пренебрегая влиянием рассеяния на расширение квантовых уровней) является совпадение уровней по энергии. Вычисленные положения уровней для КТ и квантовой ямы в одномерном приближении дают совпадение лишь верхних уровней, для которых коэффициент прохождения практически равняется единице. Для потенциального барьера с шириной a и при условии: $(2m^* E / \hbar^2) a \gg 1$, $(2m^* E / \hbar^2) a |1 - E / U_0| \gg 1$ можно использовать выражение при $E < U_0$:

$$D(E) \approx 4\sqrt{\alpha(1-\alpha)} \exp\left[-\frac{4}{3}\sqrt{\gamma(1-\alpha)^3}\right],$$

где U_0 – высота потенциального барьера; E – уровень, с которого происходит туннелирование; $\alpha = E / U_0$; $\gamma = U_0 / E_1$; $E_1 = \hbar^2 / (2m^* a)$.

Вероятность туннелирования с других уровней КТ является незначительной. В плоскости расположения КТ, параллельной плоскости канала, туннелирование между КТ может быть учтено в случае связанных квантовых точек или при наличии вертикально связанных слоев КТ. В случае изолированных КТ коэффициент прохождения также стремится к нулю. При подаче на затвор отрицательного напряжения ширина треугольного барьера между квантовой ямой и КТ будет расти, следовательно, все меньше электронов смогут туннелировать сквозь него, тогда решающим фактором ионизации КТ становится ударная ионизация.

Как было показано в [5.38], введение слоя КТ в канал гетеротранзистора способствует квантованию фононного спектра в трех направлениях, тогда как на гетерограницах происходит разделение полярных оптических фононов лишь в направлении, поперечном движению электронов, а дополнительное рассеяние на интерфейсных фононах приводит к практически одинаковым скоростям рассеяния и в объемных, и в слоистых структурах.

Рост тока при учете влияния КТ составляет 10 – 15%, что в большей степени достигается не за счет увеличения концентрации подвижных носителей, эмитированных из КТ, а за счет увеличения средней дрейфовой скорости потока носителей. Это объясняется тем, что эмитированные из КТ электроны проводимости будут иметь большие начальные скорости, по сравнению с электронами в канале. Поскольку в процессе самоорганизации КТ в плоскости канала транзистора могут располагаться неравномерно и случайно, это можно учесть при исследовании воспроизводимости характеристик гетеротранзисторов с КТ. На рис. 5.31 приведены распределения средней дрейфовой скорости носителей при одинаковой концентрации КТ, но при разном расположении [5.40].

Случайное расположение КТ в канале транзистора приводит к локальной неоднородности распределений концентрации носителей, электрического поля, а также средних значений дрейфовой скорости.

На рис. 5.32 приведены результаты расчета выходного тока транзистора с концентрацией квантовых точек в канале $3 \cdot 10^{10} \text{ см}^{-2}$

и их сравнение с экспериментом [5.41]. Размеры и форма КТ считались одинаковыми, изменялся лишь способ их расположения в канале.

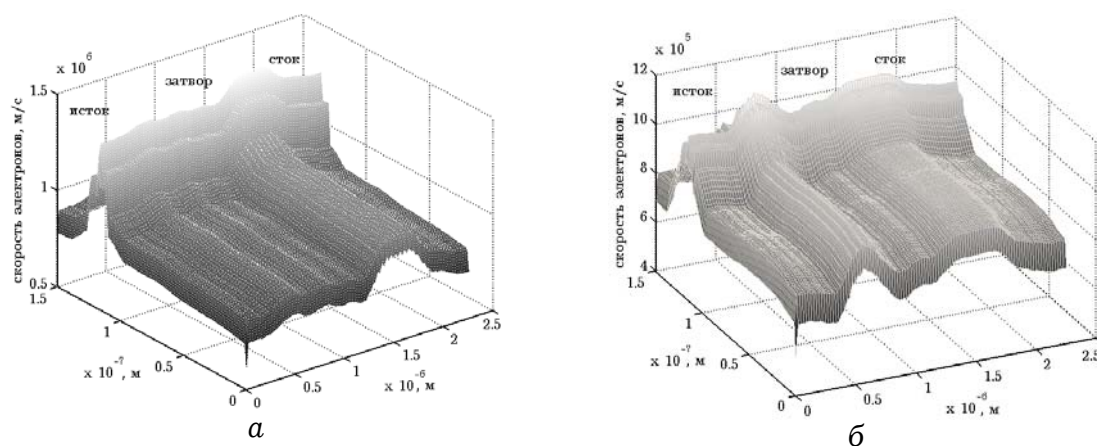


Рис. 5.31. Распределение средней дрейфовой скорости носителей для гетеротранзисторов с КТ при локализации КТ в канале у стокового (а) и у истокового (б) края затвора

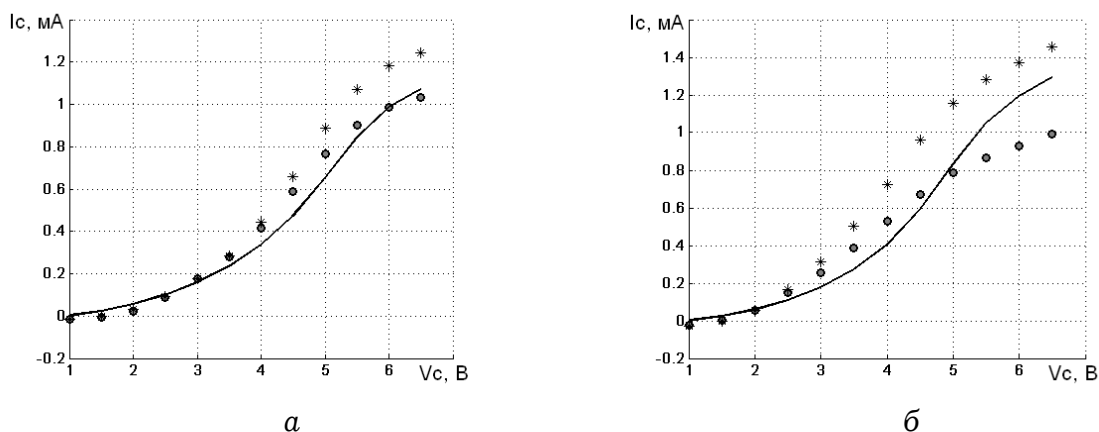


Рис. 5.32. ВАХ гетеротранзистора с КТ. Сплошная линия – эксперимент, точки – результаты расчетов для концентрации КТ $3 \cdot 10^{10} \text{ см}^{-2}$; а) $V_g = 0 \text{ В}$, б) $V_g = -0,8 \text{ В}$.

При той же концентрации КТ в канале способ их расположения заметно влияет на значение исходного тока. Таким образом, подтверждается влияние неопределенности в расположении КТ на исходные характеристики транзистора, что, очевидно, является одной из причин их отклонения от экспериментальных данных.

Литература к главе 5

- 5.1. *X. Штёрмер* Дробный квантовый эффект Холла, Нобелевская лекция. Стокгольм, 10 декабря 1998 г., Успехи физических наук, 2000, т.170, № 3 с.303-319.
- 5.2. *Minura T., Niyamizu S., Fujii T, Nanbu K.*, Jap. Appl.Phys., **19**, 5, L225 (1980).
- 5.3. *Lee J.S., Kim J.W., Lee J.H.* et al. Reduction of current collapse in AlGa_N/Ga_N heterostructure FETs. – Electron. Lett., 2003, v.39, N9, p.750–752.
- 5.4. *Delagebeaudeuf D., Nuyen T., Linh* Metal-(n) GaAlAs-GaAs two-dimensional electron gas FET// IEEE Trans. on Electron Dev.-1982.-Vol.29.-N6.-P.955-960.
- 5.5. *Lee K.,Shur M.S.* Design and fabrication of high transconductance modulation-doped (Al, Ga)As/GaAs FETs // J. of Vac. Sci. Technol. B.-1983.-Vol.1.-N2.-P.186-189.
- 5.6. *Yoshida J., Kazata M.* Analysis of high electron mobility transistors based on two-dimensional numerical model // IEEE Trans. on Electron Dev. Lett.-1984.-Vol.5.-N12.-P.508-510
- 5.7. *Widiger D.Y., Kizilyalli I.C., Hess K.*, et. al. Two dimensional transient simulation of an idealised high electron mobility transistor IEEE Trans. on Electron Dev.-1985.-Vol.32.-N6.-P.1092-1102.
- 5.8. *Мунаков В. В., Москалюк В.А.* Моделирование гетероструктурных полевых транзисторов методом частиц // Электронная техника, сер I, электроника СВЧ,1989, вып.2 (416), с.29-36.
- 5.9. *Adachi Sadao.* GaAs, AlAs and AlGaAs material parameters for use in research and devices applications // J.Appl. Phys. - 1985. - Vol.58.- N3. P.R1-R29.
- 5.10. *Мунаков В. В., Москалюк В.А.* Моделирование гетероструктурных полевых транзисторов с односторонним и двойным барьером методом частиц // „Автоматизация проектирования в электронике”, "Техника", вып. 41, с.44-50.
- 5.11. *Shur M.S.* Low Ballistic Mobility in Submicron HEMT's // IEEE Electron Device Letters.-Vol.23, №9.-2002.-P.511-513.
- 5.12. *Kasemsuwan V., Nikali A.* A Microwave Model for High Electron Mobility Transistors // IEEE Transactions on Microwave Theory and Techniques.-Vol.45, №3.-1997.-P.420-427.
- 5.13. *Rudge P., Miles R., Steer M., Snowden C.* Investigation Into Intermodulation Distortion in HEMT's Using a Quasi-2-D Physical Model // IEEE Transactions on Microwave Theory and Techniques.-Vol.49, №12.-2001.-P. 2315-2320.

- 5.14. *Verzelesi G., Mazzanti A., Basile A., Boni A., Zanoni E., Canali C.* Experimental and numerical assessment of gate-lag phenomena in AlGaAs-GaAs heterostructure field-effect transistors (FETs) // *IEEE Transactions on Electron Devices.*—Vol.50, № 8.—2003.—P.1733–1740.
- 5.15. *Москалюк В.О.* Фізика електронних процесів. Динамічні процеси. Київ: “Політехніка”.—2004. — 180 с.
- 5.16. *Москалюк В.А., Тимофеев В.И., Иващук А.В.* Расчет релаксационных параметров GaAs в сильных полях // *Технология и конструирование в электронной аппаратуре.* —2003.—№3.—С.61–64.
- 5.17. *Morton C., Rosa S., Snowden C.* Comparison of quasi-2D and ensemble Monte-Carlo simulations for deep submicron HEMTs // *Proc. MTT-S International Microwave Symposium Digest.*—1998.—Vol.1.—P.153–156.
- 5.18. *A.Akturk, G.Pennington, N.Goldsman.* Quantum Modelling and Proposed Design of CNT-Embedded Nanoscale MOSFET's // *IEEE Transactions on Electron Devices.*—2005. — V.52, №4 — P.577-584.
- 5.19. *R.Chu, Y.Zhou, J.Liu et al.* AlGaN-GaN Double-Channel HEMTs // *IEEE Transactions On Electron Devices.* — 2005.— V.52, №4 — P.438–445.
- 5.20. *Yu.G. Sadofyev, S.E. Yevstigneev, Yu.V.Kopaev, D.S. Shipitsin, S.S. Shmelev.* New approach to the independent ohmic contact formation in the structures with two parallel isotype quantum wells // *J. Vac. Sci. Technol. B.*2000.18(3). — P.1668–1671.
- 5.21. *Тимофеев В.И., Мохсен Амини* Моделирование нестационарных эффектов дрейфа в субмикронных гетеротранзисторах с двумя каналами // *Вісник ДУІКТ.* — 2005, Т.3, №3-4, с.143–146.
- 5.22. *Тимофеев В. И., Амини Мохсен.* Моделирование нестационарных эффектов дрейфа в субмикронных гетеротранзисторах с двумя каналами // *Вісник державного університету інформаційно-комунікаційних технологій.* 2005.— Том 3.— № 3-4.— С.143 — 146.
- 5.23. *Тимофеев В. И., Амини Мохсен.* Нестационарный дрейф электронов в субмикронных гетеротранзисторах с двумя гетеропереходами // *Электроника и связь.* 2006.— Часть.1. — С. 5-8.
- 5.24. *V.I.Timofeyev, M.Amini, E.M.Faleeva.* Non-stationary Drift Of Electrons in Submicron High Electron Mobility Transistor with two Heterojunctions // *Electronics and Electrical Engineering,* 2007, №4(76), P.33-36.
- 5.25. *K.H.G.Duh et al.* A super low-noise 0,1 micron T-gate InAlAs-InGaAs-InP HEMT // *IEEE Microwave Guided Wave Letter.* — 1991.— №5. — P.114–116.

- 5.26. *Di-Hui Hung C. Lin.* DC and transmission line models for a high electron mobility transistor // IEEE Transactions on Microwave Theory and Techniques. – 1989 – Vol.37, № 9 – P.1361–1369.
- 5.27. *Chang C.S., Fetterman H.R.* Analytical model for HEMT's using new velocity – field dependence // IEEE Transactions on Electron Devices. – 1987 – Vol.34. – P. 1456–1462.
- 5.28. *Chey A.J., Walter H.K.* An Analytical Current- Voltage Characteristics Model for High Electron Mobility Transistor Based on Nonlinear Charge- Control Formulation // IEEE Transactions on Electron Devices.–1989.–Vol.37.№ 4–P.908–919.
- 5.29. *Мокеров В. Г., Фёдоров Ю. В., Великовский Л. Э., Щербакова М. Ю.* Новый гетероструктурный транзистор на квантовых точках // Докл. РАН - 2000 - Т. 375, №6 – с. 754-747.
- 5.30. *Тимофеев В. И., Фалеева Е. М.* Моделирование транзистора со встроенными квантовыми точками // 16-я Международная крымская конференция: СВЧ техника и телекоммуникационные технологии. – Материалы конф. – 11-14 сентября 2006 г. Севастополь – с.222-224.
- 5.31. *Тимофеев В.И., Фалеева Е.М.* Численное моделирование субмикронных гетеротранзисторов с квантовыми точками на основе соединения нитрида галлия // Электроника и связь 2007, тематический выпуск, ч.1, С.62-65.
- 5.32. *Карпович И. А.* Квантовая инженерия: самоорганизованные квантовые точки. Соросовский образовательный журнал, 2001, Т. 7, №11. С.102-108.
- 5.33. *Брунков П. Н., Конников С. Г., Устинов В. М., Жуков А. Е., Егоров А. Ю., Максимов М. В., Леденцов Н. Н., Копьев П. С.* Емкостная спектроскопия электронных уровней в квантовых точках InAs в матрице GaAs // Физика и техника полупроводников - 1996. - Т. 30, №5 – с. 924-933.
- 5.34. *Пожела Ю. К., Мокеров В. Г.* Большое повышение максимальной дрейфовой скорости электронов в канале полевого гетеротранзистора // Физика и техника полупроводников – 2006. – Т. 40, №3. – с. 362-366.
- 5.35. *Ландау Л.Д., Лифшиц Е.М.* Теоретическая физика. Том 3. Квантовая механика. Нерелятивистская теория – М. : Наука, 1989. – 766 с.
- 5.36. *Драгунов В.П., И.Г. Неизвестный, В.А. Гридчин.* Основы нанoeлектроники : уч. пособие. – М. : Логос, 2006. – 496 с.
- 5.37. *Смирнов М.Б, Талалаев, Б.В. Новиков, С.В. Сарангов, Г.Э. Цирлин, Н.Д. Захаров, М.Б. Смирнов, В.Г.* Численное моделирование температурной зависимости спектров фотолюминесценции квантовых точек InAs/GaAs // Физика твердого тела. – 2007. – Т. 49, вып. 6. – С. 1126-1131.

- 5.38. Мокеров В.Г., Ю.В. Федоров, Л.Э. Великовский, М.Ю. Щербакова. Новый гетероструктурный транзистор на квантовых точках // ДАН: (доклады РАН). – М. : Наука/Интерпериодика, 2000. – Т. 375 №6. – С. 754-747.
- 5.39. Поздняков В.Д., В.М. Борздов, Ф.Ф. Комаров. Расчет вольт-амперных характеристик симметричных двухбарьерных резонансно-туннельных структур на основе арсенида галлия с учетом процессов разрушения когерентности электронных волн в кантовой яме // Физика и техника полупроводников. – СПб. : ФТИ им. А.Ф. Иоффе РАН, 2004. – Т. 38, вып. 9. – С. 1097– 1100.
- 5.40. Timofeyev V.I., E.M.Faleeva Model of Heterotransistor with Quantum Dots // Semiconductor Physics, Quantum Electronics and Optoelectronics. – 2010. – Vol. 13. – №2. – p. 186–188.
- 5.41. Тимофеев В.И., Е.М.Фалеева. Моделирование двухканального гетеротранзистора с квантовыми точками // СВЧ техника и телекоммуникационные технологии: 20-я Международная крымская конференция, 13-17 сент. 2010 г.: материалы конф. – Севастополь: Вебер, 2010. – С. 846–847.
- 5.42. Мокеров В.Г., Ю.Пожела, К.Пожела, В.Юцене. Гетероструктурный транзистор на квантовых точках с повышенной максимальной дрейфовой скоростью электронов // Физика и техника полупроводников. – СПб. : ФТИ им. А.Ф. Иоффе РАН, 2006. – Т. 40, вып. 40. – С. 367-371.

6. МОЩНЫЕ СВЧ ТРАНЗИСТОРЫ

Повышение мощности СВЧ транзисторов обычно обеспечивается увеличением тока стока, что достигается за счет увеличения ширины затвора [6.1, 6.2]. При этом транзистор секционируется и отдельные секции включаются параллельно. В работе [6.1] описана конструкция и характеристики гетероструктурного полевого транзистора с выходной мощностью 110 Вт, созданного на основе соединений AlGaN/GaN на подложке из карбида кремния. Увеличение мощности достигается, во-первых, «меандровой» конструкцией затвора с его общей длиной 5,52 мм. При этом размер чипа составляет 2,7 мм × 0,6 мм × 0,1мм (рис. 6.1, а). Во-вторых, каждый транзистор включает четыре чипа, включенных параллельно, и согласующие элементы (рис. 6.1, б). Одной из главных проблем при их конструировании является обеспечение необходимого теплоотвода для создания безопасного температурного режима.

6.1. МОЩНЫЕ СУБМИКРОННЫЕ ПТШ

В мощных полевых транзисторах СВЧ с барьером Шоттки, а также транзисторах, работающих в режиме большого сигнала, при некоторых значениях питающих напряжений разогрев может достигать достаточно высоких температур. Это приводит к саморазогреву транзистора и, как следствие, к изменению его электрических характеристик. Кроме того, в субмикронных полевых транзисторах имеет место разогрев электронов сильным электрическим полем и в результате температура электронного газа может достигать нескольких тысяч кельвинов. Распределение электронов по энергиям (и, следовательно, по температурам) вдоль структуры транзистора имеет ярко выраженный максимум у стокового края затвора, что связано с разогревом электрическим полем и преимущественно междолинным рассеянием с последующей релаксацией в межэлектродном пространстве затвор-сток. Поэтому представляет интерес исследование взаимного влияния электродинамических и тепловых параметров, а также распределение тепловых полей по структуре транзистора.

Хотя влияние температуры на параметры полевых приборов не так существенно, как в биполярных приборах, для мощных

субмикронных транзисторов СВЧ многие предельные параметры связаны с перегревом, вызванным протекающим током. На экспериментальном уровне такой анализ практически не реализуем, но необходим для оптимального проектирования, как отдельных транзисторов, так и монокристаллических СВЧ устройств. Самосогласованный анализ тепловых и электронных процессов в транзисторных структурах практически невозможен из-за различного масштаба моделируемых явлений: электронные процессы протекают в активной области транзистора с размером порядка 0,1...1 мкм, а анализ тепловых явлений связан с областью моделирования до 100 мкм.

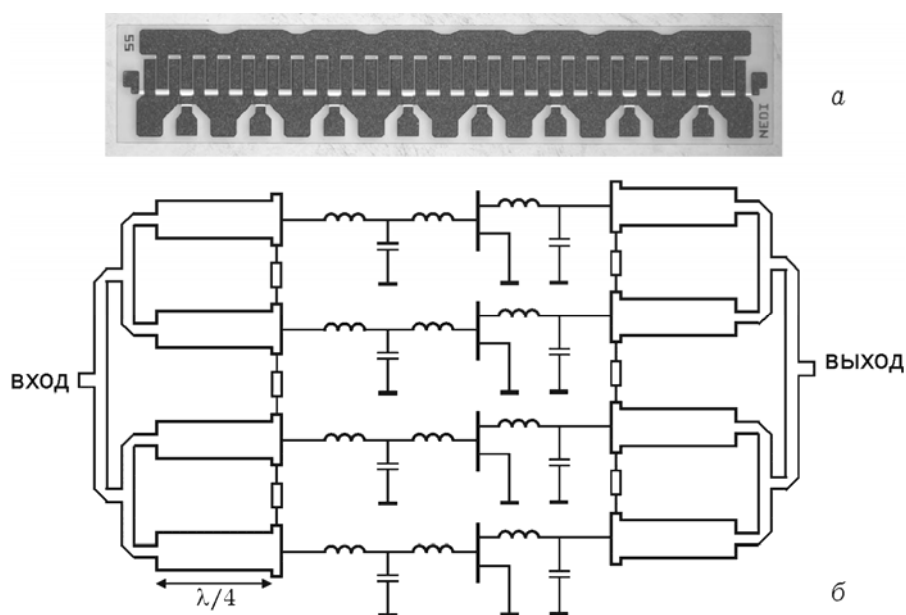


Рис. 6.1. Фотография отдельного чипа AlGaIn/GaN *HEMT* (а) и схема соединения четырех чипов (б) в конструкцию транзистора с согласующими элементами [6.1]

Для учета влияния саморазогрева протекающим током на параметры мощного полевого транзистора необходимо провести предварительный анализ общих тепловых полей всей транзисторной структуры. Результаты этого анализа позволяют в дальнейшем провести расчет теплового состояния в активной области транзистора и промоделировать их влияние на электронные процессы в канале и, соответственно, на параметры транзистора.

Для предварительного теплового анализа используется ряд допущений. Так, рассматривается лишь монокристаллическая пластина полупроводника (рис. 6.2,а) с заданием на ее поверхности контакта с подложкой граничных условий I рода (температура принимается постоянной и равной 20 °С). Поскольку основное

тепловыделение реализуется в весьма тонком слое, составляющем $\approx 0,2$ мкм, то полагается, что тепловой поток подводится к участку наружной поверхности, соответствующему зоне тепловыделений. Прочие ограничивающие поверхности считаются адиабатическими. Принимается, что материал пластины (GaAs) изотропен, теплофизические параметры материала считаются постоянными и не зависящими от температуры. Для упрощения задачи исследования необходимо обратить внимание на следующие особенности рассматриваемого объекта. Верхняя ограничивающая поверхность ($y = 0$) рассматриваемой пластины может быть условно разделена на две зоны - центральную, в которой расположены электроды и реализуются диссипативные тепловыделения, и периферийную, где таковые отсутствуют. В связи с этим представляет интерес оценить величины возможных теплоперетоков в периферийную зону и их влияние на максимальную температуру транзистора.

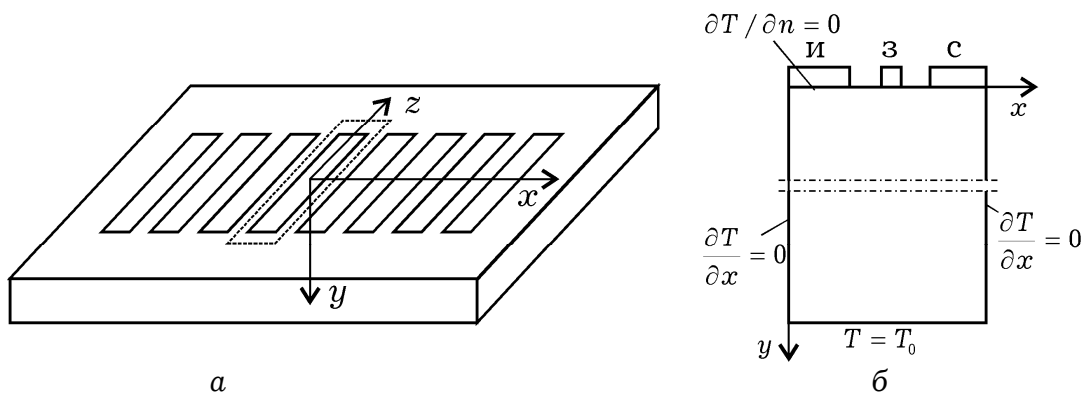


Рис. 6.2. Структурная схема (а) многосекционного транзистора и поперечное сечение моделируемого фрагмента (б)

Обращает на себя внимание определенная повторяемость (периодичность) в расположении электродов и зон тепловыделений. Ввиду этого наблюдается и соответствующая повторяемость, пространственная периодичность температурных полей. Такая картина дает возможность при решении тепловой задачи в первом приближении учитывать не всю область, а ограничиться рассмотрением лишь отдельного, специальным образом выделенного в ней элемента. Схему выделения такого элемента иллюстрирует рис. 6.2,б, где представлен фрагмент продольного сечения ($z = 0$), проходящего через середину центральной зоны. Область исследования может быть уменьшена еще вдвое, если принять во внимание, что в рамках выделенного элемента имеет место геометрическая и тепловая симметрия. Заметим, что ввиду оттоков теплоты вдоль оси абсцисс в периферийную зону, указанная тепловая симметрия выполняется,

строго говоря, только для одного центрального элемента. Однако, именно этот элемент и представляет наибольший интерес, ибо здесь реализуется максимальное значение температуры.

В самом общем случае уравнение теплопроводности относится к дифференциальным уравнениям в частных производных параболического типа:

$$\rho C \frac{\partial T}{\partial t} - \nabla \cdot (k \nabla T) + h(T - T_{ext}) = f$$

в области Ω с начальными условиями $T(x, 0) = T_0(x)$ для $x \in \Omega$.

В стационарном случае первое слагаемое равно нулю и уравнение сводится к уравнениям эллиптического типа:

$$-\nabla \cdot (k \nabla T) + h(T - T_{ext}) = f .$$

Граничные условия определены как условия Дирихле на контактах и однородные условия Неймана на свободных границах.

Решая уравнение теплопроводности в стационарном приближении, находим распределение температуры в структуре, где ρ – плотность теплового потока ($\text{Вт}/\text{м}^2$); C – теплоемкость ($\text{Дж}/\text{К}$); k – коэффициент теплопроводности ($30 \text{ Вт}/\text{мК}$); h – коэффициент конвективного теплообмена ($\text{Вт}/\text{м}^2 \text{ К}$); T_{ext} – внешняя температура (К); f – источник подводимого тепла ($\text{Вт}/\text{м}^2$) [6.3-6.6].

При такой постановке задачи решение уравнения теплопроводности может быть получено аналитически методом разделения переменных Фурье. Так как основное тепловыделение происходит в очень тонком слое, порядка $0,2 \text{ мкм}$, то принимается, что этот слой является источником тепла со стороны затвора.

Как видно из рис. 6.3, в транзисторе имеют место заметные продольные и поперечные тепловые потоки из высокотемпературной области в периферийные области кристалла, где тепловыделения отсутствуют. Вблизи затвора высокотемпературная область имеет многомерный характер, а вдали от этой зоны и в глубине кристалла почти одномерный характер.

Таким образом, тепловой поток преимущественно направлен по толщине кристалла (вдоль оси y), распространение теплового потока вдоль оси x пренебрежимо мало. Область с повышенной температурой непосредственно примыкает к проводящему каналу транзистора.

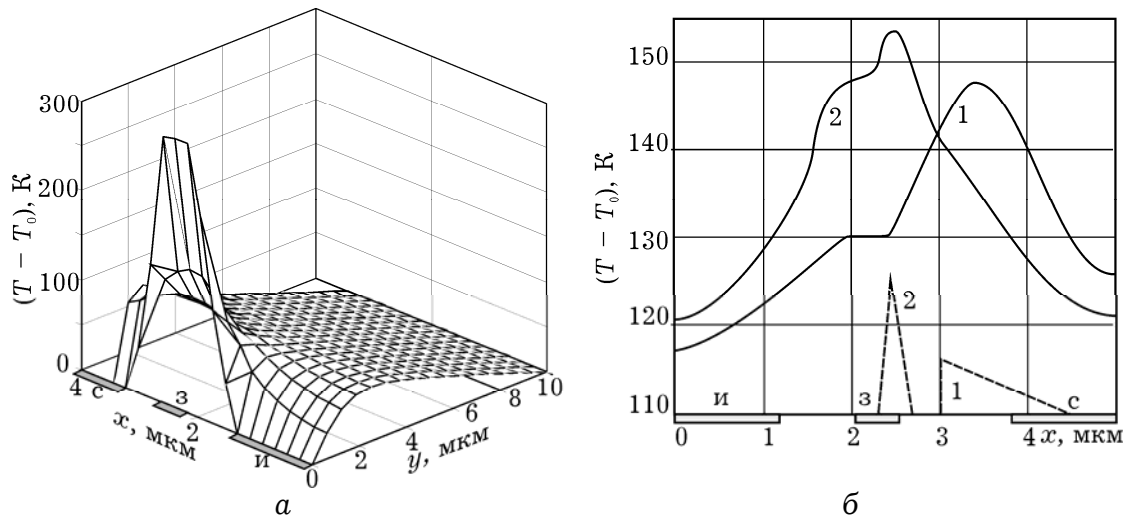


Рис. 6.3. Распределение температуры в субмикронном ПТШ

Отмеченная особенность позволяет, используя метод локализации граничных условий, существенно упростить дальнейшее решение задачи за счет поэтапного уточнения результатов [6.7]. На втором этапе рассматривается не вся исходная область, а лишь ее часть - зона локализации, включающая активную область транзистора с ближайшим окружением. Поскольку зона локализации относительно невелика, то возможно использовать достаточно мелкую ее дискретизацию, обеспечивающую высокую точность получаемых результатов. При этом на нижней границе области задается температура, определенная на первом этапе, а на верхней - диссипативные тепловыделения задаются в виде некоторой эпюры распределения плотности мощности, в той или иной мере имитирующей распределения джоулева тепла в канале транзистора. Результаты расчетов для разных эпюр приведены на рис. 6.4. Форма эпюр теплоподвода существенным образом влияет на распределение температуры лишь в некоторой подобласти, непосредственно примыкающей к месту их задания. По мере же удаления вглубь кристалла это различие все более сглаживается, многомерные поля, отвечающие разным эпюрам, становятся практически одинаковыми, и более того, на достаточно большой глубине многомерный процесс в кристалле трансформируется в одномерный.

Полученные распределения температурных полей показали, во-первых, что в поперечном сечении канала изменение температуры из-за малой толщины эпитаксиального слоя составляет доли градуса. Поэтому распределение температуры вдоль канала можно с хорошим приближением считать одномерным. Во-вторых, слой металлизации затвора «выравнивает» температуру непосредственно в подзатворной части

канала. Наибольшие градиенты температуры и соответственно пик ее распределения приходится на участок канала между стоком и затвором, где преобладают процессы скоростной и энергетической релаксации электронов.

Таким образом, область эффективного управления движением носителей полем затвора и область наибольшего тепловыделения пространственно разделены, а полученные результаты являются хорошим аргументом для использования квазидвумерной модели транзистора (гл. 4), в которой температура канала может быть принята постоянной и зависящей только от мощности диссипативных выделений.

Дальнейшая процедура моделирования влияния саморазогрева на параметры транзистора состоит в расчете статического режима и динамических малосигнальных параметров при комнатной температуре. Затем в зависимости от выделяемой мощности по результатам тепловых расчетов определялась температура канала, для которой проводилось повторное моделирование электрических параметров. Некоторые результаты таких расчетов представлены на рис. 6.4. [6.8].

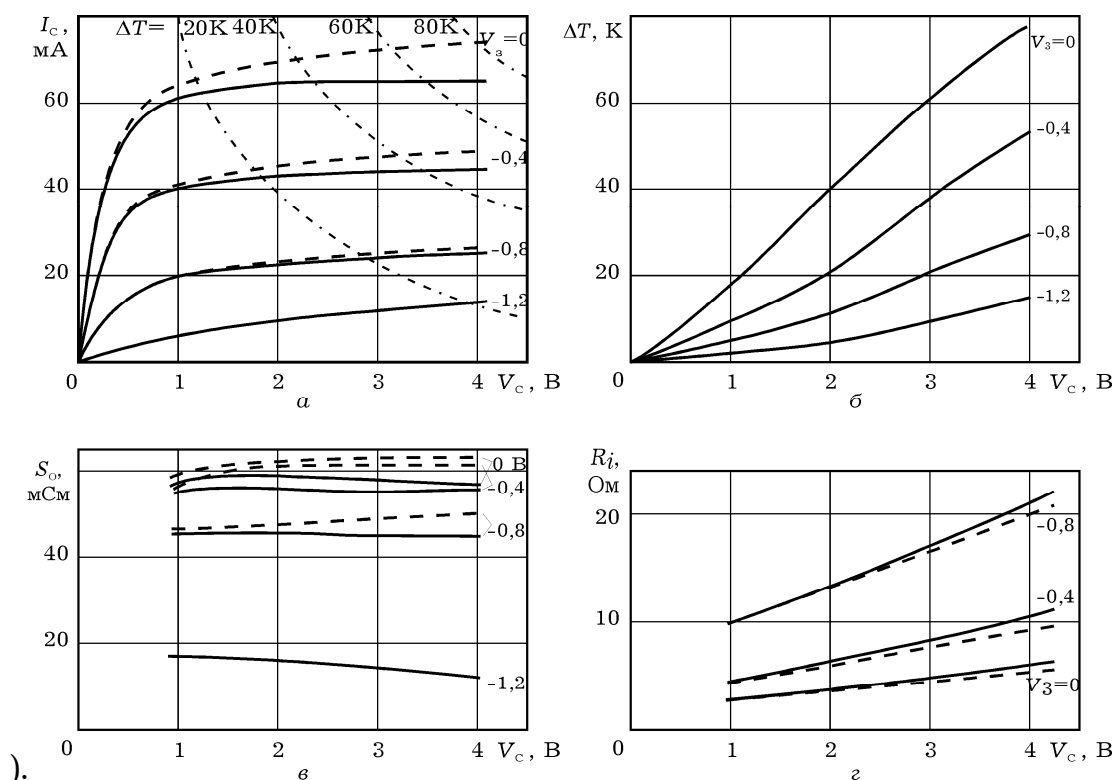


Рис. 6.4. Зависимости тока стока (а), перегрева канала (б), крутизны (в) и внутреннего сопротивления канала (г) от стокового напряжения с учетом саморазогрева (сплошные линии) и без него (штриховые линии). Значения приведены в пересчете на ширину затвора транзистора 300 мкм.

Наиболее заметно сказывается повышение температуры на значениях тока стока и некоторых «резистивных» параметров схемной модели. Уменьшение тока стока (a) и соответствующее уменьшение крутизны (b), особенно заметное в открытом канале, легко объясняется уменьшением подвижности при нагревании и влиянием на времена релаксации и поле-скоростные характеристики. Фактически по этой же причине увеличивается сопротивление канала (b). Несколько неожиданным оказалось отсутствие заметного влияния саморазогрева на выходную проводимость. Несмотря на заметное уменьшение наклона статических вольтамперных характеристик (рис. 6.4 *a*) выходная малосигнальная проводимость, определяемая наклоном динамической характеристики, практически не меняется, что вызвано несоизмеримостью времени установления теплового равновесия и периодом СВЧ колебаний. По сути этой же причиной объясняется слабая чувствительность к разогреву динамических емкостных параметров схемной модели транзистора.

Весьма полезными для оценки предельных режимов транзистора являются зависимости перегрева канала от питающих напряжений, представленные на рис. 6.4 *б*, а также линии постоянной температуры перегрева (штрихпунктирные линии на фоне стоковых вольтамперных характеристик), которые определяют допустимую рассеиваемую мощность.

Как показывают результаты моделирования, имеются области, в которых температура решетки существенно на 50-70 градусов превышает равновесную. Такими областями являются, кроме канала, области контакта истока и стока. Температура электронов под затвором может существенно отличаться от равновесной вследствие разогрева и междолинных переходов. При этом «медленные» электроны верхних долин уменьшают среднюю дрейфовую скорость коллективного движения, что приводит к процессам накопления носителей в области сильного электрического поля под затвором. Это означает, что для моделирования влияния разогрева на электронные процессы в канале транзистора необходимо учитывать отличные от равновесных значения энергии и скорости носителей.

Кроме того, предложенная модель анализа тепловых процессов позволяют исследовать предельные, с точки зрения рассеиваемой мощности, режимы работы полевых транзисторов с субмикронными размерами активной области.

Полученные распределения температур с экстремумом в приповерхностном слое у затвора свидетельствует о необходимости учета высоких градиентов температур при проектировании и выборе топологии. При этом необходимо

учитывать влияние высоких температур на электромеханические свойства затвора и барьерного контакта. Совместное использование физико-топологических и тепловых моделей субмикронных структур позволяет также оптимизировать их структуру, выходные параметры и характеристики.

6.2. МОЩНЫЕ ГЕТЕРОСТРУКТУРНЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ

В субмикронных транзисторных структурах с гетероселективным легированием в сильных полях имеются следующие условия разогрева транзисторной структуры. Общее повышение температуры транзисторной структуры связано с выделяемой мощностью. Это увеличение температуры можно ограничить за счет выбора конструкции, материалов теплоотводов и корпуса с соответствующими значениями тепловых сопротивлений. Вследствие эффектов сильного поля под затвором и локальным разогревом электронного газа до температур, достигающих несколько тысяч кельвинов, перегрев структуры увеличивается. Это увеличение конструктивными приемами устранить невозможно. Таким образом, описание саморазогрева в реальных транзисторных структурах для определения их электротепловых режимов связано как с моделированием тепловых полей, так и с описанием процессов в области сильного поля, где перегрев проявляется наиболее существенно.

Как отмечалось в работах [6.9-6.14], изменение параметров электронного газа вследствие повышения температуры приводят к возникновению высоких локальных градиентов температуры, которые влияют как на электрические характеристики транзисторной структуры, так и на параметры надежности транзистора. В отличие от типичных полевых структур, гетероструктуры (рис. 6.5) – это многослойные структуры, в которых чередуются слои (гетеропереходы, квантовые ямы и потенциальные барьеры) с высокой электропроводностью и обедненные слои, близкие по своим свойствам к диэлектрику. Моделирование тепловых полей позволяет оценить распределение температурных полей в структуре ГСПТ и выбрать режимы устойчивой работы транзистора с точки зрения температурного диапазона, а также выработать рекомендации по оптимальному выбору конструкции транзистора и обеспечению теплоотвода в монолитных ИС СВЧ. Для учета тепловых эффектов в ГСПТ можно использовать два подхода. В первом случае при моделировании эффект саморазогрева учитывается за счет введения дополнительных слагаемых к параметрам модели, определяемым в стационарных режимах. Эти слагаемые задаются

уравнениями, описывающими физические процессы эффектов саморазогрева [6.15].

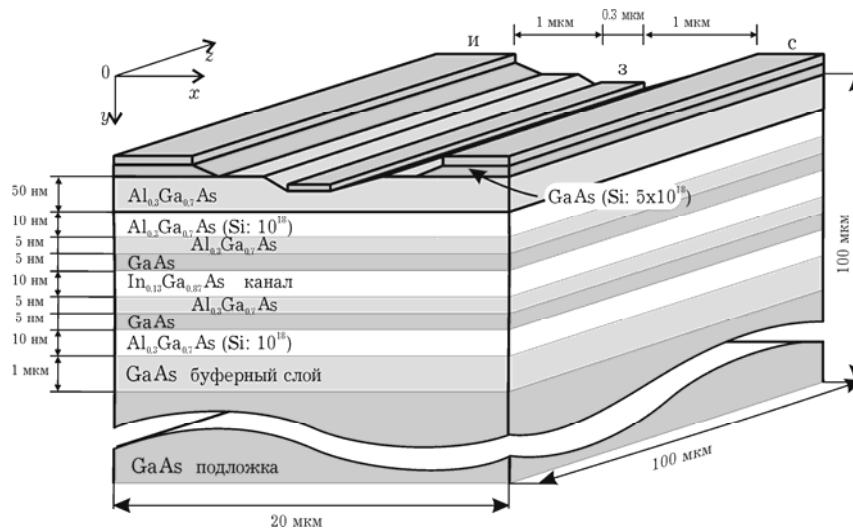


Рис.6.5. Структура транзистора с гетероселективным легированием

Для схемотехнического моделирования в эти модели включают дополнительные, зависящие от температуры источники тока (рис. 6.6). Тогда суммарный ток будет равен: $I_{\Sigma} = I_{dct} + I_{th}$, где I_{dct} – ток канала транзистора, определяемый в стационарном режиме; I_{th} – тепловой ток, являющийся функцией разности напряжения за счет изменения температуры ΔT , протекающий через «тепловое сопротивление» R_{th} и «тепловую емкость» C_{th} , зависящих от температуры и геометрических параметров прибора.

Достоинством такого моделирования является простота расчета, так как параметры элементов модели определены в стационарном режиме, а выходные характеристики получаются путем суммирования напряжений и токов модели, полученных при нормальных условиях, и дополнительных источников тока и напряжения, величины которых аппроксимируют изменения тока или напряжения, в зависимости от температуры транзисторной структуры.

Второй подход связан с учетом эффектов саморазогрева транзисторной структуры уже на стадии постановки вычислительной задачи. Вводятся обратные связи в уравнения модели между напряжениями, токами рассматриваемого элемента и изменением температуры транзисторной структуры, вследствие выделяемой мощности. Тем самым осуществляется интегральная оценка изменения характеристик транзистора в зависимости от физических процессов, вызванных эффектами саморазогрева [6.16].

Однако, введение обратной связи усиливает нелинейные свойства параметров модели, ведет к ее усложнению и к

увеличению жесткости системы уравнений, описывающих работу рассматриваемой транзисторной структуры.

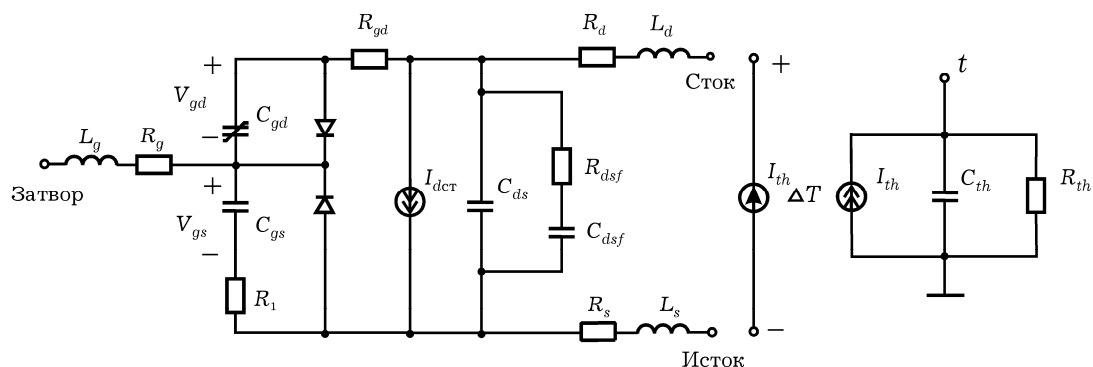


Рис. 6.6. Схемная модель гетеротранзистора с учетом цепей, учитывающих саморазогрев

Применяя методику расчета распределения температуры по уравнению теплопроводности предыдущего параграфа, получим распределения температуры для транзисторных гетероструктур с длиной затвора 1 мкм и 0,4 мкм как показано на рис. 6.7,а и 6.7,б соответственно.

Особенностью является то, что область разогрева простирается от омического контакта истока, проходит через область сильного поля под затвором и достигает омического контакта стока. В результате такого разогрева нарушается термодинамическое равновесие в приконтактных областях транзистора, что приводит к росту энергии электронов, уменьшению их подвижности и росту паразитных сопротивлений контактов.

Из температурной зависимости подвижности электронов [6.17] и распределения теплового поля транзисторной структуры в плоскости $xу$ (рис. 6.7) видно, что подвижность электронов в подзатворной области по сравнению с подвижностью электронов при 300 К может уменьшаться в несколько раз при типичных уровнях легирования областей ГСПТ.

Кроме того, за счет перераспределения электронов в потенциальных ямах возможно изменение потенциального рельефа, особенно в области проводящих слоев транзистора, что ухудшает управление транзистора полем затвора и приводит к уменьшению крутизны транзистора. Таким образом, распределение тепловых полей в многослойных гетероструктурах имеет ряд особенностей, которые необходимо учитывать при выборе топологии и конструкции транзистора с гетероселективным легированием. В частности, влияние рассмотренных эффектов можно ослабить при уменьшении активной области транзистора

как за счет уменьшения длины затвора, так и толщины активных слоев транзистора.

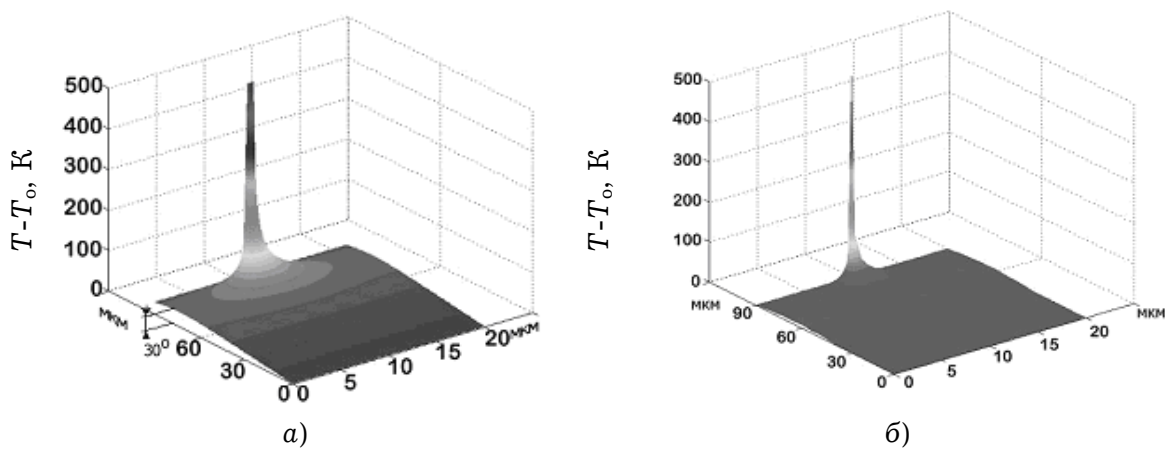


Рис. 6.7. Распределения температуры в гетеротранзисторах с длиной затвора 1 мкм (а) и 0,4 мкм (б).

По результатам двумерного моделирования тепловых полей в транзисторе с селективным легированием и температурных зависимостей подвижности можно оценить скорости дрейфа электронов в стационарном приближении для аппроксимаций и их применения для расчетов в аналитических и одномерных моделях. Кроме того, по распределениям тепловых полей с учетом процедуры усреднения можно оценить величины теплового сопротивления для схемной модели рис. 6.6.

6.3. ГЕТЕРОБИПОЛЯРНЫЕ ТРАНЗИСТОРЫ

Повышение быстродействия биполярных транзисторов возможно при использовании в качестве эмиттера материала с большей шириной запрещенной зоны, чем материал базы. Например, первые гетеротранзисторы имели эмиттер из n -GaAs, а базу из p -Ge. Позже стали использовать гетеропереходы на основе соединений $\text{Al}_x\text{Ga}_{1-x}\text{As}$, $\text{Al}_x\text{Ga}_{1-x}\text{N}$ и др.

На границе разных полупроводников возникают электрические поля, которые на энергетической диаграмме (рис.6.8) моделируются разрывом дна зоны проводимости, равным разности энергий электронного сродства $\Delta E_c = \chi_1 - \chi_2$, и разрывом потолка валентной зоны $\Delta E_v = (\Delta E_1 + \chi_1) - (\Delta E_2 + \chi_2)$. Совместно с полем контактной разности потенциалов $V_k = (\Phi_1 - \Phi_2)/e$ эти поля уменьшают высоту потенциального барьера для электронов и

увеличивают его для дырок валентной зоны. Асимметрия условий прохождения электронов и дырок через эмиттерный переход приводит к тому, что ток инжекции дырок j_p уменьшается в $\exp(\Delta E_v/kT)$ раз, а ток инжекции электронов j_n увеличивается в $\exp(\Delta E_c/kT)$ раз. В результате можно получить коэффициент инжекции электронов $\gamma = j_n/(j_n + j_p)$ практически равным единице, что делает процессы рекомбинации дырок в эмиттере несущественными.

В обычном гомопереходе отношение электронной и дырочной составляющих тока эмиттера определяется отношением концентрации доноров в эмиттере и акцепторов в базе, что вынуждало для увеличения коэффициента инжекции ограничивать степень легирования базы. В гетеропереходе возможно сильное легирование базовой области, что приводит к снижению продольного сопротивления базы и постоянной времени заряда коллекторной цепи $\tau_k = R_b C_{кб}$, которая существенно влияет на быстродействие транзистора при сверхтонких базах.

Кроме того, на границе эмиттер-база электроны приобретают дополнительную кинетическую энергию, равную ΔE_c , и при соответствующей толщине базы может быть реализован баллистический, или близкий к нему, режим переноса в узкозонном материале. Это уменьшает время пролета в базе, что является еще одним фактором повышения быстродействия.

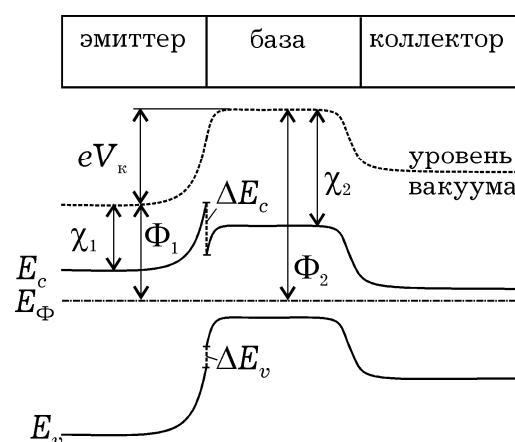


Рис. 6.8. Энергетическая диаграмма типичного *pnp*-гетеротранзистора

6.3.1. Двумерное моделирование транзисторов

Вертикальная структура субмикронного гетеробиполярного транзистора (СГБТ) показана на рис. 6.9. Предполагается, что такой транзистор размещается на относительно большой подложке, которая существенным образом влияет на распределение тепла.

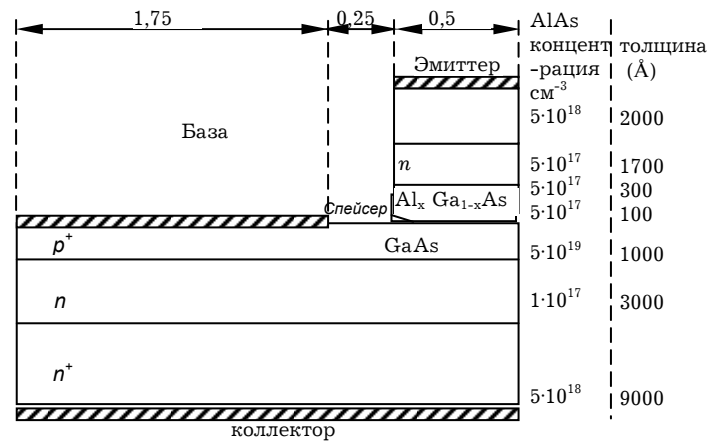


Рис. 6.9. Структура субмикронного гетеробиполярного транзистора (AlGaAs/GaAs)

К особенностям СГБТ следует отнести наличие разнородных физических процессов в эмиттерном гетеропереходе и коллекторном p - n переходе, а также процессов в области базы, которая в субмикронных транзисторах может составлять несколько десятков нанометров, и, как правило, не превышает 200 нм. Использование в эмиттере соединения $\text{Ga}_{1-x}\text{Al}_x\text{As}$ n -типа в контакте с p -легированной базой (GaAs) создает потенциальный барьер не менее 0,25 эВ, что практически исключает инжекцию дырок в эмиттере даже при высокой степени легирования базы. Это позволяет уменьшить толщину базы без увеличения ее сопротивления, а также снизить опасность смыкания переходов. Однако, как показывают результаты моделирования, СГБТ проигрывают полевым субмикронным гетеротранзисторам по уровню шума [6.18], кроме того, в гетеробиполярных транзисторах необходимо учитывать эффекты саморазогрева [6.19], ухудшающие выходные характеристики.

Благодаря новым возможностям технологии, применению новых полупроводниковых материалов и оптимизации топологии субмикронные гетеробиполярные транзисторы достигают предельных частот усиления более 200 ГГц [6.20] и используются в широком спектре устройств миллиметрового диапазона. Детальное моделирование процессов дрейфа и диффузии может быть основано на двумерной модели активной области транзистора в гидродинамическом приближении. Такие модели позволяют учесть основные эффекты биполярных субмикронных структур: разогрев электронов и дырок сильным электрическим полем и связанное с ним изменение подвижностей, процессы

рекомбинации, эффекты сильного легирования – изменение ширины запрещенной зоны и вырождения полупроводников.

В моделируемой структуре транзистора (рис. 6.9) учтена симметрия относительно вертикальной оси реального СГБТ, что позволяет рассматривать только одну половину структуры для уменьшения вычислительных затрат.

Процессы во внутренней части транзистора в самом общем случае может быть описаны уравнением Пуассона

$$\nabla^2 V = e(n - p - N_d + N_a)/\epsilon\epsilon_0, \quad (6.1)$$

и уравнениями непрерывности тока для электронов и дырок. Если для носителей заряда справедлива статистика Максвелла-Больцмана, то эффекты сильного легирования можно учесть дополнительными слагаемыми в системе диффузионно-дрейфовых уравнений:

$$\operatorname{div}[\mu_n(\nabla n - n\nabla \ln n_{ie} - n\nabla(V + \psi))] = R(p, n), \quad (6.2)$$

$$\operatorname{div}[\mu_p(\nabla p - p\nabla \ln n_{ie} - p\nabla(V + \psi))] = R(p, n). \quad (6.3)$$

Данная система с учетом уравнения Пуассона является полной. Здесь V – электрический потенциал, e – заряд электрона, n , p – концентрации электронов и дырок, $\epsilon\epsilon_0$ – диэлектрическая проницаемость, N_d , N_a – концентрации донорной и акцепторной примесей, μ_n , μ_p – подвижности электронов и дырок соответственно, R – скорость рекомбинации носителей заряда. В уравнениях (6.1) – (6.3) величины n_{ie} и ψ описывают эффекты сильного легирования. С использованием выражения для собственной концентрации носителей заряда $n_i = (N_c N_v)^{1/2} \exp(-E_g/2kT)$ может быть получено выражение для некоторой эффективной собственной концентрации носителей заряда:

$$n_{ie} = n_i \exp[(q\Omega_n + q\Omega_p)/2kT] = n_i \exp(-\Delta E_g/2kT).$$

В последней формуле величина $e\Omega_n$ трактуется как сужение ширины запрещенной зоны со стороны зоны проводимости, а величина $e\Omega_p$ – со стороны валентной зоны. Тогда величина $\Delta E_g = e\Omega_n + e\Omega_p$ является полным сужением ширины запрещенной зоны. Точное аналитическое определение величин $e\Omega_n$ и $e\Omega_p$ не представляется возможным. В дальнейшем эти величины для

моделирования выбираются на основе экспериментальных данных. Параметр ψ определяется как $\psi = e(\Omega_n - \Omega_p)/2$.

Зависимости подвижностей электронов и дырок от напряженности электрического поля рассчитываются в приближении времени релаксации для электронов и дырок.

Для решения системы уравнений используется метод прогонки в сочетании с итерационным многошаговым методом Ньютона. Граничные условия для уравнения Пуассона задавались как условия Дирихле на электродах (при электрических смещениях на эмиттере, базе и коллекторе соответственно 0,0, 1,0 и 5,0 В) и однородные условия Неймана на свободных и боковых границах.

В структуре транзистора использовались некомпенсированные полупроводники с распределением легирующей примеси как показано на рис. 6.10. Эмиттерный гетеропереход моделировался как скачок потенциала высотой около 0,4 эВ без учета квантования энергетических уровней.

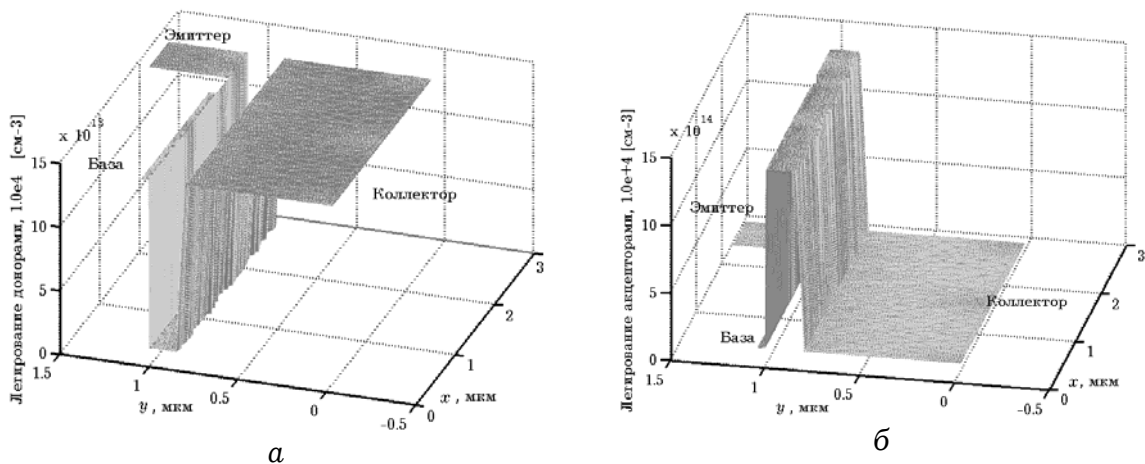


Рис. 6.10. Профиль легирования донорами(а) и акцепторами (б)

Анализ результатов моделирования (рис. 6.11, 6.12) показывает, что область сильного поля в СГБТ локализуется в межэлектродном промежутке база-коллектор и примыкает к области базы и $p-n$ перехода база-коллектор, а также вблизи эмиттерного гетероперехода. Подвижности электронов и дырок максимальны в межэлектродном промежутке эмиттер-база, что соответствует их низкополевым значениям вблизи «пороговых» значений напряженности электрического поля и почти не изменяются в сильных полях, что соответствует разогреву и междолинным переходам электронов и появлению горячих дырок.

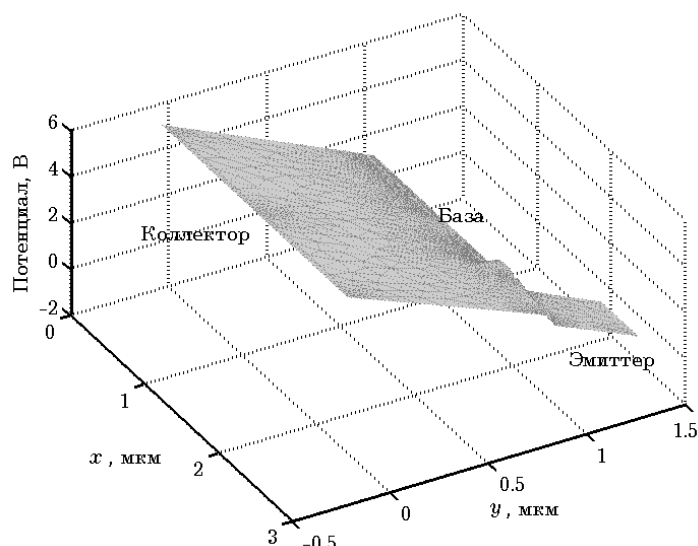


Рис. 6.11. Распределение потенциала в структуре гетеробиполярного транзистора

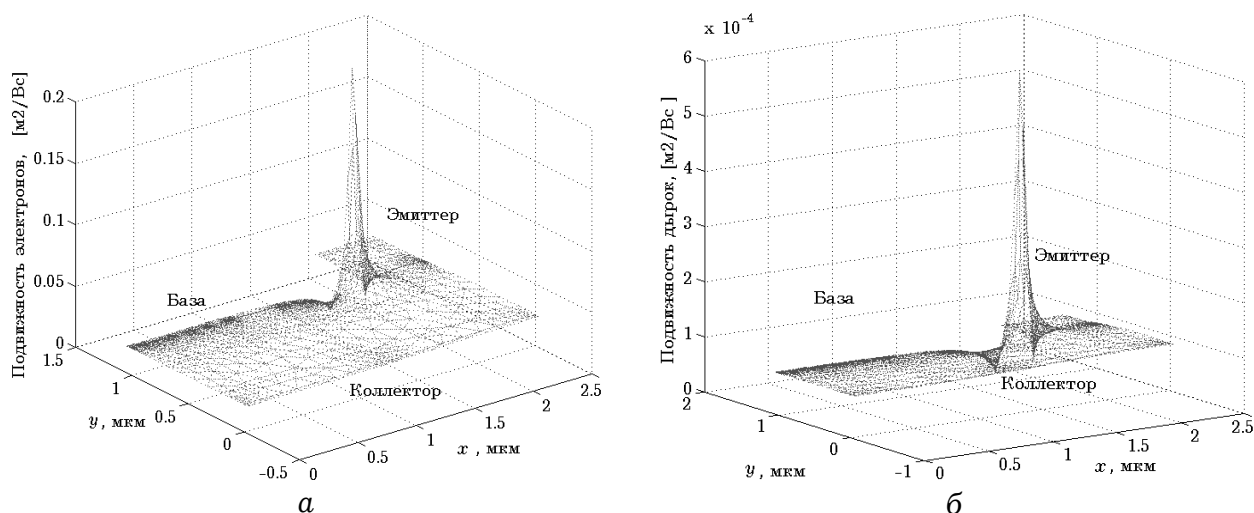


Рис. 6.12. Подвижность электронов (а) и дырок (б) в структуре транзистора

Особенности структуры и малые размеры активной области транзистора приводят к относительно высоким значениям напряженности электрического поля в области коллекторного перехода и коллектора (более $4 \cdot 10^7$ В/см), что связано с рассеиванием большой мощности и с необходимостью оптимального выбора толщины области коллектора.

6.3.2. Моделирование разогревных эффектов

Как показывают результаты анализа физических процессов в подзатворной области субмикронных ПТШ и ГСПТ и приколлекторной области СГБТ существуют области сильного поля и разогрев носителей заряда. Актуальной, особенно для

мощных структур, является задача исследования влияния разогревных эффектов. За счет дополнительного разогрева носителей возможно ухудшение параметров транзистора которое обусловлено как повышением температуры кристаллической решетки, так и разогревом электронного газа в канале, что приводит к уменьшению подвижности электронов. Представляет также интерес степень влияния разогревных эффектов на параметры схемных моделей транзисторов и частотные характеристики.

Как правило, мощные транзисторы представляют собой многосекционные структуры. При этом каждая секция может быть представлена в виде тепловых источников на поверхности подложки, которые рассчитываются с учетом функции распределения плотности мощности (потока энергии). Такое исследование возможно лишь при совместном анализе [6.21, 6.22] уравнения Пуассона, уравнений непрерывности тока для электронов и дырок, а также уравнения для потока энергии в двумерном приближении и учете температурных полей, возникающих в транзисторе за счет протекания тока.

Рассмотрим вертикальную структуру СГБТ, подобную приведенной выше. Предполагается, что транзистор располагается на относительно большой подложке, которая существенным образом влияет на распределение и отвод тепла. При этом необходимо учитывать тот факт, что размеры температурно-неравновесной области существенно превышают размеры электрически активной области, в которой происходит взаимодействие потоков электронов и дырок с электрическим полем. Эти особенности необходимо учитывать при задании граничных условий в моделируемой области для описания тепловых процессов. Область транзистора с тепловыми полями за границами активной области транзистора может быть, например, рассмотрена как некий дополнительный резистор, рассеивающий мощность. В моделируемой структуре транзистора учтена симметрия относительно вертикальной оси реального СГБТ, что позволяет рассматривать только одну половину структуры для уменьшения вычислительных затрат.

Активная область транзистора в общем случае может быть описана уравнениями Пуассона:

$$\nabla^2 V = e(n - p - N_d + N_a) / \epsilon \epsilon_0, \quad (6.4)$$

уравнениями непрерывности тока для электронов и дырок

$$\nabla \vec{J}_n = eR_n, \quad (6.5)$$

$$\nabla \vec{J}_p = eR_p, \quad (6.6)$$

а также уравнением для потока тепла

$$\nabla [c(T) \cdot \nabla T] = -P = \Delta \vec{J}_E. \quad (6.7)$$

Здесь \vec{J}_n , \vec{J}_p – плотности электронного и дырочного токов соответственно, $c(T)$ – зависящий от температуры коэффициент теплопроводности, P – мощность, рассеиваемая в единице объема, \vec{J}_E – плотность энергетического потока.

Температурная зависимость коэффициента теплопроводности для полуизолирующего арсенида галлия может быть представлена [6.23] аппроксимацией $c(T) = 0,76 - 0,001 \cdot T$ (Вт/см·К), а для арсенида галлия n -типа с концентрацией $3,5 \cdot 10^{17} \text{ см}^{-3}$ аналогичная зависимость может быть представлена как $c(T) = 1/(aT - \epsilon)$, где $a = 9,2 \cdot 10^{-3}$, $\epsilon = 0,44$.

В работе [6.22] кроме традиционных стационарных диффузионно-дрейфовых составляющих токов электронов и дырок предложен способ учета дополнительных составляющих токов. Эти составляющие учитывают эффекты сильного легирования, зависящие от уровня инжекции, и температурную зависимость коэффициента диффузии. С учетом этого уравнения для плотностей электронного и дырочного токов запишутся:

$$\vec{J}_n = \mu_n \beta_n kT \nabla n - n \mu_n \left[\nabla (eV + \chi) + \beta_n kT \frac{\Delta N_c}{N_c} - \alpha_n k \nabla T \right], \quad (6.8)$$

$$\vec{J}_p = -\mu_p \beta_p kT \nabla p - p \mu_p \left[\nabla (eV + \chi + E_g) - \beta_p kT \frac{\Delta N_v}{N_v} + \alpha_p k \nabla T \right]. \quad (6.9)$$

Как показано в работе [6.25], выражение для плотности потока энергии может быть записано как

$$\vec{J}_E = \left\{ - \left[\alpha_n kT + E_c \right] \frac{\vec{J}_n}{e} - c_n \Delta T \right\} + \left\{ \left[\alpha_p kT - E_v \right] \frac{\vec{J}_p}{e} - c_p \Delta T \right\}, \quad (6.10)$$

где коэффициенты c_n и c_p могут идентифицироваться как коэффициенты теплопроводности соответственно для электронов и дырок.

При совместном решении системы уравнений (6.4–6.10) предполагается отсутствие излучения тепла и конвективных потоков, а также независимость квазиуровней Ферми для электронов и дырок от температуры. Эти допущения позволяют рассматривать полупроводник, включая подсистемы свободных носителей заряда и решетку, как термодинамическую систему.

Для анализа эффекта разогрева возможен также подход, основанный на следующем алгоритме. На первом этапе решаются диффузионно-дрейфовые уравнения структуры транзистора и уравнение Пуассона при комнатной температуре решетки. Далее по результатам моделирования мощность рассчитывается из уравнения:

$$P = -\int_V (\varphi_n \operatorname{div} \vec{J}_n + \operatorname{grad} \varphi_n \vec{J}_n + \varphi_p \operatorname{div} \vec{J}_p + \operatorname{grad} \varphi_p \vec{J}_p) dV.$$

По выделяемой мощности рассчитывается температура активной области транзистора. Так как процессы тепловыделения связаны с областью протекания тока через транзистор, то, как правило, наибольший разогрев происходит в приповерхностных слоях. Для ГСБТ такой областью является область коллектора. Рассчитанное таким образом значение температуры задается как граничное условие Дирихле при моделировании тепловых полей в структуре транзистора при решении уравнения теплопроводности для получения функции $T(x,y)$. На заключительном шаге осуществляется повторный пересчет распределений физических величин по уравнениям Пуассона и уравнениям непрерывности для электронов и дырок, в которых температура кристаллической решетки задается неравновесной функцией $T(x,y)$, что позволяет учесть разогревные эффекты посредством температурных зависимостей коэффициентов диффузии и подвижностей для электронов и дырок.

Для интегрального учета разогрева по известной функции $T(x,y)$ путем усреднения можно рассчитать эквивалентное тепловое сопротивление R_t :

$$-c(T)\Delta T \cdot \vec{n} = \frac{(T - T_0)}{R_t},$$

где \vec{n} – единичный вектор, нормальный к внешней поверхности, T_0 – равновесная температура. Тепловое сопротивление рассчитывается с учетом распределения температурных полей в подложке транзистора и процедуры усреднения:

$$T_{\text{ср}} = \frac{\int T(x, y) dx dy}{l_x l_y}, \quad R_t = \frac{T_{\text{ср}} - T_0}{Q}, \quad \text{где} \quad Q = \frac{\oint \vec{I}_E \cdot \vec{n} \cdot da}{l_x l_y},$$

а l_x и l_y - размеры моделируемой области.

На рис. 6.13 представлен пример расчета тепловых полей для топологической структуры ГСБТ (рис. 6.3), полученных в результате решения уравнения теплопроводности при заданной температуре на электроде коллектора в соответствии с подходами расчета процессов разогрева из работы [6.26].

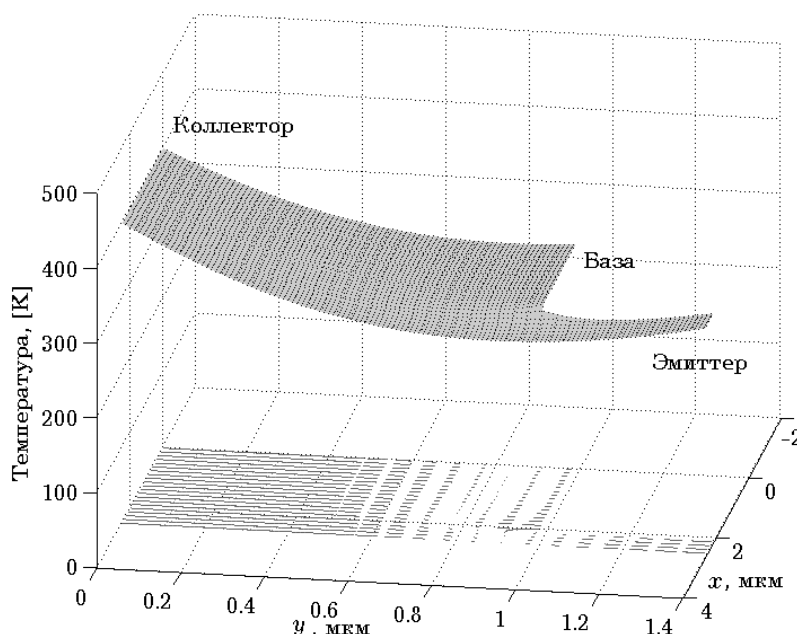


Рис. 6.13. Распределение тепловых полей в структуре ГСБТ (в проекции на горизонтальной плоскости приведено поле градиентов температуры)

6.3.3. Схемотехническое моделирование процессов саморазогрева

При схемотехническом моделировании субмикронных полупроводниковых устройств, в частности, гетероструктурных биполярных транзисторов, одним из важных аспектов является учет влияния эффектов саморазогрева на электрические характеристики исследуемых устройств. Подобные эффекты обусловлены возрастанием температуры полупроводникового устройства в результате рассеяния мощности при довольно высоких значениях рабочих токов. При этом возрастание напряженности электрического поля в активной области транзистора приводит к возникновению сильнополевых эффектов,

связанных с разогревом носителей. Дрейф и диффузия в сильном электрическом поле, особенно в области переходов, может приводить к накоплению носителей, в том числе за счет их междолинных переходов. Кроме того, в области базы могут усиливаться процессы генерации и рекомбинации. Интегрально эти процессы могут приводить к изменению рабочей точки на вольтамперной характеристике и к процессам запаздывания тока при импульсном входном воздействии. Проблема описания эффектов саморазогрева особенно актуальна для транзисторов, которые работают в режиме большого сигнала, если рабочая точка эволюционирует в широком диапазоне значений напряжения питания. В большинстве математических моделей полупроводниковых компонентов, которые входят в состав пакетов схемотехнического моделирования, эти эффекты описываются формально [6.27-6.29].

С позиций схемотехнического моделирования адаптация математической модели полупроводникового устройства для учета эффектов саморазогрева сводится к нахождению отклонения между изотепловым режимом и режимом саморазогрева с последующим определением параметров саморазогрева в зависимости от найденного теплового отклонения рабочих токов.

Для субмикронных гетеробиполярных транзисторов нелинейные отклонения исходных токов, обусловленные саморазогревом, и изменения тока коллектора можно учесть введением дополнительных составляющих уравнения токов базовой или коллекторной области. Однако подобная компенсация не является в полной мере адекватной [6.30].

Более адекватный учет влияния эффекта саморазогрева на электрические параметры субмикронных гетероструктурных транзисторов может быть связан с определением зависимостей параметров устройства от температуры и соответствующим дополнением математических моделей полученными зависимостями [6.31].

Для моделирования рассмотрим гетероструктурный биполярный транзистор, структура которого приведена на рис. 6.14, со следующими электрофизическими параметрами: толщиной базы 0,14 мкм и уровнем легирования бериллием 10^{19} см^{-3} ; толщиной коллектора 0,7 мкм с кремниевым легированием $7,5 \times 10^{15} \text{ см}^{-3}$; омической металлизацией базы – сплава золота с бериллием, эмиттера и коллектора – сплава золота с германием. Верхний слой эмиттера состоит из высоколегированного слоя n^+ - GaAs и нескольких слоев AlGaAs. База представляет собой тонкий сильно легированный акцепторами слой.

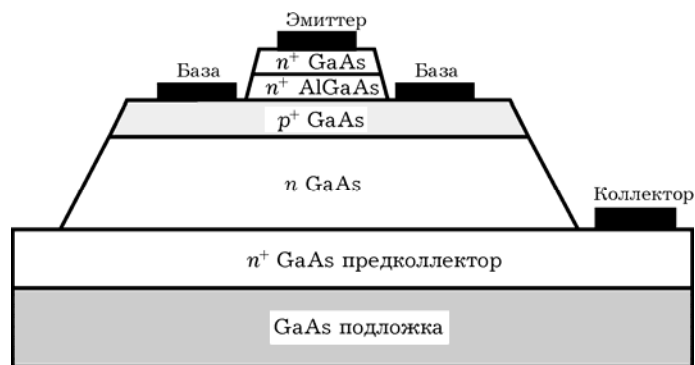


Рис. 6.14. Топологическая структура гетеробиполярного транзистора.

Предлагаемый подход заключается в схемотехническом моделировании и выборе значений балластных сопротивлений в коллекторной и базовой областях транзисторной структуры, которая, в свою очередь, приводит к определению типа обратной связи, а также типа нелинейного зависимого источника. Именно этот источник отображает влияние эффекта саморазогрева.

Введенная в модель обратная связь (ОС) может быть как «положительной» (ПОС), так и «отрицательной» (ООС). ПОС осуществляется в том случае, если в схеме как зависимый источник используется источник напряжения, управляемый напряжением (ИНУН), как показано на рис. 6.15. В этом случае с увеличением напряжения коллектор-эмиттер $V_{кэ}$ происходит увеличение тока коллектора, ООС осуществляется при использовании зависимого источника тока, управляемого напряжением (ИТУН), и тогда с увеличением напряжения $V_{кэ}$ происходит уменьшение тока коллектора.

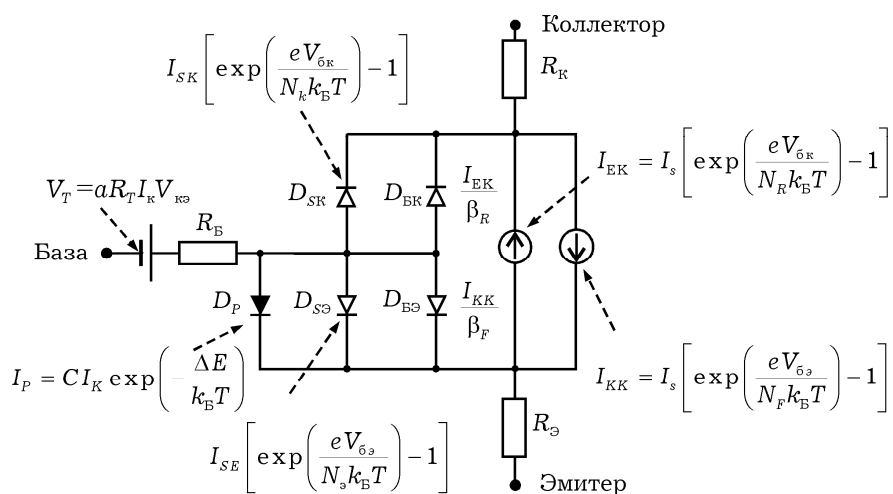


Рис. 6.15. Схемная модель гетероструктурного биполярного транзистора

Сопротивления R_k , R_b , R_e описывают сопротивления металлизации. Две пары диодов $D_{bэ}$, $D_{сэ}$, D_{bk} , $D_{ск}$ описывают нелинейные эффекты в ГСБТ. Первая пара диодов $D_{bэ}$ и $D_{сэ}$ моделирует изменение насыщения тока при изменении тока коллектора, диод $D_{bэ}$ – связан с моделированием инжекции электронов к эмиттеру, а диод $D_{сэ}$ – процессов рекомбинации в базе, образование области накопления пространственного заряда и поверхностную рекомбинацию. Вторая пара диодов D_{bk} , $D_{ск}$ моделирует зависимость обратного тока от напряжения смещения, диод D_{bk} – моделирует инжекцию электронов из коллектора в базу или дырок из базы в коллектор, диод $D_{ск}$ – рекомбинацию в базе, образование области пространственного заряда и поверхностную рекомбинацию при инверсном режиме работы. Диод D_p – моделирует дырочную составляющую тока коллектора, а зависимый источник (ИНУН) V_{th} – положительную обратную связь, температурная зависимость которой описывается уравнением: $V_{th} = \alpha R_{th} I_k V_k$, где α – коэффициент электротепловой обратной связи [В/К]; R_{th} – тепловое сопротивление [К/Вт]; I_k – ток коллектора [А]; V_k – напряжение на коллекторе [В].

Ток базы в этой модели описывается двумя парами параллельных диодов, соответственно, для перехода база-эмиттер и перехода база-коллектор. Использование двух параллельных диодов для моделирования каждого перехода позволяет рассчитать зависимость усиления тока от напряжения смещения. Первый диод каждого перехода отвечает току рекомбинации в области пространственного заряда при малых напряжениях смещения. Эта составляющая тока связана с прямосмещенным барьером с коэффициентом неидеальности n близким к 2. Второй диод описывает составляющую тока перехода между эмиттером и коллектором.

Ток коллектора связан с коэффициентом неидеальности n в прямосмещенном переходе. Так как ток коллектора транзистора является результатом инжекции электронов в базу при диффузии, то коэффициент неидеальности близок к 1.

Наиболее полно учет эффекта саморазогрева достигается при комбинированном использовании ПОС и ООС [6.31 6.32], что позволяет предложить адаптированный подход учета эффекта саморазогрева гетеробиполярной транзисторной структуры.

Так как источник тока, управляемый напряжением, может быть эквивалентно замещен источником напряжения, управляемым напряжением, то можно сделать предположение, что если бы при изменении сопротивления нагрузки ОС непрерывно изменялась от отрицательной к положительной, то при некотором оптимальном значении сопротивлении нагрузки

$R_{\text{опт}}$ глубина ОС сводится к нулю, это означает, что эффект саморазогрева учтен.

С точки зрения электрической цепи ОС, электротепловая ОС внутри ГСБТ может быть количественно описана с помощью коэффициента передачи обратной связи:

$$f = \frac{\partial I_{\text{к}}}{\partial T} \cdot \frac{\partial T}{\partial P} \cdot \frac{\partial P}{\partial I_{\text{к}}} = \frac{\partial I_{\text{к}}}{\partial T} R_{\text{th}} V_{\text{к}}, \quad (6.11)$$

где P – рассеиваемая мощность [Вт]; T [К] – температура устройства, R_{th} – сопротивление теплового резистора.

Коэффициент ОС принимает положительные и отрицательные значения, которые отвечают положительной и отрицательной обратной связи. Тепловая нестабильность наступает при $f > 1$, а эффект саморазогрева может быть учтен при сведении f к нулю. Но поскольку и R_{th} , и $V_{\text{к}}$ положительные, то из (6.11) видно, что знак коэффициента ОС определяется соотношением $(\partial I_{\text{к}}/\partial T)$.

Если в качестве управляемого источника используется ИНУН, расположенный в базовой области, то ток коллектора может быть записан:

$$I_{\text{к}} = I_{\text{с}} \exp \left\{ \frac{e}{kT_0} [V_{\text{бэ}} - I_{\text{б}} V_{\text{б}} + \alpha (T - T_0)] \right\}, \quad (6.12)$$

где $I_{\text{с}}$ – ток насыщения [А]; $I_{\text{б}}$ и $V_{\text{бэ}}$ – ток базы и напряжение база-эмиттер, соответственно; T_0 – внешняя температура [К].

Принимая во внимание значение дырочной составляющей тока:

$$I_{\text{п}} = CI_{\text{к}} \exp \left(-\frac{\Delta E}{kT} \right), \quad (6.13)$$

ток базы можно выразить через ток коллектора следующим образом:

$$I_{\text{б}} = I_{\text{к}}/\beta + I_{\text{п}} = I_{\text{к}} \left[1/\beta + C \exp \left(-\frac{\Delta E}{kT} \right) \right]. \quad (6.14)$$

где ΔE – эффективная разность в высоте потенциального барьера для дырок и для электронов [эВ]; C – отношение тока электронов к току дырок для $\Delta E = 0$; β – коэффициент усиления тока.

Подставляя (6.14) в (6.12) получаем:

$$V_{\bar{\sigma}_3} = \frac{kT_0}{e} \ln \frac{I_{\kappa}}{I_s} + I_{\kappa} R_{\bar{\sigma}_6} \left[\frac{1}{\beta} + C \exp \left(-\frac{\Delta E}{kT} \right) \right] - \alpha (T - T_0), \quad (6.15)$$

дифференцируя (6.15) относительно T и решая относительно $\partial I_{\kappa} / \partial T$ получаем:

$$\frac{\partial I_{\kappa}}{\partial T} = \frac{\alpha - \frac{\Delta E}{kT^2} R_{\bar{\sigma}_6} I_p}{\frac{kT_0}{e} + I_{\bar{\sigma}_6} R_{\bar{\sigma}_6}} I_{\kappa}. \quad (6.16)$$

Коэффициент передачи цепи ОС приобретает вид:

$$f = \frac{\alpha - \frac{\Delta E}{kT^2} R_{\bar{\sigma}_6} I_p}{\frac{kT_0}{e} + I_{\bar{\sigma}_6} R_{\bar{\sigma}_6}} R_{th} I_{\kappa} V_{\kappa}, \quad (6.17)$$

где первая и вторая части числителя описывают положительную и отрицательную ОС в ГСБТ, соответственно.

Приравнивая (6.17) к нулю ($f = 0$) с учетом (6.11) и (6.13), получаем:

$$R_T = \frac{\alpha T}{C I_{\kappa} (\Delta E / kT)} \exp \left(\frac{\Delta E}{kT} \right) = \frac{\alpha T}{(\Delta E / kT) I_p}.$$

Если выполняется неравенство $R_{\bar{\sigma}_6} < R_T$, то мы имеем дело с ПОС, если $R_{\bar{\sigma}_6} > R_T$ – ООС, если $R_{\bar{\sigma}_6} = R_T$ – ОС отсутствует.

Коэффициент передачи ОС при наличии дополнительных сопротивлений в области базы и в области эмиттера будет иметь следующий вид:

$$f = \frac{\alpha - \frac{\Delta E}{kT^2} (R_{\bar{\sigma}_6} + R_{\bar{\sigma}_3}) I_p}{\frac{kT_0}{e} + I_{\bar{\sigma}_6} R_{\bar{\sigma}_6} + I_{\bar{\sigma}_3} R_{\bar{\sigma}_3}} R_{th} I_{\kappa} V_{\kappa}.$$

При расчетах использовались следующие значения констант: $\alpha = 1,2$ мВ/К; $\Delta E = 0,28$ эВ; $k = 1,38 \cdot 10^{-23}$ Дж/К; $T_0 = 300$ К; $R_{th} = 1300$ К/Вт; $I_{\kappa} = 40$ мА; $V_{\kappa} = 3$ В; $C = 5$; $\beta = 100$.

На рис. 6.16 приведена схема однокаскадного усилителя СВЧ, которая включает индуктивности $L_{\bar{\sigma}_6}$ (10 нГн), L_{κ} (10 нГн), $L_{\bar{\sigma}_3}$ (1.49 нГн), которые обеспечивают согласование входных и выходных цепей усилителя; L_1 (1.97 нГн) и L_2 (1 нГн) – фильтрующие индуктивности в цепях питания; разделительные емкости C_1 (1 пф) и C_2 (1 пФ) на входе и на выходе усилителя.

Активным элементом каскада является ГСБТ, схема замещения которого приведена выше.

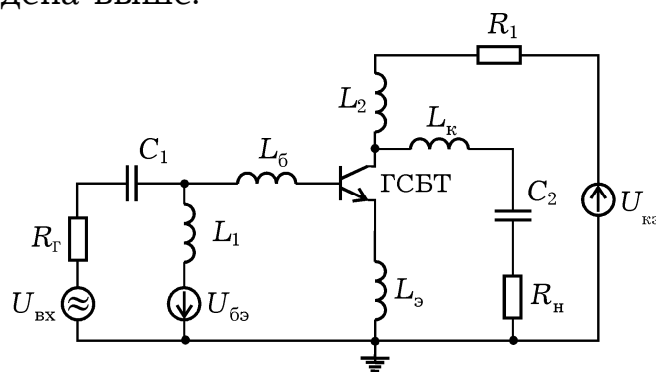


Рис. 6.16. Схема однокаскадного усилителя на гетеробиполярном транзисторе

По результатам моделирования частотных характеристик, приведенных на рис. 6.17,а и 6.17,б, можно сделать вывод, что увеличение толщины базы (то есть увеличение сопротивления базы R_6) приводит к смещению частоты усиления с соответствующим изменением коэффициента усиления по напряжению, а увеличение сопротивления $R_э$ приводит к увеличению коэффициента усиления по напряжению.

Совместное использование приведенной методики и результатов двумерного физико-топологического моделирования позволяет получить уточненные оценки величин сопротивлений базы, эмиттера и коллектора при различных температурах и степени влияния эффекта саморазогрева.

Рассмотренная методика моделирования процессов саморазогрева в субмикронных гетеробиполярных структурах праведлива для широкого спектра полевых структур.

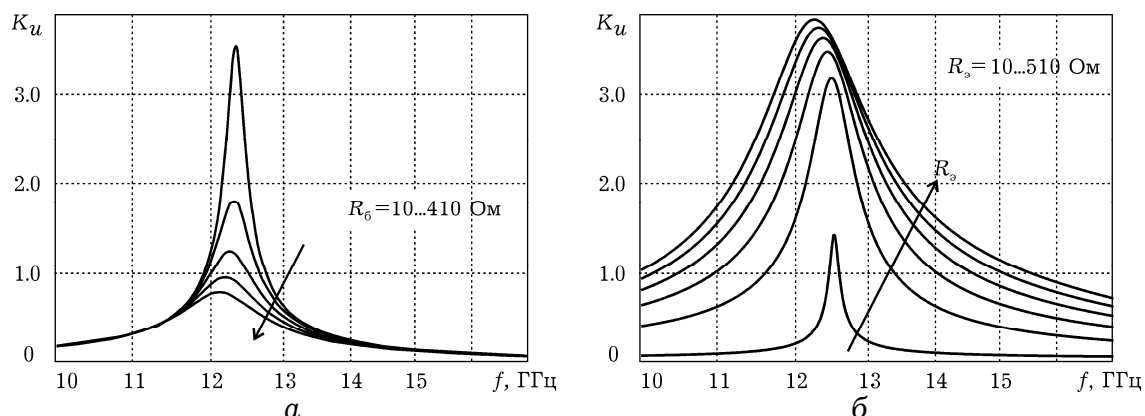


Рис. 6.17. Зависимости коэффициента усиления напряжения от частоты при различных значениях сопротивления базы R_6 и эмиттера $R_э$.

Полученные результаты свидетельствуют об особенностях распределения тепловых полей в подзатворной области транзисторной структуры, связанных с наличием приповерхностного слоя с повышенной температурой, что увеличивает эффект саморазогрева в транзисторе как за счет дополнительного разогрева электронов в канале, так и за счет их разогрева в истоковой и стоковой областях за пределами активной области. При этом повышается температура носителей, инжектированных из истока в активную область, что повышает вероятность их междолинного рассеяния. В области стока это приводит к задержке процессов релаксации энергии.

Ухудшение параметров транзисторной структуры при дополнительном разогреве (уменьшение дрейфовой скорости и подвижности, рост шума) обусловлено как увеличением паразитных сопротивлений, так и разогревом электронного газа в канале, что приводит к уменьшению подвижности электронов.

Из приведенных расчетов и результатов моделирования, видно, что для адекватного моделирования и учета эффекта саморазогрева важным является не только описание зависимого источника, который влияет на нелинейные процессы в транзисторе, но и выбор значения балластного сопротивления, которое учитывает в схемной модели ГСБТ влияние указанных эффектов.

В режимах работы транзистора, когда эффекты саморазогрева более заметны, коэффициент усиления уменьшается во всем диапазоне рабочих частот. Следует ожидать также увеличения коэффициента шума ГСБТ как вследствие влияния эффекта саморазогрева на физические процессы в области переходов, так и появлением дополнительных балластных элементов схемной модели пассивных областей базы и эмиттера.

6.3.4. Электротепловой анализ монолитных интегральных схем на субмикронных гетеробиполярных транзисторах

Предложенный выше подход может быть распространен для оценки теплового состояния монолитных интегральных микросхем (МИС), включая результаты исследований распределения теплового потока для полевого транзистора с затвором Шоттки и полевого транзистора с гетероселективным легированием [6.10, 6.12, 6.33, 6.34].

Основываясь на результатах электротеплового анализа отдельной структуры транзистора, можно рассчитать тепловые режимы МИС и распределение тепла в кристалле МИС. Для этого необходимо учитывать характеристики отдельных транзисторных

структур с учетом топологии интегральной схемы. Разбиение МИС на элементарные блоки, расчет распределения тепла по методике, приведенной в работах [6.33, 6.34], и «сшивание» результатов, позволяют получить полную картину распределения тепла в кристалле МИС, что в дальнейшем может быть использовано для выбора методик теплоотведения.

Как правило, КВЧ МИС включают диоды, транзисторы, резисторы, емкости и другие компоненты, которые могут быть как источниками тепла, так и рассеивающими элементами. Это осложняет процесс построения тепловой модели и определения распределения температуры в МИС. Процедура расчета распределения тепла в МИС состоит из двух этапов.

На первом этапе определяется тепловое сопротивление активных элементов МИС, исходя из данных о геометрических размерах и теплофизических параметрах элементов, по соотношению [6.25] $R_T = R_T^\Gamma (h_2 - h_1)(h - h_2)/h^2$, К/Вт, где h , h_1 , h_2 – высоты кристалла и изломов боковой поверхности теплового эквивалента, соответственно, мкм; $R_T^\Gamma = h / (S\lambda)$ – геометрический эквивалент теплового сопротивления, К/Вт; S – площадь вертикального сечения, м²; λ – коэффициент теплопроводности материала, Вт/(м·К). Затем, используя принцип электротепловой аналогии, определяется значение температуры каждого элемента: $T_p = T_0 + \Delta t = T_0 - P \cdot R_T$, К, где T_0 – температура теплоотвода, К; P – рассеиваемая мощность, Вт.

На втором этапе производится физико-топологическое моделирование МИС. Решается уравнение теплопроводности: $-\nabla(\lambda \nabla T) = P + h(T_{ext} - T)$ со смешанными граничными условиями Дирихле $hT = r$ и Неймана $\vec{n}(c \nabla T) + qT = g$, где h – коэффициент конвективного теплообмена с внешней средой, Вт/(м²·К); r , c , q , g – функции, определенные на замкнутой области; \vec{n} – нормаль к границе раздела двух сред.

В качестве моделируемой структуры был выбран широкополосный усилительный каскад на основе AlGaIn/GaN полевых транзисторов с гетероселективным легированием на кремниевой подложке [6.36].

Топология исследуемой структуры приведена на рис. 6.18, а результаты моделирования — на рис. 6.19.

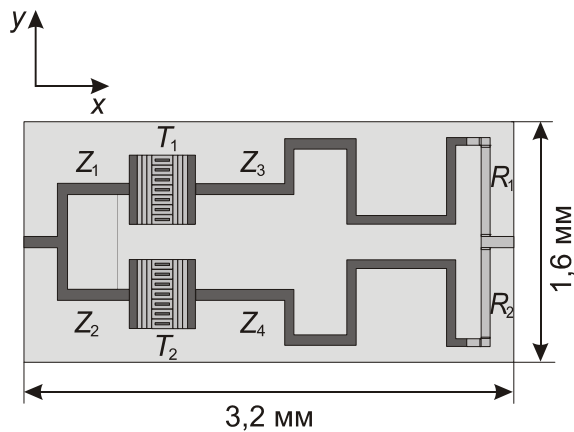


Рис. 6.18. Топология широкополосного усилителя: T_1 , T_2 – транзисторные структуры; Z_1 , Z_2 , Z_3 , Z_4 – линии передачи; R_1 , R_2 – омические контакты.

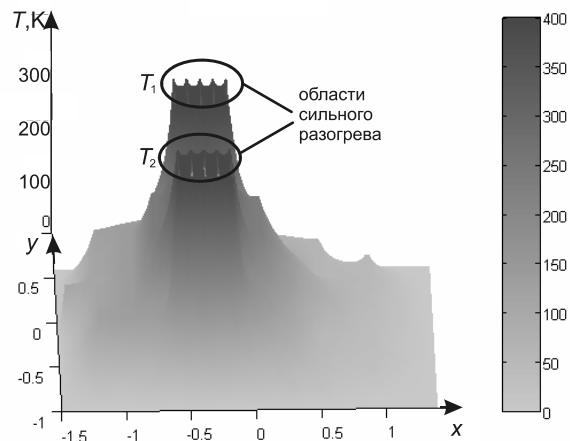


Рис. 6.19. Распределение температурных полей в МИС

Полученные распределения температур с экстремумом в области размещения транзисторных структур свидетельствуют о необходимости учета при проектировании и выборе топологии высоких градиентов температур. Совместное использование физико-топологических и тепловых моделей позволяет оптимизировать топологическую структуру МИС с точки зрения рассеиваемой мощности и эффективного теплоотвода.

6.3.5. Шумовые свойства субмикронных гетеробиполярных транзисторов

Уступая по минимальному уровню собственных шумов субмикронным транзисторам с барьером Шоттки и гетероструктурным транзисторам, СГБТ, имеющие граничные частоты усиления более 200 ГГц, незаменимы при передаче высокочастотных импульсных сигналов, например, в оптоволоконных и беспроводных каналах связи миллиметрового диапазона, требующих сверхширокой рабочей полосы пропускания. При этом актуальной становится задача моделирования шумовых характеристик таких транзисторов [6.3, 6.35].

При моделировании малозумящих устройств, как правило, ищется оптимальное соотношение сигнал/шум. В субмикронных структурах в активной области транзистора преобладают сильные поля, что приводит к эффекту «саморазогрева» структуры. Саморазогрев является дополнительной причиной усиления нелинейных эффектов, что связано в большей степени с дрейфом

и диффузией горячих электронов и дырок, а также нарастанием процессов генерации-рекомбинации.

Особенностью СГБТ является появление участка с отрицательной дифференциальной проводимостью на пологом участке вольтамперной характеристики. В работе [6.37] обсуждается возможность “компенсации” эффекта саморазогрева путем введения некоторых балластных резисторов (порядка нескольких сотен Ом) в области базы и коллектора. Однако, это, как ожидается, может приводить к уменьшению крутизны и коэффициента усиления. Кроме того, можно ожидать увеличения шума омических областей.

Для полноценного анализа и оптимизации шумовых и усилительных характеристик транзистора необходимо использовать методику расчета, как внутренних локальных источников шума, так и суммарных источников на входе и на выходе транзистора, представленного четырехполюсником. Использование волновой шумовой матрицы τ , элементы которой представляют собой суммарные спектральные плотности мощности шума на входе и на выходе с учетом их корреляционной связи, позволяет при известной матрице рассеяния рассчитать минимальный коэффициент шума и так называемый сопутствующий коэффициент усиления по мощности.

Шумовая схемная СГБТ модель может быть в общем случае представлена в виде шумовых источников внутренней части транзистора и шумовых источников контактов, которые традици-

онно описываются соотношением Найквиста: $\overline{U_n}^2 = 4kT_0\Delta fR$.

Спектральные плотности рассчитываются в предположении переноса электронов базы за счет процессов диффузии и не связаны с инжекцией через гетеропереход. Это предположение действительно, пока механизмы баллистического переноса электронов не преобладают, что, как правило, выполняется для толщин базы более 50 нм. В случае, когда инжекция через гетеропереход преобладает, то спектральные плотности мощности шума существенно другие.

К особенностям описания шумовых источников СГБТ следует отнести наличие разнородных физических процессов в двух переходах – гетеропереходе и p - n переходе, а также процессов в области базы, которая в субмикронных транзисторах может составлять несколько десятков нанометров, и, как правило, не превышает 200 нм.

Дробовой шум базы в биполярном транзисторе описывается соотношением [6.36], справедливым на низких частотах для

гетеробиполярных транзисторов: $\overline{i_{шб}}^2 = 2eI_0 C_{бб} \Delta f$. Подобное соотношение справедливо для коллектора: $\overline{i_{шк}}^2 = 2eI_k C_{кк} \Delta f$. Коэффициенты $C_{бб}$ и $C_{кк}$ на низких частотах близки к единице. На высоких частотах шум за счет протекания тока коллектора практически незаметен, в то время как дробовой шумовой ток базы становится заметным в диапазоне частот 1-40 ГГц и существенно возрастает на частотах выше 40 ГГц. Это может быть связано с тем, что на низких частотах электроны инжектируются из эмиттера в базу, где они собираются коллектором и рекомбинируют. На высоких частотах часть флуктуирующих электронов возвращается на эмиттер до того, как они рекомбинируют или попадут на коллектор. Эти возвращающиеся электроны приводят к повышению тока базы, но мало изменяют шумовой ток коллектора. Как будет показано, частотная зависимость базового шумового тока связана также с корреляцией между базовым и коллекторным шумовыми источниками.

Следует отметить, что в самом общем случае необходимо учитывать также фликкер-шум ($1/f$): $\overline{i_{шбf}}^2 = \kappa_f R_{бб} I_0^2 / f$ и взрывной шум (шум, имеющий спектр Лоренца): $\overline{i_{шбв}}^2 = \kappa_v R_{бб} I_0^2 \tau / [1 + (\omega\tau)^2]$. Здесь κ_f – лоренцовый коэффициент, пропорциональный скорости рекомбинации в приконтактном слое базы и зависящий от времени жизни τ свободных носителей на ловушках (центрах захвата): $\tau = (\tau_0/T^2) \exp(E_a/kT)$, где τ_0 – константа, E_a – энергия активации ловушечных состояний. При этом время жизни τ и лоренцевый спектр являются сильно зависимыми от температуры кристаллической решетки.

Принято считать, что фликкер-шум доминирует на низких частотах, взрывной шум – на средних, а дробовой на высоких. При анализе шума на СВЧ фликкер-шумом можно пренебречь. В настоящее время не существует единого физического механизма взрывного шума при прямых и обратных смещениях p - n перехода. Однако, некоторые экспериментальные данные свидетельствуют о случайных процессах в приповерхностном слое канала. При этом энергия активации при обратных смещениях перехода оценивается как близкая к 0,8 эВ. Учитывая современную конструкцию СГБТ, когда свободная граница между контактами сведена к минимуму, а также типичную длительность импульса взрывного шума, которая составляет более 100 нс, можно предположить, что данная составляющая шума в СВЧ диапазоне практически не проявляется. Однако, захват носителей на ловушки и изменение вследствие этого энергетических зон в

приповерхностном слое может приводить к дополнительному разогреву носителей в области канала из-за его сужения и может изменять сопротивление межконтактных областей и приводить к увеличению теплового шума пассивных областей транзистора.

Одной из наиболее существенных составляющих шума в СГБТ является дробовой шум, соотношения для которого в области базы и коллектора с учетом приведенных выше рассуждений запишутся, соответственно, для базы [6.39]:

$$\overline{i_{шб}^2} = 2eI_э \Delta f \left| \frac{C_э}{C_{э0}} \right|^2 - 2eI_к \Delta f \left| \frac{C_к}{C_{к0}} \frac{n_э}{n_к} \right| - 2eI_э \Delta f \left| \frac{C_э}{C_{э0}} - 1 \right|^2 + 2eI_к \Delta f \left| \frac{C_к}{C_{к0}} \frac{n_э}{n_к} e^{-j\omega\tau_к} - 1 \right|^2,$$

где $\tau_к = x_к / (2v_н)$ – время запаздывания фазы тока коллектора, $x_к$ – толщина коллектора, $v_н$ – скорость насыщения носителей, и для коллектора: $\overline{i_{шк}^2} = 2eI_к \Delta f$, а $n_э = \beta n_к + n_б / (1 + \beta)$, где β – коэффициент передачи тока. Коэффициенты $C_э$ и $C_к$ определяются по формулам: $C_э = \gamma x_б / \text{th}(\gamma x_б)$ и $C_к = \gamma x_б / \text{sh}(\gamma x_б)$, при этом на постоянном токе ($\omega = 0$) эти коэффициенты принимают значения соответственно $C_{э0}$ и $C_{к0}$. Постоянная распространения в базе определится как:

$$\gamma = \frac{1}{\sqrt{D_n \tau_n}} \sqrt{1 + j\omega\tau_n},$$

где D_n – коэффициент диффузии, τ_n – время жизни электронов в базе. Следует заметить, что «диффузионный» механизм описания шума в базе справедлив, когда длина свободного пробега носителей в базе меньше толщины базы. При уменьшении толщины базы и возможном баллистическом движении горячих носителей изменяются механизмы протекания тока и усиливается высокочастотная составляющая шума. Шумовые токи базы и коллектора имеют корреляционную связь, которая описывается соотношением:

$$\overline{i_{шб}^* i_{шк}} = 2eI_к \left(\frac{C_к n_э}{C_{к0} n_к} e^{-j\omega\tau_к} - 1 \right).$$

В случае сильного разогрева носителей необходимо учитывать зависимость подвижности электронов и дырок от напряженности электрического поля, которые определяются в приближении времени релаксации носителей.

В рассматриваемой шумовой модели возможно использование некоторых усредненных значений коэффициента диффузии,

определенных для средних значений электрического поля в зависимости от положения рабочей точки на вольтамперных характеристиках.

Получение минимального коэффициента шума СГБТ связано с обеспечением ряда необходимых факторов. К таким факторам следует отнести оптимальный выбор конструкции и топологии транзистора, его электрического режима, а также обеспечение оптимального режима согласования.

Анализируемая шумовая модель СГБТ представлена на рис. 6.20. Переход база-эмиттер моделируется RC -цепочкой $R_{дэ}$ и $C_{дэ}$, которая описывает параметры модельного диода база-эмиттер, а также соответствующими шумовыми источниками. При этом емкость представляет собой сумму диффузионной емкости и емкости прямого смещенного перехода. Напряжение база-эмиттер является управляющим для тока коллектора, который моделируется источником тока, управляемым напряжением (ИТУН) с управляющим параметром $g_m = g_{m0} \exp(j\omega\tau_k)$. Параметры второго диода заданы RC -цепочкой $R_{дк}$ и $C_{дк}$. Активная область транзистора моделируется также выходным сопротивлением R_0 и шумовым источником $i_{шк}$.

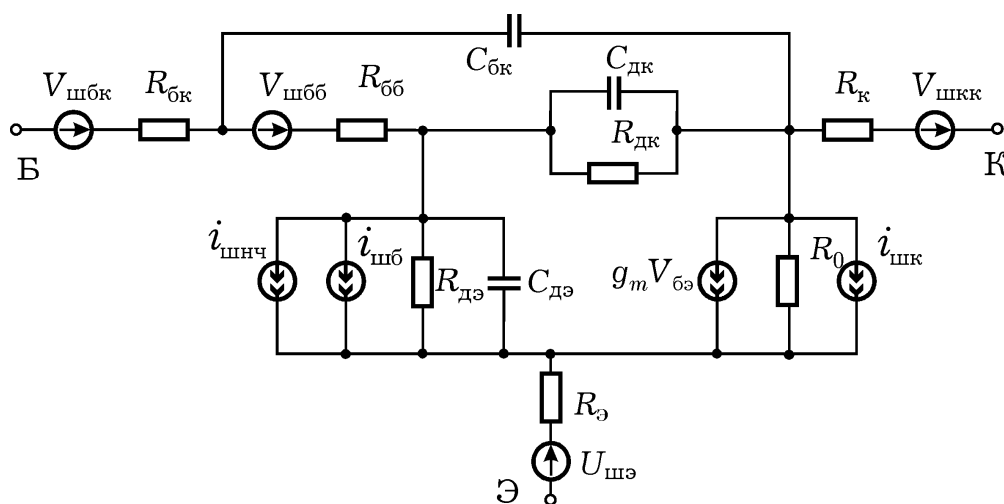


Рис. 6.20. Шумовая схемная модель гетеробиполярного транзистора

В работе [6.31] приводится анализ шумовых характеристик типичных структур $Al_xGa_{1-x}As$ СГБТ для различных конструкций эмиттера, включающих от 1 до 5 секций. Показано, что при увеличении уровня тока базы минимальный коэффициент шума возрастает во всем диапазоне частот. При этом частотная зависимость коэффициента шума тем существеннее, чем больше ток базы. Упрощенная шумовая модель транзистора приведена в

работе [6.29]. Рассчитан минимальный коэффициент шума для диапазона частот 1–20 ГГц, изменяющийся в пределах от 2 до 3 дБ для транзистора с размерами эмиттера 2×20 мкм², коэффициентом передачи тока $\beta = 50$, временем запаздывания 3,5 пс и значением сопротивлений контактов, изменяющихся в пределах от 2 до 14 Ом при токе коллектора 15 мА. По схемной модели рис. 6.20, описывающей подобную структуру и дополненной внешними согласующими индуктивными компонентами для получения минимального значения коэффициента шума, в соответствии с рассмотренным ниже алгоритмом рассчитаны частотные усилительные и шумовые характеристики.

Методика расчета минимального коэффициента шума линейной цепи связана с пересчетом локальных шумовых источников ко входу и выходу четырехполюсника и использовании волновых шумовых параметров для расчета минимального коэффициента шума, рассмотренная ранее.

Одним из путей уточнения модели является более строгий учет эффектов саморазогрева и баллистических эффектов при малых значениях толщины базы (менее 100 нм).

6.4. МОЩНЫЕ МОП-ТРАНЗИСТОРЫ

В последние годы в области мощных кремниевых СВЧ-транзисторов лидируют МОП-транзисторы, изготовленные по так называемой *LDMOS* (*Laterally Diffused Metal Oxide Semiconductors*) технологии на основе кремния. Достоинства кремниевых приборов, прежде всего в низкой стоимости пластин, отработанности технологии и относительной простоте интеграции аналоговых и цифровых схем на одном кристалле [6.40]. Причем благодаря минимизации размеров частотные характеристики кремниевых транзисторов повысились настолько, что, несмотря на более низкую подвижность электронов, успешно конкурируют с гетеробиполярными в диапазонах до 3 ГГц. Еще в 2000 году [6.41] компания *Motorola* представила *LDMOS*-транзистор с длиной затвора 0,6 мкм, толщиной подзатворного окисла 40 нм и пробивным напряжением исток-подложка 65 В. При суммарной ширине затвора 480 мм транзистор работал в двухтактном режиме с выходной мощностью 220 Вт при рабочем напряжении 28 В с КПД 46 % на частоте 2,12 ГГц.

Принципиальное отличие *LDMOS*-технологии от традиционной МОП-технологии – несимметричная структура транзистора. На p^+ -подложке (рис. 6.21) выращивается относительно слабо легированный эпитаксиальный p -слой, назначение которого –

уменьшить выходную емкость сток-исток и увеличить пробивное напряжение. Со стороны истока формируется сильнолегированная p^+ -область, которая соединяет исток с подложкой. Под высоколегированной n^+ -областью истока формируют дополнительную тонкую p -область, под областью стока – n -область. Первая определяет пороговое напряжение и параметры отсечки канала, вторая позволяет увеличить напряжение пробоя и предотвращает инжекцию горячих электронов в подзатворный окисел. Характерной особенностью *LDMOS*-транзисторов является слой металлизации, соединенный с истоком и нависающий над затвором. Этот слой является экраном, уменьшающим паразитную проходную емкость C_{gd} отрицательной обратной связи. Кроме того, экран существенно изменяет распределение поля между стоком и затвором, ослабляя напряженность поля у стокового края затвора и соответственно увеличивая пробивное напряжение стока.

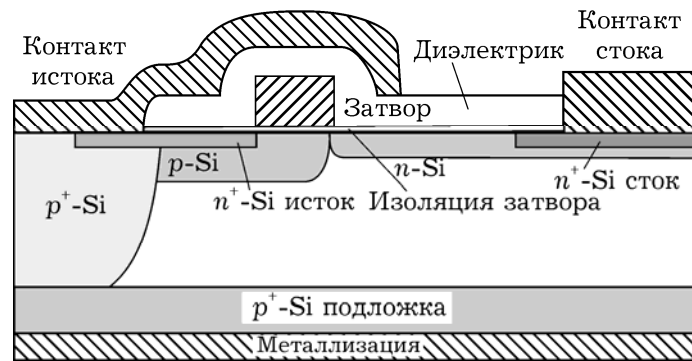


Рис. 6.21. Структура *LDMOS*-транзистора

Особенностью такой конструкции транзистора является то, что при отсутствии смещения на затворе поверхностный слой, представляющий собой планарную n^+pn^+ -структуру, не является проводящим, и транзистор закрыт. Для открывания на затвор необходимо подать некоторое положительное напряжение в несколько вольт. При этом у поверхности p -области под затвором образуется инверсионный слой, представляющий собой проводящий канал с преобладающей электронной проводимостью.

На рис. 6.22 в качестве примера представлены стоковые характеристики транзистора фирмы *Philips Semiconductor* BLA1011-10 [6.42].

В общем, эти характеристики имеют типичный для МОП-транзисторов вид. Однако при субмикронных длинах затвора в режиме насыщения наблюдается существенное увеличение тока стока. Это может быть связано с тем, что при увеличении стокового напряжения возрастает потенциал канала и в его стоковой части увеличивается степень инверсии. В результате

увеличивается толщина проводящего канала и его проводимость. Одновременно с этим напряженность тянущего электрического поля может уменьшаться, что ослабляет степень насыщения дрейфовой скорости электронов, характерной для сильного поля.

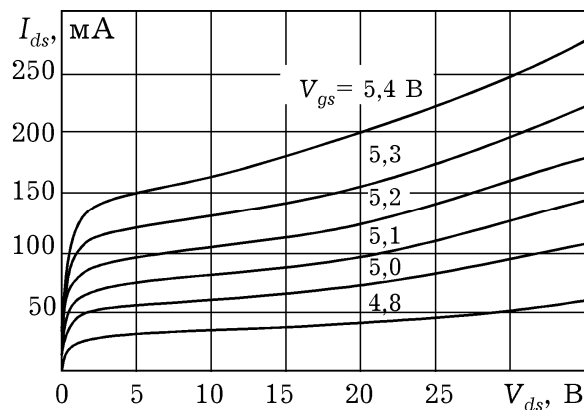


Рис. 6.22. Экспериментальные выходные вольтамперные характеристики МОП (*LDMOS*) транзистора BLA1011-10 [6.42]

Транзистор характеризуется следующими основными параметрами. Предельные значения выходного напряжения стока-исток 75 В, напряжение затвор-исток ± 15 В, максимальный выходной ток 2,2 А. Существенной особенностью данного транзистора является достаточно высокий внутренний встроенный потенциал (около 4 В), а также высокие значения крутизны в рабочем поле выходных вольтамперных характеристик.

Разработкой моделей *LDMOS*-транзисторов занимаются не столько ученые, сколько сами компании-производители, и разработчики САПР. Можно выделить две различных модели мощных *LDMOS*, широко используемых в настоящее время на практике [6.43, 6.44]. Все эти модели являются схемотехническими и могут быть описаны с помощью эквивалентных схем.

Разработчик первой из них компания *STMicroelectronics* (<http://www.st.com>). Схемотехническая модель, описываемая электрической цепью с сосредоточенными параметрами (рис. 6.23, а), имеет типичный для полевых транзисторов вид и включает следующие группы компонентов: паразитные элементы, связанные с устройством; нелинейный генератор тока; сопротивление полевого транзистора с управляющим *p-n* переходом; диод, моделирующий влияние подложки.

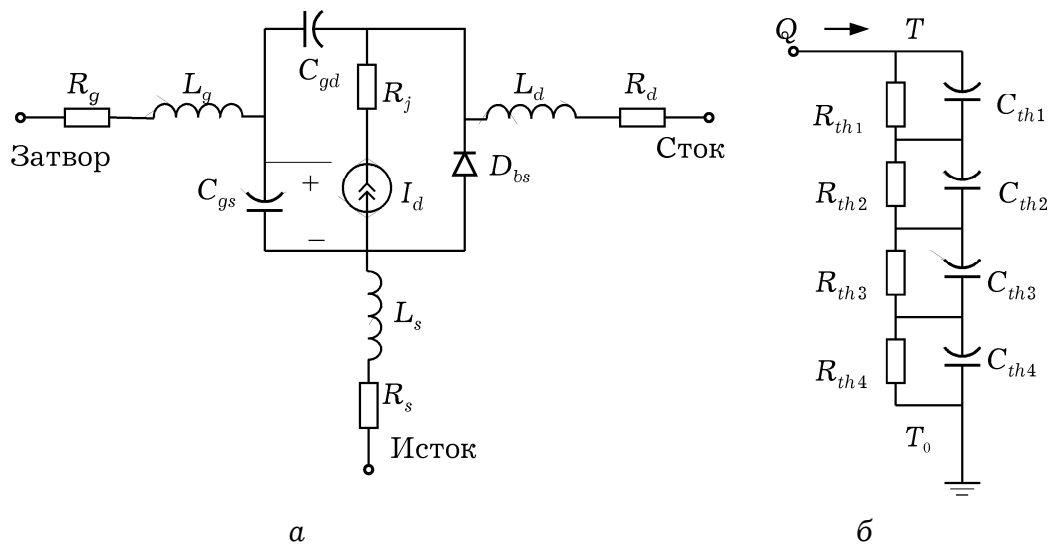


Рис. 6.23. Схемная модель *LDMOS*-транзистора [6.50].

Нелинейный источник тока, управляемый напряжениями (нелинейный ИТУН) – наиболее важный элемент для расчета статических и динамических характеристик устройства. Управляющими являются напряжения V_{gs} и V_{ds} . Кроме того, для вычисления статического тока необходимо также определить рабочий диапазон полевого транзистора. Для определения уравнений, описывающих нелинейный ИТУН, необходимы 15 параметров, таких, как, например, пороговое напряжение, наведенное стоком снижение барьера, крутизна характеристики прямой передачи, наклон подпорогового тока, длина канала, коэффициент выходной проводимости, коэффициент тепловой проводимости, тепловой коэффициент порогового напряжения, и ряд других, которые идентифицируются из экспериментальных характеристик.

Данная программа позволяет моделировать статические характеристики, описывать транзистор при помощи S -параметров при малом сигнале, рассчитывать параметры схемной модели прибора при большом сигнале. Из особенностей модели – возможность моделирования паразитных элементов корпуса, а также наличие тепловых элементов в эквивалентной цепи, учитывающих эффекты саморазогрева.

Область квазинасыщения моделируется нелинейным сопротивлением R_j , которое может быть найдено из выходных статических характеристик в линейной области изменения тока при высоких напряжениях.

Для моделирования влияния подложки и описания режима пробоя, утечки тока стока и емкости между истоком и стоком в модельную цепь введен диод. Статические и динамические

характеристики такого диода описываются обычным образом; моделируется также их зависимость от температуры.

Для учета эффекта саморазогрева вводится тепловой электрод (см. рис. 6.23, б). Напряжение между внешним тепловым электродом и стоком определяет температуру, как описано в [6.45].

Для моделирования паразитного влияния корпуса, производится его электромагнитное моделирование. Диапазон частот, в котором проводилось моделирование составляет от 1 МГц до 3 ГГц. Результаты моделирования представляются в виде S -параметров. Из рис. 6.24, 6.25 видно, что смоделированные входные и выходные статические характеристики транзистора и динамические характеристики в режиме малого сигнала (S -параметры) находятся в хорошем соответствии с экспериментальными данными.

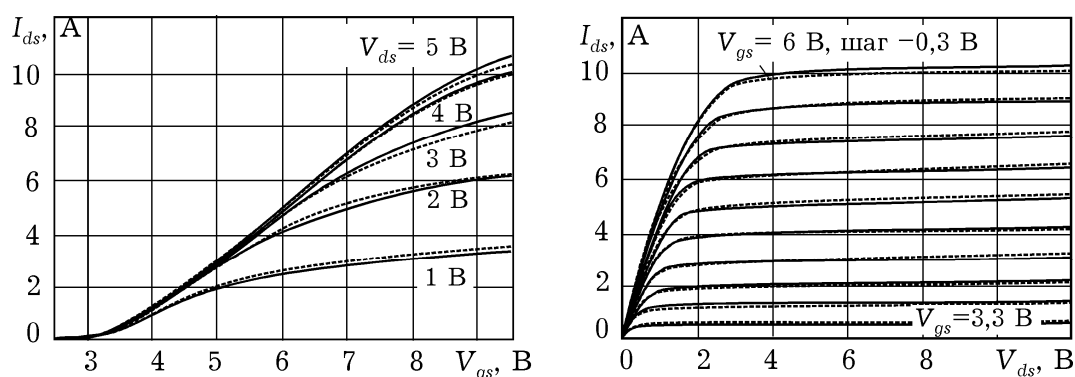


Рис. 6.24. Измеренные (сплошные линии) и смоделированные (штриховые) входные и выходные статические характеристики

Другая модель СМС (Curtice/Modelics/Cree) [6.44]. разработана группой компаний: *W.R. Curtice Consulting* (<http://www.curtice.org>), *Modelithics Inc.* (<http://www.modelithics.com>) и *Cree Microwave* (<http://www.cree.com/>) для интеграции в САПР *Agilent EEsof EDA Design & Simulation Software* компании *Agilent Technologies* (<http://www.agilent.com>).

Данная модель имеет следующие возможности: моделирование статических характеристик, динамических параметров в режиме малого сигнала, учет нелинейных эффектов, анализ во временной области. Из особенностей модели – хорошее предсказание интермодуляционных искажений, адекватное моделирование в подпороговом режиме, моделирование динамики процессов саморазогрева, простая процедура извлечения параметров модели. Модель разработана для базовой ячейки, то есть транзистора мощностью 1 Вт и может быть применена к

другим *LDMOS* путем специально разработанной процедуры масштабирования.

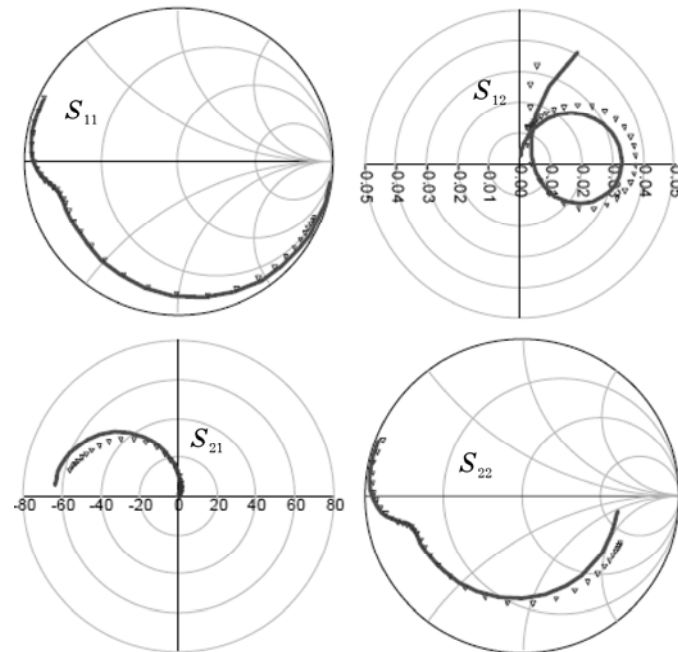


Рис. 6.25. Измеренные и смоделированные *S*-параметры для частот от 1 МГц до 3 ГГц ($V_{ds} = 7,2$ В, $I_{dq} = 100$ мА)

По сравнению с биполярными транзисторами *LDMOS* обладают лучшими характеристиками по линейности, усилению, тепловым режимам, устойчивости к рассогласованию, имеют более высокий КПД, запас по рассеиваемой мощности, надежность [6.46]. По этой технологии в настоящее время выпускается большинство мощных высокочастотных транзисторов.

Поскольку размер структуры сток-исток ячейки *LDMOS*-транзистора больше аналогичного размера структуры база-эмиттер, то температура активной области не превышает 200 °С, даже при более высокой, чем у биполярного транзистора, плотности мощности (0,5 Вт/мм). Еще одним преимуществом *LDMOS*-технологии является соединение истока с подложкой. Это обеспечивает, во-первых, низкое значение индуктивности истока, что в сочетании с малой емкостью обратной связи обеспечивает *LDMOS*-транзисторам лучшие усилительные характеристики по сравнению с биполярными аналогами. Во-вторых, такое включение позволяет создать в транзисторе лучшие условия для отвода тепла. В биполярном же транзисторе наибольший объем кремния занимает область коллектора, которая должна быть электрически изолирована от земли (радиатора), что затрудняет отвод тепла.

Еще одна особенность полевых структур вообще и *LDMOS*-транзисторов в частности связана с меньшей нелинейностью

характеристик, чем у биполярных приборов (рис. 6.26). Это приводит к тому, что усиление биполярного транзистора сильно изменяется в зависимости от мощности входного сигнала. Очевидно, что использование подобного транзистора скажется отрицательным образом на качестве всего усилителя.

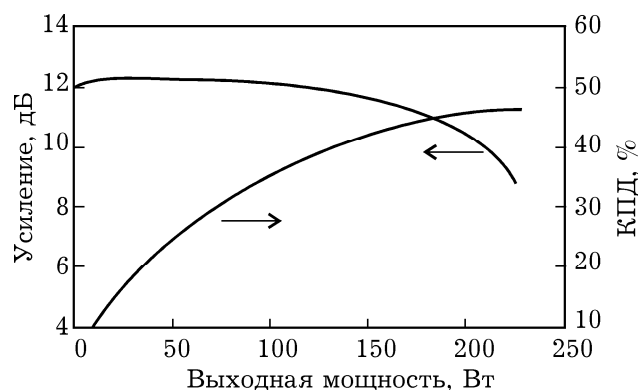


Рис. 6.26. Мощностная зависимость КПД и усиления 220-Вт транзистора компании *Motorola* [6.47]

Сегодня самая масштабная область применения *LDMOS*-приборов — усилители базовых станций сотовой связи. Будучи разновидностью кремниевой технологии, *LDMOS* обладает огромным достоинством — гораздо более низкой по сравнению с остальными материалами себестоимостью. За это рынок "прощает" *LDMOS*-приборам низкую плотность мощности и проблемы с линейностью характеристик.

Передачики большой мощности используются как в портативных мобильных, так и в стационарных системах сотовой связи. При этом мощность стационарных станций, обеспечивающих соединение "мобильников" с проводными линиями связи, достигает 600 Вт. Сегодня основные приборы, обеспечивающие высокую мощность в диапазоне 900 и 1900 МГц, — кремниевые *LDMOS*-транзисторы, на долю которых приходится 90% рынка мощных СВЧ приборов. Это объясняется самой низкой для транзисторов этого класса плотностью мощности (0,7 Вт/мм), самыми большими размерами кристалла и самой низкой стоимостью.

Перспективная область применения *LDMOS*-транзисторов — системы связи стандарта WiMAX, одного из самых перспективных стандартов связи, для которого необходима высокая линейность и КПД транзисторов.

Успехи в совершенствовании *LDMOS*-технологии не препятствуют дальнейшему развитию GaAs мощных СВЧ-транзисторов. В базовых станциях, где выходная мощность

зависит от трафика, усилители мощности на *LDMOS* транзисторах не всегда могут работать с наилучшей эффективностью. Для таких систем более подходят *GaAs*-транзисторы, имеющие в сравнении с кремниевыми транзисторами более широкий динамический диапазон. Фирма *Freescale Semiconductor*, выпустившая серьезный *LDMOS*-конкурент *GaAs pHEMT* на частоту 3,5 ГГц, тем не менее, разработала арсенидгаллиевые транзисторы на частоту 6 ГГц с выходной мощностью 100 Вт, отвечающие всем требованиям, предъявляемым к *WiMAX*-системам. Компания *Toshiba America Electronic Components (ТАЕС)* создала *GaAs* гетеропереходные полевые транзисторы (*HFET*) диапазона 13,75–14,5 ГГц для усилителей мощности наземных базовых станций спутниковых систем связи.

Но наиболее перспективными мощными СВЧ-приборами на сегодняшний день можно считать *GaN*-транзисторы, характеризующиеся, в сравнении с транзисторами на других полупроводниковых материалах, более высокой плотностью тока насыщения, большим напряжением пробоя и большой плотностью мощности в пересчете на единицу ширины затвора (рис. 6.27).

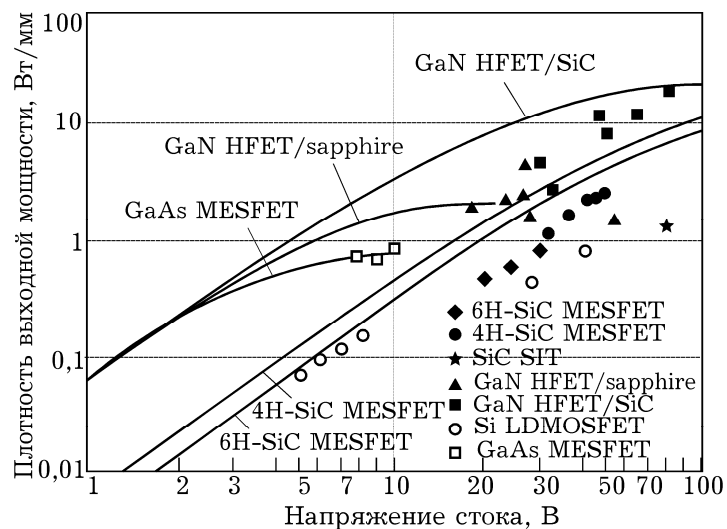


Рис. 6.27. Зависимость выходной мощности на 1 мм ширины затвора от стокового напряжения для различных материалов и структур транзисторов [6.48]

К технологии нитрида галлия обращаются все большее число полупроводниковых компаний. Всего за период с 2000 по 2005 год число компаний и исследовательских центров, занятых разработкой приборов и изучением нитрида галлия, увеличилось с 500 до 2000. Продажи *GaN*-приборов (не только транзисторов, но и

светодиодов) с 3,5–3,6 млрд. долл. в 2006 году возросли до 7,2 млрд. долл. в 2009 году [6.49].

На рынке мощных транзисторов ведущее положение в сравнении с GaAs *pHEMT* и полевыми транзисторами с затвором Шоттки (*MESFET*) занимают GaAs биполярные гетеротранзисторы (*HBT*) и *LDMOS*-транзисторы. Но в дальнейшем их, вероятно, вытеснят SiC и GaN-приборы [6.50].

Литература к главе 6

6.1. *Shi Chang Zhong, Tangsheng Chen, Chunjiang Ren, Gang Jiao, Chen Chen, Kai Shao, Naibin Yang.* AlGaIn/GaN HEMT with over 110 W Output Power. Proceedings of the 3rd European Microwave Integrated Circuits Conference, October 2008, Amsterdam, The Netherlands. P.91-94

6.2. *Eizo Mitani, Makoto Aojima, Arata Maekawa and Seigo Sano,* An 800-W AlGaIn/GaN HEMT for S-band High-Power Application, CS MANTECH Conference, May 14-17, 2007, Austin, Texas, USA/ P.213-216.

6.3. *Захаров А.Л., Асвадурова Е.И.* Расчет тепловых параметров полупроводниковых приборов. – М.: Радио и связь, 1983. – 184 с.

6.4. *Могилевский В.М., Чудновский А.Ф.* Теплопроводность полупроводников. М.: Наука, 1972. – 536 с.

6.5. Теплопроводность твердых тел: Справочник / А.С. Охотин, Р.П. Боровикова, Т.В. Нечаева, А.С. Пушкарский; Под ред. А.С. Охотина. М.: Энергоатомиздат, –1984.–320 с.

6.6. *Семеновская Е.В., Тимофеев В.И.* СВЧ методы измерения температурных полей // Электроника и связь. –2000, –№8. –Т.2. – С. 263–265.

6.7. *Прокопов В.Г., Тариелашвили Г.Т., Шеренковский Ю.В., Юрчук В.Л.* Моделирование мощных полевых транзисторов. Ч.1. Тепловой режим//Сб. докл. междунар. н.-техн. конф. «Проблемы физической и биомедицинской электроники», 18-20 мая 1996 г. Киев.- с. 240-244.

6.8. *Москалюк В.А., Тариелашвили Г.Т., Тимофеев В.И.* Моделирование мощных полевых транзисторов. Ч.2. Электрический режим//Сб. докл. междунар. н.-техн. конф. «Проблемы физической и биомедицинской электроники», 27-30 мая 1996 г. Киев.- с. 366-369.

6.9. *Cai Q., Gerber J., Rohde U., Daniel T.* HBT High-Frequency Modeling and Integrated Parameter Extraction // IEEE Transaction on Electron Devices.–1997.–Vol.45,–№12.–P.2493–2502.

6.10. *Семеновская Е.В., Приходько Н.В., Тимофеев В.И.* Физико-топологическое моделирование теплового режима мощного

- субмикронного полевого транзистора СВЧ // Электроника и связь. – 2000. – № 9. – С. 140–142.
- 6.11. Москалюк В.А., Тимофеев В.И. Моделирование процессов переноса электронов в канале и подложке субмикронного полевого транзистора // Электронная техника. Сер.1., Электроника СВЧ. – 1990. – Вып.2 (426). – С.22–27.
- 6.12. Семеновская Е.В., Тимофеев В.И. Моделирование теплового режима в субмикронных транзисторах с гетероселективным легированием // Электроника и связь. – 2001. – №13. – С.32–36.
- 6.13. Семеновская Е.В., Тимофеев В.И. Моделирование тепловых процессов в транзисторах с гетероселективным легированием // Труды междунар. конф. «Физика и технология тонких пленок». – Ивано-Франковск. – 2003. – Т.2. – С.127–128.
- 6.14. Семеновская Е.В., Тимофеев В.И. Схемотехническое моделирование тепловых эффектов и эффекта саморазогрева в субмикронных гетероструктурных транзисторах // Труды 7-го междунар. Форума «Радиоэлектроника и молодежь в XXI веке». – Харьков. – 2003. – С.81.
- 6.15. Семеновская Е.В., Тимофеев В.И., Приходько Н.В. Анализ моделей и тепловых режимов гетеробиполярных субмикронных транзисторов // Электроника и связь. – 2001. – №10. – С.133–136.
- 6.16. Mahesh B. Patil New discretization scheme for two-dimensional semiconductor device simulation on triangular grid // IEEE Transactions on computer-aided design of integrated circuits and systems. – 1998. – Vol.17, №11. – P.1160–1165.
- 6.17. Шур М. Современные приборы на основе арсенида галлия: Пер. с англ. – М.: Мир, 1991. – 632 с.
- 6.18. Тимофеев В.И. Шумовые свойства субмикронных гетеробиполярных транзисторов // Электроника и связь. – 2002. – №16. – С.112–115.
- 6.19. Тимофеев В.И. Моделирование процессов токопереноса в субмикронных гетеробиполярных транзисторах // Электроника и связь. – 2004. – №23. – С. 57–60.
- 6.20. Bester Y., Scott D., Mensa D. et al. InAlAs/InGaAs HBTs with Simultaneously High Values of F_t and F_{max} for Mixed Analog/Digital Applications // IEEE Electron Device Letters. – 2001. – Vol. 22, №2. – P.56–58.
- 6.21. Liou L.L., Ebel J., Huang C. Thermal Effects on the Characteristics of AlGaAs/GaAs Heterojunction Bipolar Transistors Using Two-Dimensional Numerical Simulation. // IEEE Transactions on Electron Devices. – 1993. – Vol.40, №1. – P.35–43.
- 6.22. Benvenuti A., Coughran W.M., Pinto M.R. A Thermal-Fully Hydrodynamic Model for Semiconductor Devices and Applications to

- III-V HBT Simulation // IEEE Transactions on electron devices. – 1997.–Vol.44, №9.–P.1349–1359.
- 6.23. *Gao G-B, Wang M-Z, Gui X., Morkoc H.* Thermal design studies of high-power heterojunction bipolar transistors // IEEE Transactions on Electron Devices. –1989.–Vol.36, –№5.– P.854–863.
- 6.24. *Marshak A.H.* Transport equations for highly doped devices and heterostructure // Solid-State Electronics, Vol.30, pp. 1089–1093.– 1987.
- 6.25 *Samelis A., Pavlidis D.* Analysis of the large-signal characteristics of power heterojunction bipolar transistors exhibiting self-heating effects // IEEE Transactions on Microwave Theory and Techniques.–1997.–Vol.45, №. 4.–P.534–541.
- 6.26. *Тимофеев В.И.* Анализ разогревных эффектов в субмикронных гетеробиполярных транзисторах // Вісник Черкаського державного технологічного університету. Сер. Радіотехніка і приладобудування.–2004.–№4.–С.72–78.
- 6.27. *Twynam J.K., Yagura M., Kishimoto K., Kinoshita T., Sato H., Shimizu M.* Thermal stabilization of AlGaAs/GaAs power HBT's using $n\text{-Al}_x\text{Ga}_{1-x}\text{As}$ emitter ballast resistors with high thermal coefficient of resistance // Solid-State Electron. – 1995. – Vol.9, №3. –P.1657–1661.
- 6.28 *Liu W., Khatibzadeh A., Sweder J., Chau H.* The use of base ballasting to prevent the collapse of current gain in AlGaAs/GaAs heterojunction bipolar transistors // IEEE Transactions on Electron Devices.–1996.–Vol.43,№2.–P.245–251.
- 6.29. *Gao G., Unlu M., Blackburn D.* Emitter ballasting resistor design for current handling capability of AlGaAs/GaAs power heterojunction bipolar transistors // IEEE Transactions on Electron Devices. –1991. –Vol.38, №2. –P.185–196.
- 6.30. *Zhang Q., Hu H., Sitch J., Surridge R., Xu J.* A new large signal HBT model // IEEE Trans. Microwave Theory Tech. –1996.– Vol.44,№5.– P.2001–2009.
- 6.31. *Zhu Y., Twynam J.K., Yagura M.* Self-heating effect compensation in HBTs and its analysis and simulation // IEEE Transactions on Electron Devices.–2001.–Vol.48, №11. –P. 2640–2645.
- 6.32. *Andlerstein M.G.* Thermal stability ballasted HBT's // IEEE Transactions on Electron Devices.–1998.–Vol.45, №2– P.1653–1655.
- 6.33 *Тимофеев В.И., Семеновская Е.В.* Электротепловой анализ монолитных интегральных схем на основе соединений A_3B_5 . „Сучасні інформаційно-комунікаційні технології” /COMINFO`2008/ Лівадія. –2008. –С.111-11
- 6.34. *Семеновская Е.В., Тимофеев В.И.* Электротепловой анализ субмикронных гетероструктур // Электроника и связь. – 2007. – № 1 – С.57–61.

- 6.35. *Wei Ce-Jun, Hwang J. C. M., Ho Wu-Jing, Higgins J. A.* Large-signal modeling of self-heating, collector, transit-time and RF-breakdown effects in power HBT's. // IEEE Transactions on Microwave Theory and Techniques. –1996.–Vol.44, № 12. –P.2641–2646.
- 6.36. *Green B.M., Tilak V., Kim H. et.al.* High-power broad-band AlGa_N/Ga_N HEMT MMICs on SiC substrates // IEEE Transactions on Microwave Theory sand Techniques. – 2001. – Vol.49, № 12 – P. 2486.
- 6.37. *Zhu Y., Twynam J.K., Yagura M.* Self-heating effect compensation in HBTs and its analysis and simulation // IEEE Transactions on electron devices. – 2001. –Vol.48, №11. –P. 2640–2645.
- 6.38. *Huber A., Huber D., Bergamaschi C., Morf T., Jackel H.* Noise model of InP-InGaAs SHBT's for RF circuit design // IEEE Transactions on microwave theory and techniques.–2002.–Vol.50,№7. –P.1675–1682.
- 6.39. *Liou J.J., Jenkins T.J., Liou L.L., Neidhard R., W.Barlage D., Fich R., Barrette J.P., Mack M., Bozada C.A., Lee H.Y., Dettmer R.W., Sewell J.S.* Bias, Frequency and Area Dependencies of High Frequency Noise in AlGaAs/GaAs HBT's // IEEE Transactions on electron devices. – 1996. –Vol.43, №1.– P.116–122.
- 6.40. *Шахнович И.* Твердотельные СВЧ-приборы и технологии. Состояние и перспективы.–ЭЛЕКТРОНИКА:НТБ, 2005, №5, с.58–64.
- 6.41. *Wayne R. Berger et al.* RF-LDMOS: A Silicon-Based, High Power, High Efficiency Linear Power Amplifier Technology. – 2000 GaAs MANTECH Conf. Dig. Ppr., 2000.
- 6.42. <http://www.semiconductors.philips.com>
- 6.43. *Wood, Pengelly, Dunleavy, Clausen, Weller and Emmadi.* New LDMOS Model Delivers Powerful Transistor Library–Part 2: Library application // High Frequency Electronics. – November, 2004. <http://www.highfrequencyelectronics.com/Archives/Archives2004.shtml>.
- 6.44. *Curtice, Dunleavy, Clausen and Pengelly.* New LDMOS Model Delivers Powerful Transistor Library–Part 1: The CMC Model // High Frequency Electronics. – October, 2004. <http://www.highfrequencyelectronics.com/Archives/Archives2004.shtml>
- 6.45. Circuits and System, 2006. MIXDES 2006. Proceedings of the International Conference, June 2006
- 6.46. Технология мощных СВЧ LDMOS-транзисторов для радарных передатчиков L-диапазона и авиационных применений. Компоненты и технологии, № 10, 2007 с.14-16.

- 6.47. *Wayne R. Berger et al.* RF-LDMOS: A Silicon-Based, High Power, High Efficiency Linear Power Amplifier Technology. – 2000 GaAs MANTECH Conf. Dig. Ppr., 2000.
- 6.48. *Okumura H.* Present Status and Future Prospect of Widegap Semiconductor High-Power Devices, Japanese Journal of Applied Physics. Vol. 45, No. 10A, 2006, pp. 7565–7586
- 6.49. www.eetimes.com/showArticle.jhtml?articleID=190300189.
- 6.50. *Лебедев А., Сбруев С.* SiC электроника. Прошлое, настоящее, будущее. – Электроника: Наука, Технология, Бизнес. № 5, 2006, с.28.

7. РЕЗОНАНСНО-ТУННЕЛЬНЫЕ ДИОДЫ

7.1. ОСНОВНЫЕ ПРЕДСТАВЛЕНИЯ

7.1.1. Структурные составляющие РТД

Резонансно-туннельный диод (РТД) является одной из самых простых структур с поперечным транспортом. Рассмотрим топологию резонансно-туннельного диода, слои которого изготовлены из $\text{Al}_{0.33}\text{GaAs}_{0.77}/\text{GaAs}$ (рис. 7.1, а).

В плоскости гетерограниц его размеры имеют значение от единиц до нескольких десятков микрометров, что значительно превышает период решетки (который для GaAs составляет $\approx 0,56$ нм), а также длину волны де Бройля $\lambda_{\text{Ф}}$, длину релаксации фазы $L_{\text{Ф}}$ и длину свободного пробега $L_{\text{р}}$.

В структуре РТД можно выделить три подсистемы (рис. 7.1, а): макроскопические области, прилегающие к контактным площадкам, а также микроскопическую область, которая находится между ними. Релевантная часть микроскопической области достаточно сильно легируется примесями n -типа (обычно не меньше 10^{23} м^{-3}), в то время как микроскопическую область оставляют свободной от примесей. Эти области имеют устоявшиеся в литературе названия:

- *резервуары* – это высоколегированные области, которые прилегают к эмиттерному и коллекторному электродам (рис. 7.1, а);
- *квантовая (активная) область* – нелегированная область между двумя резервуарами, часть которой в увеличенном масштабе показана на рис. 7.1, б.

Такое разделение, как будет показано ниже, дает возможность создавать так называемые «комбинированные» модели [7.1], называемые так, поскольку при их построении для разных участков одного устройства используются разные способы описания электронного коллектива: для резервуаров – полуклассический, а для квантовой области – разнообразные квантово-механические формализмы, в том числе метод волновых

функций. Создание комбинированных моделей является альтернативой описания в рамках единственного квантово-механического формализма всего устройства, которое приводит к неоправданно высокой интенсивности использования математических и численных ресурсов без существенного повышения точности модели.

Заметим, что известные на сегодня устройства наноэлектроники, включая нанотрубки, одноэлектронные транзисторы и другие перспективные компоненты, непременно содержат классические области. Любой компонент или система наноэлектроники должны содержать «классическую» область [7.2 стр. 15]. В частности, классическая область необходима для проведения измерения [7.3 стр. 77].

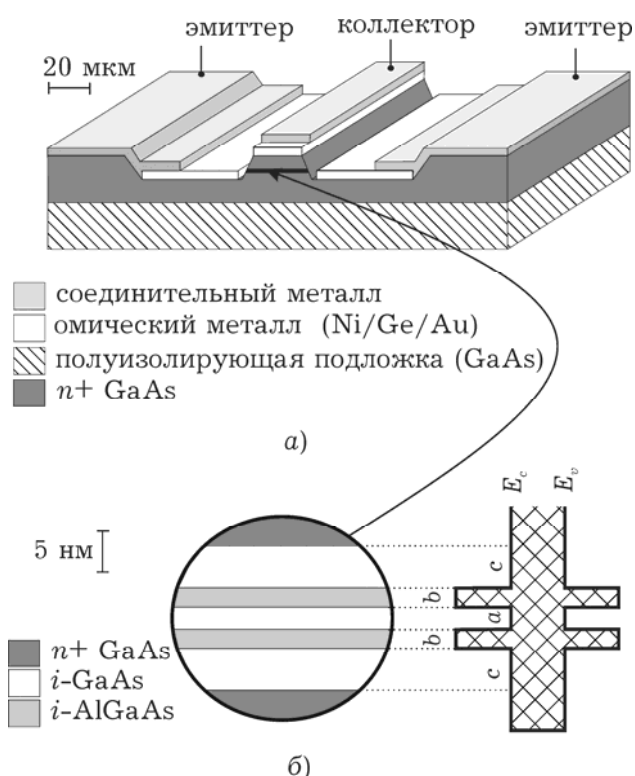


Рис. 7.1. Структура резонансно-туннельного диода (а) и активной области и зонная диаграмма (б)

7.1.2. Принцип работы резонансно-туннельного диода

Характерным свойством резонансно-туннельного диода является наличие на ВАХ участка отрицательной дифференциальной проводимости (ОДП), который имеет квантовую природу и определяет особенности его использования. РТД был впервые предложен советским ученым Йогансенем в 1963 году [7.4; 7.5] как аналог резонатора Фабри-Перо для электронных волн (волн де Бройля). В 1974 году Чанг впервые наблюдал участок ОДП на практике при температуре, близкой к абсолютному нулю [7.6]. Наличие участка ОДП недвусмысленно указывало на реализацию в созданном образце резонансного туннелирования. Дальнейшему развитию резонансно-туннельных диодов способствовало совершенствование технологий молекулярно-лучевой эпитаксии (МЛЭ) и химического осаждения из газовой фазы металлоорганических соединений, что позволило Шевчуку

[7.7] уже в 1985 году наблюдать участок ОДП на ВАХ РТД при комнатной температуре.

Принцип работы РТД легко понять, следуя объяснением *S. Lurie* [7.8]. При низкой температуре в сильно легированных резервуарах электроны занимают состояния, начиная со дна зоны проводимости E_C , и заканчивая энергией Ферми E_F : $E_C < E < E_F$ (рис. 7.2). Все возможные комбинации волновых чисел, таким образом, находятся в т.н. сфере Ферми радиуса $k_F = \sqrt{2m^* E_F} / \hbar$ (рис. 7.3).

В резервуарах, где электронный газ трехмерный, закон дисперсии имеет вид:

$$E_{3D} = E_C + \frac{\hbar^2 k_z^2}{2m^*} + \frac{\hbar^2 k_{\parallel}^2}{2m^*}, \quad (7.1)$$

где E_C – дно зоны проводимости, $k_{\parallel}^2 = k_x^2 + k_y^2$; $k_{\parallel(z)}$ – "продольное" ("поперечное") волновое число электрона, которому соответствует "продольная" ("поперечная") энергия.

В квантовой яме поперечная энергия электрона E_z квантуется, и может приобретать дискретные значения E_i . Закон дисперсии для электронов модифицируется, поскольку электронный газ становится двумерным:

$$E_{2D} = E_i + \frac{\hbar^2 k_{\parallel}^2}{2m^*}. \quad (7.2)$$

В дальнейшем будем считать, что в яме существует всего один уровень энергии $E_z = E_0$. В силу закона сохранения энергии при туннелировании электрона из трехмерного электронного газа эмиттера в двумерный – квантовой ямы должно быть:

$$E_{2D} = E_{3D}. \quad (7.3)$$

Из условия (7.3) и с учетом сохранения волнового числа k_{\parallel} получим условие, которое налагается на поперечное волновое число электронов, туннелирующих из эмиттера:

$$k_z = \sqrt{2m^* (E_0 - E_C)} / \hbar \equiv k_0. \quad (7.4)$$

Условие (7.4) называют *условием резонансного туннелирования*. Количество электронов в эмиттере, для которых оно

выполняется, зависит от положения уровня энергии в квантовой яме (рис. 7.2, а). Так, если $k_0 > k_\Phi$ ни один электрон не удовлетворяет условию (7.4); соответственно и ток через ДБКС отсутствует. Начиная с $k_0 = k_\Phi$ и до $k_0 = 0$ количество электронов, способных к прохождению ДБКС, увеличивается и достигает максимума при $k_0 = 0$. При этом ток растет за счет увеличения количества электронов, для которых выполняется условие резонансного туннелирования. После прохождения экватора сферы Ферми волновое число k_0 , в соответствии с (7.4), становится комплексным, что трактуется как отсутствие электронных состояний с резонансной энергией. Ток при этом резко падает до нуля (рис. 7.3, а).

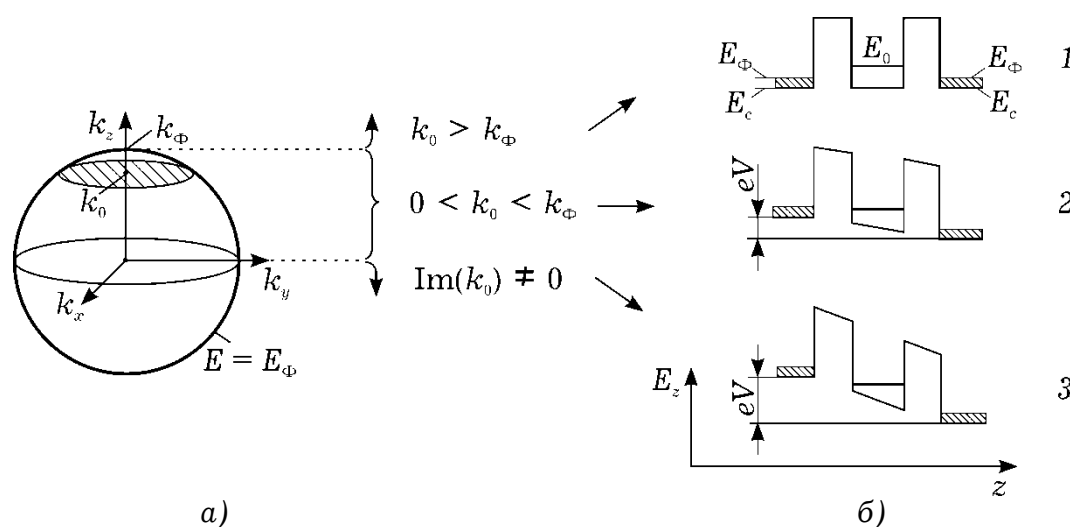


Рис. 7.2. К принципу работы РТД: а) сфера Ферми, внутри которой находятся все возможные значения волнового числа электронов из эмиттера. Заштрихованный круг обозначает состояния электронов, которые способны туннелировать при данном положении уровня в квантовой яме; б) потенциальный рельеф ДБКС при разных смещениях: 1 – $V = 0$, 2 – "пиковое" напряжение; 3 – напряжение, при котором уровень в квантовой яме опустился ниже дна зоны проводимости в эмиттере

Положение уровня поперечной энергии в квантовой яме можно регулировать, прикладывая напряжение (рис. 7.2, б). Пусть при нулевом напряжении единственный уровень энергии лежит выше уровня Ферми. При повышении потенциала коллектора, уровень энергии "опускается" и может пробегать все значения от E_Φ к E_c и ниже. Соответственно, ток через диод в зависимости от напряжения принимает значение от нулевого, когда $E_0 > E_\Phi$ до некоторого максимального значения, когда $E_0 = 0$. При дальнейшем увеличении напряжения количество электронов резко уменьшается, что приводит к падению тока: так

формируется участок отрицательной дифференциальной проводимости G_d , на котором $G_d = dI/dV < 0$.

На рис. 7.3, б приведена ВАХ экспериментального образца РТД, который описан в [7.9]. Принципиальным отличием от теоретической кривой является то, что с увеличением напряжения ток сразу же начинает расти, и уже не опускается до нуля, а участок отрицательной дифференциальной проводимости имеет конечный размах по напряжению. Такие отличия объясняются прежде всего следующим:

- уровни энергии в квантовой яме в действительности не дискретны, а имеют конечную ширину Γ . Такое уширение происходит оттого, что время жизни τ электрона в яме за счет выхода через барьеры оказывается конечным, а, следовательно, в силу принципа неопределенности для энергии [7.2 р. 185], энергетические уровни "размываются" на величину неопределенности энергии, имеющей порядок $\Gamma \sim \hbar/\tau$. Следовательно, ширина участка ОДП по напряжению конечна, а ВАХ становится более плавной.
- при ненулевой температуре граница между незаселенными и заселенными состояниями в эмиттере размыта (максимальная энергия электрона не ограничивается E_F). Следовательно, при изменении напряжения количество туннелирующих электронов и создаваемый ими ток изменяется плавно;
- процессы рассеивания и хаотизации фазы электрона при туннелировании делают условие резонанса (7.4) не единственным, при котором возможен перенос заряда из левого к правому резервуару. Существуют механизмы нерезонансного переноса, которые существенно увеличивают ток "долины" ВАХ.

7.1.2. Механизм токопереноса

Подход к расчету тока через структуры с поперечным электронным транспортом аналогичен рассмотренному в разд. 1.5: "квантовая область", являющаяся по существу микроскопической системой, описывается как центр рассеивания. Описание последнего можно осуществлять, пользуясь волновыми матрицами рассеивания или передачи электронных волн. Покажем как в рамках формализма Ландауэра, зная матрицу передачи квантовой системы, можно рассчитать ток.

В резервуарах, где концентрация электронов высока, а падение напряжения пренебрежительно мало, принимается предположение о равновесности электронного газа. Считается, что при некотором напряжении V электроны в левом и правом резервуарах подчиняются статистике Ферми-Дирака с парамет-

рами E_{Φ_1} и E_{Φ_2} , соответственно, где $E_{\Phi_2} = E_{\Phi_1} - eV$ (рис. 7.4, а) Кроме того, считается, что резервуары поглощают все электроны, которые в них попадают (то есть отсутствует отражение от границы резервуара и квантовой области).

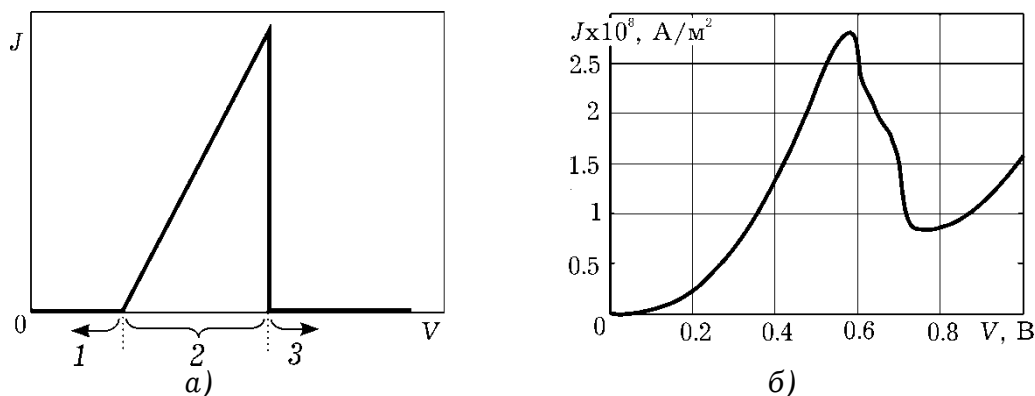


Рис. 7.3. Теоретическая (а) и экспериментальная ВАХ резонансно-туннельного диода (б), адаптированная из (9). Обозначенные участки теоретической ВАХ соответствуют характерным значениям резонансного волнового числа:

1 - $k_0 > k_{\Phi}$; 2 - $0 < k_0 < k_{\Phi}$; 3 - $\text{Im}(k_0) < 0$ (см. также рис. 7.2, а)

Поскольку квантовая область является нелегированной (для GaAs концентрация электронов в ней $n_i = 2,1 \cdot 10^{12} \text{ м}^{-3}$ при температуре 300 K), а резервуары легируются сильно ($N_d \sim 10^{24} \text{ м}^{-3}$), основную роль в явлениях переноса играют электроны левого и правого резервуаров, которые попадают в квантовую область и двигаются в ней баллистически. Этот процесс напоминает процесс термоэлектронной эмиссии, имеющей место в вакуумных диодах [7.10 с. 194], однако в данном случае электроны эмитируются не только из "катода", но и с "анода".

Суммарный ток, протекающий через квантовую область, можно рассчитать как суперпозицию токов, которые создаются электронами, попадающих в нее из левого и правого резервуара (далее - "левые" и "правые" электроны). Плотности последних обозначим, соответственно, через J_{12} и J_{21} , а суммарную плотность тока - через J :

$$\mathbf{J} = \mathbf{J}_{12} + \mathbf{J}_{21}.$$

Считается, что потоки "левых" и "правых" электронов между собой не взаимодействуют. Взаимодействие между электронами может быть учтено эффективно, если считать, что каждый электрон движется в поле кристаллической решетки и поле, создаваемого всеми остальными электронами - так называемое

приближение Хартри [7.2 с. 311], [7.11 с. 71]. Считается, что площадь слоев гетероструктуры настолько велика, что ток равномерно распределяется по поперечному перерезу площадью S , поэтому плотность тока $\mathbf{J}_{12(21)} = \mathbf{e}_z \frac{I_{12(21)}}{S} \equiv \mathbf{e}_z J_{12(21)}$.

Для токов "левых" и "правых" электронов должны отдельно выполняться законы непрерывности: $dJ_{12(21)}/dz = 0$, поэтому для «левых» электронов можно рассчитать ток в одной точке, а для «правых» – в другой. Удобнее всего рассчитать ток по формуле:

$$J = J_{12}(z_R) + J_{21}(z_L), \quad (7.5)$$

где z_R и z_L – правая и левая граница квантовой области (рис. 7.4). Точки z_L и z_R – это граничные точки классической области, поэтому в них электрон можно рассматривать как квазиклассическую частицу.

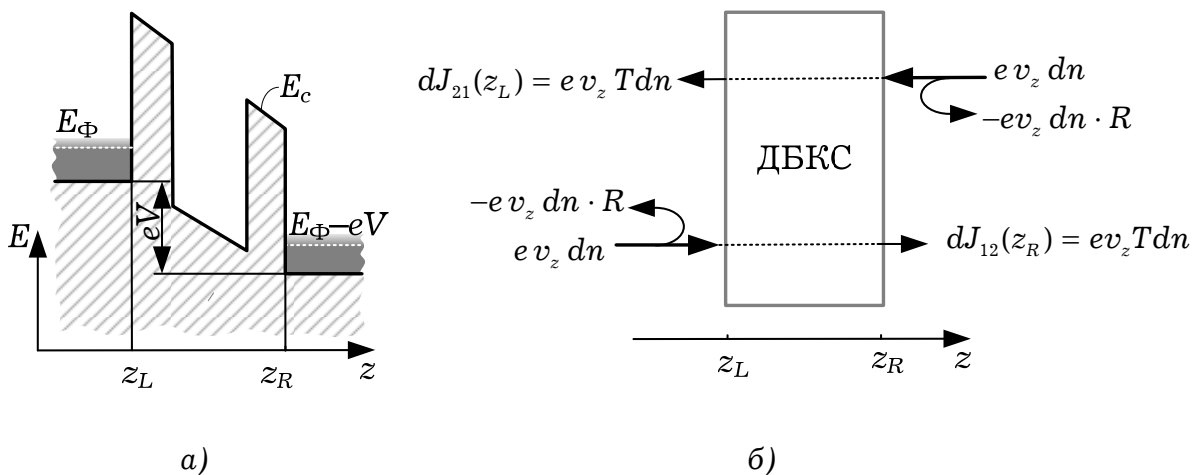


Рис. 7.4. Зонная диаграмма двухбарьерной квантовой системы (а) и составляющие тока через нее (б)

Электрон, который находится в точке z_R и имеет скорость v , создает элементарный ток $j = ev_z$. Плотность тока, создаваемого электронами, скорости которых находятся в пределах от $v_z = \hbar k_z / m^*$ до $v_z + dv_z$, определяется формулой:

$$dJ_{12} = e v_z dn(z_R) = e \frac{\hbar k_z}{m^*} dn(z_R), \quad (7.6)$$

где $dn(z_R)$ – концентрация электронов (скорости которых находятся в вышеупомянутых пределах) в точке z_R .

Предположим, что нам известен коэффициент прохождения электроном квантовой системы $T(E_z)$, который имеет смысл вероятности прохождения. Тогда концентрацию "левых" электронов, достигших правого резервуара, можно определить по формуле $dn(z_R) = dn(z_L) \cdot T(E_z)$ (рис. 7.4). Если электрон преодолевает квантовую систему без рассеивания, вероятность прохождения зависит лишь от поперечной энергии: $T = T(E_z)$.

Именно нахождение коэффициента прохождения $T(E_z)$ является главной проблематикой моделирования структур с поперечным транспортом, поскольку он определяет электрические свойства микроскопической системы.

Для нахождения J_{12} воспользуемся тем, что концентрация электронов dn в элементарном объеме волновых чисел $dV_k = dk_x dk_y dk_z$ составляет:

$$dn = g(\mathbf{k})f(\mathbf{k})dV_k, \quad (7.7)$$

где $g(\mathbf{k})$ и $f(\mathbf{k})$ – плотность состояний и функция распределения электронов в k -пространстве в левом резервуаре, соответственно; $g(\mathbf{k}) = 2/(2\pi)^3$; $f(\mathbf{k}) = 1/(1 + e^\varepsilon)$, где, в свою очередь:

$$\varepsilon(E_\Phi, \mathbf{k}) \equiv (E - E_\Phi) / k_B T = \left(\underbrace{\frac{\hbar^2 k_x^2}{2m^*}}_{\equiv E_x} + \underbrace{\frac{\hbar^2 k_y^2}{2m^*}}_{\equiv E_y} + \underbrace{\frac{\hbar^2 k_z^2}{2m^*}}_{\equiv E_z} - eV - E_\Phi \right) / k_B T.$$

С учетом (7.7), уравнение (7.6) в раскрытом виде переписывается так:

$$dJ_{12} = \frac{e\hbar k_z}{m^*} g(\mathbf{k})f(\mathbf{k})T(E_z)dV_k = \frac{e\hbar}{m^*} \frac{2}{(2\pi)^3} \frac{k_z T(E_z)}{1 + \exp\{\varepsilon(\mathbf{k}, E_\Phi)\}} dV_k. \quad (7.8)$$

Поскольку коэффициент прохождения зависит лишь от поперечной энергии, в формуле (7.8) удастся провести интегрирование по поперечным проекциям волнового вектора (k_x и k_y). В то же время вид функции $T(E_z)$ даже для самых простых форм квантовой области обычно не позволяет провести интегрирование по k_z аналитически.

Интегрирование (7.8) следует производить по части k -пространства, определяемую неравенством $k_z > 0$:

$$J_{12} = \frac{e\hbar}{m^*} \frac{2}{(2\pi)^3} \int_0^\infty T(E_z) \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{k_z}{1 + \exp\{\varepsilon(\mathbf{k}, E_\Phi)\}} dk_x dk_y dk_z \quad (7.9)$$

(эти пределы охватывают все электроны, движущиеся слева направо и имеющие произвольные продольные скорости).

Интегрируя (7.9) по k_x и k_y и переходя к интегрированию по энергии, получим (см. приложение П7.1):

$$J_{12} = \frac{2m^* ek_B T}{(2\pi)^2 \hbar^3} \int_{eV}^\infty T(E_z) \cdot \ln \left(1 + \exp \left(-\frac{E_z - E_\Phi - eV}{k_B T} \right) \right) dE_z, \quad (7.10)$$

где нижний предел интегрирования учитывает отсутствие в левом резервуаре электронов, энергия которых меньше eV .

Аналогично найдем J_{21} , (см. приложение П7.1):

$$J_{21} = \frac{2m^* ek_B T}{(2\pi)^2 \hbar^3} \int_0^\infty T(E_z) \cdot \ln \left(1 + \exp \left(-\frac{E_z - E_\Phi}{k_B T} \right) \right) dE_z. \quad (7.11)$$

Подставив (7.10) и (7.11) в (7.5), получим известную формулу Цу-Эсаки [7.12]:

$$J = \frac{2m^* ek_B T}{(2\pi)^2 \hbar^3} \int_{eV}^\infty T(E_z) D(E_z) dE_z, \quad (7.12)$$

где $D(E_z)$ – так называемая функция снабжения электронов (англ. "supply function"):

$$D(E) \equiv \ln \left(\frac{1 + \exp \left(-\frac{E - E_\Phi}{k_B T} \right)}{1 + \exp \left(-\frac{E - (E_\Phi - eV)}{k_B T} \right)} \right).$$

Из двух нижних пределов 0 и eV был выбран наибольший, потому что для состояний с $E_z < eV$ $T(E_z) = 0$.

Из формулы Цу-Эсаки следует, что задача по нахождения вольт-амперной характеристики $J(V)$ сводится к определению коэффициента прохождения $T(E_z, V)$, которая при известном потенциальном рельефе лишь немногим сложнее, чем классическая квантово-механическая задача о прохождении

электроном некоторого потенциального рельефа (например, прямоугольного потенциального барьера).

7.2. УСЛОВНО-АНАЛИТИЧЕСКИЕ МОДЕЛИ

Со времен формирования первых представлений о функционировании полупроводниковых структур с поперечным транспортом было разработано множество аналитических методов для вычисления коэффициента прохождения электроном квантовой структуры. Однако, сложность получающихся выражений обычно не позволяет применять их без дальнейших упрощений. По большей части они используются для построения численных моделей. По этой причине мы использовали приставку "условно-" в их названии. Исключением является *метод лоренцовой модельной функции* (разд. 7.2.2), однако его применимость ограничена исключительно резонансно-тоннельным диодом при условии малых смещений.

Общий подход всех методов заключается в том, чтобы найти матрицу передачи квантовой системы, то есть матрицу, которая связывает падающие и отраженные электронные волны с обеих сторон квантовой системы. Зная элементы матрицы передачи, находят коэффициент прохождения.

Как вводится матрица передачи? Согласно принятым предположениям, за пределами квантовой системы падение прикладываемого напряжения отсутствует, потому электрон находится исключительно в поле кристаллической решетки, и огибающая его волновой функции рассматривается как суперпозиция падающих и отраженных плоских волн. Поэтому решение уравнения Шредингера для левого и правого резервуара (на рис. 7.5, а это I и V участки) запишется так:

$$\psi_1 = A_1 e^{ik_1 z} + B_1 e^{-ik_1 z}, \quad (7.13)$$

$$\psi_5 = A_5 e^{ik_5 x} + B_5 e^{-ik_5 x}, \quad (7.14)$$

где A_1, B_1, A_5, B_5 – постоянные интегрирования; $k_{1(5)} = \sqrt{2m_{1(5)}^*(E_z - U_{1(5)})}/\hbar$, где $m_{1(5)}^*$ – эффективная масса электрона в I(V) области, $U_{1(5)}$ – потенциальная энергия электрона ("потенциальный рельеф") в I(V) области. Введем обозначения для членов выражений (7.13), (7.14):

$$\psi_1^{\rightarrow} \equiv A_1 e^{ik_1 z}, \quad \psi_1^{\leftarrow} \equiv B_1 e^{-ik_1 z}, \quad \psi_5^{\rightarrow} \equiv A_5 e^{ik_5 z}, \quad \psi_5^{\leftarrow} \equiv B_5 e^{-ik_5 z},$$

которые традиционно называются, соответственно, падающими и отраженными волнами в первой и пятой области. Эти волны описывают электрон, который налетает или отражается от квантовой области слева или справа.

Матрица передачи всей квантовой системы \mathbf{M} вводится так:

$$\begin{bmatrix} A_5 \\ B_5 \end{bmatrix} = \mathbf{M} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} \equiv \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix}, \quad (7.15)$$

Зная \mathbf{M} , можно найти коэффициент прохождения $T(E_z)$ квантовой системы. По определению [7.3 с. 419]:

$$T(E_z) = j_5^{\rightarrow} / j_1^{\rightarrow}, \quad (7.16)$$

где j_5^{\rightarrow} – поток плотности вероятности в электронной волне, которая прошла в пятую область; j_1^{\rightarrow} – поток плотности вероятности в падающей волне в первой области. Согласно определению потока плотности вероятности:

$$j_{5(1)}^{\rightarrow} = \frac{i\hbar}{2m_{5(1)}^*} \left(\{\psi_{5(1)}^{\rightarrow}\}^* \frac{d\psi_{5(1)}^{\rightarrow}}{dz} - \psi_{5(1)}^{\rightarrow} \frac{d\{\psi_{5(1)}^{\rightarrow}\}^*}{dz} \right).$$

Простые вычисления дают:

$$j_5^{\rightarrow} = |A_5|^2 \frac{\hbar k_5}{m_5^*}, \quad j_1^{\rightarrow} = |A_1|^2 \frac{\hbar k_1}{m_1^*}.$$

Подставляя эти значения в формулу (7.16), получим:

$$T(E_z) = \frac{|A_5|^2}{|A_1|^2} \frac{k_5}{k_1} \frac{m_1^*}{m_5^*}. \quad (7.17)$$

Из формулы (7.17) следует: если потенциальная энергия электрона в левом или правом резервуаре меньше полной энергии электрона (волновой вектор приобретает мнимое значение), то такой электрон не вносит вклад в ток, поскольку вероятность прохождения оказываются мнимыми. Последнее означает невозможность такого события.

Коэффициент отражения от квантовой системы определяется как:

$$R = -j_1^{\leftarrow} / j_1^{\rightarrow}, \quad (7.18)$$

где $j_1^{\leftarrow} = \frac{i\hbar}{2m_1^*} \left(\{\psi_1^{\leftarrow}\}^* \frac{d\psi_1^{\leftarrow}}{dz} - \psi_1^{\leftarrow} \frac{d\{\psi_1^{\leftarrow}\}^*}{dz} \right)$ — поток плотности вероятности в отраженной в первую область электронной волне.
Простые вычисления дают:

$$R = |B_1|^2 / |A_1|^2. \quad (7.19)$$

Найдем коэффициент прохождения для электрона, который налетает на квантовую систему из левого резервуара. Поскольку резервуары являются идеально адсорбирующими, волна, отраженная от правого резервуара $\psi_5^{\leftarrow} \equiv B_5 e^{-ik_5 z}$, будет отсутствовать. Без потери общности можно нормировать волновую функцию в левом резервуаре так, что $A_1 = 1$. Тогда амплитуды A_5 и B_1 приобретут смысл комплексных коэффициентов прохождения и отражения, которые обозначим $t \equiv A_5$, $r \equiv B_1$. Таким образом, для электрона, который налетает на квантовую систему слева, (7.17) и (7.19) переписутся так:

$$T(E_z) = |t|^2 \frac{|k_5| m_1^*}{|k_1| m_5^*}, \quad (7.20)$$

$$R = |r|^2. \quad (7.21)$$

Формула (7.15) во введенных обозначениях будет иметь вид:

$$\begin{bmatrix} t \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} 1 \\ r \end{bmatrix}. \quad (7.22)$$

Здесь неизвестными являются коэффициенты t и r , найдя которые, будем иметь возможность вычислить $T(E_z)$ и R согласно (7.20) и (7.21). Проведя элементарные вычисления, из (7.22) получим:

$$t = m_{11} + m_{12} m_{21} / m_{22}; \quad (7.23)$$

$$r = -m_{21} / m_{22}. \quad (7.24)$$

Следовательно, если матрица передачи квантовой системы известна, формулы (7.20), (7.21), (7.23), (7.24), позволяют вычислить коэффициенты прохождения и отражения.

Напоследок отметим, что закон сохранения плотности потока вероятности в огибающей волновой функции, как и в самой волновой функции [7.3 с. 419] диктует выполнение условия:

$$T(E_z) + R(E_z) = 1;$$

кроме того, в [7.2 с. 108] показано, что для электронов, которые налетают справа или слева, коэффициенты отражения и прохождения одинаковы: $T_{12} = T_{21}$, $R_{12} = R_{21}$. Именно поэтому в формуле Цу-Эсаки фигурирует один коэффициент прохождения $T(E_z) \equiv T_{12} = T_{21}$. Системы, которые удовлетворяют указанным условиям, называются *обратимыми системами без потерь*. Для компонентов матрицы \mathbf{M} , которая описывает указанные системы, справедливы равенства: $m_{11} = m_{22}^*$, $m_{12} = m_{21}^*$ [7.13 с. 250].

7.2.1. Метод плоских волн

Метод плоских волн применяется для нахождения коэффициента прохождения через квантовые системы, потенциальная энергия электрона в которых может быть аппроксимированная кусочно-постоянной функцией (рис. 7.5, а, б). Аппроксимация, которая показана на рис. 7.5, а сплошной линией, является оправданной, когда к системе прикладывается небольшое смещение, а влиянием пространственного заряда можно пренебречь.

На любых двух i -м и $(i + 1)$ -ом смежных участках с постоянным потенциальным рельефом решением уравнения Шредингера будет суперпозиция плоских волн:

$$\psi_i = A_i e^{ik_i z} + B_i e^{-ik_i z}, \quad (7.25)$$

$$\psi_{i+1} = A_{i+1} e^{ik_{i+1} z} + B_{i+1} e^{-ik_{i+1} z}, \quad (7.26)$$

где $k_i = \sqrt{2m_i^*(U_i - E_z)} / \hbar$, m_i^* — эффективная масса электрона в i -ом слое, U_i — потенциальная энергия электрона в зоне проводимости i -го слоя.

На границе между i -й и $i + 1$ -й областями, в точке $z_{i, i+1}$ волновая функция и поток вероятности в ней должны быть

непрерывными, то есть необходимо обеспечить выполнение соотношений (см. также разд. 1.5):

$$\Psi_i(z_{i,i+1}) = \Psi_{i+1}(z_{i,i+1}), \quad (7.27)$$

$$1 / m_i^* \cdot \Psi_i'(z_{i,i+1}) = 1 / m_{i+1}^* \cdot \Psi_{i+1}'(z_{i,i+1}), \quad (7.28)$$

где штрихом обозначено дифференцирование по z .

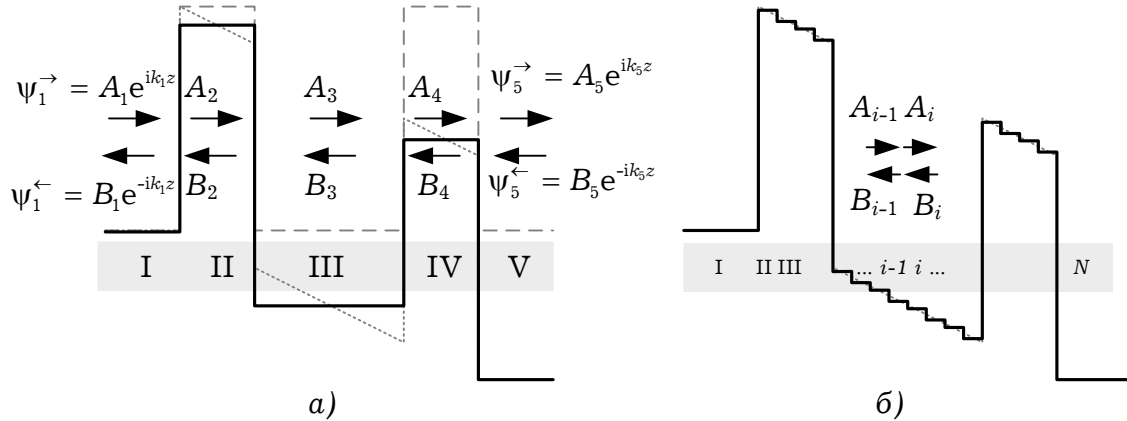


Рис. 7.5. Аппроксимации потенциального рельефа ДБКС, которые можно описать матрицами передачи плоских волн (сплошные линии)

С учетом (7.25) и (7.26) из (7.27) и (7.28) можно получить матричное уравнение, которое связывает $[A_{i+1} \ B_{i+1}]^T$ с $[A_i \ B_i]^T$:

$$\begin{bmatrix} A_{i+1} \\ B_{i+1} \end{bmatrix} = \mathbf{M}_{i+1,i} \begin{bmatrix} A_i \\ B_i \end{bmatrix}, \quad (7.29)$$

где $\mathbf{M}_{i+1,i} = \begin{bmatrix} e^{ik_{i+1}z} & e^{-ik_{i+1}z} \\ \frac{k_{i+1}}{m_{i+1}^*} e^{ik_{i+1}z} & -\frac{k_{i+1}}{m_{i+1}^*} e^{-ik_{i+1}z} \end{bmatrix}^{-1} \begin{bmatrix} e^{ik_i z} & e^{-ik_i z} \\ \frac{k_i}{m_i^*} e^{ik_i z} & -\frac{k_i}{m_i^*} e^{-ik_i z} \end{bmatrix}$ - матрица передачи электронных волн (или просто \mathbf{M} -матрица), которая выражает амплитуды электронных волн в $(i+1)$ -ом слое через амплитуды в i -ом слое.

Аналогично можно получить уравнение, которое связывает соответствующие амплитуды в $(i-1)$ -ом и i -ом слое:

$$\begin{bmatrix} A_i \\ B_i \end{bmatrix} = \mathbf{M}_{i,i-1} \begin{bmatrix} A_{i-1} \\ B_{i-1} \end{bmatrix}. \quad (7.30)$$

Пользуясь (7.29) и (7.30), можно выразить $[A_{i+1}, B_{i+1}]^T$ через $[A_{i-1}, B_{i-1}]^T$:

$$\begin{bmatrix} A_{i+1} \\ B_{i+1} \end{bmatrix} = \mathbf{M}_{i+1,i} \times \mathbf{M}_{i,i-1} \begin{bmatrix} A_{i-1} \\ B_{i-1} \end{bmatrix},$$

откуда следует, что \mathbf{M} -матрицы мультипликативны.

Легко убедиться, что указанное свойство \mathbf{M} -матриц позволяет выразить амплитуды падающих и отраженных волн в левом резервуаре через аналогичные величины правого резервуара:

$$\begin{bmatrix} A_N \\ B_N \end{bmatrix} = \mathbf{M}_{N,N-1} \times \mathbf{M}_{N-1,N-2} \times \dots \times \mathbf{M}_{2,1} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} = \prod_{i=2}^N \mathbf{M}_{i,i-1} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix}.$$

Матрица $\mathbf{M} = \prod_{i=2}^N \mathbf{M}_{i,i-1}$ является искомой матрицей передачи всей квантовой системы.

Применим метод плоских волн к ДБКС модельного РТД, состоящего из AlAs-барьеров шириной $b = 1,7$ нм и GaAs-ями шириной $a = 4,5$ нм. Относительные эффективные массы в соответствующих слоях имеют значение: $m_{\text{AlAs}} = 0,15$, $m_{\text{GaAs}} = 0,067$. Разрыв зоны проводимости в месте контакта $U_0 = 1$ эВ. Концентрация донорных примесей в резервуарах $N_D = 2,5 \cdot 10^{24} \text{ м}^{-3}$ (рис. 7.6).

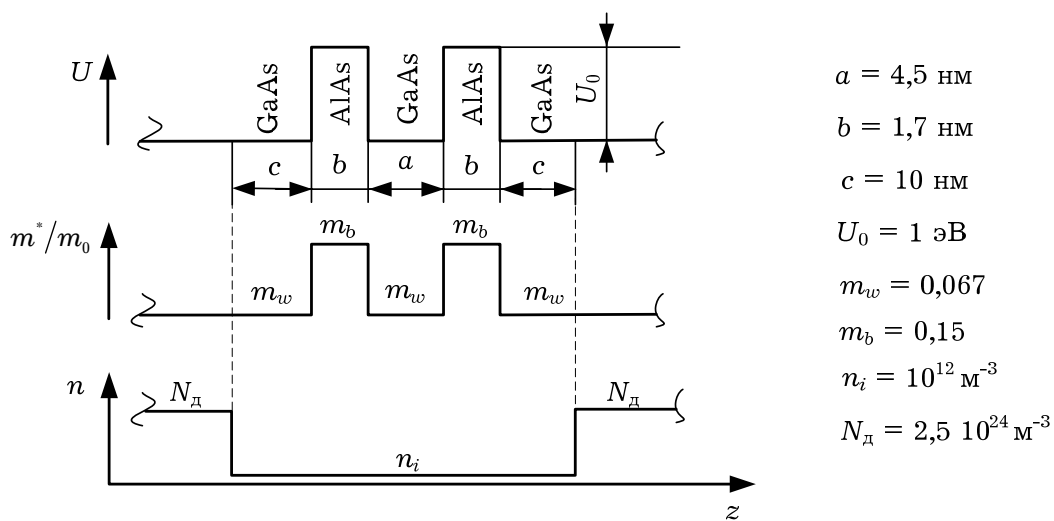


Рис. 7.6. Топология активной части модельного резонансно-туннельного диода

На рис. 7.7 показана зависимость коэффициента прохождения от поперечной энергии электрона для нескольких значений напряжения. Потенциальный рельеф в зависимости от прилагаемого напряжения аппроксимировался согласно рис. 7.5, а.

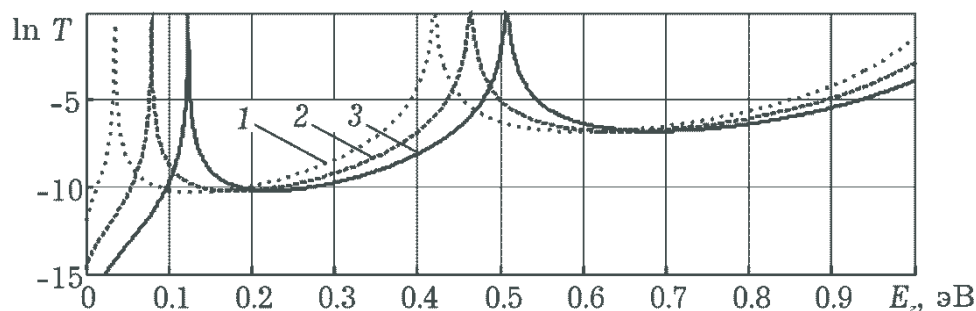


Рис. 7.7. Зависимость коэффициента прохождения модельного резонансно-туннельного диода, рассчитанная методом матриц передачи для напряжений: 1 – 0 В; 2 – 0,1 В; 3 – 0,2 В

7.2.2. Использование лоренцевой модельной функции

Использование матриц передачи на практике требует численной реализации. Умножение уже пяти матриц приводит к сложному выражению для коэффициента передачи с точки зрения возможности анализа (хотя такие попытки осуществлялись [7.14]). Например, как объяснить резонансный характер зависимости коэффициента прохождения для резонансно-туннельного диода (рис. 7.7)? Оказывается, что для РТД можно пользоваться функциями, суперпозиция которых удовлетворительно аппроксимирует резонансный характер зависимости $T(E_z)$. Называются они функциями лоренцового типа, и описываются формулой:

$$T_i(E_z) = \frac{\Gamma_i^2}{\Gamma_i^2 + 4(E_z - E_i)^2}. \quad (7.31)$$

где Γ_i – это величина расширения резонансного уровня, которое обусловлено принципом неопределенности Гейзенберга для энергии; E_i – положение i -го метастабильного уровня в квантовой яме. Легко видеть, что $T_i(E_i) = 1$, а при $|E_z - E_i| \gg \Gamma_i$ $T_i(E_i) \rightarrow 0$, то есть эта функция описывает некоторую резонансную зависимость.

Иногда формулу (7.31) также называют *формулой Брейта-Вигнера*; она впервые была получена и использована для описания процессов резонансного рассеивания нейтронов на атомных ядрах.

Для трехмерного сферического потенциального барьера одна из форм этой формулы выводится, например, в [7.3 с. 426].

В квантовой теории возмущений особое место занимает задача, когда возмущение мало, однако его вид таков, что спектр квантовой системы из дискретного становится непрерывным [7.2 с. 202], [7.3 с. 284]. Такая задача возникает, когда невозмущенной системой является яма, окруженная бесконечно широкими потенциальными барьерами (рис. 7.8, а), а возмущенной – такая же яма, но окруженная достаточно широкими барьерами конечной ширины (рис. 7.8, б), настолько широкими, чтобы возмущение можно было считать малым. Для резонансно-туннельных диодов, которые имеют более-менее качественные электрические характеристики, это всегда выполняется.

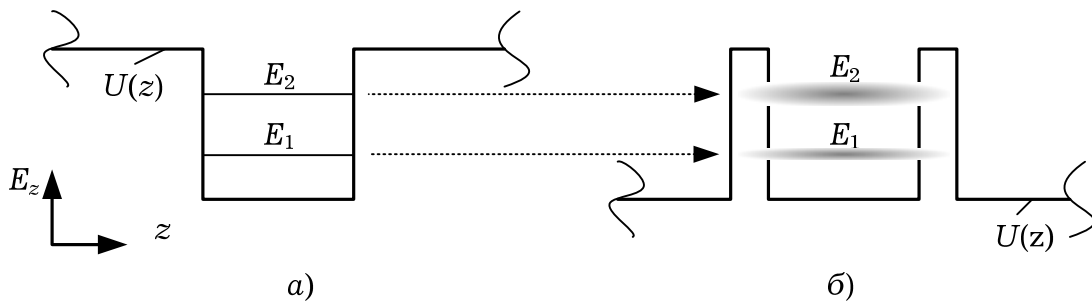


Рис. 7.8. Образование метастабильных уровней энергии: а) невозмущенная система (ширина барьеров бесконечна); б) возмущенная система (ширина барьеров конечна)

В системе, которая изображена на рис. 7.8, а при энергии меньше высоты барьера спектр энергии электрона дискретен. Говорят, что *электрон локализован в квантовой яме*. Вероятность найти его на бесконечности равна нулю, а энергия – фиксирована (неопределенность энергии $\Delta E = 0$). Действительно, согласно принципу неопределенности Гейзенберга для энергии [7.2 с. 202]:

$$\Delta E \Delta t \sim \hbar, \quad (7.32)$$

а поскольку электрон находится в квантовой яме бесконечно долго ($\Delta t \rightarrow \infty$), (7.32) выполняется, если $\Delta E = 0$.

Ситуация принципиально изменяется, когда барьеры имеют конечную ширину (рис. 7.8, б). Движение электрона в этом случае будет *инфинитным*, а спектр *непрерывным*. Однако и в этом случае вероятность найти электрон в яме будет наибольшей вблизи дискретных уровней энергии невозмущенной системы, уменьшаясь при удалении от них. Говорят, что дискретные уровни энергии "размываются". Если величина ΔE значительно меньше

расстояния между дискретными уровнями, такой спектр называют "квазидискретным". Последний является особым случаем *непрерывного* спектра.

Таким образом, в формуле (7.31) E_i - это собственные энергии невозмущенной системы (рис. 7.8, а) для дискретного спектра.

Из формулы (7.31) следует, что при $E_z = E_i$ функция $T(E_z)$ имеет максимумы, которые равняются единице. Если собственными частотами системы считать частоты ДБКС $\omega_{0,i} = E_i / \hbar$, а внешними - частоты волны де Бройля электронов $\omega_{\text{внешн.}} = E_z / \hbar$, то равенство $E_z = E_i$ представляет собой классическое определение резонанса. Явление прохождения ДБКС электроном, обладающего резонансной энергией, называется *резонансным туннелированием*.

Размытые уровни энергии, о которых шла речь, называют метастабильными, а среднее время движения электрона на них Δt называют временем жизни данного метастабильного состояния.

Величина Γ_i в формуле (7.32) - это так называемая ширина i -го метастабильного уровня. Ее называют "*естественным расширением*" соответствующего уровня; оно определяется как:

$$\Gamma_i = \frac{\hbar}{\tau_i}, \quad (7.33)$$

где τ_i - это среднее время жизни данного метастабильного состояния.

Величина Γ_i имеет следующий физический смысл. Для системы, коэффициент прохождения которой описывается лоренцевой модельной функцией, часть электронов, которая прошла через систему в диапазоне шириной Γ_i с центром в $E_z = E_i$, будет составлять $1 / \sqrt{2} \approx 0.707$ от общего количества электронов, которые прошли через ДБКС при всех возможных энергиях (при условии, если распределение электронов по энергиям равномерно). Условно говоря, "окно прозрачности" квантовой системы имеет порядок Γ_i .

Время жизни в (7.33) находят из квазиклассических соображений. При каждом отражении от i -го барьера часть $T_i = 1 - R_i$ электронной волны "утекает" из ямы. Частота столкновений электрона с барьерами ν_i приблизительно равняется $2a/v_i$, где v_i - это скорость движения электрона на i -ом метастабильном уровне; $v_i = \sqrt{2E_i / m^*}$; a - ширина квантовой ямы. Вероятность того, что при столкновении с барьером электрон

покинет квантовую яму, составляет $T_i^{(1)}$ для левого и $T_i^{(2)}$ для правого барьера. Указанные оценки справедливы при условии $|T_i^{(1)}| \ll 1$, $|T_i^{(2)}| \ll 1$; тогда величину τ_i можно приближенно рассчитать по формуле:

$$\tau_i = \frac{2a}{v_i} \left(\frac{1}{T_i^{(1)} + T_i^{(2)}} \right). \quad (7.34)$$

Если в квантовой яме резонансно-туннельного диода имеется N метастабильных уровней энергии, то функция $T(E_z)$ может быть аппроксимирована суммой функций лоренцевого типа:

$$T(E_z) = \sum_{i=1}^N \frac{\Gamma_i^2}{\Gamma_i^2 + 4(E_z - E_i)^2}. \quad (7.35)$$

Если в квантовой яме существуют другие механизмы, позволяющие электрону покинуть данный энергетический уровень, например процессы рассеивания, то в рамках лоренцевой аппроксимации их тоже можно учесть. Пусть суммарное время релаксации по всем механизмам релаксации равняется Γ_p и определяется обычным образом:

$$1 / \tau_p = \sum_n \lambda_n,$$

где λ_n — частоты n -го вида рассеивания. Со временем τ_p так же, как и со временем τ_i , можно связать дополнительное расширение метастабильного уровня:

$$\Gamma_p = \frac{\hbar}{\tau_p}, \quad (7.36)$$

которое называется "релаксационным".

По аналогии с резонатором Фабри-Перо [7.15] или следуя квантовой теории можно показать, что при учете рассеивания формулу (7.35) можно уточнить и получить:

$$T(E_z) = \sum_{i=1}^N \frac{\Gamma_i^2}{\Gamma_{\Sigma,i}^2 + 4(E_z - E_i)^2}, \quad (7.37)$$

где $\Gamma_{\Sigma,i} = \Gamma_i + \Gamma_p$ – суммарное расширение i -го энергетического уровня.

7.2.3. Полуаналитическая упрощенная модель РТД

С помощью формулы (7.37) и формулы Цу-Эсаки можно построить упрощенную аналитическую модель резонансно-туннельного диода [7.16]. Преимуществами такой модели по сравнению с численными являются простота и интуитивно понятный смысл величин, которые вычисляются на пути получения ВАХ, а также частичная аналитичность.

Исходя из разд. 7.2.2, поиск зависимости $J(V)$ сводится к следующим шагам:

- поиск суммарного количества N и энергии E_i дискретных уровней в невозмущенной системе (задача на собственные энергии для потенциального рельефа из рис. 7.8, а);

- поиск коэффициентов прохождения для каждого из двух потенциальных барьеров, $T^{(1)}$ и $T^{(2)}$ (рис. 7.8, б);

- поиск величин "естественного" и "релаксационного" расширения метастабильных энергетических уровней в квантовой яме по формулам (7.33) и (7.36);

- использование рассчитанных величин для получения ВАХ из формулы Цу-Эсаки.

Найти количество и энергию дискретных уровней в квантовой яме, окруженной бесконечно широкими барьерами (рис. 7.8, а), можно, решив соответствующую квазиквантовомеханическую задачу, которая лишь немногим сложнее, чем рассмотренная в [7.2 с. 93].

Расчеты позволяют получить следующие формулы [7.17]. Количество метастабильных уровней в квантовой яме:

$$N = \left[\frac{2\sqrt{m_b}}{\pi\gamma} \right] + 1 = \left[\sqrt{\frac{m_b}{m_0} \cdot \frac{m_w}{m_0} \frac{2}{\pi\gamma_0}} \right] + 1, \quad (7.38)$$

где $[x]$ помечает целую часть x ; m_0 , m_w , m_b – масса покоя электрона и эффективные массы электрона в яме и барьере;

$\gamma^2 = \frac{2\hbar^2}{m_w U a^2}$, $\gamma_0^2 = \frac{2\hbar^2}{m_0 U a^2}$; U – высота потенциального барьера, a –

ширина ямы.

Собственные энергии электрона в квантовой яме могут быть найдены из трансцендентного уравнения:

$$(-1)^{n+1} \cos\left(\frac{\pi}{2}(n-1) - \xi\right) = (-1)^{n+1} \sqrt{1 + \frac{m_w}{m_b} \left(\frac{1}{\gamma^2 \xi^2} - 1\right)}, \quad (7.39)$$

где

$$\xi = ka / 2, \quad k = \sqrt{2m_w E} / \hbar. \quad (7.40)$$

При $m_w = m_b$ правая часть уравнения (7.39) является уравнением прямой $f(\xi) = \pm \gamma \xi$, (если же $m_w \neq m_b$, то это будут некоторые кривые, которые выходят из точки $(0;0)$), левая – в зависимости от четности n – синусом или косинусом. Точки пересечения двух упомянутых функций являются искомыми ξ_i , которым соответствуют искомые E_i согласно (7.40). Один из возможных способов графического решения показан на рис. 7.9.

Коэффициент прохождения через прямоугольные потенциальные барьеры можно

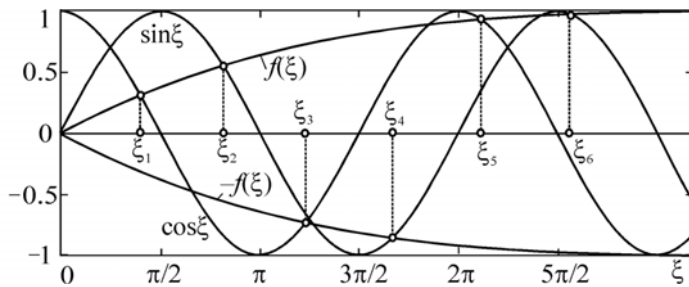


Рис. 7.9. Графическое решение уравнения (7.39). Показанные точки пересечения кривых $f(\xi)$ и кривых $\sin(\xi)$ или $\cos(\xi)$ – его корни

рассчитать, например, используя метод мат-риц передачи. Формула для электрона с энергией $E_z = E_i$ имеет вид:

$$T_i = \frac{4C_i^2}{(C_i + 1)^2 \text{sh}^2 b \chi_i + 4C_i^2},$$

где b – ширина барьера,

$$C_i = \frac{k_i}{\chi_i} \cdot \frac{m_2^*}{m_1^*},$$

$\hbar k_i = \sqrt{2m_b^* E_i}$, $\hbar \chi_i = \sqrt{2m_2^* (U_0 - E_i)}$, U_0 – высота потенциального барьера.

В простейшем случае можно пренебречь зависимостью $T^{(2)}$ от прилагаемого к ДБКС напряжения V , и считать, что под воздействием напряжения уровни энергии опускаются на величину $eV / 2$:

$$E_i(V) = E_i(0) - \frac{eV}{2}.$$

После того как при энергиях E_i проницаемости левого и правого барьеров ($T_i^{(1)}$ и $T_i^{(2)}$) найдены, естественное расширение i -го резонансного уровня Γ_i можно определить из уравнения (7.33).

Наиболее существенным механизмом рассеивания в GaAs при комнатной температуре является полярное оптическое рассеивание. Приблизительно можно считать, что время релаксации импульса в квантовой яме совпадает со временем релаксации в трехмерном полупроводнике, которое определяется формулой [7.18 с. 71]:

$$\frac{1}{\tau_p} = \alpha \omega_0 \sqrt{\frac{2}{\pi}} \frac{t^{1/2}}{\text{sh}t} K_0(t), \quad (7.41)$$

где α – полярная постоянная для GaAs; $t = \frac{\hbar \omega_0}{2k_B T}$; K_0 – функция Бесселя нулевого порядка второго рода.

При нулевом смещении суперпозиция лоренцевых функций является хорошей аппроксимацией функции прохождения, особенно для низких энергетических уровней, о чем свидетельствует рис. 7.10, где на одном графике изображена $T(E_z)$, рассчитанная методом матриц передачи, и $T(E_z)$, рассчитанная в соответствии с (7.35) для все той же структуры из рис. 7.6. Для высших энергетических уровней (второй пик на рис. 7.10) аппроксимация менее точна, поскольку для энергии $E_z = E_2$ условие для проницаемости барьеров $T^{(i)} \ll 1$ выполняется хуже.

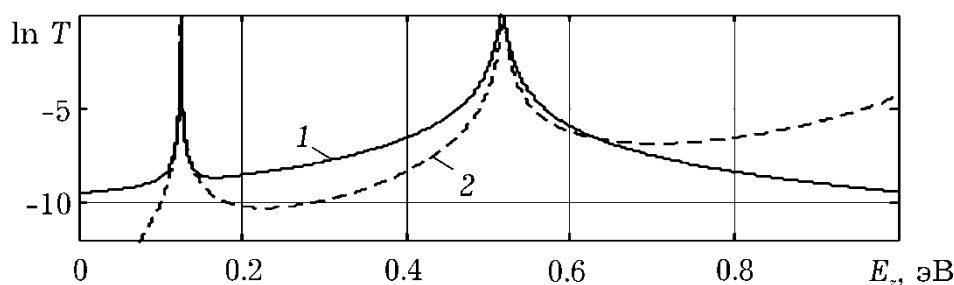


Рис. 7.10. Коэффициент прохождения модельного РТД при нулевом напряжении, рассчитанный: 1 – с использованием суммы лоренцевых модельных функций (формула (7.35)); 2 – методом матриц передачи

Одна из проблем, которую приходится решать при проектировании резонансно-туннельных диодов, заключается в малой контрастности (отношение пикового тока к току в долине РТД) вольтамперных характеристик. Причина заключается в наличии рассеивания и хаотизации фазы электронов в квантовой

яме. Модели, которые не учитывают эти факторы, дают сильно заниженные значения тока в долине ВАХ.

Неупругое рассеивание в яме резонансно-туннельного диода приводит к тому, что электрон с резонансной энергией ($E_z = E_i$) имеет конечную вероятность ее изменить за счет взаимодействия с полярными оптическими фононами, которые в GaAs имеют энергию $\hbar\omega \approx 36$ мэВ. Вероятность прохождения электроном квантовой системы с энергией, близкой к резонансной, падает, и даже при $E_z = E_i$ не равняется единице. Процессы рассеивания моделируются введением релаксационного расширения уровня, ширина которого складывается Γ_p , в дополнение к естественному расширению каждого i -го уровня, Γ_i .

Изменение формы $T(E_z)$ по аналогии с затуханием электромагнитной волны в резонаторе Фарби-Перо [7.15] позволяет учесть формула (7.37). Отметим, что такой подход является внутренне противоречивым, поскольку механизм рассеивания электрона не сопровождается затуханием волновой функции (что имеет место для электромагнитной волны). Последнее означало бы фактически исчезновение части электронов, что, конечно же, не наблюдается в квантовой яме; напротив, должен выполняться закон непрерывности тока.

Также критике поддаются использование для расчета Γ_p времени релаксации импульса. Согласно квантовой статистике имеется больше оснований использовать в этой роли время релаксации фазы [7.19 с. 14; 7.20].

Несмотря на указанные недостатки, формула (7.37) позволяет максимально просто учесть рассеивание в яме РТД. Время релаксации импульса, будучи рассчитанным по формуле (7.41), увеличивается при повышении температуры, что приводит к увеличению Γ_p , то есть дополнительному "размытию" функции $T(E_z)$ (рис. 7.11, а), и, в конечном итоге, снижение контрастности ВАХ – рис. 7.11, б. Другие механизмы, в частности "надбарьерное" прохождение и "размытие" функции распределения, приводят к повышению тока долины. Последнее особенно заметно для РТД с "низкими" барьерами (например $\text{Al}_{0.33}\text{Ga}_{0.77}\text{As}$); такие диоды сегодня, впрочем, имеют лишь историческое значение. Для AlAs/GaAs -РТД надбарьерным проходом можно пренебречь.

Это коррелирует с экспериментальными ВАХ, на которых с момента создания первых РТД (1973 г.) до 1985 г. не удавалось получить участка с отрицательной дифференциальной проводимостью при комнатной температуре.

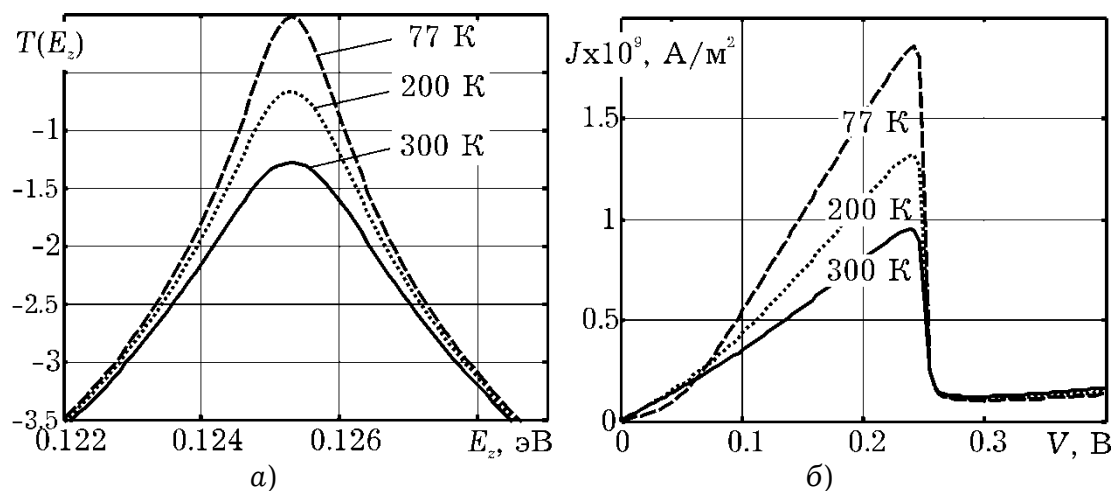


Рис. 7.11. Функция прохождения ДБКС $T(E_z)$ (а) и ВАХ резонансно-туннельного диода (б) при различных температурах

Как уже упоминалось, для второго энергетического уровня модельной структуры аппроксимация лоренцевыми функциями дает завышенные значения проницаемости двухбарьерной квантовой системы (рис. 7.10). Следовательно, и ток, обусловленный туннелированием через второй метастабильный уровень, превышает значение, рассчитанное методом матриц передачи (рис. 7.12). Кроме того, погрешности возникают при аппроксимации изменения потенциального рельефа под действием напряжения смещения V (считается, что дно потенциальной ямы вместе с уровнями энергии в ней "опускается" на величину $eV/2$).

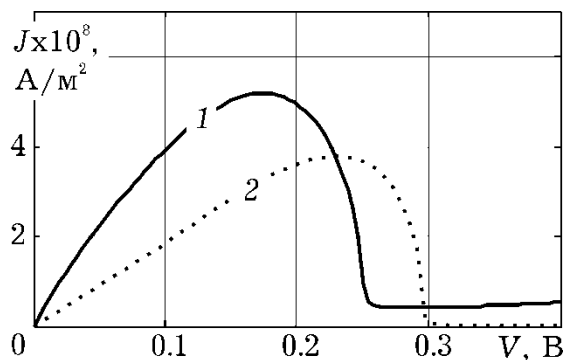


Рис. 7.12. ВАХ модельного РТД, рассчитанная с использованием:
1 – лоренцевой аппроксимации $T(E_z)$; 2 – метода матриц передачи.
Рассеивание не учитывается

7.2.4. Другие условно-аналитические методы

Для кусочно-постоянного потенциального рельефа вид решения уравнения Шредингера на каждом участке известен – это плоские волны. Кроме случая кусочно-постоянного потенциала существуют еще как минимум два: кусочно-линейный и кусочно-квадратичный, для которых вид решения уравнения Шредингера известен. Для первого случая решениями будут функции Эйри

[7.21; 7.22], для второго – функции параболического цилиндра [7.23].

Однако существенным препятствием на пути использования методов матриц передачи является численная неустойчивость таких методов. Так, если увеличить количество слоев в структуре, которая изображена на рис. 7.4 (десятибарьерная квантовая система), метод матриц передачи теряет численную стойкость за счет плохой обусловленности результирующей матрицы передачи системы (рис. 7.13).

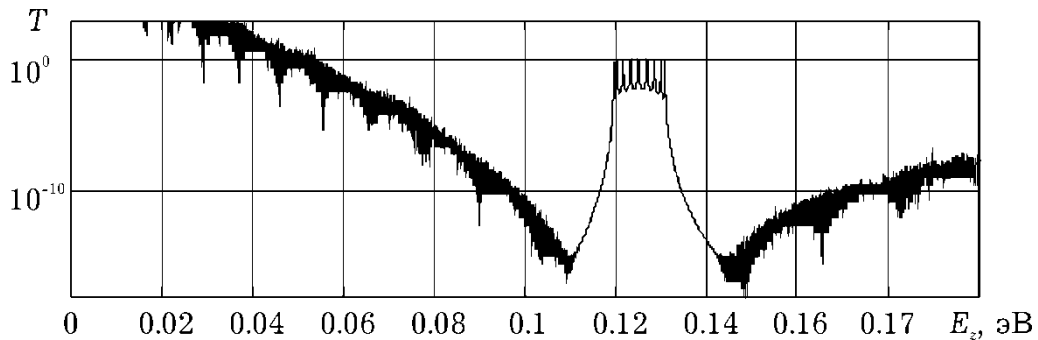


Рис. 7.13. Коэффициент прохождения десятибарьерной квантовой системы, рассчитанный методом матриц передачи. При энергии менее 0,5 эВ коэффициент передачи из-за численной неустойчивости рассчитывается неправильно и поэтому превышает единицу

Даже если процедура умножения матриц не приводит к появлению плохо обусловленной матрицы, она все же занимает неоправданно много вычислительного времени, если количество слоев достаточно большое. Это будет иметь место, если рельеф сильно нелинеен, что наблюдается, например, при учете пространственного заряда. Следовательно, возникает потребность в создании быстрых и стабильных методов нахождения коэффициента передачи. Этого можно достичь, используя численные методы интегрирования уравнения Шредингера.

7.3. ЧИСЛЕННЫЕ МОДЕЛИ

Количественное моделирование устройств с поперечным электронным транспортом возможно лишь с использованием численных методов. Предположения, которые принимаются при численном моделировании, аналогичны описанным в разд. 7.1. Взаимодействие между электронами учитывается в приближении Хартри, то есть считается, что электроны двигаются в поле

кристаллической решетки и потенциала, который создают все другие электроны (так называемое "самосогласованное поле").

Построение численной модели включает шаги, которые направлены на нахождение таких величин:

- концентрация электронов $n(z)$;
- самосогласованный потенциал $V_s(z)$;
- локальная плотность электронных состояний $g(E_z, z)$;
- локальная функция распределения $N(E_z, z)$;
- коэффициент передачи за поперечной энергией $T(E_z)$;
- плотность тока через резонансно-тоннельный диод $J(V)$.

7.3.1. Концентрация и волновые функции

Исходным положением для нахождения концентрации электронов является разделение области моделирования на квантовую и классическую области, то есть "устройство" и "резервуары" (рис. 7.14, а).

К квантовой области принадлежат ДБКС и спейсери. К классическим областям относят участки высоколегированных областей такой длины, которая необходима для того, чтобы на ее краях концентрация и потенциал выравнивались к своим равновесным значениям с нужной точностью. На практике такие участки выбираются длиной в десятки нанометров.

На рис. 7.14, б квантовой области соответствует область $z \in [z_0; z_5]$, где $z_0 = (i_0 - 1)\Delta$, $z_5 = (i_5 - 1)\Delta$, Δ – шаг дискретной сетки. Области $z \in [0; z_0]$ и $z \in [z_5; L]$ относятся к левому и правому резервуару соответственно.

Считается, что электронный газ во всем левом резервуаре подчиняется статистике Ферми-Дирака с параметром:

$$E_{\Phi}^L = U(0) + E_{\Phi} = U_1 + E_{\Phi},$$

в правом:

$$E_{\Phi}^R = U(L) + E_{\Phi} = U_N + E_{\Phi},$$

где $U = E_c + U_s$, то есть сумма потенциальной энергии электрона в кристаллической решетке E_c и самосогласованного потенциала $U_s = -eV_s$; L – длина участка, который моделируется; $U_i = (i-1)\Delta$; E_{Φ} – приведенный к дну зоны проводимости уровень Ферме

электронов в высоколегированных областях. Он определяется как корень уравнения

$$N_D^+ = 4\pi(2m_c^* / h^2)^{3/2} \int_0^\infty \frac{\sqrt{E}}{1 + \exp\left(\frac{E - E_\Phi}{k_B T}\right)} dE,$$

где N_D^+ – концентрация ионизированных примесей.

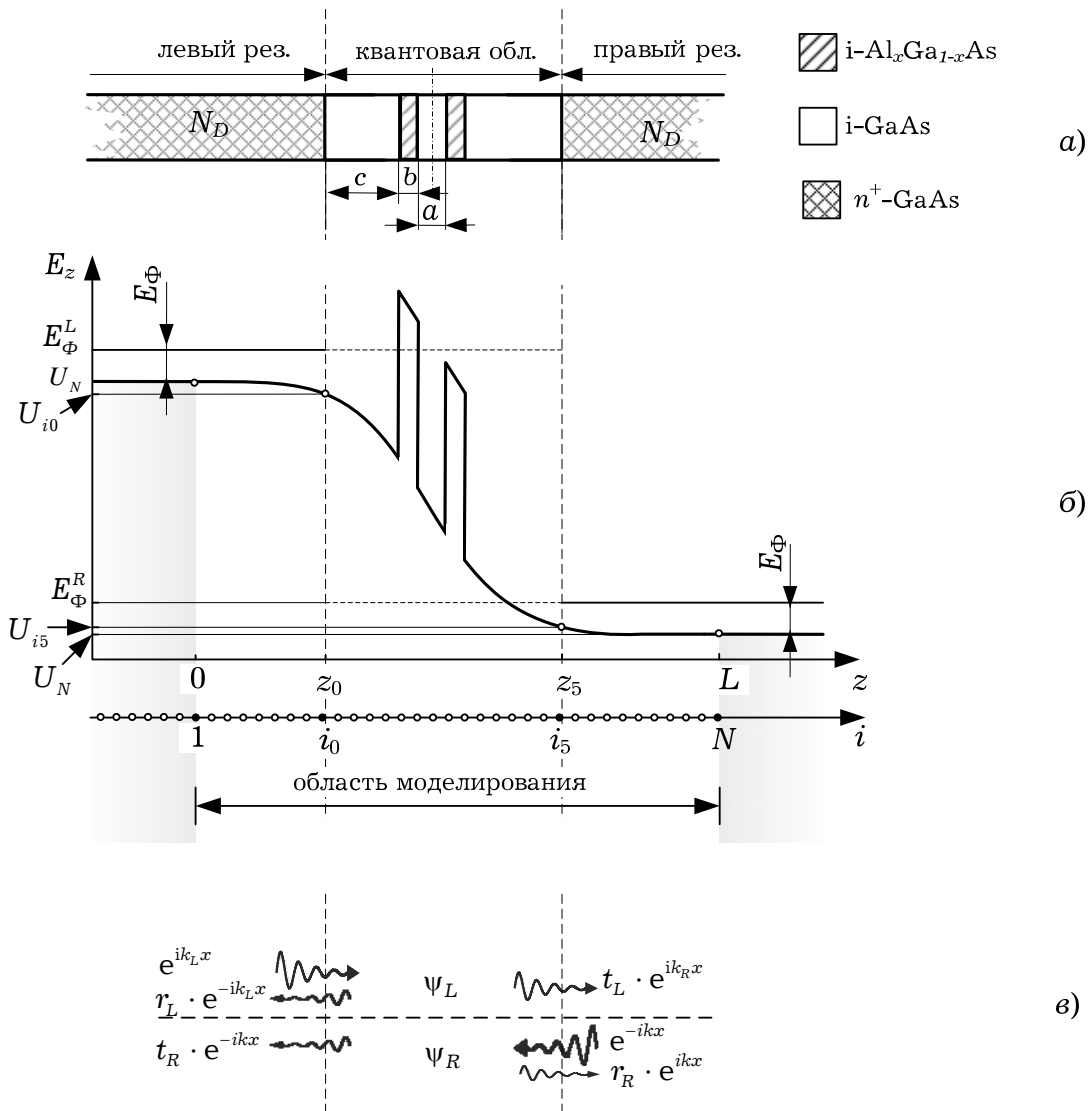


Рис. 7.14. Предположение и обозначения, принятые при моделировании: а) разделение на 3 области; б) дискретизация и условные обозначения; в) волновые функции в области моделирования

Профиль дна зоны проводимости в классических областях $U(z)$ находится из уравнения Пуассона для всего участка, кото-

рый моделируется. Он определяет плотность состояний трехмерного электронного газа для классических участков. В произвольной точке i (которой соответствует координата $z_i = \Delta(i - 1)$) левой или правой классической области:

$$g_i(E) = 4\pi(2m_i^* / h^2)^{3/2}(E - U_i);$$

функция Ферми для левого (правого) резервуара:

$$f_{L(R)}(E) = \frac{1}{1 + \exp\left(\frac{E - E_{\Phi}^{L(R)}}{k_B T}\right)} = \frac{1}{1 + \exp\left(\frac{E - (E_{\Phi} + U_{1(N)})}{k_B T}\right)}.$$

Таким образом, концентрацию в i -й точке левого (правого) резервуара можно найти по формуле:

$$n_i = 4\pi(2m^* / h^2)^{3/2} \int_{U_i}^{\infty} \frac{\sqrt{E - U_i}}{1 + \exp\left(\frac{E - (E_{\Phi} + U_{1(N)})}{k_B T}\right)} dE. \quad (7.42)$$

Считая, что электроны попадают в квантовую область либо из правого, либо из левого резервуаров и между собой не взаимодействуют и, руководствуясь теми же суждениями, что и в разд. 7.1.2, легко убедиться, что в общем виде концентрацию электронов в неоднородной области можно определить, взяв тройной интеграл:

$$n = \int_{-\infty}^{\infty} |\psi(\mathbf{k})|^2 N(\mathbf{k}) dk_x dk_y dk_z, \quad (7.43)$$

который в случае одномерного квантового транспорта может быть сведен к двум обычным интегралам по поперечной энергии:

$$n = n_R + n_L. \quad (7.44)$$

При переходе от (7.43) к (7.44) считается, что:

Вероятность заполнения электронных состояний, то есть функция распределения электронов по энергиям, совпадает с равновесной в соответствующем резервуаре, умноженной на соответствующую вероятность $|\psi(\mathbf{k})|^2$. Напомним, что для классических областей выше было принято предположение, что

электроны в резервуарах подчинены равновесной статистике Ферми.

Соответствующие уровни Ферми в квантовых областях обозначены на рис. 7.14, б пунктирными линиями, которые становятся сплошными в классических областях, где энергия Ферми имеет непосредственное физическое содержание.

Плотность состояний такая же, как и в граничащих с квантовой областью точках. Последние суть $z_0 = (i_0-1)\Delta$ для электронов, которые инжектируются из левого резервуара и $z_5 = (i_5-1)\Delta$ для электронов, которые инжектируются из правого резервуара. Это одно из наименее обоснованных предположений, принятых в модели.

Принятые для квантовых областей начала отсчета функции плотности состояний (то есть энергии, при которых плотность состояний равняется 0) обозначены на рис. 7.14, б пунктирными линиями. В классических областях они обозначены сплошной линией; понятие "плотность состояний" там имеет непосредственное содержание, чего нельзя сказать о квантовой области.

Слагаемые формулы (7.44) получены, исходя из следующих суждений.

Интегрирование проводится за всеми значениями проекций k_x и k_y , и по положительной (отрицательной) составляющей k_z для "левых" ("правых") электронов:

$$n_{L(R)} = \int_{0(-\infty)}^{\infty(0)} |\Psi_{L(R)}(E_z(k_z))|^2 dk_z \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} N(\mathbf{k}) dk_x dk_y \quad (7.45)$$

В формуле (7.45) $N(\mathbf{k}) = g(\mathbf{k})f(\mathbf{k})$, где $g(\mathbf{k}) = 2/(2\pi)^3$ – плотность k -состояний для трехмерного кристалла; $f(\mathbf{k}, k_\Phi)$ – функция Ферми-Дирака, которая определяется так:

$$f(\mathbf{k}, k_\Phi) = \frac{1}{1 + \exp\left(\frac{E(\mathbf{k}) - E_\Phi^{L(R)}(k_\Phi)}{kT}\right)} = \begin{cases} \frac{1}{1 + \exp\left(\left\{\frac{\hbar^2}{2m^*}(k_x^2 + k_y^2 + k_z^2) + U_{i_0} - \left(U_1 + \frac{\hbar^2}{2m^*}k_\Phi^2\right)\right\} / kT\right)}, & k_z > 0 \\ \frac{1}{1 + \exp\left(\left\{\frac{\hbar^2}{2m^*}(k_x^2 + k_y^2 + k_z^2) + U_{i_5} - \left(U_N + \frac{\hbar^2}{2m^*}k_\Phi^2\right)\right\} / kT\right)}, & k_z < 0 \end{cases}$$

Последнее равенство обусловлено тем, что закон дисперсии для электронов, которые инжектируются из левого (правого) резервуара¹:

$$E(\mathbf{k}) = \begin{cases} \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2) + U_{i0} = E_x + E_y + E_z, k_z > 0 \text{ (из правого)}, \\ \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2) + U_{i5} = E_x + E_y + E_z, k_z > 0 \text{ (из левого)} \end{cases}$$

Считается, что $E_z = \frac{\hbar^2 k_z^2}{2m^*} + U_{i0(i5)}$, то есть поля, в которых находится электрон, изменяют лишь z -составляющую его энергии².

Уровень Ферми в левом (рис. 7.14, б):

$$E_{\Phi}^{L(R)} = U_{1(N)} + \frac{\hbar^2 k_{\Phi}^2}{2m^*} = U_{1(N)} + E_{\Phi}.$$

Подставляя эти значения $g(\mathbf{k})$ и $f(\mathbf{k}, k_{\Phi})$ в (7.45) и производя интегрирование по двум поперечным проекциям \mathbf{k} , получим формулы для $n_{L(R)}$:

$$n_{L(R)} = N_c' \int_{U_{i0}(U_{i5})}^{\infty} \frac{|\Psi_{L(R)}|^2}{\sqrt{E_z - U_{i0(i5)}}} \ln \left(1 + e^{-\frac{E_z - (E_{\Phi} + U_{1(N)})}{k_B T}} \right) dE_z \quad (7.46)$$

где $N_c' = \frac{\sqrt{2}(m^*)^{3/2} k_B T}{(2\pi)^2 \hbar^3}$.

Для получения формулы (7.46) использовался метод интегрирования, аналогичный описанному в приложении П7.1.

Подынтегральное выражение (7.46) имеет смысл полной функции распределения электронов по z -компоненте энергии в точке z :

¹ Этим мы приравниваем плотность состояний в квантовой области к плотности состояний в точке i_0 . Понятно, что в квантовой области уже нельзя говорить о плотности электронных состояний, поскольку это противоречит возможности туннелирования. Наиболее естественно предположить, что полная энергетическая функция распределения электронов $N(E)$ в квантовой области совпадает с $N(E)$ в крайней точке i_0 (где электроны еще считаются распределенными равномерно), умноженной на соответствующую вероятность.

² Такое предположение принималось с тем, чтобы обеспечить единое начало отсчета энергии для левого и правого резервуара (в численных моделях $|\psi_L(E_z(k_z), z)|^2$ и $|\psi_R(E_z(k_z), z)|^2$ рассчитываются в одном цикле)

$$N(E_z, z) = N_c' \frac{|\Psi_{L(R)}|^2}{\sqrt{E_z - U_{i_0(i_5)}}} \ln \left(1 + \exp \left(-\frac{E_z - (E_\Phi + U_{1(N)})}{k_B T} \right) \right), \quad (7.47)$$

Функция $N(E_z, z)$ зависит от распределения электронов в резервуарах. Следовательно, из-за экспоненциального уменьшения населенности резервуаров информация о распределении электронов на высших уровнях в квантовой яме теряется. Чтобы получить информацию обо всех уровнях энергии в квантовой яме, а не только о заселенных, удобно пользоваться лишь частью функции $N(Ez, z)$:

$$g(E_z, z) = N_c' \frac{|\Psi_{L(R)}|^2}{\sqrt{E_z - U_{i_0(i_5)}}}, \quad (7.48)$$

которая имеет смысл, аналогичный функции плотности состояний в статистической физике, но зависит от координаты, поэтому $g(E_z, z)$ будем называть локальной плотностью состояний.

В формулу для концентрации электронов (7.46) входят пока что неизвестные волновые функции электронов из правого (левого) резервуара $\Psi_{L(R)}(E_z(k_z), z)$. В численных моделях возникает необходимость дискретизации уравнения Шредингера

$$\frac{d}{dz} \frac{1}{m^*(z)} \frac{d\Psi_{L(R)}(z)}{dz} + \frac{2}{\hbar^2} (E - U(z)) \Psi_{L(R)}(z) = 0 \quad (7.49)$$

Консервативная конечно-разностная схема, полученная *интегро-интерполяционным методом* [7.24 с. 167; 7.25 с. 553] для внутренних точек сетки имеет вид (см. приложение П7.2):

Если точка i не лежит на границе двух материалов, $i \notin \{1, i_1, i_2, \dots, i_5, N\}$:

$$\Psi_{i+1} - 2 \left(1 - \frac{\Delta^2 m \cdot m_i}{\hbar^2} (E_z - U_i) \right) \Psi_i + \Psi_{i-1} = 0, \quad (7.50)$$

где m_i — относительная эффективная масса, m — масса покоя электрона;

Если точка i совпадает из гетерограницей, $i \in \{i_1, i_2, \dots, i_5\}$:

$$\psi_{\Gamma,i+1} + \frac{m_{i+1}}{m_{i-1}} \psi_{\Gamma,i-1} - \left(1 + \frac{m_{i+1}}{m_{i-1}} - \frac{2\Delta^2 m_{i+1} m}{\hbar^2} (E_z - V_i) \right) \psi_i = 0; \quad (7.51)$$

Для внешних точек, пользуясь общими требованиями к волновой функции и ее производной, можно получить следующие уравнения, которые замыкают конечно-разностные схемы для ψ_L та ψ_R (см. приложение П7.3):

- для ψ_L слева:

$$(ik_L\Delta - 1)\psi_1 + \psi_2 = 2ik_L\Delta; \quad (7.52)$$

- для ψ_L справа:

$$(ik_R\Delta - 1)\psi_N + \psi_{N-1} = 0; \quad (7.53)$$

- для ψ_R слева:

$$(ik_L\Delta - 1)\psi_1 + \psi_2 = 0; \quad (7.54)$$

- для ψ_R справа:

$$(ik_R\Delta - 1)\psi_N + \psi_{N-1} = 2ik_R\Delta, \quad (7.55)$$

где $k_{L(R)} = \sqrt{2m^*(E_z - U_{i1(i5)})} / \hbar$; причем в уравнениях (7.52), (7.53) под ψ следует понимать ψ_L , а в (7.54), (7.55) – ψ_R .

Уравнения (7.52)–(7.55) были получены, считая, что граница между квантовой областью и резервуарами не совпадает с гетерограницей. Метод получения этих уравнений называется методом квантовой передачи через границу (англ. "Quantum transmitting boundary method" [7.26]).

Уравнение (7.50)–(7.55) формируют конечно-разностные схемы уравнений Шредингера для ψ_L и ψ_R для всех N_K точек, на которые разбита активная область $[z_0; z_5]$. Для их решения используется метод Томаса [7.27 с. 227].

7.3.2. Влияние пространственного заряда

Учесть межэлектронное взаимодействие, не выходя за пределы одноэлектронного приближения, можно благодаря предположению о том, что каждый электрон двигается под

действием поля, создаваемого всеми остальными электронами и ионами. Такое приближение называется приближением Хартри [7.2 стр. 311; 7.11 стр. 71]. Потенциал этого поля – так называемый самосогласованный потенциал V_s – можно найти из уравнения Пуассона:

$$\frac{d}{dz} \varepsilon(z) \frac{dV_s}{dz} = \frac{e}{\varepsilon_0} [n(z) - N_d^+(z)], \quad (7.56)$$

где ε_0 – диэлектрическая постоянная;

ε – относительная диэлектрическая проницаемость;

N_d^+ – концентрация ионизированных доноров.

Величины V_s и $n(z) = n_L(z) + n_R(z)$ априори неизвестны. Однако известно, что $n_{L(R)}(z)$ определяются формулами (7.46) и является функциями V_s :

$$U(z) = E_c(z) + U_s(z),$$

где $U_s(z) = -eV_s$ – потенциальная энергия электрона в поле кристаллической решетки и пространственного заряда, величина, которую мы до этого времени считали данной.

Таким образом, (7.44) и (7.56) представляют собой систему интегро-дифференциальных уравнений, которую необходимо решить, чтобы найти самосогласованный потенциал. Выпишем эту систему в таком виде, чтобы была ясна формально-математическая сторона вопроса:

$$\frac{d}{dz} \varepsilon(z) \frac{dV_s}{dz} = \frac{e}{\varepsilon_0} [n(z) - N_d^+(z)], \quad (7.57)$$

$$n(z) = \begin{cases} \sum_{i=L,R} \int |\psi_i(E_z(k_z), V_s, z)|^2 f_i(E_z) dE_z, z \in [z_1, z_5] \\ 4\pi(2m^*/h^2)^{3/2} \int_{U_i}^{\infty} \frac{\sqrt{E - U(z)}}{1 + e^{\frac{E - (E_\Phi + U_{1(N)})}{k_B T}}} dE, z \notin [z_1, z_5] \end{cases}, \quad (7.58)$$

$$\text{где } f_i(E_z) \equiv \underbrace{\frac{\sqrt{2}(m^*)^{3/2} k_B T}{(2\pi)^2 \hbar^3}}_{N_c} \frac{\ln \left(1 + \exp \left(-\frac{E_z - (E_\Phi + U_{1(N)})}{k_B T} \right) \right)}{\sqrt{E_z - U_{i0(i5)}}}.$$

Поскольку (7.58) зависит от V_s преимущественно в части волновых функций $\psi_{L(R)}$, которые, в свою очередь, являются решениями уравнения Шредингера, (7.57)–(7.58) называют самосогласованной системой Шредингера-Пуассона.

Метод решения системы Шредингера-Пуассона заключается в следующем. Вначале задают начальное приближение для самосогласованного потенциала $V_s(z)$. В качестве начального приближения $V_s^{old}(z)$ лучше всего выбрать линейное падение потенциала в квантовой области, а в резервуарах считать потенциал неизменным, то есть:

$$V_s^{old}(z) = \begin{cases} 0, & z < z_1, \\ \frac{V}{L}z, & z \in [z_1, z_5] \\ V, & z > z_5, \end{cases}$$

Это начальное приближение используется для поиска концентрации $n^{old}(z)$ по всей длине модельной структуры согласно (7.58). Подставляя найденную функцию $n^{old}(z)$ в уравнение Пуассона (7.57), можно найти следующее приближение к самосогласованному потенциалу, $V_s^{new}(z)$.

Самый простой способ найти самосогласованный потенциал с абсолютной точностью ε_V – это организовать итерационный процесс так, чтобы найденное $V_s^{new}(z)$ использовалось в качестве следующего приближения для поиска концентрации $n^{new}(z)$ и так далее. Это – так называемый метод простых итераций. Однако практика показывает, что условие $\max |V_s^{new}(z) - V_s^{old}(z)| < \varepsilon_V$ не выполняется ни при каком количестве итераций, то есть итерационный процесс не сходится.

Поэтому альтернативно используются два метода [7.28]:

Релаксационный алгоритм. Он состоит в том, что для поиска $n^{new}(z)$ используется не непосредственно новое приближение V_s^{new} , а величина, определяемая по формуле:

$$V_s^{new'} = V_s^{old} + \alpha(V_s^{new} - V_s^{old}), \quad (7.59)$$

которая сводится к методу простых итераций, если $\alpha = 1$. Однако при $\alpha \ll 1$ (7.59) избавляется от неустойчивости, свойственной предыдущему методу. Типичные значения α составляют 0,01 и

меньше; в этих условиях метод обычно уверенно, однако очень медленно сходится.

Метод Гумеля. Идея ускоренных методов самосогласования основывается на учете характера зависимости концентрации от самосогласованного потенциала. Если считать, что электронный газ в области моделирования невырожден и находится в локальном равновесии, то концентрацию можно по формуле [7.29 стр. 125]:

$$n(z) = N_c \exp\left(\frac{E_\Phi - E_c(z) + eV_s(z)}{k_B T}\right) = n_0 \exp\left(\frac{eV_s}{k_B T}\right).$$

$$\text{Для } V_s^{old} \text{ получим: } n^{old}(z) = n_0 \exp\left(\frac{eV_s^{old}}{k_B T}\right).$$

$$\text{Для } V_s^{new}: n^{new}(z) = n_0 \exp\left(\frac{eV_s^{new}}{k_B T}\right).$$

Если имело бы место локальное равновесие и невырожденность, новые и старые приближения связывались бы между собой соотношением:

$$n^{new}(z) = n^{old}(z) \exp\left(\frac{e(V_s^{new} - V_s^{old})}{k_B T}\right), \quad (7.60)$$

что непосредственно следует из двух предыдущих формул. В случае, когда указанные условия не выполняются, (7.60) является хорошим приближением для концентрации n^{old} , поскольку учитывает ее нелинейную зависимость от потенциала. Метод Гумеля [7.30] заключается в том, чтобы находить новые приближения для потенциала из уравнения:

$$\frac{d}{dz} \varepsilon(z) \frac{dV_s^{new}}{dz} = \frac{e}{\varepsilon_0} \left[n^{old}(z) \exp\left(\frac{V_s^{new} - V_s^{old}}{V^{ref}}\right) - N_D^+(z) \right], \quad (7.61)$$

где $V^{ref} = k_B T / e$ называют опорным потенциалом (англ. "reference potential"), он является по существу тепловым потенциалом.

Это уравнение нелинейно относительно V_s^{new} , поэтому конечно-разностная схема, которая ему соответствует, будет

состоять из нелинейных уравнений, которые можно решить, например, методом Ньютона.

С точки зрения экономии численных ресурсов более привлекательным является линеаризованный метод Гумеля, который можно получить из (7.61), разложив $\exp\left(\frac{e(V_s^{new} - V_s^{old})}{k_B T}\right)$ в ряд Маклорена и ограничившись первыми двумя членами: $\exp(x) \approx 1 + x$. Результат имеет вид:

$$\frac{d}{dz} \varepsilon(z) \frac{dV_s^{new}}{dz} - n^{old}(z) \frac{e}{\varepsilon_0} \frac{V_s^{new}}{V^{ref}} = \frac{e}{\varepsilon_0} \left[n^{old}(z) \left(1 - \frac{V_s^{old}}{V^{ref}} \right) - N_D^+(z) \right] \quad (7.62)$$

Уравнению (7.62), в отличие от (7.61), будет соответствовать конечно-разностная схема, которая состоит из линейных алгебраических уравнений. Способ их получения полностью аналогичен ранее применявшемуся к уравнению Шредингера (7.49) (см. приложение П7.2), поскольку математически (7.49) и (7.62) суть одно и то же уравнение.

Запишем конечно-разностную схему, которая соответствует уравнению (7.62) в i -й точке дискретной сетки, в виде:

$$a_i V_{i-1} + b_i V_i + c_i V_{i+1} = d_i, \quad (7.63)$$

(здесь и далее чтобы не усложнять восприятие, вместо V_s употребляется просто V). Для коэффициентов уравнений (7.63) можно получить следующие формулы (II порядок точности):

Для всех внутренних точек, которые не являются гетерограницами $i \notin \{1, i_1, i_2, \dots, i_5, N\}$:

$$a_i = c_i = 1; \quad (7.64)$$

$$b_i = -2 - \frac{1}{\varepsilon_i} \Delta^2 \frac{e n_i^{old}}{\varepsilon_0 V^{ref}}; \quad (7.65)$$

$$d_i = \frac{\Delta^2}{\varepsilon_i} \frac{e}{\varepsilon_0} \left[n_i^{old} \left(1 - \frac{V_i^{old}}{V_i^{ref}} \right) - N_i^+ \right]. \quad (7.66)$$

Для всех гетерограниц $i \in \{1, i_1, i_2, \dots, i_5, N\}$:

$$a_i = 1; \quad (7.67)$$

$$b_i = -1 - \frac{\varepsilon_{i-1}}{\varepsilon_{i+1}} - \frac{1}{\varepsilon_{i+1}} \Delta^2 \frac{e n_i^{old}}{\varepsilon_0 V^{ref}}; \quad (7.68)$$

$$c_i = \frac{\varepsilon_{i-1}}{\varepsilon_{i+1}}; \quad (7.69)$$

$$d_i = \frac{\Delta^2}{\varepsilon_{i+1}} \frac{e}{\varepsilon_0} \left[n_i^{old} \left(1 - \frac{V_i^{old}}{V_i^{ref}} \right) - N_i^+ \right]. \quad (7.70)$$

Что касается внешних точек, то они выбираются настолько далеко от активной части устройства, что электрическое поле в них отсутствует: $dV_s/dz = 0$, откуда получаем условия Дирихле:

$$V_1 = 0, V_N = V. \quad (7.71)$$

Уравнения (7.63)–(7.71) формируют конечно-разностную схему, которая соответствует уравнению (7.62). Самосогла-сованный потенциал считается найденным с точностью ε_V , если выполняется условие $\max |V_s^{new}(z) - V_s^{old}(z)| < \varepsilon_V$.

Распределение концентрации и потенциального рельефа в модельном диоде (рис. 7.6), полученные в результате самосогла-сованного решения системы Шредингера-Пуассона, приведено на рис. 7.15. Из рисунка видно, что с увеличением напряжения накопления заряда в области левого спейсера и квантовой ямы увеличивается. Также понятным является влияние перераспределения заряда на рассчитанные характеристики: весь потенциальный рельеф квантовой области "поднимается" по оси энергии, чем замедляет достижение пикового напряжения.

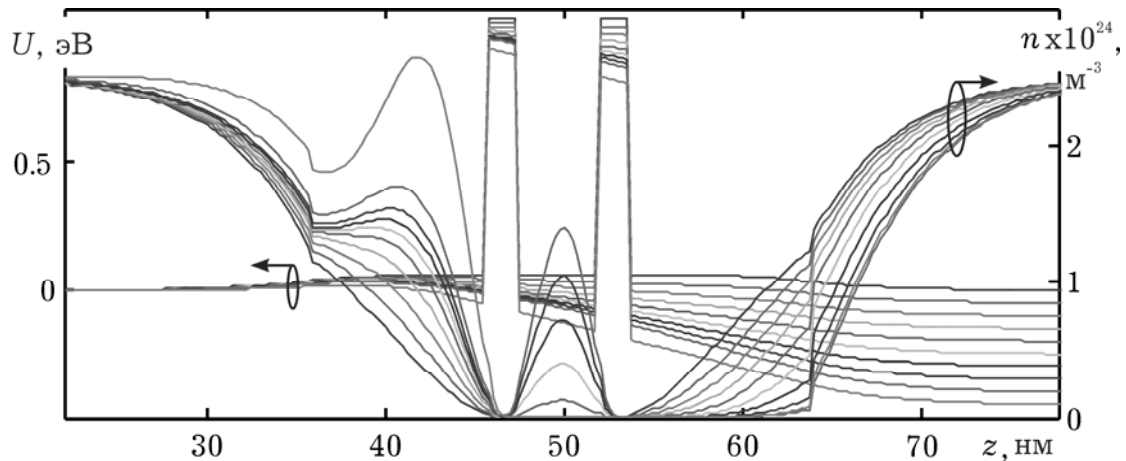


Рис. 7.15. Распределения концентрации и потенциальной энергии электронов в модельном резонансно-тоннельном диоде (рис. 7.6), рассчитанные самосогла-сованно для напряжений от 0 до 0,45 В

Выдающейся чертой численного моделирования является возможность объяснения природы квантово-размерных эффектов

в активной части РТД с помощью анализа распределения концентрации $n(z)$, локальной плотности состояний $g(E_z, z)$ и функции распределения $N(E_z, z)$. На рис. 7.16, а плотность состояний как функция двух аргументов изображена с помощью градаций серого. Темные горизонтальные полосы на характеристике $g(E_z, z)$ соответствуют метастабильным уровням энергии в квантовой яме; характерная интерференционная картина распределения плотности состояний в спейсерах указывает на проявление волновой природы электрона.

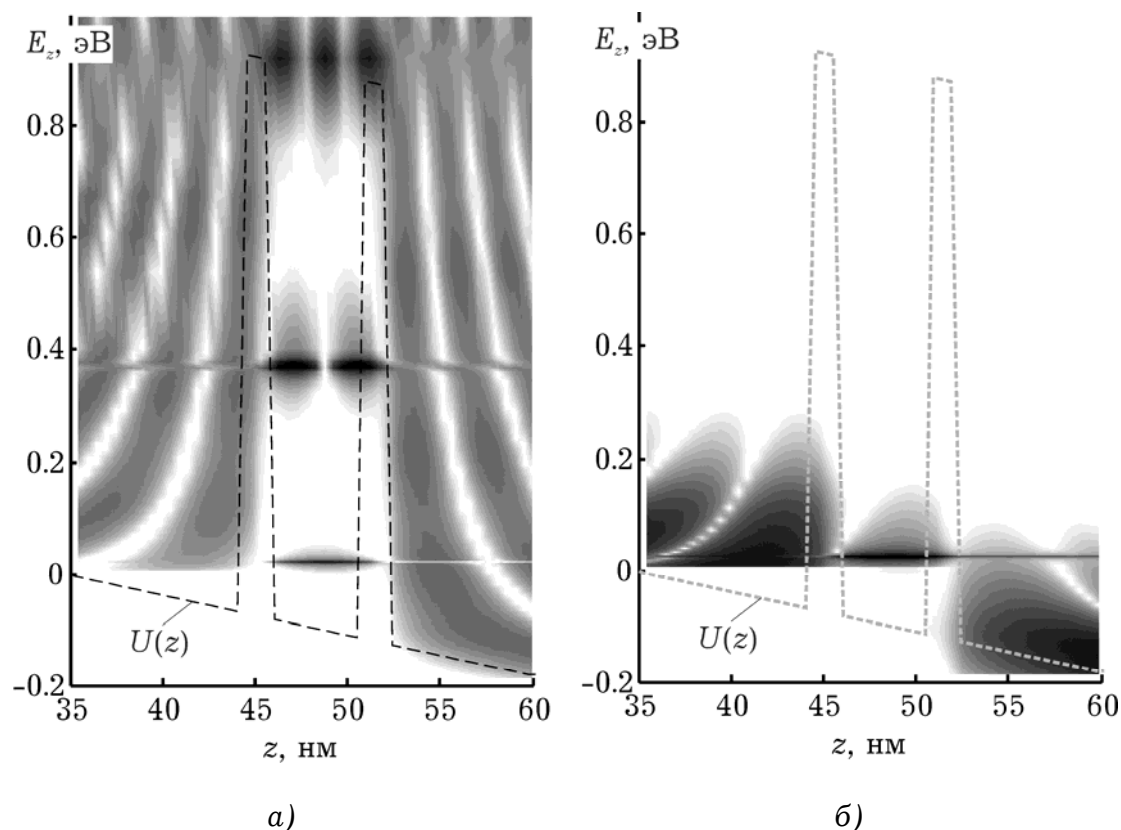


Рис. 7.16. Локальная плотность состояний (а) и функция распределения электронов в квантовой области РТД (б). Приложенное напряжение $V = 0.2$ В. Более темному тону соответствует большее значение функций в логарифмическом масштабе. Штриховой линией обозначен потенциальный рельеф

Плотность состояний на первом метастабильном уровне имеет максимум по центру квантовой ямы, на втором – два симметричных максимума, на третьем уровне, который частично находится над потенциальными барьерами, – три. Природу такого распределения $g(z)$ на i -ом уровне в яме можно объяснить в рамках теории возмущений. Действительно, количество максимумов $g(z)$ на i -ом уровне должно быть таким же, как у $|\psi_i(z)|^2$ для невозмущенной системы, то есть равняться номеру

уровня. Наблюдается также увеличение величины расширения уровней с ростом энергии, которое объясняется уменьшением времени жизни электрона на высших уровнях.

Из рис. 7.16, б очевидно, что накопление заряда в яме происходит лишь на самом низком уровне. Этот факт объясняет, почему концентрация при малых смещениях имеет максимум в центре квантовой ямы (рис. 7.15).

7.3.3. Обобщенная формула Цу-Эсаки

Плотность тока через РТД может быть рассчитана в соответствии с обычными правилами расчета средних значений физических величин в статистической физике:

$$\mathbf{J} = e \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{v}(\mathbf{k}, z) N(\mathbf{k}) dk_x dk_y dk_z, \quad (7.72)$$

если считать, что функция распределения электронов в квантовой области

$$N(\mathbf{k}, z) \equiv \begin{cases} N_L(\mathbf{k}) \cdot |\psi_L(\mathbf{k}, z)|^2, & \text{если } k_z > 0, \\ N_R(\mathbf{k}) \cdot |\psi_R(\mathbf{k}, z)|^2, & \text{если } k_z < 0; \end{cases} \quad (7.73)$$

$\mathbf{v}(\mathbf{k}, z)$ – средняя скорость электрона в точке z , которая определяется как плотность потока вероятности, деленная на саму плотность вероятности [7.3 с. 123]:

$$\mathbf{v}(\mathbf{k}, z) = \frac{\mathbf{j}(\mathbf{k}, z)}{|\psi(\mathbf{k}, z)|^2} = \frac{i\hbar}{2m^*} (\psi \nabla \psi^* - \psi^* \nabla \psi) / |\psi(\mathbf{k}, z)|^2. \quad (7.74)$$

С учетом того, что ток протекает вдоль оси $0z$, для разных знаков z -проекции волнового вектора единственная ненулевая проекция тока $j_z(k_z, z) \equiv j$ запишется так:

$$j(k_z) \equiv \begin{cases} j_{12} = \frac{i\hbar}{2m^*} \left(\psi_L \frac{d\psi_L^*}{dz} - \psi_L^* \frac{d\psi_L}{dz} \right), & \text{если } k_z > 0; \\ j_{21} = \frac{i\hbar}{2m^*} \left(\psi_R \frac{d\psi_R^*}{dz} - \psi_R^* \frac{d\psi_R}{dz} \right), & \text{если } k_z < 0; \end{cases} \quad (7.75)$$

Учитывая (7.73) и (7.74), придадим (7.72) вид:

$$J(z) = J_{12} + J_{21}, \quad (7.76)$$

где $J_{12(21)} \equiv e \int_{0(-\infty)}^{\infty(0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} j_{12(21)}(k_z, z) N_{L(R)}(k) dk_z dk_x dk_y$.

Проведя интегрирование по продольным проекциям волновых чисел k_x та k_y , а также заменив интегрирование по k_z на интегрирование по E_z , получим:

$$J_{12} = \frac{\sqrt{2}(m^*)^{3/2} e k_B T}{(2\pi)^2 \hbar^3} \int_{U_{i0}}^{\infty} \frac{j_{12}(E_z(k_z), z)}{\sqrt{E_z - U_{i0}}} \ln \left(1 + e^{\frac{E_z - (E_{\Phi} + U_1)}{k_B T}} \right) dE_z, \quad (7.77)$$

$$J_{21} = \frac{\sqrt{2}(m^*)^{3/2} e k_B T}{(2\pi)^2 \hbar^3} \int_{U_{i5}}^{\infty} \frac{j_{21}(E_z(k_z), z)}{\sqrt{E_z - U_{i5}}} \ln \left(1 + e^{\frac{E_z - (E_{\Phi} + U_N)}{k_B T}} \right) dE_z. \quad (7.78)$$

В силу выполнения закона сохранения заряда в квантовой механике

$$\frac{\partial e |\psi|^2}{\partial t} + \text{div}(\mathbf{j}(k_z, z)) = 0,$$

выполняется также и закон сохранения заряда в классической электродинамике, потому что он есть ничто другое как усреднения первого по всему ансамблю:

$$\frac{\partial q}{\partial t} + \text{div} \mathbf{J} = 0. \quad (7.79)$$

Для стационарных процессов и транспорта вдоль оси Oz уравнение (7.79) сводится к уравнению $dJ/dz = 0$. Этим подтверждается, что (7.77) и (7.78) обеспечивают выполнение законов сохранения в любой точке структуры.

Покажем, как из (7.77) и (7.78) следует обычная формула Цу-Эсаки. Для этого найдем j_{12} в точке $z = z_5$ и j_{21} в точке $z = z_0$. Отметим, что для $z \geq z_5$ $\psi_L = t_L \exp(ik_R z)$, а для $z \leq z_1$ $\psi_R = t_R \exp(ik_L z)$. Из (7.75) имеем:

$$j_{12}(z_5) = \frac{\hbar k_R}{m^*} |t_L|^2, \quad (7.80)$$

$$j_{21}(z_0) = -\frac{\hbar k_L}{m^*} |t_R|^2. \quad (7.81)$$

Поскольку в (7.77) и (7.78) $1/\sqrt{E_z - U_{i0(5)}} = \sqrt{2m^*}/\hbar k_{L(R)}$, то с учетом (7.80) и (7.81):

$$\frac{j_{12(21)}}{\sqrt{E_z - U_{i0(i5)}}} = \pm \frac{\sqrt{2}}{\sqrt{m^*}} \frac{k_{R(L)}}{k_{L(R)}} |t_{L(R)}|^2 = \pm \frac{\sqrt{2}}{\sqrt{m^*}} T(E_z), \quad (7.82)$$

где $T(E_z) \equiv \frac{k_{L(R)}}{k_{R(L)}} |t_{R(L)}|^2$ – коэффициент прохождения.

Подставляя (7.82) в (7.77) и (7.78) и суммируя последние, получим формулу Цу-Эсаки :

$$J = \frac{2m^* e k_B T}{(2\pi)^2 \hbar^3} \int_{\max(U_{i5}, U_{i0})}^{\infty} T(E_z) D(E_z) dE_z,$$

где $T(E_z) \equiv \frac{k_{L(R)}}{k_{R(L)}} |t_{R(L)}|^2$ – коэффициент прохождения.

Для численного расчета лучше воспользоваться тем, что $|t_L(E_z)|^2 = |\Psi_N|^2$.

Подынтегральное выражение в формуле Цу-Эсаки имеет очень резкий резонансный характер за счет функции $T(E_z)$, поэтому интегрирование обычными численными методами крайне неэффективно. Следовательно, нужно использовать методы с адаптивным подбором шага, например алгоритм Симпсона (в *MatLab* реализуется функцией *quad*). График функции $T(E_z)$ в "адаптивных" точках изображен на рис. 7.17.

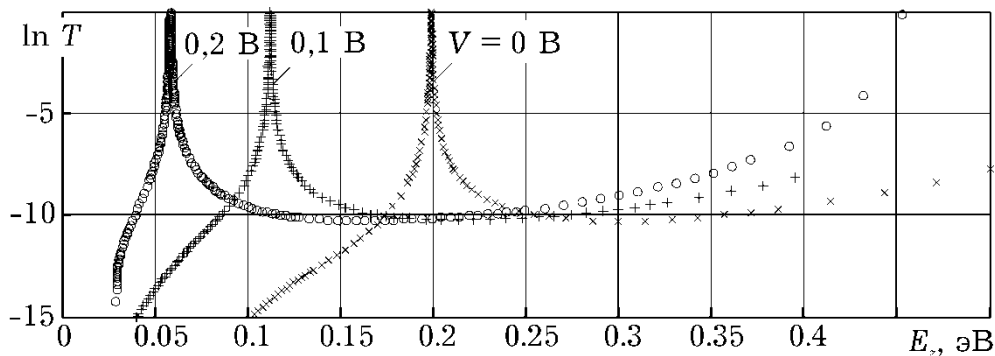


Рис. 7.17. Коэффициент прохождения через квантовую систему в релевантных для расчета ВАХ точках

Вольтамперная характеристика модельного РТД изображена на рис. 7.18. Ориентировочное время расчета 100 точек ВАХ в

MatLab без самосогласования составляет около 90 с, с самосогласованным – 3000 с.

Пользуясь формулой Цу-Эсаки, были рассчитаны ВАХ модельного РТД для случаев: а) напряжение на активной области падает линейно, б) потенциальный рельеф рассчитан самосогласованно. Результаты приведены на рис. 7.18.

Как видим, учет пространственного заряда приводит к увеличению пикового напряжения и уменьшению пикового тока. Это объясняется накоплением отрицательного пространственного заряда в спейсерах и квантовой яме, которая увеличивает потенциальную энергию электрона, тем самым "поднимая" все энергетические уровни в квантовой яме, а это смещает условие резонанса в сторону высших напряжений.

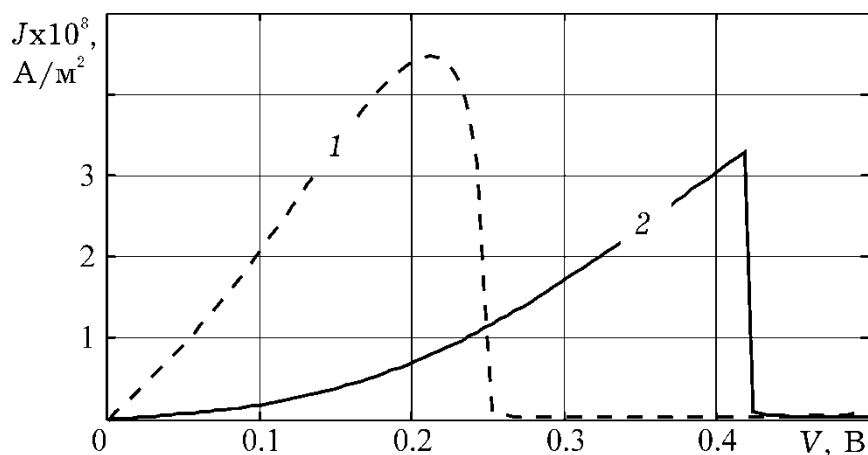


Рис. 7.18. Вольтамперные характеристики модельного РТД: 1 – рассчитана в предположении линейного падения напряжения, 2 – рассчитанная с привлечением метода Хартри

7.3.4. Рассеивание в квантовой яме РТД

Представление об электронном транспорте в квантовой области, описанные в текущем разделе, иногда классифицируются как "динамическая модель" [7.31; 7.32], чтобы отличать от "кинетических" [7.31; 7.33], то есть таких, что основываются и выводятся из положений квантовой кинетики. Действительно, в активной области рассматривается движение каждого отдельного электрона в поле, которое обусловлено полем самой кристаллической решетки и пространственного заряда. Это задача на движение отдельной частицы, и относится она к динамике, которая является разделом квантовой механики. При этом явления, которые обусловлены качественными отличиями между

поведением отдельного электрона и большого электронного ансамбля, что является предметом кинетики, не рассматриваются.

По очерченным выше причинам учет процессов рассеивания в разработанной модели является чем-то не совсем естественным (поскольку выходит за рамки динамики), за что такие модели поддаются справедливой критике [7.34; 7.35]. Однако концептуальная простота модели делает привлекательным учет рассеивания в феноменологический, и, можно сказать, немного неестественный способ. Один из таких подходов описан ниже.

Если небольшая часть электронов испытывает рассеивание за время пребывания в квантовой области, то можно разделить ток J , который протекает через указанную область, на ток электронов, которые не испытали рассеивания или "когерентную" составляющую J_{coh} , и ток электронов, которые такие столкновения испытали или "некогерентную" составляющую J_{incoh} :

$$J = J_{coh} + J_{incoh}. \quad (7.83)$$

Электрон, который испытал рассеивание, выходит из когерентного канала и пополняет последовательный канал. Но как отобразить в рамках используемых представлений процессы такого перехода?

Для ответа на этот вопрос вспомним, что эрмитовость³ оператора Гамильтона обеспечивает выполнение закона сохранения потока плотности вероятности и закона непрерывности заряда (7.79). Суть перехода из когерентного в последовательный канал заключается в том, что когерентная составляющая тока должна уменьшаться в направлении протекания тока. Если ток протекает в положительном направлении оси $0z$, то должно быть:

$$dJ_{coh}/dz < 0. \quad (7.84)$$

Учитывая (7.77) и (7.78), для выполнения (7.84) необходимо, чтобы было:

$$dj_{12}/dz < 0, \quad (7.85)$$

$$dj_{21}/dz > 0, \quad (7.86)$$

что возможно, только если оператор Гамильтона не является эрмитовым.

³ С точки зрения математики, квантовая механика — это алгебра самосопряженных («эрмитовских») операторов

Сделать из эрмитового оператора неэрмитов очень просто: достаточно ввести в него мнимую величину $-iW$. Тогда оператор Гамильтона из уравнения Шредингера (7.49) приобретет вид:

$$\hat{H} = -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m^*(z)} \frac{d}{dz} + U(z) - iW. \quad (7.87)$$

Для выяснения сути введенной величины $-iW$, которая называется *оптическим потенциалом*⁴, рассмотрим максимально простую задачу, а именно случай свободного электрона: $U(z) = U = \text{const}$. Поскольку любой потенциальный рельеф можно описать кусочно-постоянными функциями, то полученные результаты, можно считать универсальными и применимыми к произвольной форме $U(z)$.

Решением уравнения Шредингера в отмеченном случае является суперпозиция плоских волн:

$$\psi(z) = \underbrace{A \exp\{i\gamma z\}}_{\psi_{\rightarrow}} + \underbrace{B \exp\{-i\gamma z\}}_{\psi_{\leftarrow}}, \quad (7.88)$$

где $\pm\gamma \equiv \pm(k + i\alpha) = \frac{\sqrt{2m^*(E - U + iW)}}{\hbar}$, то есть k — это действительная часть волнового вектора; α — мнимая. Для электрона, который движется в положительном направлении оси $0z$, $\psi_{\rightarrow} = A \exp\{i\gamma z\}$, — волна, соответствующая падающему электрону, $\psi_{\leftarrow} = B \exp\{-i\gamma z\}$, — волна, соответствующая отраженному электрону.

Простые вычисления дают следующее выражение для плотности вероятности в падающей $\psi_{\rightarrow}(z)$ и отраженной $\psi_{\leftarrow}(z)$ волнах, а также в волновой функции $\psi(z)$:

$$j_{\rightarrow} = \frac{\hbar |A|^2}{m^*} e^{-2\alpha z}, \quad j_{\leftarrow} = \frac{\hbar |B|^2}{m^*} e^{+2\alpha z},$$

$$j = \frac{\hbar}{m^*} \left\{ |A|^2 e^{-2\alpha z} + |B|^2 e^{+2\alpha z} \right\}.$$

⁴ Такое название обусловлено областью науки, в которой был использован подобный прием для учета поглощения света; позже этот способ был использован Фешбахом, Портером и Вайскопфом при описании процессов протекания ядерных реакций. Он вполне аналогичный рассматриваемому здесь

Элементарный электрический ток, который создает падающий и отраженный электроны в направлении движения соответствующей частицы уменьшаются, о чем свидетельствуют знаки соответствующих производных:

$$\frac{dj_{\rightarrow}}{dz} = -2\alpha j_{\rightarrow}, \quad \frac{dj_{\leftarrow}}{dz} = +2\alpha j_{\leftarrow}$$

Таким образом, мы получили желаемый характер пространственной зависимости тока (см. (7.84)). Отметим, что на расстоянии $L = 1/(2\alpha)$ элементарный ток, создаваемый падающей и отраженной частицей, уменьшается в $e \approx 2.7$ раз. Это расстояние можно назвать длиной релаксации фазы.

Мнимая часть волнового вектора

$$\alpha(W) = \sqrt[4]{a^2 + b^2} \sin\left(\frac{\text{Arcsin}(b / \sqrt{a^2 + b^2})}{2}\right),$$

где $a = \sqrt{2m^*(E_z - U)} / \hbar$; $b = \sqrt{2m^*W} / \hbar$, монотонно возрастает с ростом оптического потенциала W , что следует из приведенной формулы.

Метод расчета величины W основывается на том, что ей сопоставляют некоторое характерное время (чаще всего – время релаксации импульса) в соответствие с формулой: $W = \hbar / 2\tau_p$.

Таким образом, с увеличением скорости рассеивания $v \equiv 1 / \tau_p$ оптический потенциал тоже растет, что уменьшает длину L , то есть увеличивает интенсивность процессов перехода электронов из когерентного канала в некогерентный.

Для оценки времени релаксации импульса электрона в активной области устройства с поперечным транспортом можно использовать время релаксации импульса в трехмерном полупроводнике.

Таким образом, введение в гамильтониан оптического потенциала iW позволяет моделировать процессы уменьшения когерентной составляющей электрического тока по мере движения электронов от одного электрода к другому. Однако это нарушает закон непрерывности заряда, если считать, что когерентный канал – единственный канал переноса. Поэтому вводят уже упомянутый некогерентный канал, плотность тока в котором обозначена через J_{incoh} (7.83). Закон сохранения тока требует выполнения равенства $j = j_{incoh}(z) + j_{incoh}(z)$ в каждой точке устройства, где j –

элементарный ток, создаваемый одним электроном. Однако найти способ нахождения $j_{incoh}(z)$, который бы удовлетворял этому условию, пока что не удалось⁵. Обычно используется подход, который базируется на модели единичного рассеивателя [7.36], впоследствии развитый в работах [7.37; 7.38; 7.1]. Согласно этому подходу при расчете тока считается, что в яме резонансно-туннельного диода содержится один единственный центр рассеивания, при прохождении которого часть когерентного тока скачкообразно уменьшается на величину A :

$$A = 1 - T_{coh} - R_{coh}, \quad (7.89)$$

где R_{coh} – коэффициент отражения от квантовой системы. Отметим, что при $W \neq 0$ $T_{coh} < 1 - R_{coh}$, что всегда имеет место для диссипативных систем; если же рассеивание отсутствует ($W = 0$) $T_{coh} = 1 - R_{coh}$ и $A = 0$, что справедливо для систем без потерь.

В дальнейшем полагают, что каждый электрон, испытавший рассеивание, будет некоторое время двигаться в квантовой яме до тех пор, пока не выйдет через один из барьеров. Вероятность его выхода через правый барьер равняется $T_R/(T_L + T_R)$, где $T_{R(L)}$ – коэффициент прохождения правого (левого) барьера. Положения центра рассеивания во всех известных работах выбирают посередине квантовой ямы. Эта точка условно делит квантовую систему на левую и правую части. Коэффициенты прохождения $T_{R(L)}$ суть коэффициенты прохождения левой и правой подсистемы отдельно. Способ нахождения этих коэффициентов не отличается от уже обсуждавшегося.

Таким образом, вероятность для электрона, который эмитировал из левого резервуара, достичь правого при условии рассеяния в некогерентный канал $A T_R/(T_L + T_R) \equiv T_{incoh}$ ⁶. Вероятность того, что электрон достигнет правого резервуара через когерентный канал, равняется T_{coh} . Поскольку эти два события взаимоисключающие, вероятность достижения правого резервуара любым из двух способов определяются выражением:

$$T(E_z) = T_{coh} + (1 - R_{coh} - T_{coh}) \frac{T_R}{T_R + T_L}. \quad (7.90)$$

⁵ этот факт в большинстве работ, где используется оптический потенциал, замалчивается, что можно рассматривать как нарушение научной этики

⁶ вероятность события «А» при условии события «Б» $p_{A|B}$ равна произведению безусловных вероятностей: $p_{A|B} = p_A p_B$, что и было здесь использовано

Следует еще раз отметить, что такой способ учета рассеяния нарушает физичность модели, а именно не обеспечивает выполнения закона непрерывности полного тока (7.79) по всей длине квантовой области. Его использование объясняется простотой, однако не отличается высокой адекватностью, поэтому время рассеивания τ_p обычно играет роль подгоночного параметра.

На рис. 7.19, а изображена зависимость суммарного коэффициента передачи $T(E_z)$ и его составляющих - вероятности когерентного $T_{coh}(E_z)$ и некогерентного $T_{incoh}(E_z)$ прохождения. Время релаксации τ рассчитывалось по формуле [7.1]:

$$\frac{1}{\tau} = S_{op} \left[1 + \frac{1}{\exp(\hbar\omega_{op} / k_B T) - 1} \right],$$

где $\hbar\omega_{op} = 0.036$ эВ - энергия полярных оптических фононов в GaAs; $S_{op} = 36 \cdot 10^{12}$ с⁻¹ - скорость рассеивания.

Видно, что функция $T(E_z)$ размывается, то есть рассеивание расширяет границы спектра электронов, которые могут пройти сквозь ДБКС.

Учет рассеивания приводит к ожидаемому изменению формы ВАХ. На рис. 7.19, б приведена ВАХ, рассчитанная с учетом рассеивания (параметры которого указаны выше). Заметно, что рассеивание приводит к увеличению тока приблизительно на порядок и росту пикового тока на 20%. Кроме того, уменьшается величина отрицательной дифференциальной проводимости. Эти изменения объясняются появлением нового канала передачи с более широкой спектральной полосой.

Следует отметить, что сходимость линеаризованного метода Гумеля значительно замедляется в области отрицательной дифференциальной проводимости (ОДП) по сравнению с областями положительной дифференциальной проводимости (ПДП). Для достижения точности 10^{-4} эВ в области ОДП нужно в среднем 30 итераций, тогда как в области ПДП среднее количество таких итераций составляет 10. При этом если не уменьшить абсолютную погрешность вычисления интеграла концентрации до 10^{17} м⁻³, алгоритм Гумеля вообще может не сойтись; при этом одни и те же приближения самосогласованного потенциала повторяются через одну (иногда - две, три) итерации, так и не достигая заданной точности. Это значит, что незначительные изменения концентрации электронов в активной области РТД приводят к значительным изменениям самосогласованного потенциала, который определяет ток. Отсюда

можно сделать вывод, что в переходных процессах, где распределение объемного заряда во времени изменяется, область ОДП будет областью неустойчивости, поскольку при одном и том же напряжении при почти одинаковых концентрациях ток может сильно отличаться. Описанное явление действительно наблюдается на практике [7.39].

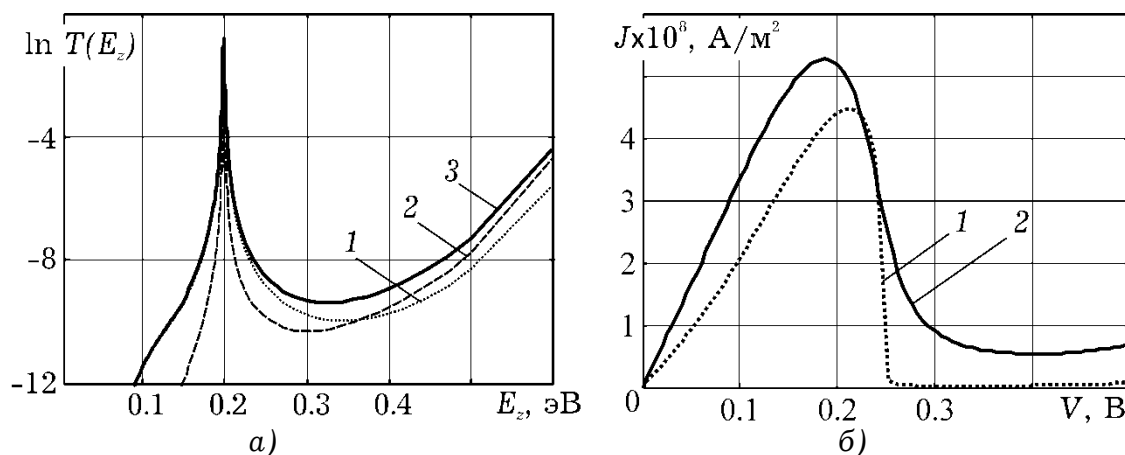


Рис. 7.19. Влияние рассеивания на характеристики РТД: а) коэффициенты прохождения ($V = 0$): 1 – T_{incoh} , 2 – T_{coh} , 3 – $T \equiv T_{incoh} + T_{coh}$; б) вольтамперные характеристики без (1) и с учетом рассеивания (2). Моделирование проводилось несамосогласованно

7.3.5. Влияние вторичных факторов на характеристики РТД

Сопrotивление приконтактных областей. При рассмотрении физических процессов в квантовой области достаточно обоснованным является предположение о том, что высоколегированные приконтактные и контактные области – идеально проводящие. Однако даже поверхностный анализ экспериментальных вольтамперных характеристик резонансно-тоннельных диодов убеждает, что сопротивление пассивных областей РТД, включая сопротивление высоколегированных слоев и контактов, существенно влияет на вид ВАХ. В частности, пик тока на экспериментальных ВАХ наблюдается при значительно высших напряжениях, чем это допускают любые оценки, которые осуществлены при пренебрежении влияния указанного фактора [7.40; 7.41]. Так, Цу отмечал [7.42 с. 61], что первый пик ВАХ должен наблюдаться при напряжении $V = 2E_1 / e$, где E_1 – положение первого метастабильного уровня в квантовой яме. Применяя эту приближенную формулу к РТД, описанному в [7.41], яма и барьеры которого изготовлены из слоя GaAs шириной 45 \AA и слоев AlAs шириной 17 \AA , соответственно, получаем оценку

пикового напряжения порядка $2 \cdot E_1 \approx 0,3 \text{ В}$ в то время как экспериментальная ВАХ подтверждает наличие такого пика при напряжении 2,5 В при одной полярности напряжения и около 1 В – при другой, что подтверждает тезис о необходимости учета особенностей топологии, которые влияют на сопротивление контактных и приконтактных слоев.

Поскольку сопротивление самих контактов зависит от топологии конкретного РТД, его следует рассчитывать для каждого устройства отдельно. Унифицировано можно попытаться описать лишь приконтактные высоколегированные области.

Обычно можно считать, что высоколегированные области имеют такую же площадь, что и слои РТД. Пусть имеем всего M таких областей; для каждой i -й области известна ее ширина (L_i) и концентрация ионизированных доноров (N_i). Падение напряжения на i -й области составляет:

$$V_i = J \cdot R_{s,i},$$

где $R_{s,i}$ – погонное сопротивление, J – плотность тока, который протекает через РТД; $R_{s,i} = \rho_i L_i$, где ρ_i – удельное сопротивление. Удельное сопротивление однородного массивного полупроводника может быть определено из известного соотношения:

$$\rho_i = (e \cdot \mu_i \cdot N_{d,i})^{-1},$$

где μ_i – подвижность электронов в i -м слое. Воспользовавшись ею, получим уравнение, в соответствии с которым следует пересчитывать теоретические вольтамперные характеристики с учетом влияния высоколегированных областей:

$$V_{\Sigma} = V_{\text{DBQS}}(J) + J \sum_{i=1}^M R_s,$$

где V_{Σ} – напряжение, измеряемое в эксперименте, $V_{\text{DBQS}}(J)$ – рассчитанная ВАХ активной части РТД. Из этой формулы следует, что с увеличением сопротивления высоколегированных областей пиковое напряжение смещается в область высоких напряжений быстрее, чем напряжение долины. Это приводит к увеличению отрицательного дифференциального сопротивления. Это также может привести к возникновению гистерезиса, если величина паразитного сопротивления превысит определенное значение. Гистерезис ВАХ действительно наблюдается на

практике [7.43; 7.44], однако описанный механизм его формирования не является единственно возможным.

Влияние паразитных сопротивлений на ВАХ РТД отображено на рис. 7.20. При плохом качестве контактов сопротивление паразитных областей может оказаться нелинейным, что, согласно экспериментальным данным, имеет место, например в [7.40] и [7.41], в которых пиковое напряжение на ВАХ при одной из полярностей напряжения источника питания на 1,4 В превышает это же напряжение при другой полярности и достигает 2,5 В. При этом характер зависимости не может быть объяснен, если считать, что сопротивление контактных областей линейно.

Форма разрывов зон. Выше принималось предположение о том, что потенциальная энергия электрона в зоне проводимости испытывает скачкообразное изменение на гетерограницах и приводит к формированию прямоугольных потенциальных ям и барьеров. Такое предположение достаточно хорошо обосновано, если размеры слоев значительно больше периода кристаллической решетки. Однако это условие не выполняется для рассматриваемых структур.

Другой способ аппроксимации разрывов зон можно предложить, исходя из результатов, полученных методом псевдопотенциалов [7.45], в рамках которого было показано, что более физической является двухступенчатая аппроксимация. При этом считается, что потенциал изменяется резко, однако в два этапа (рис. 7.21, а). Такая аппроксимация, вероятно, является более корректной, поскольку основывается на более адекватных методах расчета зонной структуры, чем метод огибающей волновых функций позволяет в принципе.

Учет реальной формы барьеров приводит к увеличению плотности тока через резонансно-туннельный диод за счет повышения проницаемости барьеров (рис. 7.21, б).

Точность и адекватность описанных моделей можно повысить за счет реализации следующих шагов:

- 1) для описания кинетики электронного коллектива в резервуарах использовать кинетическое уравнение Больцмана вместо предположения о локальном равновесии [7.46];

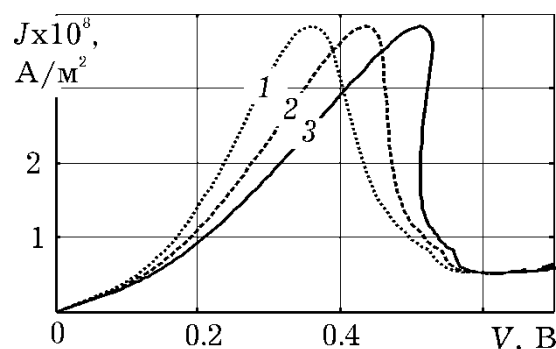


Рис. 7.20. Влияние величины погонного сопротивления пассивных областей РТД R_s на ВАХ РТД:
 1 – $R_s = 0$; 2 – $R_s = 40$ Ом мкм²,
 3 – $R_s = 20$ Ом мкм²

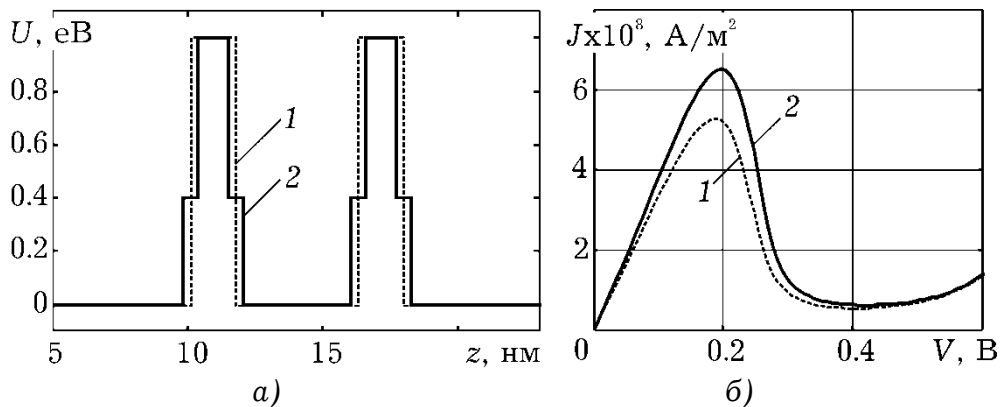


Рис. 7.21. Стандартная (1) и "сверхрешеточная" (2) модели гетерограницы: а) профиль зоны проводимости в Γ -долине, б) результаты моделирования ВАХ

- 2) учет непараболичности закона дисперсии [7.41];
- 3) упразднение предположения о мгновенной термализации электронов в резервуарах;
- 4) учет токопереноса через метастабильные состояния в эмиттерной квантовой яме [7.47];
- 5) если эффективные массы в эмиттере и квантовой яме отличаются (или если учитывается непараболичность законов дисперсии), выражение для расчета тока не сводится к формуле Цу-Эсаки; возникает необходимость проводить численное интегрирование также по составляющей $k_{\parallel} = \sqrt{k_x^2 + k_y^2}$ [7.48] или же использовать приближенные формулы для плотности тока [7.49], которые учитывают отличие эффективных масс.

Отдельно следует выделить способы повышения адекватности, которые базируются на учете транспорта с участием высших долин. В случае если направлением роста гетероструктуры является [100], для GaAs/AlAs РТД релевантной будет X-долина.

7.4. ДВУХДОЛИННОЕ МОДЕЛИРОВАНИЕ

Все полученные выше результаты основывались на предположении, что участие в переносе заряда принимают исключительно электроны Γ -долины, а высшие долины незаселены.

Известно, что междолинные переходы в полупроводниках становятся существенными в сильных полях [7.18 с. 89], которые

преимущественно не имеют места в резонансно-туннельных диодах. Однако в полупроводниковых гетероструктурах возможен другой тип междолинных переходов – *упругое междолинное рассеивание на гетерограницах* или «смешивание электронных состояний».

Для гетероструктур $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ при $x > 0,45$ в барьерных слоях самой низкой долиной становится X-долина (рис. 2.14). Для двухбарьерной квантовой системы AlAs/GaAs потенциальный рельеф каждой из долин приобретет вид, изображенный на рис. 7.22.

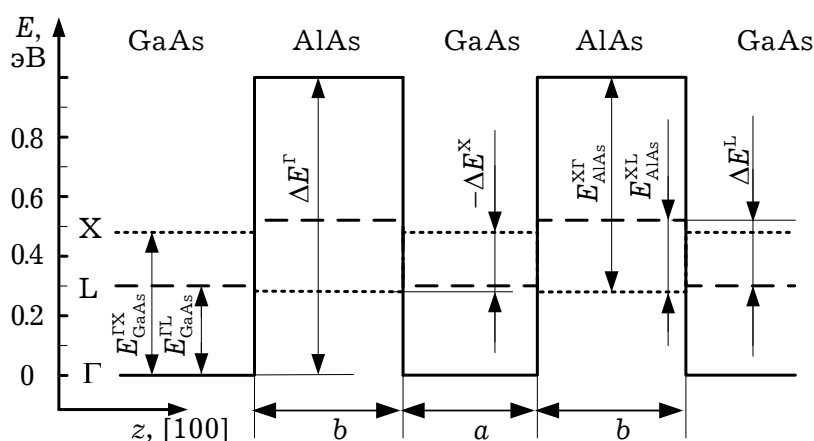


Рис. 7.22. Параметры зоны проводимости AlAs/GaAs РТД

Величины междолинных расстояний и разрывов зон представлены в табл. 7.1.

Таблица 7.1.

Параметры зонной структуры гетероперехода AlAs/GaAs , взятые преимущественно из [7.50] и [7.51].

Разрывы зон, эВ	Междолинные расстояния, эВ	
	GaAs	AlAs
$\Delta E_{\Gamma} = 1$		
$\Delta E_X = -0.24$	$E_{\text{GaAs}}^{\Gamma X} = 0.49$	$E_{\text{AlAs}}^{X\Gamma} = 0.73$
$\Delta E_L = 0.21$	$E_{\text{GaAs}}^{\Gamma L} = 0.29$	$E_{\text{AlAs}}^{XL} = 0.23$

Слои РТД чаще всего выращиваются в направлении [100]. Именно в направлении [100] электроны ускоряются внешним полем, которое увеличивает компоненту $k_{[100]}$ волнового вектора (в наших обозначениях – k_z).

Из рис. 2.11 видно, что на гетерограницах наиболее благоприятные условия создаются для перехода электронов из Γ -долин GaAs в X-долины AlAs, но не во все, а лишь в две из них, которые ориентированы вдоль кристаллографического

направления $[100]$ ⁷. Переход в X-долины, которые ориентированы в направлениях $[001]$ и $[010]$, маловероятен, поскольку распределение электронов по импульсам $\hbar k_{[001]}$ и $\hbar k_{[010]}$ остается равновесным (продольное поле отсутствует). По тем же причинам невыгодны также переходы в L-долины, ориентированные в направлении $[111]$ и эквивалентных. Это подтверждается расчетами в рамках полуэмпирического метода сильной связи [7.4] и свидетельствует в пользу пренебрежения всеми междолинными переходами кроме Γ -X. Все последующие выводы справедливы для гетероструктуры со слоями AlAs-GaAs, выращенной на подложке с кристаллографической ориентацией $[100]$.

Большинство предположений, принятых в однодолинной модели, остаются в силе и для двухдолинной модели. Отличие заключается в рассмотрении электронов X-долины. Для упрощения считается, что X-долины в резервуарах не заселены. Электроны из X-долины, которые попадают в резервуары из квантовой области, мгновенно переходят в Γ -долину и термализируются.

Считая ψ^Γ и ψ^X известными, построим самосогласованную модель, аналогичную описанной в разд. 7.3. Распределение концентрации и ток через РТД изменяются с учетом возможности движения электронов в квантовой области как в Γ -, так и в X-долинах. Поскольку мы учитываем электроны из двух долин, то концентрация всех электронов в квантовой области является суммой концентраций электронов в Γ - и X-долинах (n^Γ и n^X соответственно) которые находятся по формулам, которые являются простым обобщением формул, полученных для однодолинной модели:

$$n = n^\Gamma + n^X \quad (7.91)$$

где

$$n^{\Gamma(X)} = \int_{-\infty}^{\infty} |\psi^{\Gamma(X)}(\mathbf{k})|^2 N(\mathbf{k}) dk_x dk_y dk_z, \quad (7.92)$$

Здесь $N(\mathbf{k}) = N^\Gamma(\mathbf{k}) + N^X(\mathbf{k})$, где $N^{\Gamma(X)}$ – функция распределения электронов в Γ (X) долине левого (если $k_z > 0$) или правого (если $k_z < 0$) резервуаров. Согласно принятым предположениям $N^X = 0$.

⁷ каждая из этих X-долин лишь наполовину находится в первой зоне Бриллюэна

Формулы для концентрации и тока в двухдолинной модели имеют вид:

$$n = n_L^\Gamma + n_R^\Gamma + n_L^X + n_R^X,$$

где:

$$n_{L(R)}^\Gamma = N'_c \int_{U_{i0}(U_{i5})}^{\infty} \frac{|\Psi_{L(R)}^\Gamma|^2}{\sqrt{E_z - U_{i0(i5)}^\Gamma}} \ln \left(1 + e^{-\frac{E_z - (E_\Phi + U_{i1(N)}^\Gamma)}{k_B T}} \right) dE_z, \quad (7.93)$$

$$n_{L(R)}^X = N'_c \int_{U_{i0}(U_{i5})}^{\infty} \frac{|\Psi_{L(R)}^X|^2}{\sqrt{E_z - U_{i0(i5)}^\Gamma}} \ln \left(1 + e^{-\frac{E_z - (E_\Phi + U_{i1(N)}^\Gamma)}{k_B T}} \right) dE_z; \quad (7.94)$$

$$J = \frac{2m^* e k_B T}{(2\pi)^2 \hbar^3} \int_{\max(U_{i5}, U_{i0})}^{\infty} T(E_z) D(E_z) dE_z, \quad (7.95)$$

Неизвестными в (7.93)–(7.95) пока еще являются волновые функции и коэффициент передачи.

Поскольку X-долины в резервуарах считаются незаселенными

$$\psi_I^X = 0, \text{ если } z < z_1$$

где z_1 – левая граница квантовой области (правую будем обозначать через z_5). Индексами I, II, III будем обозначать, соответственно, левый резервуар, квантовую область и правый резервуар. Под ψ , пока не будет оговорено другое, следует понимать волновую функцию электронов, которые налетают на квантовую область из левого резервуара.

Электрону в левом резервуаре (I область) соответствует суперпозиция падающей и ой плоских волн:

$$\psi_I^\Gamma = \exp(ik_L^\Gamma z) + r \exp(-ik_R^\Gamma z).$$

Если учитывать смешивание состояний в квантовой системе, на выходе из нее в точке $z = z_5$ электрону будет соответствовать уже две волны:

$$\psi_I^{\text{III}} = t_{\Gamma X} \exp(ik_{R,\Gamma} z), \quad (7.96)$$

$$\psi_X^{\text{III}} = t_{X\Gamma} \exp(ik_{R,X} z), \quad (7.97)$$

где $k_R^{\Gamma(X)} = \sqrt{2m_{III}^{\Gamma(X)}(E_{\Gamma(X)} - U_{\Gamma(X)})} / \hbar$ – волновой вектор электрона в $\Gamma(X)$ -долине, $m_{III}^{\Gamma(X)}$ – эффективная масса электрона в $\Gamma(X)$ -долине в III области $t_{\Gamma X \Gamma}$, $t_{\Gamma X}$

Вероятность того, что электрон, пройдя через квантовую систему, окажется в Γ -долине, уже рассчитывалась в однодолинной модели (формула (7.20)) и называлась просто «коэффициент прохождения». Для двухзонной модели (7.20) запишется так:

$$T_{\Gamma X \Gamma} = |t_{\Gamma X \Gamma}|^2 \frac{|k_{III}^{\Gamma}|}{|k_I^{\Gamma}|} \frac{m_I^{\Gamma}}{m_{III}^{\Gamma}},$$

Вероятность того, что электрон, пройдя через квантовую систему, выйдет через X-долину, рассчитывается аналогично. Формула имеет вид:

$$T_{\Gamma X} = |t_{\Gamma X}|^2 \frac{|k_{III}^X|}{|k_I^{\Gamma}|} \frac{m_I^{\Gamma}}{m_{III}^X}.$$

Поскольку события «электрон выйдет через Γ -долину» и «электрон выйдет через X-долину» взаимоисключающие, соответствующие вероятности суммируются:

$$T = T_{\Gamma X \Gamma} + T_{\Gamma X}. \quad (7.98)$$

В формулах (7.96) и (7.97) остаются неизвестными постоянные интегрирования уравнения Шредингера. Кроме того неизвестными остаются волновые функции, необходимые для расчета концентрации. Их поиск является основной проблематикой двухдолинного моделирования.

В работе [7.52], основываясь на экспериментальных данных и расчетах, проведенных в рамках модели сильной связи, предложен метод расчета огибающих одноэлектронных волновых функций ψ_{Γ} и ψ_X . Последующее развитие он получил в [7.53], был теоретически обоснован в [7.54 с. 50] и в дальнейшем применялся в «комбинированных» моделях РТД [7.55].

Для учета Γ -X интерфейсного смешивания в уравнение Шредингера для электрона из $\Gamma(X)$ -долины

$$\hat{H}_{\Gamma(X)} \Psi_{\Gamma(X)} = E_{z, \Gamma(X)} \Psi_{\Gamma(X)} \quad (7.99)$$

слева вводится член $\alpha^{(*)} \cdot \delta(z_k) \cdot \Psi_{\Gamma(X)}$, что дает:

$$\hat{H}_{\Gamma(X)}\Psi_{\Gamma(X)} + \alpha^{(*)}\delta(z_k)\Psi_{X(\Gamma)} = E_{z,\Gamma(X)}\Psi_{\Gamma(X)}, \quad (7.100)$$

где $\hat{H}_{\Gamma(X)} = U_{\Gamma(X)} - \left(E - \frac{\hbar^2 k_{\parallel}^2}{2m_{\Gamma(X)}^*} \right) \equiv U_{\Gamma(X)} - E_{z,\Gamma(X)}$, $k_{\parallel}^2 = k_x^2 + k_y^2$, m_{Γ}^* –

эффективная масса электрона в Γ -долине, m_X^* – продольная⁸ эффективная масса в X -долине; z_k – координаты гетерограниц, δ – дельта-функция Дирака; α [эВ·Å] – константа междолинной связи, в общем случае комплексная величина. Для пары GaAs/AlAs значение α , согласующееся с расчетами методом псевдопотенциалов, составляет: $\alpha = 0.1$ эВ·Å [7.52]. Также использовались значения $\alpha = 0.15$ эВ·Å [7.53] и $\alpha = 0.45$ эВ·Å [7.55]; объективные причины выбора таких значений не указываются.

Чтобы понять сущность этого метода, запишем оба уравнения (7.100) в матричном виде:

$$\begin{aligned} & \begin{bmatrix} -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m_{\Gamma}^*} \frac{d}{dz} & 0 \\ 0 & -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{m_X^*} \frac{d}{dz} \end{bmatrix} \begin{bmatrix} \Psi_{\Gamma} \\ \Psi_X \end{bmatrix} + \\ & + \begin{bmatrix} U_{\Gamma} - E + \frac{\hbar^2 k_{\parallel}^2}{2m_{\Gamma}^*} & \alpha \delta(z_k) \\ \alpha^* \delta(z_k) & U_X - E + \frac{\hbar^2 k_{\parallel}^2}{2m_X^*} \end{bmatrix} \begin{bmatrix} \Psi_{\Gamma} \\ \Psi_X \end{bmatrix} = 0. \end{aligned} \quad (7.101)$$

Видно, что при $\alpha = 0$ (7.101) сводится к двум независимым уравнениям Шредингера для X - и Γ -долин. Но при $\alpha \neq 0$ в точках, которые есть гетерограницами, «включается» взаимодействие между X - и Γ -долинами, и происходит смешивание электронных состояний

Из уравнения (7.101) следует, что на волновую функцию на гетерограницах налагаются такие граничные условия:

⁸ напомним, что X -долины, в которые возможны переходы, ориентированы вдоль направления [100], поэтому масса в z -направлении, которая должна использоваться для расчетов, равна продольной массе в X -долине m_X^l . Исчерпывающее обоснование см. в [7.17]

$$\Psi_{\Gamma(X)}(z_{k-}) = \Psi_{\Gamma(X)}(z_{k+}), \quad (7.102)$$

$$\frac{\hbar^2}{2} \left[\frac{1}{m_{\Gamma(X)}^*} \Psi'_{\Gamma(X)}(z_{k-}) - \frac{1}{m_{\Gamma(X)}^*} \Psi'_{\Gamma(X)}(z_{k+}) \right] + \alpha^{(*)} \Psi_{X(\Gamma)}(z_k) = 0, \quad (7.103)$$

где штрихом обозначено дифференцирование по z . Выражения (7.104) и (7.103) при $\alpha = 0$ сводятся к обычным граничным условиям отдельно для Ψ_{Γ} и Ψ_X .

Используя (7.105) и (7.103) можно построить метод матриц передачи и другие условно-аналитические методы нахождения коэффициента прохождения, — так же, как это было сделано для однодолинной модели в разд. 7.2.

В приложении П7.2 описана трехточечная конечно-разностная схема, которая соответствует (7.101) для всех внутренних точек. Уравнения, замыкающие схему, для Ψ_{Γ} такие же, как для однодолинной модели (см. приложение П7.3). Для Ψ_X способ получения уравнений, замыкающих КРС, аналогичен; следует лишь учесть, что в X-долине амплитуда волны, соответствующая налетающему электрону, нулевая, что отображает предположение о незаселенности X-долины в резервуарах.

Неизвестные постоянные интегрирования $t_{\Gamma X \Gamma}$ и $t_{\Gamma X}$, использующиеся для расчета коэффициента прохождения, находят, исходя из того, что $|t_{\Gamma X \Gamma}(E_z)|^2 = |\Psi_{\Gamma}^{\Gamma}|^2 |t_{\Gamma X}(E_z)|^2 = |\Psi_N^X|^2$, где $\Psi_N^{\Gamma(X)}$ — волновые функции электрона из левого резервуара на границе квантовой области и правого резервуара.

Процесс нахождения самосогласованного потенциала для двухдолинной модели принципиально не отличается от алгоритма, который описан для однодолинной (разд. 7.3.3), и является его простым обобщением.

Функция прохождения электрона через квантовую область РТД $T(E_z)$ при учете Γ -X рассеивания испытывает качественные изменения, которые представляют интерес как с фундаментальной, так и с прикладной точки зрения (рис. 7.23). Резонансный характер, свойственный однодолинной функции передачи, в целом сохраняется (рис. 7.24). Однако появляются дополнительные резонансы и антирезонансы с несимметричным профилем. Известно, что резонансы такого типа могут проявляться в физических системах разнообразной природы; их общее название — резонансы Фано [7.]56). Они возникают, когда имеет место интерференция двух волновых процессов. В данном случае речь идет об интерференции состояний электрона из Γ - и

X-долин (в терминах квантовой механики можно сказать, что электрон с энергией E_z находится в состоянии, которое является суперпозицией состояний в Γ - и X-долинах).

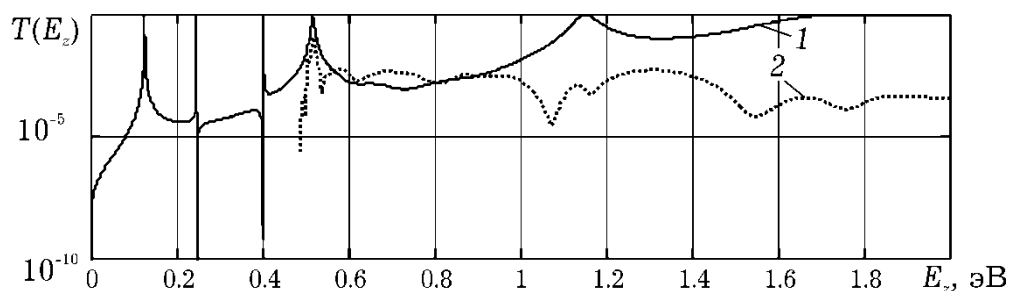


Рис. 7.23. Коэффициент прохождения $T(E_z)$ в двухдолинном приближении (1 – $T_{\Gamma X \Gamma}$, 2 – $T_{\Gamma X}$) при нулевом напряжении, моделирование несамосогласованное

Ключом к пониманию зависимости $T(E_z)$ являются графики, изображающие распределение плотности электронных состояний в Γ - и X-долинах, $g_{\Gamma(X)}(E_z)$ – рис. 7.26, а(б). Из них видно, что более широкие максимумы коэффициента прохождения (1-й и 4-й) соответствуют метастабильным состояниям в квантовой яме, которая формируется в Γ -долине; узкие несимметричные резонансы и антирезонансы (2-й и 3-й максимумы и минимумы, следующие за ними) наблюдаются при энергиях, которые соответствуют метастабильным уровням в квантовой яме, которая формируется в X-долине.

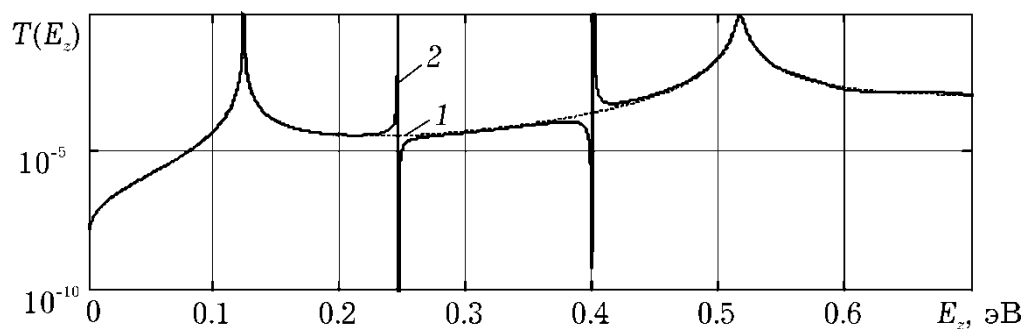


Рис. 7.24. Коэффициент прохождения из Γ в Γ долину: 1 – рассчитанный в рамках однодолинной, 2 – двухдолинной модели. Расчеты осуществлялись методом матриц передачи при нулевом напряжении

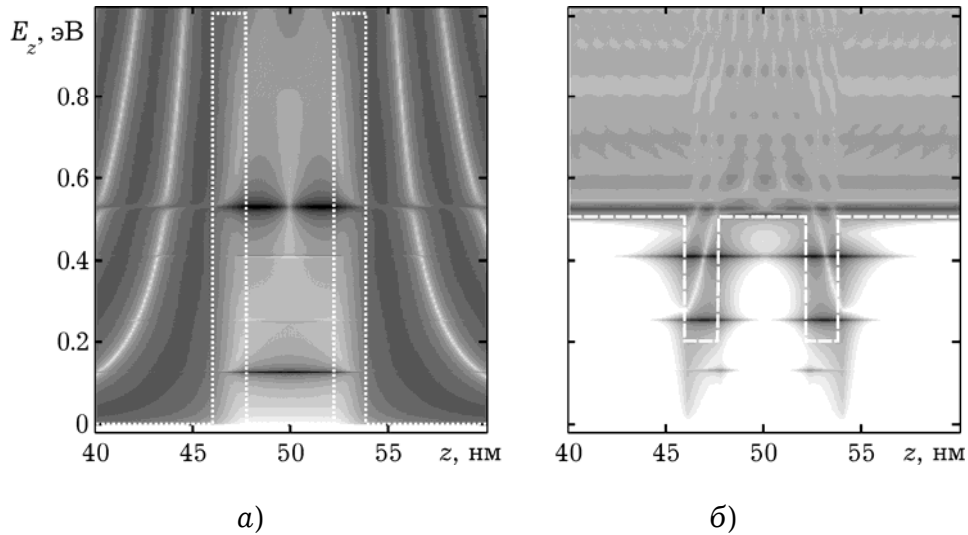


Рис. 7.25. Плотность электронных состояний в Γ - (а) и X -долинах (б). Более темные градации серого соответствуют большей плотности состояний в логарифмическом масштабе

Обнаруженные метастабильные уровни являются дополнительными каналами переноса тока через указанную структуру. На вольт-амперной характеристике учет Γ - X -рассеивания не вызывает появления дополнительных пиков, хотя если бы параметр α имел большее значение, это вполне могло наблюдаться, как это показано в [7.55]. Напротив, наблюдается повышение плотности тока как в долине, так и в пике ВАХ, что обычно приближает результаты моделирования к значениям тока, которые наблюдаются на практике (рис. 7.26).

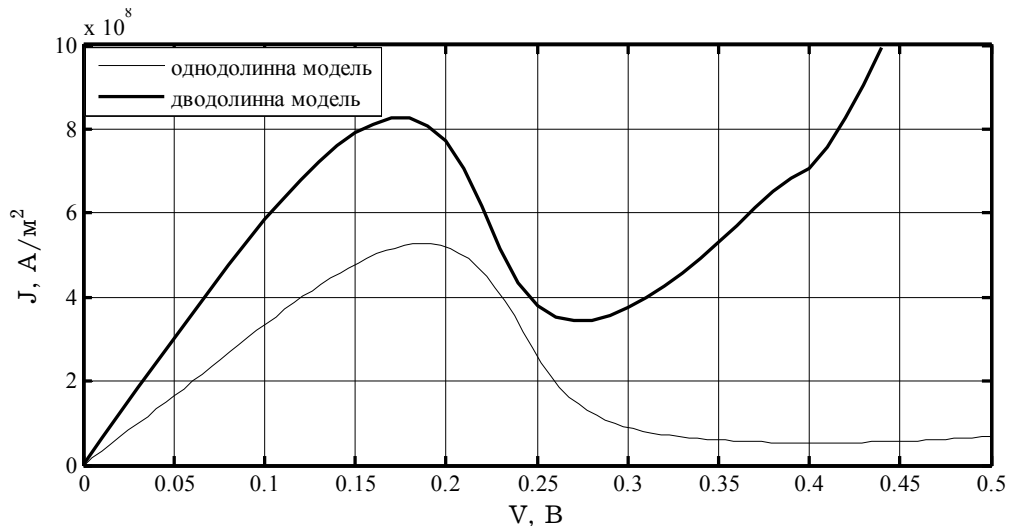


Рис. 7.26. ВАХ РТД при однодолинном и двухдолинном несамосоогласованном моделировании, параметр $\alpha = 1 \text{ эВ \AA}$

7.5. ЭКСПЕРИМЕНТАЛЬНАЯ ВЕРИФИКАЦИЯ МОДЕЛЕЙ

В [7.57] был описан процесс изготовления резонансно-тоннельного диода (далее – РТД1) методом молекулярно-лучевой эпитаксии и результаты измерения его вольтамперных характеристик при температуре 77 К. РТД1 изготовлен на GaAs-подложке, легированной донорной примесью, концентрация которой составляет $2 \cdot 10^{24} \text{ м}^{-3}$. На подложку нанесен высоколегированный слой GaAs толщиной 1 мкм, за которым следуют нелегированные слои AlAs/GaAs. Верхний слой состоит из высоколегированного GaAs толщиной 0,5 мкм. Нелегированная часть состоит из спейсеров (5 и 10 нм *i*-GaAs), которые предотвращают диффузию примеси в барьерные слои и накопление электронов вблизи двухбарьерной квантовой системы. Последняя образована одинаковыми барьерными AlAs-слоями толщиной по 3 нм, окружающими 3-нм слой GaAs, в котором формируется квантовая яма (рис. 7.27). Площадь мезоструктуры составляет $16 \times 16 \text{ мкм}$.

Экспериментальная вольтамперная характеристика (рис. 7.28), измеренная при температуре 77 К, демонстрирует пиковое значение тока $J_p = 2,4 \cdot 10^6 \text{ А/м}^2$ при напряжении $V_p = 0,42 \text{ В}$. Значения тока и напряжения в долине ВАХ составляют, соответственно, $1 \times 10^6 \text{ А/м}^2$ и 0,46 В.

Для сравнения ВАХ РТД1 с данными эксперимента использована численная двухдолинная модель. Исходные данные для нее взяты из рис. 7.28, а также из табл. 7.2. Параметры выбирались преимущественно из электронного справочника Физико-технического института им. Иоффе [7.51] и не изменялись с целью лучшего согласования с экспериментальными данными. Однако было использовано другие два согласующих параметра, а именно: для согласования ВАХ по пиковому току все рассчитанные значения тока делились на 4,3. Для согласования по пиковым напряжениям вся экспериментальная

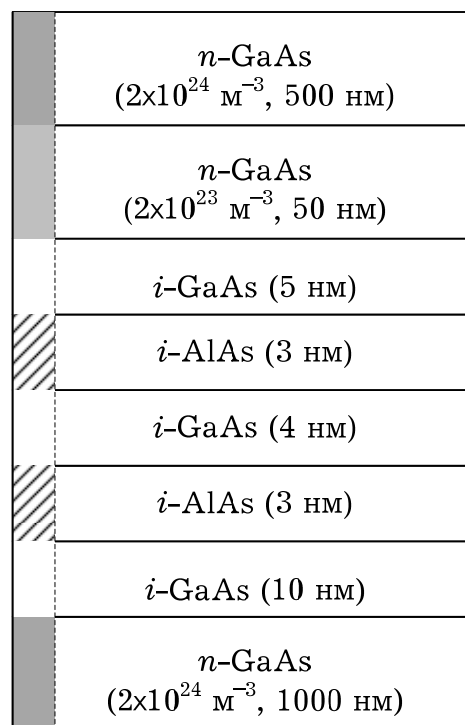


Рис. 7.27. Толщина, химический состав слоев и концентрация донорной примеси в РТД1

кривая сдвинута на 0,1 В вправо. Аргументацией введения этих параметров является, соответственно, неизвестная точно площадь и уровень легирования резервуаров и неопределенность величины и зависимости от напряжения сопротивления пассивных областей. Однако авторы не настаивают, что именно эти факторы вызывают расхождение экспериментальных и теоретических ВАХ.

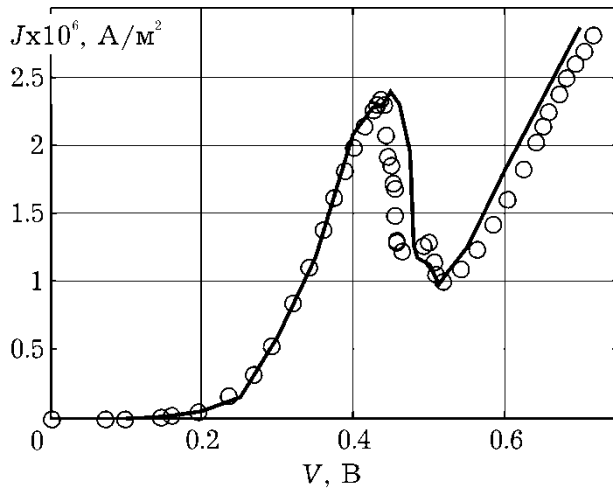


Рис. 7.28. Вольтамперная характеристика РТД1, рассчитанная с помощью двухдолинной самосогласованной модели с использованием двух «подгоночных» параметров (сплошная линия) отлично совпадает с экспериментальной кривой (кружочки)

Таблица 7.2

Физико-топологические параметры AlAs/GaAs РТД

параметр	значение	параметр	значение
T , К	77	m_{GaAs}^{Γ}	0,067
$\hbar\omega_{\text{оп}}$, эВ	0,036	$m_{\text{GaAs}}^{\text{X}}$	1,3
$S_{\text{оп}}$, пс ⁻¹	36	m_{AlAs}^{Γ}	0,15
ϵ_{GaAs}	12,9	$m_{\text{AlAs}}^{\text{X}}$	0,97
ϵ_{AlAs}	10,06	α , эВ·Å	0,15

Рассмотренный выше РТД1 имеет традиционную топологию активной области: однородные спейсери, между которыми находятся три слоя, формирующих ДБКС. Рассмотрим РТД, описанный в [7.58] (далее – РТД2). Особенностью этого диода является сложная структура его верхнего спейсера (рис. 7.29). Последний состоит из четырех слоев: верхний слой – нелегированный GaAs; следующие слои – $\text{Al}_{0,02}\text{Ga}_{0,98}\text{As}$, $\text{Al}_{0,05}\text{Ga}_{0,95}\text{As}$, $\text{Al}_{0,08}\text{Ga}_{0,92}\text{As}$, которые формируют так называемый «ступенчатый эмиттер», называемый так, поскольку от резервуара к верхнему барьерному слою потенциальный рельеф зоны проводимости растет ступенчато. Приконтактные слои РТД2

состоят из GaAs, легированного примесью донорного типа с концентрацией $1 \cdot 10^{24} \text{ м}^{-3}$. Двухбарьерная квантовая система включает AlAs-барьеры и GaAs-яму, толщина которых равна, соответственно, 3,4 и 5,9 нм. Нижний спейсер является однородным нелегированным слоем GaAs. Площадь выращенной мезоструктуры составляет $16 \times 16 \text{ мкм}^2$. В отличие от РТД1, измерения ВАХ проводились при температуре 300 К.

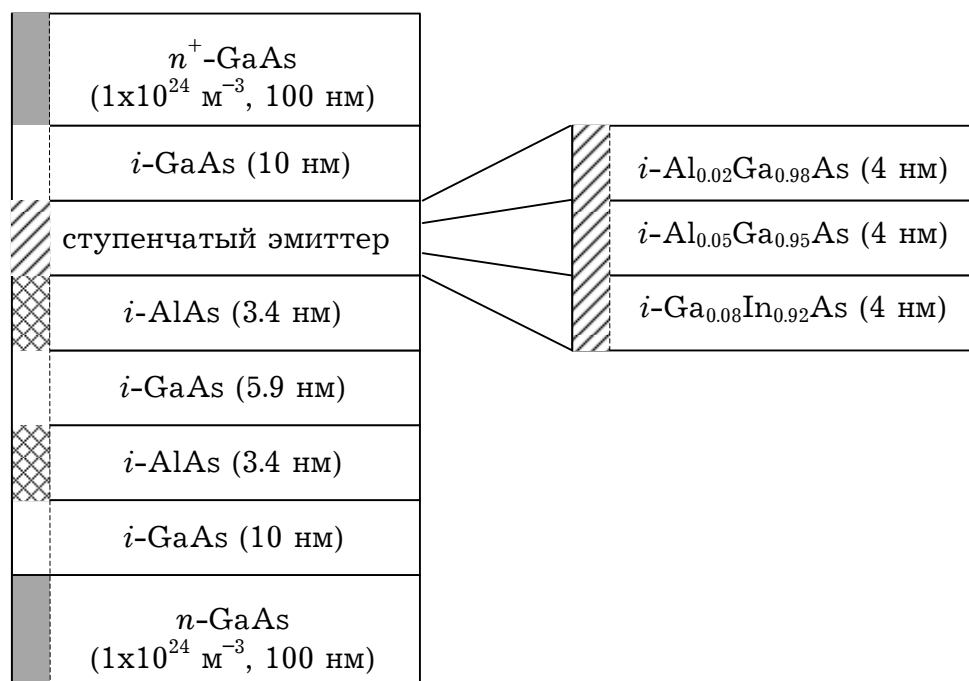


Рис. 7.29. Толщина, химический состав слоев и концентрация донорной примеси в РТД2. Справа показана структура ступенчатого эмиттера

Благодаря введению ступенчатого эмиттера в топологию диода на ВАХ возникает выраженный экстремум, следующий за основным пиком (рис. 7.30). Этот пик обусловлен особенностью взаимодействия уровней энергии в эмиттерной квантовой яме и основной квантовой яме благодаря ступенчатой структуре эмиттера.

Моделирование вольтамперных характеристик проводилось с использованием однодолинной несамосогласованной модели. Последняя была выбрана, поскольку, в отличие от самосогласованной, предусматривает существование экстремума. Самосогласованная модель пик не предусматривает, хотя дает лучшее согласование по пиковому напряжению. Такой парадокс, возможно, объясняется тем, что для слоев ступенчатого эмиттера учета заряда по методу Хартри является более грубым приближением, чем линейное падение потенциала (через пренебрежение межэлектронным обменным взаимодействием).

Входные параметры для моделирования брались из рис. 7.29 и табл. 7.2. Для более качественного сравнения с экспериментальными данными рассчитанные значения плотности тока умножались на 1,3 – этот коэффициент уравнивает характеристики за пиковым током. Кроме того, вся рассчитанная ВАХ смещалась вправо на 0,03 В. Такое смещение призвано учесть полностью проигнорированные контактные разности потенциалов, а также сопротивление пассивных областей.

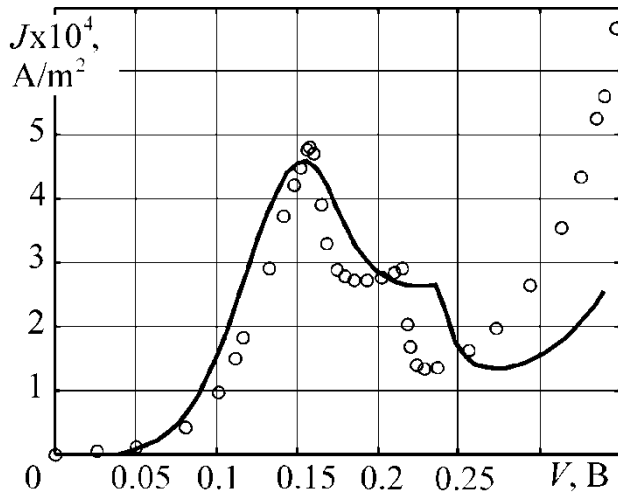


Рис. 7.30. Вольтамперная характеристика РТД2, рассчитанная с помощью однодолинной несамосогласованной модели (сплошная линия) с использованием двух «подгоночных» параметров качественно согласуется с экспериментальной (кружочки)

Происхождение участка плато на рис. 7.30 можно понять, анализируя результаты моделирования с помощью численной модели и сравнивая с ВАХ диода, который имеет обычный (*i*-GaAs), а не ступенчатый эмиттер (рис. 7.31).

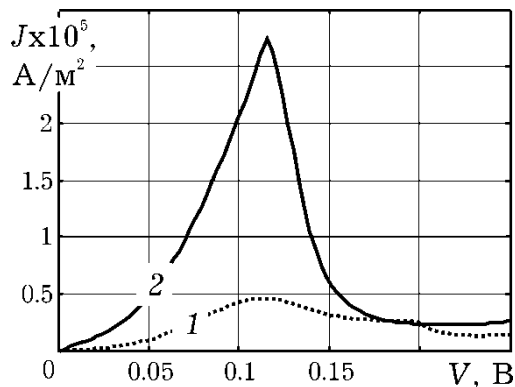


Рис. 7.31. Вольтамперные характеристики двух диодов, которые отличаются лишь структурой эмиттерного спейсера:
1 – «ступенчатый» эмиттер (рис. 7.29)
2 – обычный эмиттер (22 нм *i*-GaAs)

Для обеих структур основной пик наблюдается, когда метастабильный уровень в основной квантовой яме лежит между дном зоны проводимости и уровнем Ферми в левом резервуаре (рис. 7.32, а). При этом для ступенчатого эмиттера пиковый ток в несколько раз меньше, чем для обычного, поскольку между левым резервуаром и основной квантовой ямой находятся барьеры, образованные слоями ступенчатого эмиттера. Характер видоизменения этих барьеров при увеличении напряжения

(рис. 7.32, б-в) и предопределяет образование области плато на ВАХ. Ток, который протекает через РТД, определяется поведением коэффициента прохождения при энергиях, лежащих между уровнем Ферме и дном зоны проводимости. Можно считать, что для таких энергий коэффициент прохождения через квантовую систему $T(E_z)$ пропорционален коэффициенту прохождения через ступенчатый эмиттер T_L и коэффициенту прохождения через оставшуюся часть квантовой системы T_R , которая включает ДБКС и правый спейсер. Коэффициент прохождения T_R имеет резонансный характер, а его максимумы совпадают с центрами метастабильных уровней основной квантовой ямы. Коэффициент прохождения T_L при увеличении напряжения монотонно возрастает. В указанных пределах изменения энергии их произведение при данной структуре ступенчатого эмиттера в диапазоне напряжений 0,15...0,20 В почти не изменяется: рост T_L компенсируется уменьшением T_R . При дальнейшем увеличении напряжения T_R уменьшается значительно быстрее, чем T_L , поэтому T резко уменьшается. Это приводит к падению плотности тока. Динамика изменения подынтегральной функции в формуле Цу-Есаки показана на рис. 7.33, а. Видно, что в отличие от монотонного уменьшения площади под подынтегральной функцией, как это имеет место для диода с обычным эмиттером (рис. 7.32, б), для диода со ступенчатой структурой эмиттера за счет описанных эффектов подынтегральная функция ведет себя так, что точка максимума в области ОДП описывает резонансную кривую, а площадь под подынтегральной кривой почти постоянна (рис. 7.32, а).

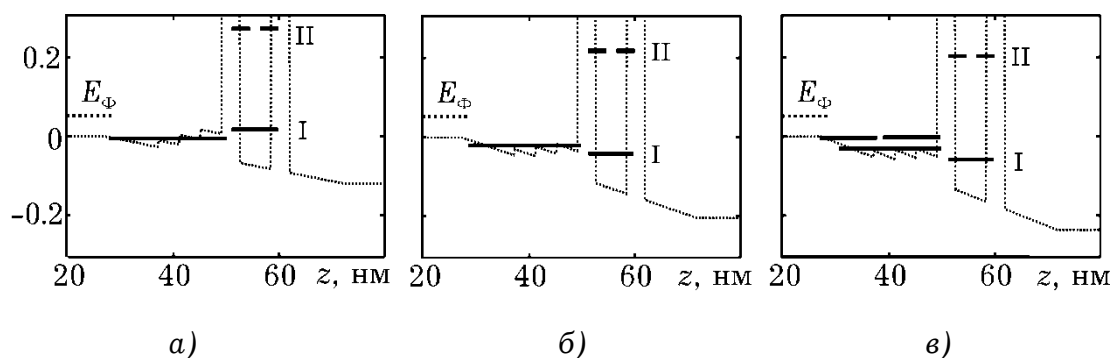


Рис. 7.32. Положение уровней энергии и уровня Ферме в ДБКС диода со ступенчатым эмиттером: а) пиковое напряжение, $V = 0,119$ В, б) напряжение экстремума $V = 0,197$ В, в) напряжение долины $V = 0,23$. Наибольшая плотность тока через диод наблюдается, когда уровень энергии в основной квантовой яме лежит между E_F и E_C ; экстремум и долина на ВАХ возникают, когда уровень в ОКЯ оказывается ниже дна зоны проводимости

Резюмируя вышеприведенные результаты верификации и учитывая опыт авторов, следует заметить, что описанные модели РТД из всех параметров ВАХ лучше всего предсказывают величину контрастности (величина, равная отношению пикового и долинного токов). По этому параметру отличие от экспериментальных данных составляет не больше 5%.

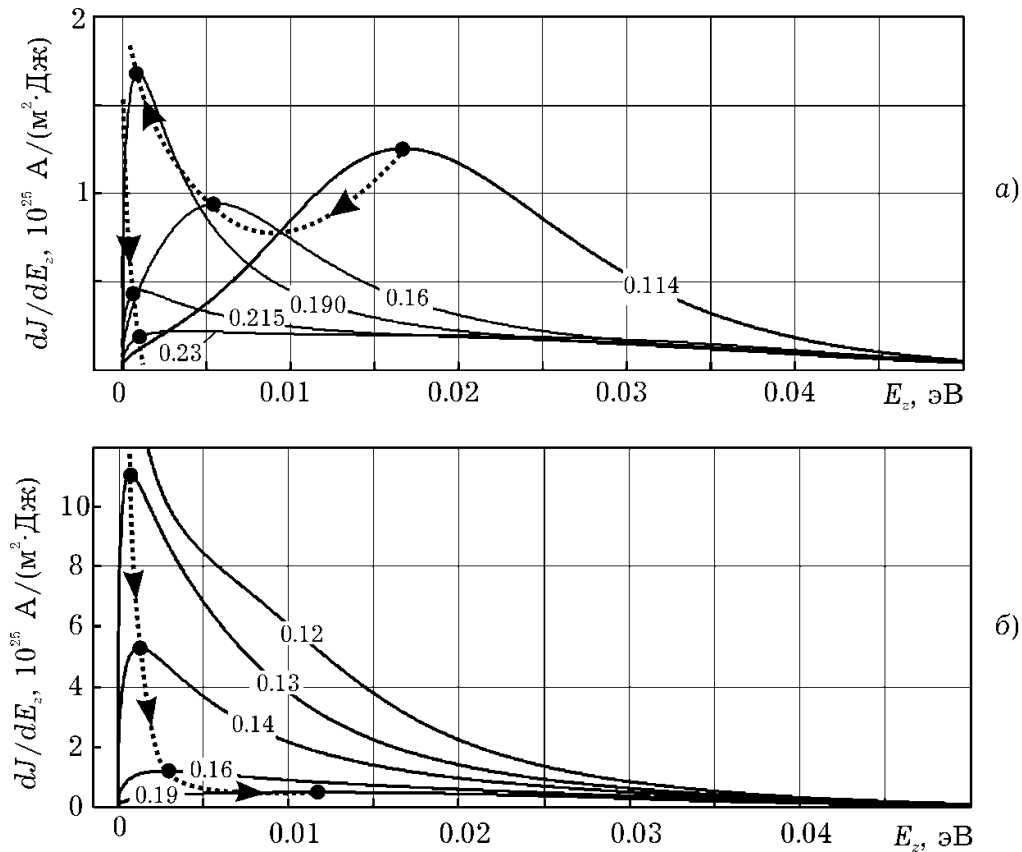


Рис. 7.33. Динамика изменения подынтегральной функции при напряжениях от пиковой к долинной (показаны цифры в Вольтах). Пунктирной линией показана траектория точки максимума подынтегральной функции в формуле Цу-Эсаки: а) резонансно-тоннельный диод со ступенчатым эмиттером; б) РТД с обычным эмиттером. В случае "а" при напряжениях от 0,115 до 0,197 В площадь под кривой почти не изменяется; в случае "б" – монотонно уменьшается по мере роста напряжения от пикового к долинному

Абсолютные же значения измеренного тока отличаются от теоретических в несколько раз (обычно не больше, чем вдвое), причем закономерности этого расхождения нами не обнаружены. Наиболее вероятной причиной может быть то, что в реальных диодах профиль легирования резервуаров может отличаться от номинального. В [7.9] говорится, что концентрация донорной примеси в резервуарах может отличаться в два раза; кроме того, площадь мезоструктуры не поддается точному измерению.

Известно также, и это подтверждается экспериментально, что зависимость пикового тока от толщины барьера имеет экспоненциальный характер. Поэтому вариации толщины барьерных слоев в пределах одного монослоя могут приводить к изменению тока в пределах одного порядка.

Даже при самосогласованном моделировании теоретически предсказываемое пиковое напряжение не достигает значений, измеренных в эксперименте. Это принято объяснять наличием последовательно включенного сопротивления пассивных областей. Зависимость этих сопротивлений от напряжения нелинейна. Обычно даже в диодах с симметричной квантовой системой пиковые напряжения при разных полярностях расходятся не менее, чем на 0,1 В. Насколько нам известно, системные работы по анализу влияния пассивных областей на ВАХ отсутствуют, хотя некоторые попытки осуществлялись [7.59].

Сложнее всего предусмотреть форму зависимости $J(V)$ в диапазоне между пиковым и долинным напряжениями. На этом участке дифференциальная проводимость преимущественно отрицательна, однако во многих РТД ее величина может изменяться так, что образуются один или больше участков, на которых дифференциальная проводимость почти нулевая. Наличие таких участков прослеживается для всех РТД независимо от материалов и технологий. Обычно их называют участками "плато", поскольку они имеют вид почти горизонтальных линий на склоне N -образной ВАХ. Существующие модели РТД, включая вышеописанные, предусматривают монотонное падение тока на участке ОДП⁹. Отсюда следует, что "плато" на ВАХ обусловлено механизмами транспорта, которые не учитываются в модели. Из этого предположения возникло ряд теорий, которые вводили в существующие модели факторы, которые, по мнению авторов упомянутых моделей, ответственны за образование области «плато». Высказывались мысли о том, что этот участок связан с внутренней [7.60] и внешней [7.61] нестабильностью, с Γ - X - Γ переносом в рамках двухдолинной модели [7.55]. В то же время в [7.62] были получены экспериментальные подтверждения того, что формирование этих областей обусловлено метастабильными состояниями в квазитреугольной эмиттерной квантовой яме.

⁹ Исключение составляют модели, разработанные в рамках формализма функций Вигнера, которые стойко предусматривают такую область. Однако наличие области плато является едва ли не единственным экспериментальным фактом, который они предсказывают, а сам формализм функций Вигнера критикуют как теоретически недостаточно обоснованный [7.63]

Однако учет указанного механизма транспорта для моделей РТД стало настоящим вызовом, поскольку необходимо было описать процессы неупругого рассеивания с учетом размерного квантования. Пока что это удалось осуществить лишь с помощью феноменологической модели [7.47]. Такая модель действительно предусматривает формирование участка плато по экспериментально верифицированному механизму, однако количественного согласования получить не удалось, отчасти потому, что модель не предусматривает область плато в рамках самосогласованной модели, а также не обеспечивает выполнения закона сохранения тока.

Также замечено, что самосогласованные модели могут оказываться менее адекватными, чем модели, в которых принимается предположение о линейном падении потенциала в активной области. Два случая описаны выше:

1) на ВАХ РТД₂, рассчитанной самосогласованно, отсутствует область "плато", обусловленная наличием ступенчатого эмиттера;

2) участок «плато» на ВАХ резонансно-туннельных диодов с однородным стехиометрическим составом эмиттера, обусловленный транспортом между эмиттерной и основной квантовыми ямами [7.47], также предусматривается лишь в рамках несамосогласованной модели.

Кроме того, совпадение по пиковым и долинным токам лучше достигается преимущественно в рамках несамосогласованной модели; лучше также согласуется и формы ВАХ. Таким образом, едва ли не единственным аргументом, который указывает на целесообразность использования метода Хартри для учета пространственного заряда есть то, что предсказуемое им пиковое напряжение ближе к экспериментально наблюдаемому.

Литература к главе 7

7.1. *Абрамов И.И.* Комбинированная модель резонансно-туннельного диода / И.И. Абрамов, И.А. Гончаренко, Н.В. Коломейцева // Физика и техника полупроводников. – 2005. – Том 39, Вып. 9. – С. 1138–1145.

7.2. *Ландау Л.Д.* Теоретическая физика: учеб. пособ. для вузов в 10 томах. Том 3. Квантовая механика (нерелятивистская теория) / Лев Давыдович Ландау, Евгений Михайлович Лифшиц. – 5-е изд. стереот. – М.: Физматлит, 2002. – 808 С.

- 7.3. *Блохинцев Д.И.* Основы квантовой механики, учебное пособие, изд. 5-е переработанное / Д.И. Блохинцев. – М.: Наука, 1976. – 664 с.
- 7.4. *Иогансен Л.В.* О возможности резонансного прохождения электронов в кристаллах через системы барьеров / Л.В. Иогансен // Журнал экспериментальной и теоретической физики. – 1963. – Т. 45, вып. 2. – С. 207–213.
- 7.5. *Иогансен Л.В.* Тонкопленочные электронные интерферометры / Л.В. Иогансен // Журнал экспериментальной и теоретической физики. – 1965. – Т.86, вып. 1. – С. 175–179.
- 7.6. *Chang L.L.* Resonant tunneling in semiconductor double barriers / L.L. Chang, L. Esaki, R. Tsu // Appl. Phys. Lett. – 1974. – Vol. 24. P. 593–595.
- 7.7. *Shewchuk T.J.* Resonant tunneling oscillations in a GaAs-Al_xGa_{1-x}As heterostructure at room temperature / T.J. Shewchuk, P.C. Chapin, P.D. Coleman // Applied Physics Letters. – 1985. – Vol. 46. – P. 508–510.
- 7.8. *Luryi S.* Mechanism of operation of double-barrier resonant-tunneling oscillators / S. Luryi // Electron Devices Meeting, 1985 International. – 1985. – Vol. 31. – P. 666–669.
- 7.9. *Boykin T.B.* Resonant tunneling diodes with emitter prewells / T.B. Boykin, R.C. Bowen, G. Klimeck // Appl. Phys. Lett. – 1999. – Vol. 75. – P. 1302–1304.
- 7.10. *Герасимов С.М.* Физические основы электронной техники / С.М. Герасимов, М.В. Белоус, В.А. Москалю. – К.: "Вища школа", 1981. – 368 с.
- 7.11. *Хьюеу Н.В.* Основы метода вторичного квантования / Нгуен Ван Хьюеу. – М.: Энергоатомиздат, 1984. – 208 с.
- 7.12. *Tsu R.* Tunneling in a finite superlattice / R. Tsu, L. Esaki // Appl. Phys. Letters. – 1973. – Vol. 22. – P. 562–564.
- 7.13. *Saleh B.E.A.* Fundamentals of photonics / B.E.A. Saleh, M.C. Teich. New Jersey: John Wiley & Sons Inc, 2007. – 1178 p.
- 7.14. *Araki K.* Analysis of barrier transmission in resonant tunneling diodes / Kinichiro Araki // Journal of Applied Physics. – Vol. 62, No. 3, P. 1059–1069.
- 7.15. *Тагер А.С.* Размерные эффекты в субмикронных полупроводниковых структурах и перспектива их использования в электронике СВЧ / А.С. Тагер // Электронная техника. Сер. „Электроника СВЧ”. – 1987. – Вып. 9. – С. 21–33.
- 7.16. *Moskaliuk V.* Simplified analytical model of resonant-tunneling diode / V. Moskaliuk, A. Fediai // Abstracts Proceedings of 32nd International Spring Seminar on Electronics Technology "ISSE 2009" . – Brno, Czech Republic, 2009. – P. 1–5.

- 7.17. Москалюк В.А. Метод нахождения резонансных уровней энергии при многозонном моделировании гетероструктур / В.А. Москалюк, А.В. Федяй // Техника и приборы СВЧ. – 2011. – №1. – С. 43–48.
- 7.18. Москалюк В.О. Фізика електронних процесів. Динамічні процеси / В.О. Москалюк. Київ: "Політехніка", 2004. – 180 с.
- 7.19. Менский М.Б. Квантовые измерения и декогеренция / М.Б. Менский. М: Физматлит, 2001. – 232 с.
- 7.20. Поздняков Д.В. Расчет вольт-амперных характеристик симметричных двухбарьерных резонансно-туннельных структур на основе арсенида галлия с учетом процессов разрушения когерентности электронных волн /. Д.В. Поздняков, В.М. Борздов, Ф.Ф. Комаров // ФТП. – 2004. – Т. 38, № 9. – С. 1097–1100.
- 7.21. Brennan K.F. Theory of resonant tunneling in a variably spaced multiquantum well structure: An Airy function approach / K.F. Brennan, C.J. Summers // J. Appl. Phys. – 1987. – Vol. 61. – P. 614–623.
- 7.22. Федяй А.В. Моделирование резонансно-туннельного диода методом функций Эйри / А.В. Федяй, И.С. Тисный // Н.-техн. сб. «Электроника и связь», тематический выпуск «Электроника и нанотехнологии», ч.1. – 2009. – № 2-3, С. 19–21.
- 7.23. Мар'єнков В.В. Моделювання РТД з параболічними потенціальними ямами: бакалаврська робота. Режим доступу до ресурсу: <http://phbme.ntu-kpi.kiev.ua/~feduay/2b.pdf>.
- 7.24. Самарский А.А. Введение в численные методы / А.А. Самарский. – М.: «Наука», 1987. – 271 с.
- 7.25. Самарский А.А. Введение в теорию разностных схем / А.А. Самарский. – М.: «Наука», 1971. – 553 с.
- 7.26. Lent C.S. The quantum transmitting boundary method / C.S. Lent, D.J. Kirkner // Journal of Applied Physics. 1990. – Vol. 67. – P. 6353–6359.
- 7.27. Хокни Р. Численное моделирование методом частиц: Пер. с англ. / Р. Хокни, Дж. Иствуд. – М.: Мир, 1987. – 640 с.
- 7.28. Pinaud O. Transient simulations of a resonant tunneling diode / O. Pinaud // Journal of Applied Physics. – 2002. – Vol. 92. P. 1987–1994.
- 7.29. Москалюк В.О. Фізика електронних процесів, частина І: навчальний посібник / В.О. Москалюк. – Київ: "Політехніка", 2004. – 148 с.
- 7.30. Gummel H.K. A self-consistent iterative scheme for one-dimensional steady state transistor calculation / H.K. Gummel // IEEE Transactions on Electron Devices – 1964. – Vol. 11, P 455–465.

- 7.31. *Sun J.P.* Resonant tunneling diodes: models and properties / J.P. Sun, R.K. Mains, G.I. Haddad // Proc. of IEEE. – 1998. – Vol. 86. – P. 641–661.
- 7.32. *Федяй А.В.* Моделі резонансно-тунельних діодів (аналітичний огляд). Когерентні моделі / А.В. Федяй // Н.-техн. Ж. «Електроника и связь» – 2007. – №4. – С. 17–29.
- 7.33. *Федяй А.В.* Моделі резонансно-тунельних діодів (аналітичний огляд). Кінетичні моделі / А.В. Федяй // Н.-техн. Ж. «Електроника и связь». – 2007. – №5. – С. 11–16.
- 7.34. *Anantram M.P.* Modeling of Nanoscale Devices / M.P. Anantram, M.S. Lundstrom, D.E. Nikonov // Proceedings of the IEEE. – 2008. – Vol. 96, P. 1511–1550.
- 7.35. *Москалюк В.О.* Проблеми моделювання компонентів наноелектроніки на прикладі резонансно-тунельного діода / В.О. Москалюк, А.В. Федяй. // Н.-техн. ж. «Вісник ДУІКТ». – 2009. – Т. 7, №4. – С. 424–429.
- 7.36. *Buttiker M.* Coherent and sequential tunneling in double barrier / M. Buttiker // IBM. J. Res. Develop. – 1988. – Vol. 32. – P. 63–75.
- 7.37. *Zohta Y.* Improved optical model for resonant tunneling diode / Y. Zohta, T. Tanamoto // J. Appl. Phys. – 1993. – Vol. 74. – P. 6996–6998.
- 7.38. *Sun J.P.* Self-consistent scattering calculation of Resonant Tunneling Diode Characteristics / J.P. Sun, G.I. Haddad // VLSI Design. – 1998. – Vol. 6. – P. 83–86.
- 7.39. *Mains R.K.* Observation of intrinsic bistability in resonant tunneling diode modeling / R.K. Mains, J. P. Sun, G.I. Haddad // Appl. Phys. Lett. – 1988. – Vol. 55, P. 371–373.
- 7.40. *Diamond S.K.* Fabrication of 200-GHz f_{max} Resonant-Tunneling Diodes for Integrated Circuit and Microwave Applications / S. Diamond, E. Ozbay, M.J.W. Rodwell et. al. // IEEE Electron device letters. – 1989. – Vol. 10, No. 3. – P. 104–106.
- 7.41. *Wolak E.* The design of GaAs/AlAs resonant tunneling diodes with peak current densities over 2×10^5 A cm⁻² / E. Wolak, E. Ozbay, B.G. Park et. al. // J. Appl. Phys. – 1991. – Vol. 69. – P. 3345–3350.
- 7.42. *Tsu R.* Superlattice to nanoelectronics / Ralf Tsu // London: Elsevier, 2005. – 325 P.
- 7.43. *Martin A.D.* Observation of intrinsic tristability in a resonant tunneling structure / A.D. Martin; M.L.F. Lerch; P.E. Simmonds; L. Eaves // Applied Physics Letters. – 1994. – Vol. 64. – P. 1248–1250.

- 7.44. *Wei T.* Capacitance and hysteresis study of AlAs/GaAs resonant tunneling diode with asymmetric spacer layers / T. Wei, S. Stapleton // *J. Appl. Phys.* – 1995. – Vol. 77. – P. 4071–4076.
- 7.45. *Караваяев Г.Ф.* "Сверхрешеточная" модель плавной гетерограницы GaAs/AlAs (001) / Г.Ф. Караваяев, С.Н. Гриняев // *Физика твердого тела.* – 2006. – Том. 48, Вып. 5. – С. 893–901.
- 7.46. *Baro M.A.* 1D coupled Schrodinger drift-diffusion model including collisions / M. Baro, N. Ben Abdallah, P. Degond, A. El Ayyadi // *Journal of Computational Physics.* – 2005. – Vol. 203. – P. 129–153.
- 7.47. *Москалюк В.А.* Моделирование артефактов на ВАХ резонансно-туннельного диода / В.А. Москалюк, А.В. Федяй // 21-я Международная конференция «СВЧ-техника и телекоммуникационные технологии». – Севастополь. – 12–16 сентября 2011 г. – с. 787–788.
- 7.48. *Bowen R.C.* Quantitative simulation of resonant tunneling diode / R.C. Bowen, G. Klimeck, R.K. Lake, W.R. Frensley, T. Moise // *J. Appl. Phys.* – 1997. – Vol. 81. – 3207–3213.
- 7.49. *Schulman J.N.* Extension of Tsu-Esaki model for effective mass effects in resonant tunneling / J.N. Schulman // *Appl. Phys. Lett.* – 1998. – Vol. 72. – P. 2829–2831.
- 7.50. *Vurgaftman I.* Band parameters for III-V compound semiconductors and their alloys / I. Vurgaftman, J.R. Meyer // *J. Appl. Phys.* – 2001. – Vol. 89, №11. – P. 5815–5862.
- 7.51. Электронный архив «Новые полупроводниковые материалы: Характеристики и свойства». Режим доступа до ресурсу: <http://www.ioffe.ru/SVA/NSM/rintroduction.html>.
- 7.52. *Liu H.C.* Resonant tunneling through single layer heterostructure / H.C. Liu // *Appl. Phys. Lett.* – 1987. – Vol. 51 – P. 1019–1021.
- 7.53. *Sun J.P.* A self-consistent model of Γ -X mixing in GaAs/AlAs/GaAs quantum well using quantum transmitting boundary method / J.P. Sun, R.K. Mains, K. Yang, G.I. Haddad // *J. Appl. Phys.* – 1993. – Vol. 74. – P. 5053–5060.
- 7.54. *Васько Ф.Т.* Электронные состояния и оптические переходы в полупроводниковых гетероструктурах / Федор Трофимович Васько. – К.: Наукова Думка, 1993. – 181 с.
- 7.55. *Абрамов И.И.* Комбинированная двухзонная модель резонансно-туннельного диода / И.И. Абрамов, И.А. Гончаренко // *Физика и техника полупроводников.* – 2007. – Т. 41, № 11. – С. 1395–1400.
- 7.56. Резонанси Фано (матеріал з Вікіпедії). Режим доступа до ресурсу: http://ru.wikipedia.org/wiki/Резонанс_Фано.

- 7.57. *Evstigneev S.V.* Multiple-barrier resonant tunneling structures for application in a microwave generator stabilized by microstrip resonator / S.V Evstigneev, A.L. Karuzskii, Yu.A. Mityagin et.al. / 8th Int. Symp. "Nanostructures: Physics and Technology". *St Petersburg, Russia, June 19–23, 2000.:2000.* – P. 494–497.
- 7.58. *Yatskiv R.* Analysis of the resonant tunneling diode with the stepped pre-barrier / R. Yatskiv, J. Voves // *Journal of physics: conference series.* – 2009. – Vol. 193. – P. 1–4.
- 7.59. *Обухов И.А.* Моделирование переноса заряда в мезоскопических структурах / И.А. Обухов. – Севастополь: «Вебер», 2005. – 226 С.
- 7.60. *Wolak E.* The design of GaAs/AlAs resonant tunneling diodes with peak current densities over 2×10^5 A cm⁻² / E. Wolak, E. Ozbay, B.G. Park et. al. // *J. Appl. Phys.* – 1991. – Vol. 69. – P. 3345–3350.
- 7.61. *Young J.F.* Effect of circuit oscillations on dc current-voltage characteristics of double barrier resonant tunneling structures / J.F. Young, B.M. Wood, H.C. Liu et. al. // *Appl. Phys. Lett.* – 1988. – Vol. 52.– P. 1398–1400.
- 7.62. *Qiu Z.J.* Experimental verification of origin of plateau-like current-voltage characteristics of resonant tunneling diodes / Z.J. Qiu, Y.S. Gui, S.L. Guo et. al. // *Appl. Phys. Lett.* – 2007. – Vol. 4. – P. 1961–1963.
- 7.63. *Татарский В.И.* Вигнеровское представление квантовой механики / В.И. Татарский // *Успехи физических наук.* – 1983. – Т. 139, вып. 4. – С. 587–619.

8. ПЕРСПЕКТИВНЫЕ НАНОСТРУКТУРЫ И НАНОКОМПОНЕНТЫ

8.1. СРАВНИТЕЛЬНЫЕ ХАРАКТЕРИСТИКИ НАНОЧАСТИЦ ДЛЯ ПРИМЕНЕНИЯ В ЭЛЕКТРОНИКЕ

Одно из направлений повышения быстродействия электронных компонентов – использование новых физических эффектов в перспективных материалах и наноразмерных структурах.

Обычно для создания нанoeлектронных компонентов используются освоенные современными микроэлектронными технологиями так называемые структуры низкой размерности: двумерные в виде слоев наноскопических размеров, одномерные – квантовые нити или провода, и нульмерные – квантовые точки. Но в качестве современных наноконструктов электроники могут также использоваться различные по своим физическим и химическим свойствам структуры и наноматериалы. Это, как правило, различные по составу вещества, состоящие из атомов и молекул, имеющие различные химические связи и строение кристаллической решетки, а также некристаллические соединения и биообъекты – нуклеиновые кислоты, протеины, вирусы, клетки (рис. 8.1). Современные наноструктуры в качестве электронных компонентов используют различные сочетания приведенных материалов. Уже известны практические структуры наносенсоров, использующие свойства молекул изменять проводимость канала полевого транзистора на основе нанопроводника (нанопроволоки) [8.1].

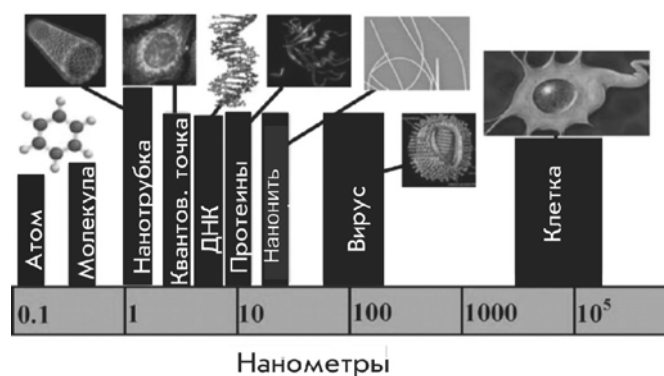


Рис. 8.1 Сравнительные размеры наноструктур и нанообъектов

Современные наноструктуры в качестве электронных компонентов используют нанобъекты в различных сочетаниях. Более подробное представление о свойствах частиц, молекул и нанобъектов дает таблица 8.1. В таблице представлены их характеристики, а также различные механизмы изменения их состояний. Интерес представляет анализ скоростных свойств (времен и энергий переключения), влияния шумов (термальной ошибки и различного рода флуктуаций), плотность компоновки структур, мощности рассеяния, методы модуляции и др.

Известно, что в электронных компонентах, использующих эффект поля, ограничивающими факторами являются: конечное время релаксации энергии (около 1 пс), термальные эффекты и квантовые эффекты. При этом «пороговая энергия» составляет $3k_B T \ln 2$ (около 0,05 эВ). К классу спинтронных компонентов могут быть отнесены структуры, использующие прецессию спина одиночного электрона, релаксационные процессы спиновой монообласти или спиновые волны, возникающие под воздействием магнитного поля. В первом случае время переключения ограничивается граничной частотой прецессии спина электрона, во втором – временем релаксации спина домена (около 10 пс), а в последнем (до 10 пс) – максимальной частотой изменения магнитного поля. Третья группа представленных в таблице 8.1 возможных для реализации компонентов использует физические и молекулярные изменения – движение молекул (одной из перспективных представляется молекула ротаксана [8.2]), вращение молекул, фазовые изменения в материалах. Скорость переключения в таких структурах ограничивается частотой вибрации молекул, которая для первого и второго случая составляет около 0,1 пс, а для случая фазовых изменений – 1 нс. Поскольку физические размеры молекул существенно больше электронных и спинтронных структур, требуются более существенные энергетические затраты («энергия переключения» возрастает на порядки), связанные с модуляцией характеристик молекулярных структур, при этом, однако, возможно получение более мощных компонентов. Следует также отметить, что вероятность термальной ошибки у таких компонентов на несколько порядков меньше, чем у электронных и спинтронных структур.

Существуют физические факторы масштабных ограничений для наноструктур с различными механизмами модуляции характеристик. Для электронных структур и структур с «одиночным электронным спином» это температурные и квантовые эффекты, для структур со спиновой монообластью и спинволнами – это проявление эффекта сверхпарамагнетизма и ограничение скорости распространения спиновых волн.

Таблица 8.1. Характеристики частиц и нанобъектов [8.2]

объект	электрон	электронный спин				физические молекулярные изменения			
		одиночный	спиновая монообласть	спинволна	движение молекул	вращение молекул	фазовые изменения		
скорость переключения	1 пс время релаксации энергии	1пс частота прецессии	10 пс, время релаксации домена	10 пс, внешнее магнитное поле	0,1 пс, частота вращения	0,1 пс, частота вращения	1 нс, частота вибрации		
Энергия переключения	0,05 эВ 1 электрон	0,05 эВ 1 спин	0,05 эВ	0,03 эВ	0,8 эВ 521 атом	0,7 эВ 115 атомов	3 эВ, 54 атома		
вероятность термальной ошибки	0.6	0.6	0.6	0.75	0,0001	0,0001	0,0001		
плотность компоновки элем./см ²	10 ¹⁴	10 ¹⁴	10 ¹¹	10 ⁹	10 ¹²	10 ¹²	10 ¹³		
мощность рассеяния	мВт/см ²	мВт/ см ²	кВт/ см ²	Вт/ см ²	мВт/ см ²	мВт/ см ²	100 кВт/ см ²		
масштабный фактор ограничения	температ. и квантовые эффекты	и температур. квантовые эффекты	Сверхпарамагнитность	скорость распространения	стабильность молекул, размеры, сборка	стабильность молекул,размеры, сборка	рассеяние тепла		
Метод модуляции	эффект поля	магн. поле, спин-орбит. взаимодействие.	магн. поле, спинорбит. и магн.-электр. взаимодействие.	электрическое притяжение	фотоны, тепло, электроны	фотоны, тепло, электроны	Джоулево тепло		
преимущественный материал	A ³ B ⁵ , Si и др.	неизвестно	ферромагнетики	ферромагнетики	ротаксан	металлоуглероды	калгониды		

Для молекулярных структур это проблемы стабильности молекул и сборки (технологии получения структур с повторяемыми характеристиками), а для структур, использующих фазовые изменения – проблема отвода тепла. Технологической проблемой является также проблема компоновки и сборки наноконструкций интегральных схем с существенно различными свойствами – это могут быть полупроводники и углеродные материалы (фуллерены), ферромагнетики, молекулы органических и неорганических соединений, металлоуглероды, калгониды и др.

В работе [8.3] приводятся результаты экспериментальных исследований контакта «органическая молекула-металлическая поверхность». Как предполагается, имеет место химический механизм взаимодействия между молекулой и металлическим электродом, связанный с изменением электронной плотности и появлением молекулярного электростатического потенциала. Рассматривается система из двух связанных атомов углерода, соединенных с кольцами фенила (рис. 8.2). При взаимодействии происходят изменения в геометрии молекулы, что характерно для химически связанных элементов, которые могут происходить на расстоянии около 1 нм. Рассмотренные модели процессов предполагают образование квантовомеханических кластеров у поверхности металлического электрода. При этом считается, что волновые функции атомов металла локализуются, т.е. взаимодействие определенного слоя металла не изменяет волновые функции соседних слоев металлического электрода.

Модуляция проводимости контакта «металл-молекула» происходит при изменении электрического смещения на металлической поверхности, что приводит к изменению длины связей и переносу электронов вследствие туннелирования. Механизм модуляции проводимости вследствие поверхностной поляризации углеродной нанотрубки, представляющей собой канал полевого транзистора, реализован в нанобиосенсоре (рис. 8.3)

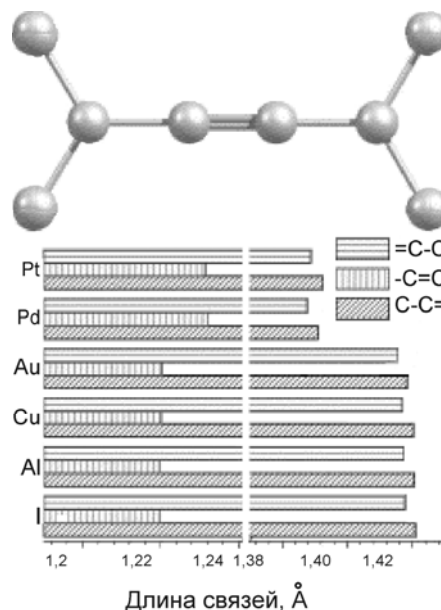


Рис. 8.2 Длины связей контакта металл-молекула для различных материалов

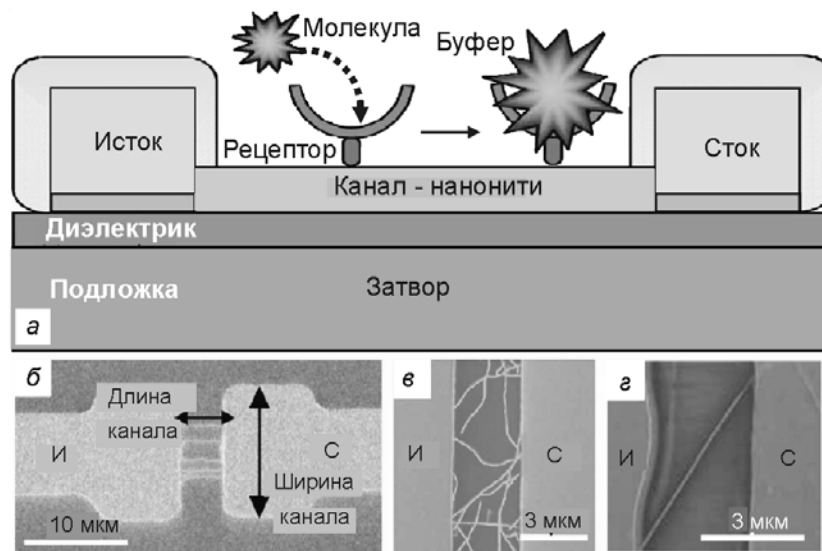


Рис. 8.3 Принцип работы нанобиосенсора – (а), топология– (б), технологии реализации канала транзистора – на углеродной нанотрубке (в), на нанопроводнике из оксида индия – (г).

При захвате органических молекул на поверхности нанотрубки происходит обеднение (или обогащение) носителями канала транзистора в зависимости от типа проводимости нанотрубки. Такие биосенсоры могут использоваться для исследования протеинов, ДНК, различного вида антител и других биообъектов.

8.2. НАНОКОМПОНЕНТЫ С ИСПОЛЬЗОВАНИЕМ ТУННЕЛЬНОГО ЭФФЕКТА

Одноэлектронный транзистор с уникальными характеристиками, выполненный в структуре «кремний-на-изоляторе» с туннельными переходами и встроенными квантовыми точками, представлен в работе [8.4]. При этом туннельные переходы и квантовые точки существенно меньше поликристаллических кремниевых гребенчатых затворов с длиной около 70 нм, а межэлектродные расстояния составляют около 240 нм. Топология и параметры транзистора приводятся на рис. 8.4. Данная технология позволяет получать квантовые точки с контролируемыми параметрами. Область под затвором имеет профиль легирования в виде обратного гауссиана, что приводит к созданию потенциального барьера. Эффективная работа транзистора связана с необходимостью превышения энергии кулоновского взаимодействия над энергией термальных флуктуаций. Если за счет теплового движения частица приобрела достаточно большую энергию, она может прорвать кулоновскую блокаду. Поэтому для каждого одноэлектронного устройства существует своя критическая температура, выше которой она

перестает работать. Эта температура обратно пропорциональна площади перехода: чем меньше его емкость, тем больше скачок напряжения и тем выше барьер кулоновской блокады.

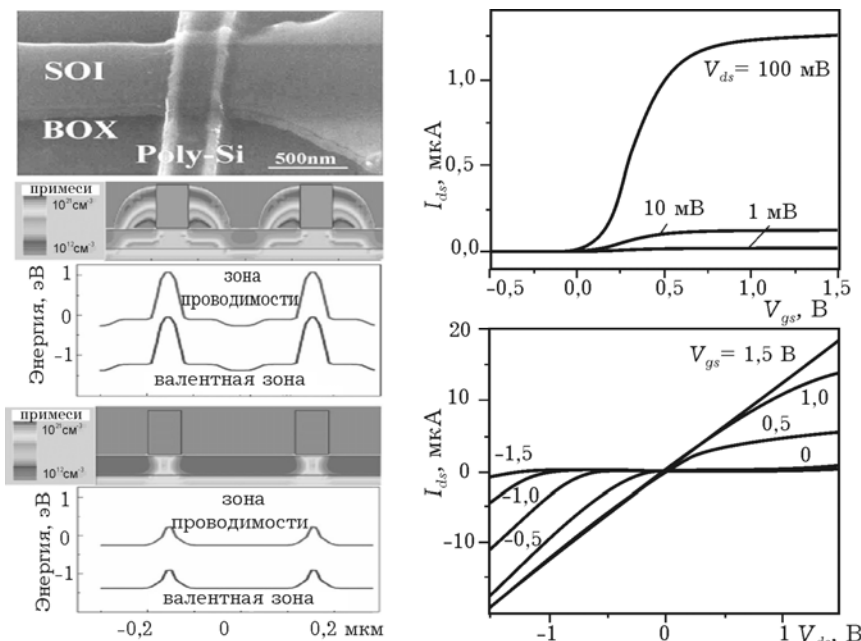


Рис. 8.4 Топология одноэлектронного транзистора, энергетические диаграммы и вольтамперные характеристики при $T = 300$ К.

Основные характеристики одноэлектронного транзистора представлены на рис. 8.4. Уникальность рассматриваемого в данной работе транзистора заключается в получении выходных характеристик при комнатной температуре.

В работе [8.5] представлена упрощенная аналитическая модель для расчета тока кремниевого одноэлектронного транзистора с дискретными квантовыми уровнями энергии и параболическими барьерами туннельных переходов. В модели рассматриваются только три самых вероятных состояния от основного уровня и первый возбужденный уровень для каждого электрона. Считается также, что сопротивление туннельного перехода больше, чем квантовое сопротивление $h/e^2 = 25,8$ кОм и квантовые флуктуации подавляются. Кроме того, отсутствует туннелирование более высокого порядка – «сотуннелирование», при котором сохраняется энергия лишь между начальным и конечным состояниями всего массива переходов, а также не учитывается неупругое туннелирование, которое сопровождается генерацией или рекомбинацией электронно-дырочных пар. Как правило, в упрощенных моделях считается, что время туннелирования пренебрежимо мало по сравнению с другими характерными временами, включая интервалы между туннельными переходами отдельных частиц. Квантовые энергетические уровни дискретны (рис. 8.5) с интервалом ΔE_n , а

электростатическая энергия равна $E_c = e^2/C_\Sigma$, где C_Σ – суммарная емкость транзистора.

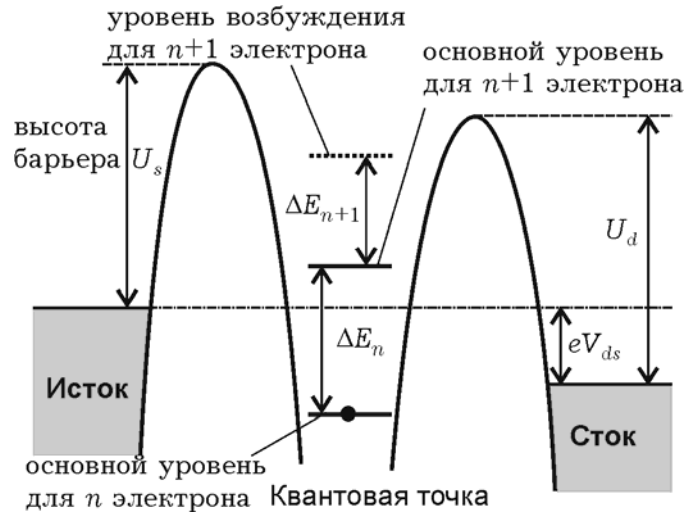


Рис. 8.5. Энергетическая диаграмма одноэлектронного транзистора.

Вероятности заполнения каждого из трех рассматриваемых состояний определяются по формулам:

$$P_{G,n} = \frac{\Gamma_{G,(n+1,n)} - \Gamma_{E,(n+1,n)}}{A}, P_{G,n+1} = \frac{\Gamma_{G,(n,n+1)} - \Gamma_{E,(n+1,n)}}{A}, P_{E,n+1} = \frac{\Gamma_{E,(n,n+1)} - \Gamma_{G,(n+1,n)}}{A},$$

где $A = \Gamma_{E,(n,n+1)}\Gamma_{G,(n+1,n)} + \Gamma_{G,(n,n+1)}\Gamma_{E,(n+1,n)} + \Gamma_{G,(n+1,n)}\Gamma_{E,(n+1,n)}$.

Здесь Γ – скорость туннелирования электронов из квантовой точки:

$$\Gamma = \Gamma_0 \left\{ [1 + \exp(-\Delta F / k_B T)] [1 + \exp(U - E) / T_0] \right\}^{-1},$$

где Γ_0 – нормированное значение скорости туннелирования, ΔF – разность свободной энергии системы, U – высота туннельного барьера, E – энергия электрона, T_0 – параметр формы туннельного барьера и T – температура кристалла.

Суммарная скорость туннелирования электронов на электроды истока и стока определяется по формуле $\Gamma_{G,(n,n+1)} = \Gamma_{G,(n,n+1)}^S + \Gamma_{G,(n,n+1)}^D$. Полный ток стока транзистора

определится как $I_{DS} = \sum_{k=0}^n I_{DS,k}$, где $I_{DS,n}$ – компоненты тока для каждого n -го электрона в квантовой точке, которые определяются по известным вероятностям и скоростям переходов как:

$$I_{DS,n} = \pm \left[e \left(P_{G,n} \Gamma_{G,(n,n+1)}^S - P_{G,n+1} \Gamma_{G,(n+1,n)}^S \right) + e \left(P_{G,n} \Gamma_{G,(n,n+1)}^S - P_{E,n+1} \Gamma_{E,(n+1,n)}^S \right) \right]$$

Энергетический зазор ΔF определяется по известным значениям суммарной емкости транзистора при заданных электрических смещениях на затворе и стоке транзистора. Приведенные в работе [8.5] выходные вольтамперные характеристики содержат участок с отрицательной дифференциальной проводимостью. Модель транзистора при известных параметрах квантовой точки и туннельного перехода может использоваться в пакете схемотехнического проектирования HSPICE для расчета быстродействующих ключей.

Для создания быстродействующей одноэлектронной структуры необходимо, чтобы сопротивление туннельного перехода было существенно выше квантового сопротивления. При выполнении этого условия можно пренебречь когерентными квантовыми процессами, связанными с одновременными туннельными переходами.

Когда энергия зарядки сравнима с энергией термических флуктуаций может проявляться эффект кулоновской блокады (кулоновской лестницы). Для приборов с размерами квантовой точки (кластера) до 10 нм изменение энергии переходов может достигать 0,1 эВ и одноэлектронные эффекты могут проявляться при комнатных температурах. При туннелировании возможно прохождение дробного заряда (дробного квантования), что может быть связано с тем, что волновая функция электрона может распределяться между электродами, влияя на энергетическую картину.

Для моделирования таких приборов недостаточно квазиклассического описания на уровне дрейфово-диффузионных или кинетических моделей – необходимо использование аппарата квантовой механики. Для описания неравновесных процессов необходимо самосогласованно решать уравнения Пуассона и Шредингера, что возможно, например, с привлечением функции Грина.

В работе [8.6] представлен транзистор, в котором реализован резонансно-туннельный эффект. Туннельные переходы индуцируются внешним полем как показано на рис. 8.6, где приведены энергетические диаграммы «туннельного» транзистора для различных смещений на затворе: на рис. 8.6,а туннельные переходы происходят, а на рис. 8.6,б – туннелирование отсутствует. Транзистор с длиной затвора 20 нм и шириной 50 нм выполнен по технологии «кремний на изоляторе» с n -каналом в виде нанопроволоки.

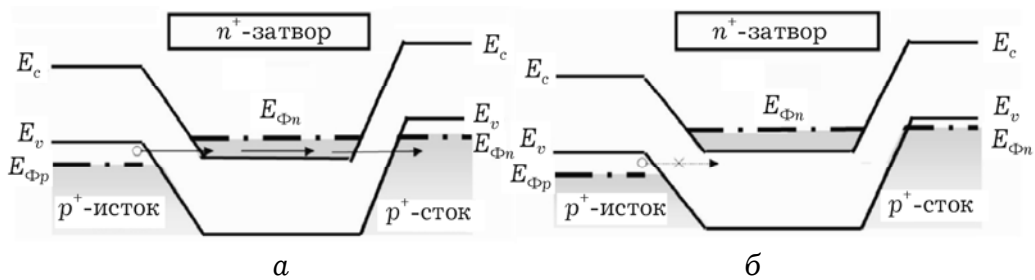


Рис.8.6. Энергетические диаграммы «туннельного» транзистора для различных смещений на затворе

8.3. НИЗКОРАЗМЕРНЫЕ СТРУКТУРЫ НА ОСНОВЕ ГРАФЕНА

В последнее время все больший интерес представляет использование уникальных свойств графена (рис. 8.7). В отличие от металлов и полупроводников, у которых закон дисперсии в зоне проводимости можно с хорошей точностью считать квадратичным (энергия квазичастиц прямо пропорциональна квадрату импульса), у графена закон дисперсии существенно другой. Графен нельзя отнести ни к металлам, ни к полупроводникам, ни, тем более, к диэлектрикам. Графен – полуметалл (не единственный, к такому классу веществ принадлежат еще бор, кремний, германий, мышьяк, полоний, графит и др.). С точки зрения зонной теории графен – это вещество, у которого в отличие от металлов валентная зона и зона проводимости соприкасаются (запрещенная зона отсутствует, как и в металлах). Точки соприкосновения в графене образуют так называемые точки Дирака (рис. 8.7, б). В этих точках закон дисперсии имеет линейный вид: энергия электронов прямо пропорциональна их импульсу.

Подобной зависимостью энергии от импульса обладают еще и безмассовые частицы – фотоны. Получается, что реальные частицы – электроны – в графене ведут себя как фотоны, то есть имеют нулевую массу и их энергия не зависит от скорости движения, поэтому закон дисперсии квазичастиц в графене называют релятивистским. Однако, в отличие от фотонов, электроны в графене имеют еще заряд и полуцелый спин, то есть являются фермионами, поэтому квазичастицы графена называют также *дираковскими фермионами*. Среди известных в настоящее время веществ нет аналогов с подобным «графенным» законом дисперсии.

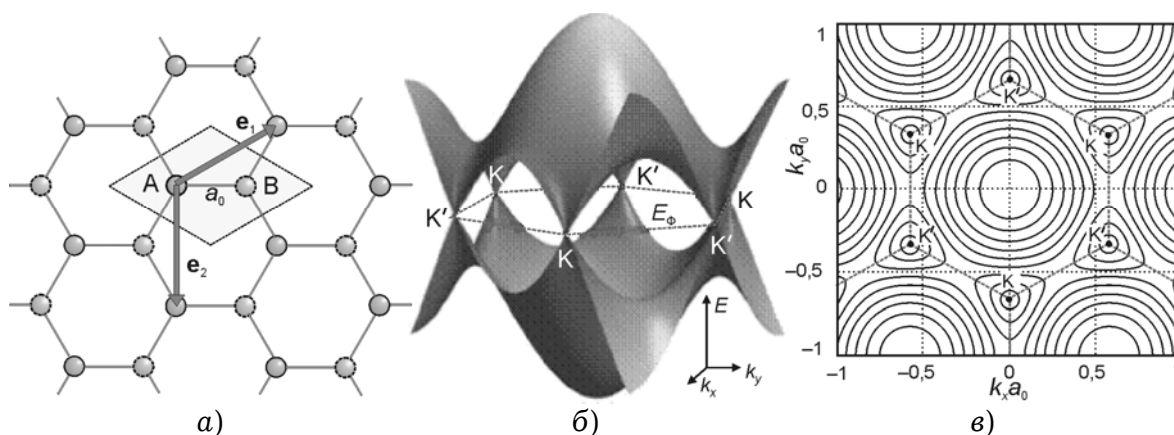


Рис. 8.7. Кристаллическая структура графена с двумя подрешетками (а), выделено элементарную ячейку и векторы трансляции. Закон дисперсии в трехмерном изображении (б) и с помощью изоэнергетических линий (в) с точками соприкосновения зоны проводимости и валентной зоны в точках K' и K ; штриховая линия соответствует энергии Ферми

Уровни Ландау впервые обнаружены в двумерной структуре проводника в магнитном поле. При этом электроны в таком проводнике будут двигаться по замкнутым орбитам с так называемой циклотронной частотой $\omega_c = eV/m$, зависящей от индукции магнитного поля. Энергия такого движения квантуется $E_n = (n + 1/2)\hbar\omega_c$ и пропорциональна квантовому числу n , обозначающему номер орбиты или уровня. Множество таких замкнутых орбит образуют уровни Ландау. Расстояние по энергетической шкале между соседними орбитами одинаковое $\hbar\omega_c$ – уровни Ландау эквидистантны. В случае нулевого уровня ($n = 0$) энергия электронов не становится равной нулю.

Квантование Ландау приводит к возникновению как минимум двух явлений: эффекта де Газа-ван Альфена и эффекта Шубникова-де Гааза. Первый заключается в осцилляциях намагниченности материала в сильном магнитном поле (порядка 1 Тл) при температурах жидкого гелия, второй – в осцилляциях проводимости при тех же условиях. Низкая температура и сильное магнитное поле позволяют избежать значительного влияния температуры на обнаружение квантования Ландау. В этих эффектах пики на осцилляциях физических величин и являются признаком образования уровней Ландау.

В графене, из-за линейности закона дисперсии, уровни Ландау $E_n = \pm\sqrt{2\hbar v_F^2 |n|} V$ квантуются совершенно иначе (рис.8.8,а):. Энергия каждого уровня отсчитывается от энергии в дираковской точке E_D – точке соприкосновения зон валентной и проводимости. Во-первых, энергия квазичастиц в графене зависит от номера орбиты как \sqrt{n} , поэтому уровни Ландау в нём не эквидистантны, и, во-вторых, при $n = 0$ энергия квазичастиц в графене равняется нулю даже в присутствии магнитного поля.

В работе [8.7] найдено подтверждение этих особенностей квантования Ландау в графене. Графен был выращен на подложке из карбида кремния SiC. В отличие от традиционного «фундамента» для получения графена (из оксида кремния), использование SiC позволило избежать ненужного влияния подложки на дальнейшие измерения. Фиксирование уровней Ландау в обсуждаемой работе происходило с помощью эффекта Шубникова-де Гааза. При изменении напряжения на исследуемом образце измерялась его проводимость в присутствии сильного магнитного поля при температуре 4,3 К. Результат одного из таких измерений – в магнитном поле с индукцией 5 Тл – показан на рис. 8.8,б.

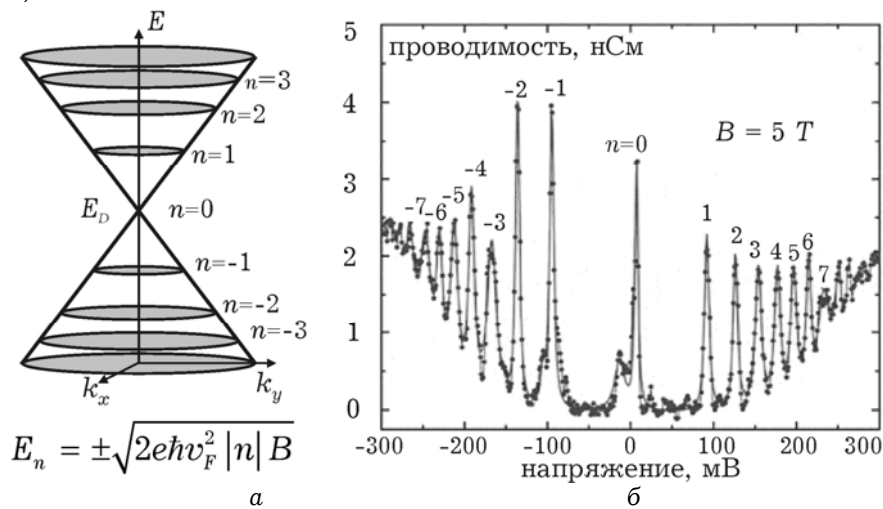


Рис. 8.8. Схематическое изображение закона дисперсии квазичастиц в графене в магнитном поле (а), или уровней Ландау для электронов (верхние окружности) и дырок (нижние окружности). Непосредственное измерение квантования Ландау в графене во внешнем магнитном поле 5 Тл (б). Экспериментально полученные данные показаны точками. Кривая – интерполяция экспериментальных результатов. Каждый всплеск на осциллирующей кривой соответствует уровню Ландау (нумерация указана на кривой).

По оси абсцисс отложено прикладываемое напряжение в энергетических единицах (мэВ), по оси ординат – измеренная проводимость графена (в наносименсах). Ярво выраженные пики на приведенной осциллирующей зависимости свидетельствуют о том, что экспериментально наблюдаются уровни Ландау. На этом графике видна и неэквидистантность расположения уровней Ландау.

Квазичастицы в графене формально описываются гамильтонианом дираковского вида, что обусловлено симметрией кристаллической решетки, которая состоит из двух эквивалентных углеродных подрешеток [8.8]:

$$H = -i\hbar v_F \sigma \nabla,$$

где $v_F \approx 10^6$ м/с – скорость фермионов, $\sigma = (\sigma_x, \sigma_y)$ – матрица Паули. Такое описание является теоретически строгим в пренебрежении многочастичными эффектами и подтверждено экспериментальными измерениями циклотронной массы носителей в графене в зависимости от их энергии.

В графене при положительных энергиях (выше дираковской точки) токонесущие состояния подобны электронам и заряжены отрицательно [8.8]. При отрицательных энергиях, если валентная зона целиком не заполнена, квазичастицы ведут себя как положительно заряженные частицы (дырки) и могут рассматриваться как твердотельный аналог позитронов. В противоположность тому, что электроны и дырки в физике твердого тела описываются уравнениями Шредингера со своими эффективными массами, в графене электроны и дырки оказываются взаимосвязанными, проявляя свойства зарядово-сопряженной симметрии. В графене это обусловлено симметрией его кристаллической решетки и тем, что квазичастицы в графене описываются двухкомпонентной волновой функцией, построенной на основе двух подрешеток.

В работах [8.9, 8.10] уникальные физические характеристики графена в структуре с *p-n* переходом используются для генерации и передачи плазмонных колебаний (волн плотности заряда) в терагерцовом диапазоне. Генерация плазмонов в графене подобна оптической генерации в результате стимулированной эмиссии фотонов в лазерах на многодолинных полупроводниках в терагерцовом диапазоне (рис. 8.9). Основное отличие заключается в низкой групповой скорости плазмонов в графене в терагерцовом диапазоне и сильном ограничении электромагнитного поля связанного с плазмонами вблизи слоя графена, что дает существенно большие значения усиления, чем в традиционных полупроводниковых лазерах. Энергия электромагнитного поля в графене ограничена вследствие малых поперечных размеров его слоя.

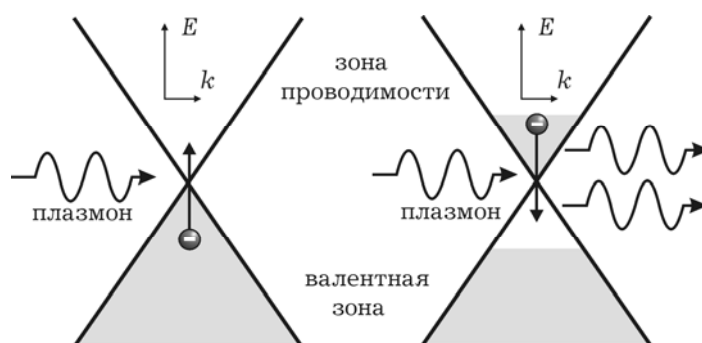


Рис. 8.9. Энергетическая модель, иллюстрирующая механизм генерации плазмонов в графене

Представленная в работе [8.9] модель распространения плазмонов в графене учитывает внутридолинное рассеяние и плазмонное усиление вследствие стимулированной эмиссии. Здесь же приводятся результаты моделирования характеристик линии передачи, представляющей собой следующую структуру (рис. 8.10). Неоднородное распределение электронов и дырок в слое графена обеспечивается как за счет их инжекции из n - и p -контактов структуры, так и за счет внутридолинного туннелирования в графене (при отсутствии запрещенной зоны), а управление осуществляется за счет электростатического смещения на затворах. Плазмонные волны, возникающие в слое графена локализуются в диэлектрической полоске, нанесенной на слой графена.

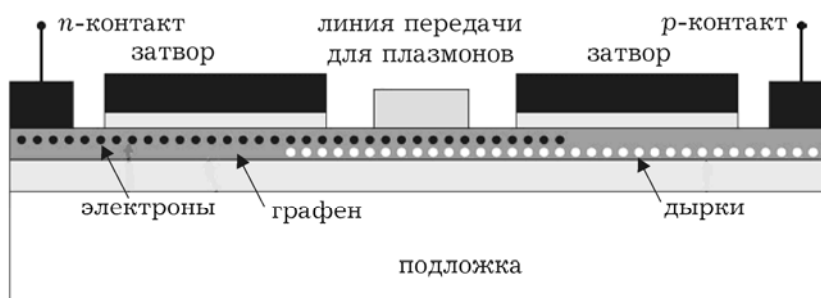


Рис. 8.10. Волноведущая активная плазмонная структура

В подобной структуре (ширина диэлектрика до 0,5 мкм при длине линии до 5 мкм, электронно-дырочной концентрации 10^{10} см^{-2} и при низких температурах до 77 К), возможно получение усиления сигнала в диапазоне частот 1–6 ТГц.

8.4 НАНОКОМПОНЕНТЫ С ИСПОЛЬЗОВАНИЕМ НАНОТРУБОК И НАНОПРОВОЛОК

Структура нанотрубки после сворачивания полосы графенового слоя в цилиндрическую поверхность без разрывов характеризуется хиральным вектором (n, m) , указывающим координаты бензольного кольца, совмещаемого с кольцом, находящимся в начале координат, при сворачивании слоя. Два типа нанотрубок – зигзагообразные (*zig-zag*) с хиральным вектором $(n, 0)$ и в конфигурации кресла (*armchair*) с хиральным вектором (n, n) – обладают кроме оси симметрии еще и плоскостью симметрии, перпендикулярной к оси нанотрубки. При этом нанотрубки могут обладать металлической или полупроводниковой проводимостью.

На рис. 8.11 показана структура транзистора на нанопроволоках [8.11]. Толщина нанопроводников (до 10 в структуре) в транзисторе на поликристаллическом кремнии

достигает 50 нм [8.11]. Такие структуры находят свое применение в трехмерных интегральных схемах, жидкокристаллических экранах, а также в «системах на кристалле».

Транзистор на нанотрубках схематически представлен на рис. 8.12. Транзистор представляет собой структуру [8.13] на углеродной трубке (CNT – carbon nanotube), встроенную в акцепторно легированные исток и сток. Фуллерены или нанотрубки представляют собой каркасные структуры (молекулы фуллерена C_{60}), которые при определенных условиях упорядочиваются и образуют молекулярный кристалл, в узлах которого находится молекула фуллерена. Нанотрубки имеют поперечный размер от 1 до 10 нм. Модуляция проводимости в однослойной структуре нанотрубки, имеющей характер проводимости полупроводника, осуществляется внешним поперечным электрическим полем.

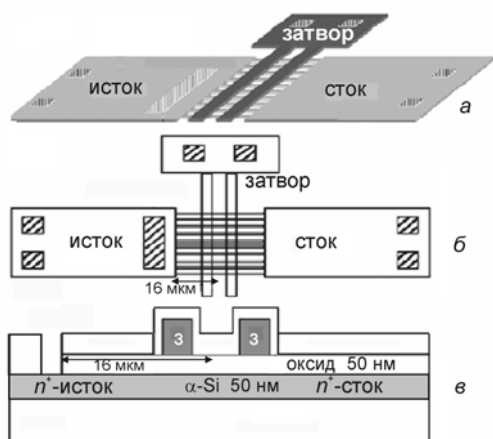


Рис. 8.11 Двухзатворный транзистор на нанопроволоках

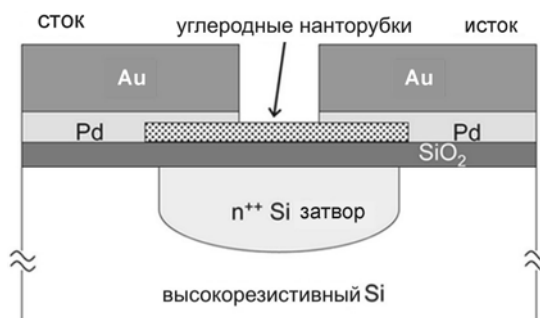


Рис. 8.12 Транзистор на нанотрубках

Колебания электронной плазмы в нанотрубке квантованы и могут быть описаны квазичастицами – плазмонами, в предположении отсутствия межэлектронных взаимодействий, когда коллективные колебания плазмы рассматриваются как колебания свободного электронного газа.

Технологически канал транзистора на нанотрубках может быть выполнен как одно-, двух- или трехмерная система. Так в работе [8.14] исследуются свойства транзисторов на нанотрубках, включающих одиночную нанотрубку (одномерная система), а также структур с нанотрубками, которые образуют двух- и трехмерные системы. При этом физические процессы для электронного газа могут рассматриваться как независимые в каждой отдельной нанотрубке. Трехмерные структуры позволяют получать большие значения плотности тока в транзисторе.

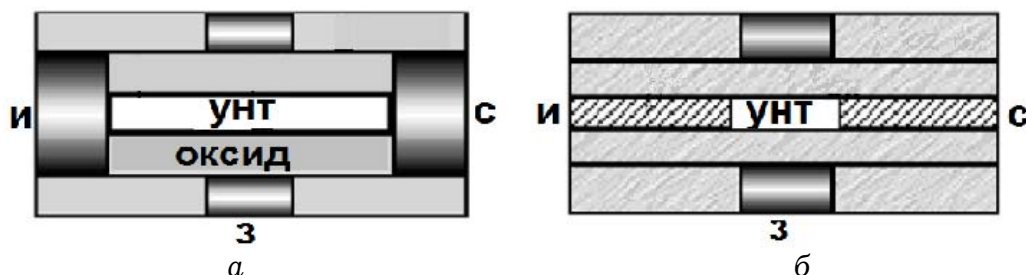


Рис. 8.13 Структуры полевых транзисторов с каналом на нелегированной (а) и ступенчато легированной нанотрубке (б).

В работе [8.15] предложены оригинальные конструкции полевых транзисторов на нанотрубках с донорным легированием. Полученные химическим путем нанотрубки со ступенчатым профилем легирования (минимум находится в середине нанотрубки) длиной 50 нм встроены в структуру транзистора так, что это позволяет достичь эффективной длины канала до 5 нм. Приводятся две конструкции транзисторов (рис. 8.13) с длиной канала 5 и 10 нм, свойства которых (табл. 8.2), позволяют получить уникальные частотные и временные характеристики.

Таблица 8.2

емкость затвора (аФ)	время пролета (фс)	длина затвора (нм)	ширина затвора (нм)	крутизна (мСм)	Граничная частота f_T (ТГц)
0,7	22	5	50	63	14,5
1,25	32	10	100	88	11

В частности, при достаточно высоких значениях крутизны граничная частота f_T превышает 10 терагерц при времени пролета носителей в канале 22 фс.

В последнее время перспективными как с точки зрения технологии, так и получения привлекательных выходных характеристик являются полевые транзисторы на нанотрубках на кремниевых подложках [8.16-8.18].

Система нанотрубок (рис. 8.14) в транзисторе позволяет увеличить ток через канал и увеличить напряжение отсечки тока стока, а также крутизну ВАХ. Использование палладиевых контактов уменьшает паразитные сопротивления стока и истока. Плотность нанотрубок в канале составляет примерно 1-3 на мкм, однако, на снимке со сканирующего электронного микроскопа обращает на себя внимание неупорядоченность нанотрубок. Дефекты нанотрубок ограничивают повторяемость параметров транзисторов из-за флуктуаций проводимости.

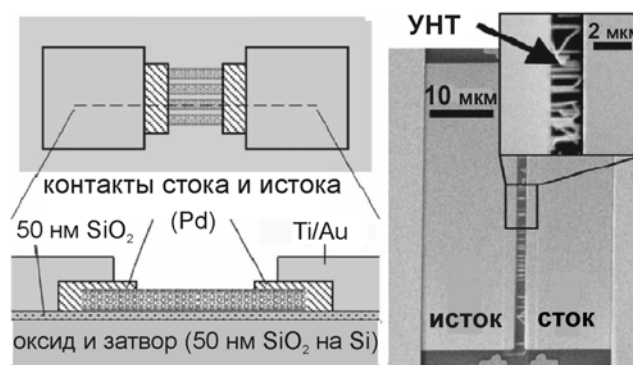


Рис. 8.14 Транзистор с системой нанотрубок [8.17]

Еще один пример использования сочетания легированных и нелегированных нанотрубок в транзисторных структурах приведен в работе [8.17]. Здесь используется система легированных нанотрубок для формирования канала транзистора в межэлектродном пространстве и нелегированных нанотрубок диаметром до 2 нм под управляющим затвором длиной около 30 нм и шириной 60 нм. Такая двухзатворная структура (рис. 8.15) реализует быстродействующий инвертор. Плотность нанотрубок в слое может достигать 250 на 1 мкм.

Критическим параметром для получения стабильных характеристик остается количество дефектных самоорганизованных нанотрубок в системе. Как уже отмечалось, деформации нанотрубок в процессе формирования могут приводить к заметному изменению величины проводимости, а в некоторых случаях и изменению характера проводимости.

В работе [8.19] приводится структура, экспериментальные характеристики и аналитическая модель для пакета схемотехнического проектирования PSPICE полевого транзистора на кремниевых нанонитях с барьерными контактами Шоттки стока и истока. Кроме того, особенностью конструкции транзистора является размещение канала в виде нанонити на лицевой, а изолированного затвора на обратной стороне транзистора. Кремниевые нанонити характеризуются высокой подвижностью носителей и проводимостью, поэтому в предлагаемой модели пренебрегается падением напряжения в канале. Для описания физических процессов на барьерных контактах используется аналитическая модель термоэлектронной полевой эмиссии для обратных электрических смещений и термоэлектронной эмиссии для прямых смещений.

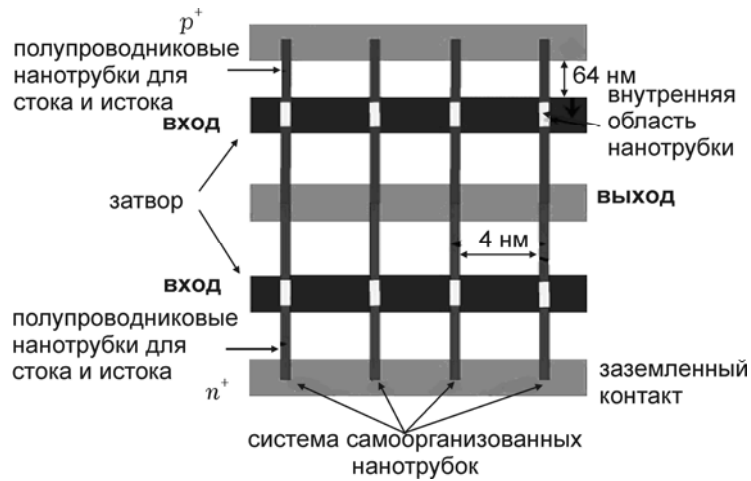


Рис. 8.15. Инвертор на транзисторах [8.18] с системой легированных и нелегированных нанотрубок

Анализируемая структура транзистора и соответствующая ей энергетическая диаграмма представлена на рис. 8.16.

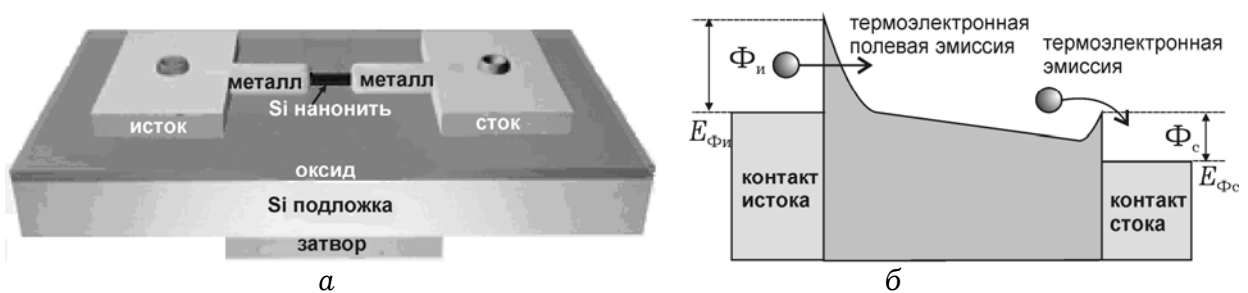


Рис. 8.16. Структура и энергетическая диаграмма транзистора с кремниевыми нанонитями

Уравнения для описания термоэлектронной полевой эмиссии включают:

$$I = S J_s \exp(-eV / E_t),$$

$$E_t = E_0' [E_0' / k_B T - \text{th}(E_0' / k_B T)],$$

$$J_s = \frac{A^* T^2 (\pi E_0')^{1/2}}{k_B T} [-eV + \phi_b / \text{ch}(E_0' / k_B T)]^{1/2} \exp(-\phi_b / E_0'),$$

$$E_0' = \hbar / 2 (N_d / m^* \varepsilon \varepsilon_0)^{1/2},$$

$$E_0 = E_0' \text{cth}(E_0' / k_B T),$$

где e – заряд электрона, V – падение напряжения на барьере Шоттки, S – площадь канала, A^* – константа Ричардсона, ϕ_b – высота барьера Шоттки, N_d – концентрация донорной примеси,

m^* – эффективная масса электронов в нанонити, $\epsilon\epsilon_0$ – диэлектрическая проницаемость нанонити.

Изменение высоты барьера Шоттки вследствие внешнего электрического поля определяется по формуле:

$$\varphi_b = \varphi_{b0} - \sqrt{e\mathcal{E}_m / 4\pi\epsilon\epsilon_0},$$

где φ_{b0} – высота барьера Шоттки в отсутствие электрического поля, \mathcal{E}_m – максимальное значение электрического поля контакта металл-полупроводник.

Одномерное уравнение Пуассона для определения φ_b для донорно легированной нанонити записывается в виде:

$$\frac{d^2\varphi(x)}{dx^2} - \frac{\varphi(x) - V_g + V_{bi}}{\lambda} = -\frac{eN_d}{\epsilon\epsilon_0},$$

где V_{bi} – встроенный потенциал, $\lambda = (\epsilon_s\epsilon_{ox}/t_{ox}t_s)^{1/2}$ – эффективная длина экранирования, ϵ_{ox} и t_{ox} – диэлектрическая проницаемость и толщина оксидного слоя, t_s – толщина нанонити.

Если потенциал аппроксимировать параболической функцией, решение уравнения Пуассона может быть представлено выражением:

$$\varphi(x) = [(V_{ds} + V_{bi} - \varphi_{s0})(e^{x/\lambda} - e^{-x/\lambda}) + (V_{bi} - \varphi_{s0}) \times (e^{(L-x)/\lambda} - e^{-(L-x)/\lambda})] / (e^{L/\lambda} - e^{-L/\lambda}),$$

где L – длина канала, $\varphi_{s0} = V_{gs} - V_{FB} - eN_d\lambda^2/\epsilon\epsilon_0$, V_{FB} – потенциал плоских зон.

Дифференцируя выражение для потенциала, можно найти распределение электрического поля $\mathcal{E}(x) = -d\varphi/dx$.

Максимальное значение электрического поля на контакте металл-полупроводник при $x = 0$ можно найти по соотношению:

$$\mathcal{E}_m = \mathcal{E}(0) = C_1V_{ds} + C_2V_{gs} + C_3$$

Здесь C_1 , C_2 , C_3 – подстроечные параметры. Подставляя последнее выражение в формулу для φ_b , получим:

$$\varphi_b = \varphi_{b0} - (AV_{ds} + BV_{gs} + C)^{1/2},$$

где A , B , C – константы, которые с учетом поверхностных состояний [8.20] могут быть определены как:

$$A = e^3 C_1 / 4\pi\epsilon\epsilon_0, \quad B = e^3 C_2 / 4\pi\epsilon\epsilon_0, \quad C = e^3 C_3 / 4\pi\epsilon\epsilon_0$$

Экспериментальные характеристики и результаты моделирования транзистора на нанонити в виде выходных вольтамперных характеристик и зависимости высоты потенциального барьера контакта металл-полупроводник от выходного напряжения представлены на рис. 8.17. Экспериментальные результаты показаны в виде символьных обозначений. Структура транзистора характеризуется следующими параметрами: длина канала 20 нм, ширина канала 2 мкм, концентрация донорной примеси $1,1 \cdot 10^{17} \text{ см}^{-3}$, толщина слоя оксида $t_{\text{ox}} = 100$ нм. Напряжение отсечки тока транзистора составило $V_t = -2$ В. Обращает на себя внимание заметная зависимость высоты потенциального барьера (рис.8.17) от приложенных питающих напряжений транзистора.

8.5 ПАССИВНЫЕ КОМПОНЕНТЫ НАНОИНТЕГРАЛЬНЫХ СХЕМ

В работах [8.21-8.23] приводятся примеры использования углеродных наноструктур в качестве межсоединений и линий передачи.

В работе [8.21] рассматривается линия передачи в виде

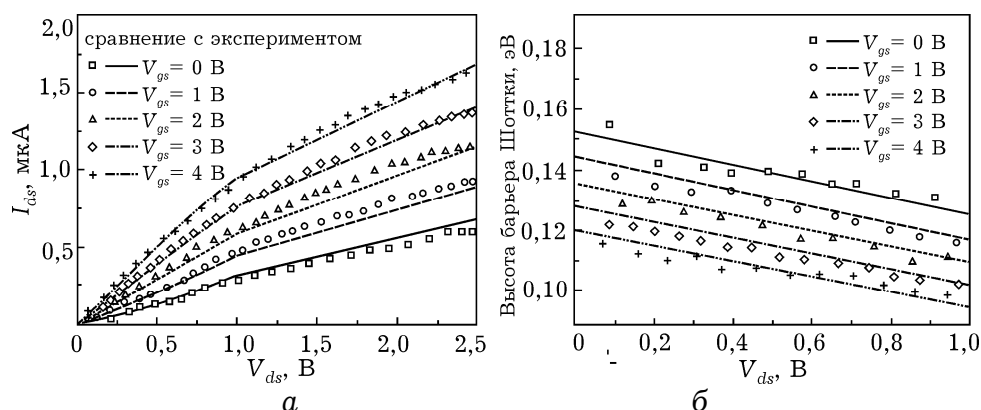
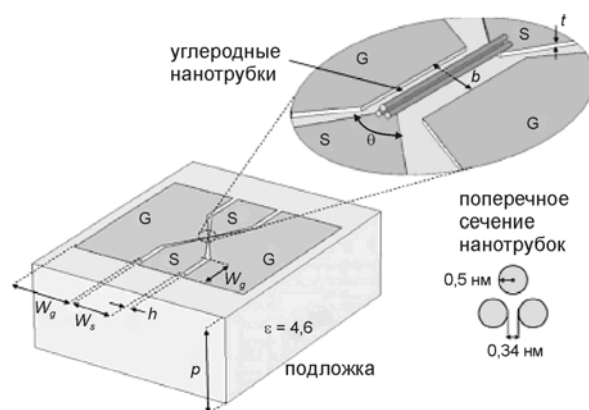


Рис. 8.17. Выходные ВАХ и зависимость высоты потенциального барьера металл-полупроводник транзистора на нанонити [8.20]

«пачки» углеродных нанотрубок, что с характерными размерами схематически показано на рис. 8.18. При этом используется сочетание металлических и полупроводниковых нанотрубок в пачке, общее их количество может достигать десятков. Изменяя количество и тип нанотрубок, добиваются изменения их

проводимости, а также емкостных и индуктивных свойств. Здесь же предложена одномерная модель распространения волн в описанной системе нанотрубок, аналогичная уравнениям длинных линий, в предположении отсутствия туннельных переходов электронов между отдельными нанотрубками в пачке, что, как правило, выполняется, поскольку проводимость каждой отдельной нанотрубки существенно больше проводимости между трубками.

Для распределенной одномерной модели линии передачи с металлической и квазиметаллической проводимостью для одиночной нанотрубки в предположении спинового вырождения могут быть записаны следующие уравнения для заряда на единицу длины и для тока вдоль линии [8.24]:



$$\frac{\partial Q_c}{\partial t} = - \frac{\partial I_c}{\partial z},$$

$$\frac{\partial I_c}{\partial t} = - \frac{4ev_F F}{h} - v_F^2 \frac{\partial Q_c}{\partial z},$$

где $v_F = 8 \cdot 10^5$ м/с – скорость Ферми, e – заряд электрона, h – постоянная Планка, F – результирующая сила, действующая на электроны. Эти силы можно представить в виде суммы электромагнитных сил F_{em} и сил

Рис.8.18. Структура с линией передачи в виде «пачки» нанотрубок

торможения F_a , связанных с действием различных механизмов рассеяния: $F = F_{em} - F_a$, где $F_a = mv/\tau$, τ – время релаксации. При известном токе в нанотрубке скорость может быть найдена как

$$v = - \frac{hI_c}{4emv_F},$$

а с учетом последних соотношений:

$$F_a = - \frac{hI_c}{4ev_F \tau}.$$

Обобщая приведенные выше соотношения, можно получить уравнение для тока вдоль линии в следующем виде:

$$\frac{\partial I_c}{\partial t} = - \frac{4ev_F F_{em}}{h} - \frac{I_c}{\partial z} - v_F^2 \frac{\partial Q_c}{\partial z}.$$

В цитируемой работе [8.24] получены также уравнения для пачки нанотрубок, аналогичные по форме известным телеграфным уравнениям в виде:

$$\frac{\partial I}{\partial z} = -C' \frac{\partial V}{\partial t},$$

$$\frac{\partial V}{\partial z} = -\frac{R'}{2n} I - \left(L'_e + \frac{L'_k}{2n}\right) \frac{\partial I}{\partial t}.$$

Здесь используется подход, использующий распределение электрохимического потенциала V вдоль линии, а величины, входящие в уравнения, определяются по следующим соотношениям.

Электрохимический потенциал:

$$V = V_s + \frac{Q'}{2nC'_q},$$

где V_s – электростатический потенциал, n – количество нанотрубок в пачке, $C'_q = \frac{2e^2}{hv_F}$, $L'_k = \frac{h}{2e^2v_F}$, $R' = \frac{h}{2e^2v_F\tau}$.

Здесь используются определенные выше понятия квантовой емкости и кинетической индуктивности. По аналогии электромагнитные параметры линии – емкость и индуктивность линии в виде пачки нанотрубок определяются по соотношениям:

$$C'_e = \frac{2\pi\epsilon\epsilon_0}{ch^{-1}(d/a)}, \quad L'_e = \frac{\mu_0}{2\pi} ch^{-1}(d/a).$$

Здесь a – радиус нанотрубки, d – расстояние пачки нанотрубок до заземляющей пластины (геометрические размеры показаны на рис. 8.19).

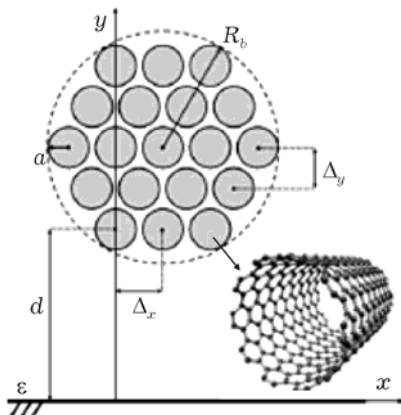


Рис. 8.19 Пачка нанотрубок с металлической проводимостью

Кроме того, в последних соотношениях используются магнитная и диэлектрическая проницаемость среды вне пачки нанотрубок. В работе [8.22] предложены конструкции межсоединений в виде «пачки» однородных нанотрубок, пригодных, в том числе, для коммутации не только планарных, но и многослойных (вертикальных) наноструктур (рис. 8.20). Диаметр нанотрубок в такой структуре может достигать десятков нанометров, а высота до сотен микрометров, что позволяет осуществлять многослойную компоновку наносхем.

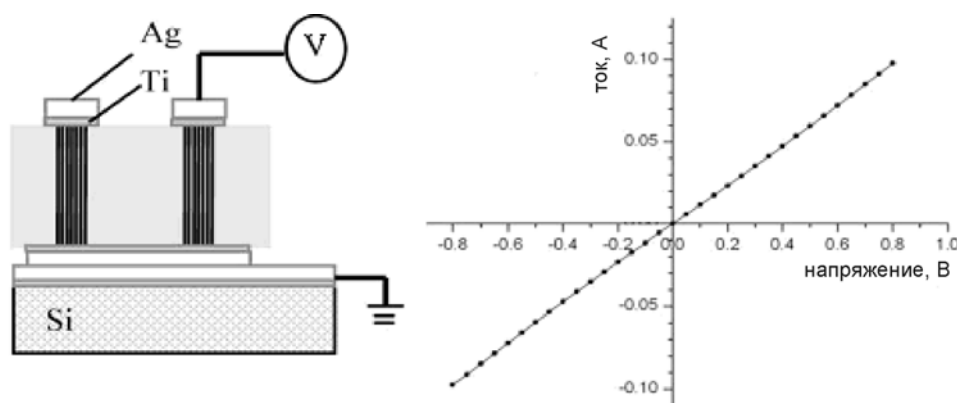


Рис. 8.20. Тестовая структура в виде вертикальных нанотрубок и ее ВАХ

В тестовой структуре на рис. 8.20 среда заполнения вокруг системы нанотрубок между контактами представляет собой полимерный материал бензоциклобутан. Обращает на себя внимание абсолютно линейная вольтамперная характеристика системы вертикальных нанотрубок. При длинах порядка сотен микрометров характерные значения сопротивления приведенной системы нанотрубок составляют единицы и десятки Ом.

Пример использования нановолокон для межсоединений, а также изготовления контактов на кристалле приводится в работе [8.23]. Принципиальным отличием является использование планарных графитовых структур в виде нановолокон (рис. 8.21). Одним из основных требований к межсоединениям является постоянство проводимости в широком диапазоне температур. Приведенные на рис. 8.22 температурные зависимости проводимости для четырех тестовых нановолоконных структур на подложке из окиси кремния демонстрируют монотонный рост до примерно 150 К и относительное ее постоянство до 350 К. Тестовые структуры изготовлены методом плазменно-химического выпаривания. Нановолокно может быть получено в виде слоя толщиной 30–40 нм на контакте со слоем титана, напыленного на кремниевую подложку.

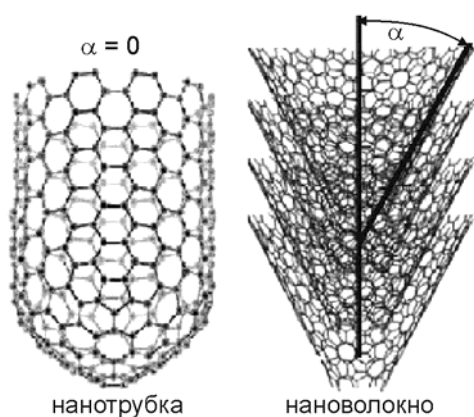


Рис. 8.21. Структурные особенности углеродной нанотрубки и нановолокна

диапазона. Квантовые конденсаторы представляют собой структуру (рис. 8.23), подобную транзисторной. Компоновка структуры предполагает размещение на подложке до 100 углеродных нанотрубок. Посредством управляющего электрода имеется возможность реализовать перестраиваемый конденсатор (варактор) в широком диапазоне питающих напряжений различной полярности. Приводятся данные [8.25] о возможности перестройки емкости конденсатора в

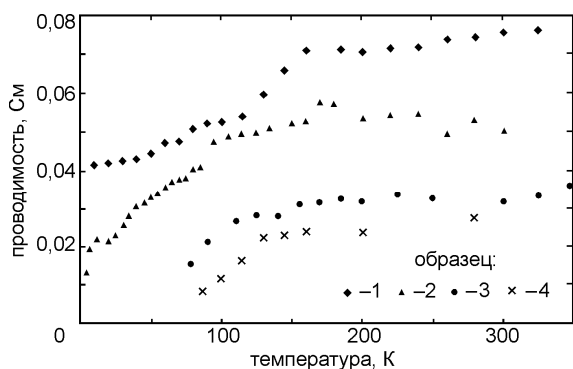


Рис. 8.22. Температурные зависимости проводимости образцов конденсатора углеродного нановолокна [8.24]

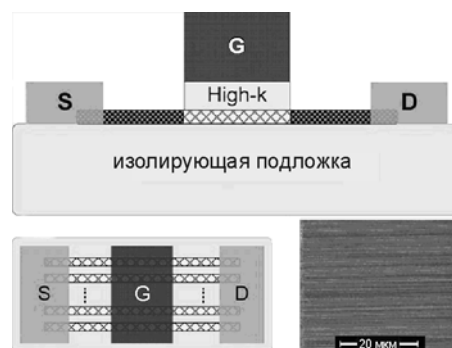


Рис.8.23 Структура квантового конденсатора

аттофарадах почти на порядок при изменении электрического смещения от -2 В до $+2$ В. Отличительными особенностями конденсаторов являются малые тепловые потери и высокие скорости переключения вследствие баллистического переноса электронов в нанотрубке, однако остается проблема минимизации паразитных параметров структуры. Технологически этому способствует применение так называемых “*high-k*”-материалов – диэлектриков, которые имеют статическую диэлектрическую проницаемость больше, чем у оксида кремния ($\epsilon = 3,9$). К таким материалам относятся Al_2O_3 , Ti_2O_3 , ZrO_2 , Ta_2O_5 , HfO_2 и др.

В работе отмечается, что в качестве физической модели в подобных структурах может использоваться прыжковый механизм проводимости. Двумерная матрица из нановолокон по температурным характеристикам демонстрирует преимущества по сравнению с межсоединениями на нанотрубках и нанонитях.

В работе [8.25] приводятся примеры использования углеродных нанотрубок в качестве квантового конденсатора для нелинейных цепей терагерцового

В работе [8.26] рассматривается применение углеродных нанотрубок с характерными размерами (рис. 8.24) в качестве дипольной антенны, использующей «пачки» нанотрубок круглой и прямоугольной формы, где радиус пачки R (или высота и ширина для прямоугольной «пачки») имеют размеры десятков нанометров, а длина L порядка десятков микрометров. Для создания антенных устройств используются нанотрубки типа «кресло» с металлической проводимостью.

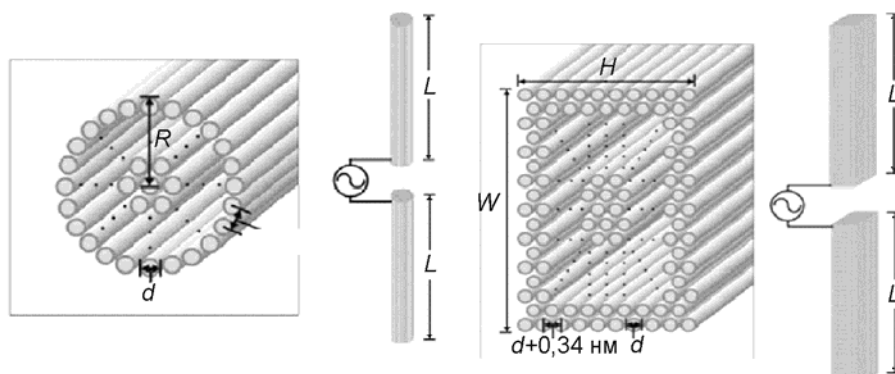


Рис. 8.24. Дипольные антенны на нанотрубках

Одним из важнейших параметров квантовых систем, как уже отмечалось при рассмотрении туннельных эффектов, является квантовая единица сопротивления: $h/e^2 = 25,8$ кОм. В отличие от обычного резистора квантовое сопротивление не является обратно пропорциональным площади поперечного сечения и много меньше, чем у традиционного металлического резистора с наноразмерами. Кроме того, в таких волноводах можно пренебречь скин-эффектом вплоть до терагерцового диапазона частот, поскольку ток электронов связан с π -зонами атома углерода, подобно тому, как это происходит в графитовом листе. Таким образом, общие потери мощности в описываемых антеннах существенно меньше, чем в традиционных. Одним из недостатков таких антенн является высокий характеристический импеданс (до 100 кОм), что может затруднять согласование со стандартными устройствами с волновым сопротивлением 50 Ом. В работе [8.26] показано, что антенны в виде «пачки» углеродных нанотрубок обладают излучательной способностью на 30-40 дБ выше, чем антенны на одиночных нанотрубках.

Одной из нерешенных проблем в области нанотехнологий остается обеспечение электрического контакта наноэлектронных устройств с макроскопическим уровнем без существенных потерь в потенциальной плотности тока, достижимой на наноуровне. Пока контакт со всеми нанотрубчатými и нанопроводными устройствами реализуют посредством создания контактных площадок. Но этот путь не может быть эффективным для объединения множества наносистем. Действительно, топологиче-

ские размеры наносистем существенно меньше, чем разрешение современного литографического оборудования. И относительно низкая плотность размещения контактных площадок перечеркивает достоинства возможной плотности упаковки нанoeлектронных элементов.

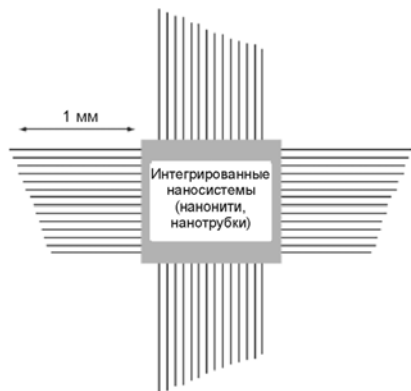


Рис. 8.25. Наносистема с решетками наноантенн для беспроводной связи с взаимодействующими макроустройствами

Одно из возможных решений проблемы состоит в организации беспроводной связи с наноустройствами [8.27]. Если каждый соединительный провод связан с нанотрубкой различной длины (следовательно, различной резонансной частоты), то проблема мультиплексирования сигналов входа/выхода может быть перенесена из пространственной области в спектральную. Следовательно, ослабляя требования к высокому разрешению литографического процесса для соединительных проводов, можно существенно снизить стоимость нанoeлектронных изделий. Идея беспроводных соединений

достаточно универсальна и легко масштабируема (рис.8.25). По такому же принципу возможна организация электропитания наноустройств, причем наноантенны будут использоваться в качестве ректенн (выпрямляющих антенн), в которых роль выпрямителей выполняют все те же нанотрубки с внедренным дефектом атомной сетки. Подобные модификации наноструктур позволяют перейти от пассивных наноантенн к активным решениям, обладающим в дополнение к пространственно-избирательным и частотно-селективным свойствам возможностью прямого усиления принятых сигналов.

Анализ волновых процессов в нанотрубках можно проводить, используя теорию длинных линий [8.28] и представление этих линий в виде схемной модели с сосредоточенными параметрами. Особенность электрических свойств углеродных нанотрубок может быть учтена включением в их эквивалентную электрическую схему, помимо распределенных по длине проводника традиционных емкости и индуктивности, двух дополнительных распределенных элементов: кинетической индуктивности и квантовой емкости.

Кинетическая индуктивность характеризует эффект "запасания" тока в нанопроводнике не за счет магнитного поля, а вследствие инерции движения носителей заряда. Электроны реагируют на приложенное электрическое поле не мгновенно, а с некоторой задержкой. Поэтому при воздействии периодически

изменяющихся электрических полей соответствующие вариации скорости электронов отстают по фазе от изменений напряженности электрического поля. Внешне это проявляется аналогично наличию индуктивности. Кинетическая энергия, обусловленная перемещением электронов при прохождении тока, может быть описана выражением [8.29]:

$$K = \int_V n \frac{mv^2}{2} dV,$$

где n – концентрация электронов в проводнике, m и v – масса и скорость их движения. В результате интегрирования по объему проводника несложно получить:

$$K = n \frac{mv^2}{2} V = n \frac{mv^2}{2} Sl,$$

где S – площадь поперечного сечения проводника, l – его длина.

Энергия, связанная с инерционностью движения электронов, подобна той, которая запасается в магнитном поле проводника с током. Это позволяет сопоставить магнитную энергию индуктивности с инерцией зарядов:

$$\frac{1}{2} L_k i^2 = \int_V n \frac{mv^2}{2} dV = n \frac{mv^2}{2} Sl.$$

Если учесть выражение для тока $i = envS$, то получим выражение для кинетической индуктивности в виде:

$$L_k = \frac{ml}{ne^2 S}.$$

В случае макроскопических проводников кинетическая индуктивность всегда много меньше обычной «магнитной» индуктивности. В тонких проводниках с малым поперечным сечением кинетическая индуктивность преобладает.

Это обстоятельство, а также то, что размеры поперечного сечения проводника сопоставимы с квантовомеханической длиной волны электрона, по существу превращает проводник в одномодовый волновод для электронных волн.

Кинетическая индуктивность для одномерных волноводных систем, канализирующих электронные волны, в пересчете на

единицу длины проводника описывается упрощенной зависимостью [8.30]:

$$L_{\kappa} = \frac{h}{2e^2 v_F},$$

где h – постоянная Планка, v_F – скорость Ферми для носителей заряда: $v_F = \sqrt{2E_F / m}$.

Для углеродных нанотрубок в [8.29] указана величина $v_F = 8 \cdot 10^5$ м/с, соответственно, удельная линейная (погонная) кинетическая индуктивность составит $L_{\kappa} = 16$ нГ/мкм.

Понятие *квантовой емкости* было введено в 1987 году Сержем Лурье (*Luryi*) [8.30] для определения емкости слоя двумерного электронного газа: $C_Q = e^2 \cdot D$, где D – двумерная плотность состояний системы. В свою очередь, для двумерного электронного газа $D = 4m^* \pi / h^2$. Чтобы внести в квантовую ферми-систему новую частицу (в соответствии с принципом Паули), ее энергия должна превышать уровень Ферми. В одномерных системах это может быть приравнено к удельной энергии на единицу длины. От этой удельной энергии может быть рассчитана погонная емкость на единицу длины, получившая наименование квантовой емкости. Применительно к одномерному случаю нанотрубок в работе [8.29] квантовая емкость на единицу длины определяется как:

$$C_q = 2e^2 / h v_F.$$

При скорости Ферми для графена и углеродных нанотрубок удельная квантовая емкость имеет порядок 100 аФ/мкм.

Формирование дипольной антенны из углеродных нанотрубок иллюстрирует рис. 8.26. В двухпроводной линии на нанотрубках магнитной индуктивностью можно пренебречь вследствие ее малой величины. Схемная модель такой линии приведена на рис. 8.27,а [8.29].

Электростатическую емкость двухпроводной линии можно рассчитать по формуле [8.29]:

$$C_{ES} = \pi e / ch^{-1}(W/d) \approx \pi / \ln(W/d).$$

Следует принять во внимание, что в углеродных нанотрубках, благодаря их зонной структуре и спиновым моментам электронов, возможны четыре квантовых канала распространения электронов [8.31] (четыре энергетических состояния электронов проводимости): по два канала для двух состояний спина. Каждый из этих каналов имеет собственную

кинетическую индуктивность и квантовую емкость. Все четыре канала характеризуются общей электростатической емкостью относительно металлической подложки, что существенно влияет на радиочастотные свойства нанопроводников (рис. 8.27б) [8.31].

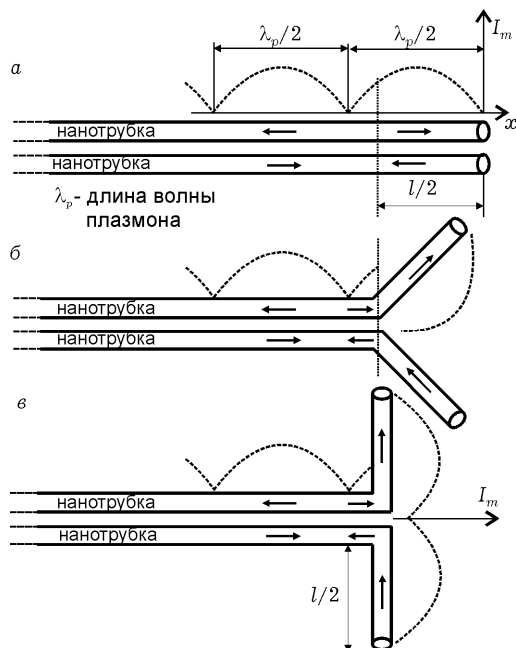


Рис. 8.26. Формирование дипольной антенны из двухпроводной линии на нанотрубках. Штриховая линия – распределение амплитуды стоячей волны тока

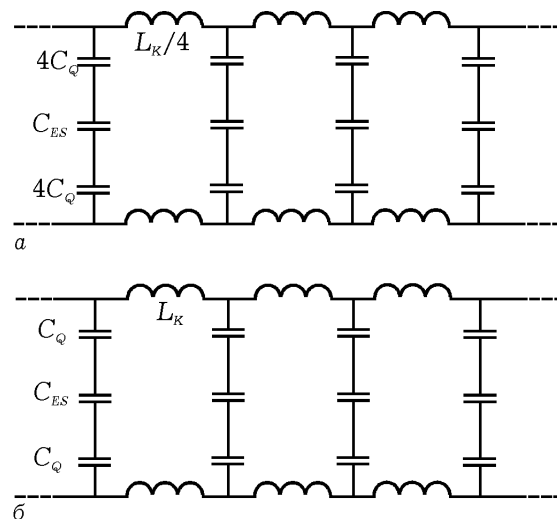


Рис. 8.27. Схемные модели двухпроводной линии передачи на нанотрубках а) – без учета спина электронов б) – с учетом спина

Поскольку квантовые каналы параллельны, результирующая кинетическая индуктивность линии будет в четыре раза ниже, чем в одноканальной трактовке, а квантовая емкость – в четыре раза выше. Соответственно, примерно до 6,45 кОм ($h/4e^2$) уменьшится и квантовое сопротивление переменному току. С учетом этого, характеристический импеданс Z_c и волновую скорость v_p можно представить в виде [8.29]:

$$Z_c = \frac{1}{\sqrt{2}} \sqrt{\frac{R + j\omega L_k / 4}{j\omega C_{\text{общ}}}}, \quad v_p = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{(L_k / 4) C_{\text{общ}}}},$$

где $(C_{\text{общ}})^{-1} = (C_{\text{ES}})^{-1} + (8C_Q)^{-1}$, R – активное сопротивление отдельно взятой нанотрубки на единицу длины. Индекс "p" для волновой скорости обозначает плазмон – квазичастицу, соответствующую коллективным колебаниям свободного электронного газа, в данном случае – одномерного. В соответствии с работой [8.29], длина плазмонной волны приблизительно в 100

раз меньше, чем длина электромагнитной волны в свободном пространстве для той же частоты. Как известно, скорость волны для традиционной двухпроводной линии передачи не зависит от расстояния между проводами и равна скорости света. При использовании же нанотрубок интервал между ними оказывает заметное влияние на скорость распространения электромагнитной волны вследствие изменения общей емкости линии передачи. Это существенно изменяет текущее распределение тока по сравнению с тонкопроводной ситуацией.

Одним из путей снижения потерь в наноантеннах является использование достаточно высокой частоты электромагнитных колебаний. Расчеты условий распространения волны в пассивной двухпроводной нанолнии передачи без дисперсии согласно [8.29] позволяют математически сформулировать ограничение $\omega L_k/4 \gg R$. При $R = 10$ кОм/мкм низкие потери достижимы для частотного диапазона выше 400 ГГц.

8.6. СПИНТРОННЫЕ СТРУКТУРЫ

В отличие от электроники, использующей свойства электрона как заряженной частицы, спинтроника (от *SPIN-Transport electrONICS*) базируется на использовании квантово-механической характеристики электронов - спина. В компонентах спинтроники ток создается электронами с однонаправленными спинами (спиновый ток), для получения которого необходимо поляризовать спины, упорядочив их в одном направлении. При этом необходимо, чтобы время жизни спина было достаточно большим для передачи его на нужные расстояния. Чаще эти свойства связаны с ферромагнетиками, в которых собственное (внутреннее) магнитное поле может в сотни и тысячи раз превосходить вызвавшее его внешнее магнитное поле. Это может быть связано с существованием так называемого обменного взаимодействия, возникающего при перекрытии волновых функций электронов, принадлежащих соседним атомам кристаллической решетки, а также нескомпенсированных спиновых магнитных моментов валентных электронов. Компоненты спинтроники могут обладать свойствами памяти, высокой скоростью переключения (несколько пикосекунд) и малым потреблением энергии, так как переворот спина практически не требует затрат энергии, а в промежутках между операциями спинтронное устройство может отключаться от источника питания. В компонентах спинтроники может быть использован эффект гигантского магнитосопротивления (*Giant Magneto Resistance - GMR*), наблюдающийся в искусственных тонкопленочных материалах, составленных из чередующихся

ферромагнитных и немагнитных слоев. Также используют явление магнитного туннелирования (*Magnetic Tunnel Junction* - *MTJ*), проявляющееся в элементах из двух слоев ферромагнетика, разделенных изолятором. В качестве диэлектрика обычно служит окись алюминия (ее толщина настолько мала, менее 2 нм, что электрон может просачиваться через этот барьер), а сопротивление изменяется внешним магнитным полем.

На рис. 8.28 показаны трехслойные пленки с эффектом *GMR* и механизм спинового рассеивания, приводящий к различной проводимости для параллельной (а) и антипараллельной (б) ориентации векторов намагниченности [8.32].

Применяя различные материалы и структуры, можно использовать эффект управления потоком электронов, различающихся по спину. Механизм спинового эффекта Холла иллюстрирует рис. 8.29,а: в присутствии электрического поля E_x , когда электронное распределение Ферми смещено вдоль p_x , спины электронов с $p_y > 0$ поворачиваются вверх, а с $p_y < 0$ вниз [8.33]. Спин-орбитальное взаимодействие (рис. 8.29,б) позволяет «сортировать» поток электронов по спину: электроны со спином «вверх» рассеиваются на дефектах и примесях кристаллической решетки преимущественно влево, а электроны со спином «вниз» – вправо по ходу движения.

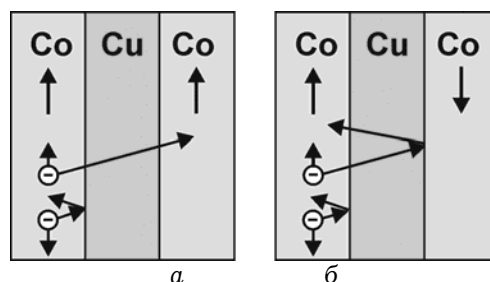


Рис. 8.28. Механизм проводимости в многослойных магнитных пленках

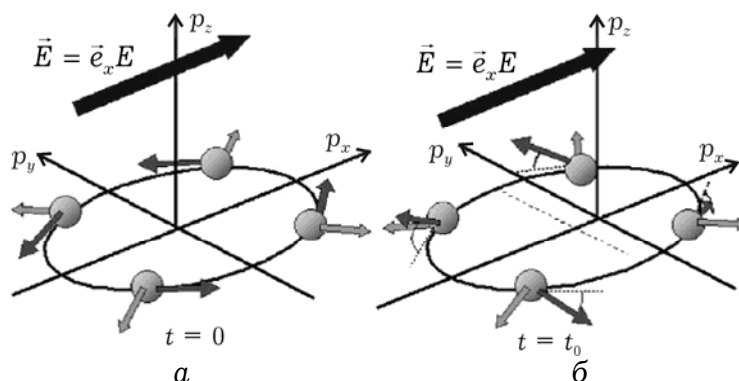


Рис. 8.29. Механизмы спинового эффекта Холла и «сортировки» электронов по спину

Развитие спинтроники связано с поиском материалов, которые обладали бы как свойствами ферромагнетиков, так и полупроводников, что затруднительно из-за большого различия в

кристаллической структуре и характере химических связей. Ферромагнитные полупроводники, с одной стороны, были бы источниками спин-поляризованных электронов, а с другой - легко интегрировались бы с традиционными полупроводниковыми устройствами. Идеальный ферромагнитный полупроводник должен иметь точку Кюри (температура, при которой ферромагнетик теряет свои свойства) выше комнатной температуры и допускать создание зон с n - и p -проводимостью в одном монокристалле.

Задача изменения магнитного момента электрона с помощью электрического поля без приложения магнитных полей в полупроводниках была решена в работе [8.34] для GaAs с толщиной пленки несколько микрон при температуре около 30 К. Это открывает возможности создания нового элемента для преобразования информации, когда намагниченность изменяется при приложении электрического напряжения и наоборот. Развитие работ привело к обнаружению эффекта Холла [8.35] в пленках толщиной 1,5 мкм в полупроводнике ZnSe, легированного атомами хлора при комнатных температурах.

Магнитные свойства магнитных атомов обусловлены d -орбиталями, в то же время полупроводники образуются из атомов с внешними s - или p -орбиталями. Поэтому создание полупроводников со свойствами ферромагнетика затруднено вследствие того, что помещение магнитных атомов в кристаллическую решетку полупроводника резко ухудшает их магнитные характеристики. Для создания гибридного спин-электронного устройства [8.36] можно использовать структуру, показанную на рис. 8.30.

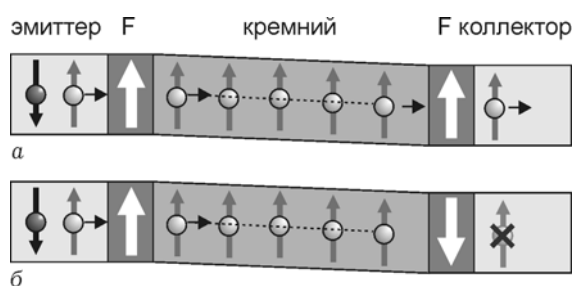


Рис. 8.30. Инжекция спиновых токов в кремнии. F -слои магнитного вещества, действующие как спиновые поляризатор и анализатор.

намагниченности в анализаторе.

Таким образом, для создания спинового транзистора необходимо использовать три основных элемента: эмиттер для электрического инжектирования спин-поляризованных электронов

Показано как электроны, инжектируемые из эмиттера сквозь слой магнитного материала F входят в кремний уже поляризованными по спину. В коллектор электроны попадают, проходя сквозь еще один слой магнитного вещества, который играет роль спинового анализатора: электрическое сопротивление перехода резко возрастает, если спиновая поляризация электронов противоположна направлению намагниченности в анализаторе.

в полупроводник, внешнее напряжение, управляющее спиновым током в полупроводнике, и прецизионный анализатор результирующего спинового тока.

В работах [8.37, 8.38] объектом исследования является так называемый спиновый вентиль (*spin valve*) (рис. 8.31). В простейшем случае он состоит из двух ферромагнитных слоев, разделенных сверхтонким промежуточком - спейсером - из немагнитного металла или диэлектрика (спейсер необходим, чтобы предотвратить прямое обменное взаимодействие между ферромагнитными слоями. В случае металлического спейсера перенос электронов через него

происходит в баллистическом - бесстолкновительном режиме, в случае диэлектрического - в режиме туннелирования. Один из ферромагнитных слоев (он именуется свободным, *free*) обладает малой энергией анизотропии и легко меняет направление своей намагниченности под действием внешнего магнитного поля соответствующей ориентации или при протекании тока. Другой ферромагнитный слой (его называют закрепленным, *pinned*) имеет существенно большую энергию анизотропии и переключается (меняет направление намагниченности) в гораздо более сильных полях.

Когда поток электронов течет из закрепленного слоя в антипараллельно ориентированный свободный, последний при превышении указанного порога скачком перемагничивается с возникновением параллельной конфигурации [8.38]. Это проявляется в скачкообразном уменьшении сопротивления магнитного перехода, что при заданной величине тока означает также скачкообразное понижение напряжения на переходе. Перенос электронов при протекании электрического тока в ферромагнитном проводнике сопровождается переносом потока спинового момента количества движения и связанного с ним потока магнитного момента. В таком потоке направление спинового магнитного момента электронов параллельно направлению магнитного момента решетки, поскольку это соответствует минимуму энергии *sd*-обменного взаимодействия. При переходе электронов из одного ферромагнитного слоя спинового вентиля в другой, имеющий иную ориентацию магнитного момента решетки, параллельность нарушается и электрону приходится менять ориентацию спина. Параллельность восстанавливается на малом расстоянии от границы раздела между ферромагнетиками, в слое

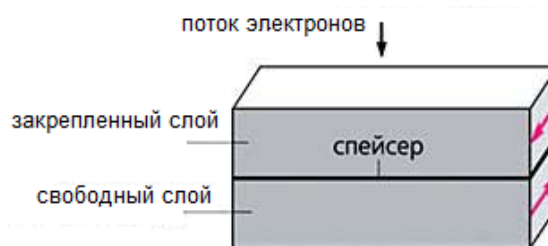


Рис. 8.31. Схема спинового вентиля. Стрелками показано направление намагниченности в слоях.

толщиной ~ 1 нм. Этот процесс сопровождается передачей спинового момента количества движения от электронов к решетке. При этом даже при коллинеарной ориентации намагниченности обоих слоев в свободном слое имеются флуктуации намагниченности. По отношению к таким моментам спин электронов, входящих в свободный слой, имеет как продольную, так и поперечную компоненты. Поскольку величина передаваемого в процессе взаимодействия момента пропорциональна плотности тока через переход, при определенной плотности тока возникает неустойчивость флуктуаций намагниченности решетки, которая и приводит к переключению.

В работе [8.38] дана физическая интерпретация указанному эффекту, которая сводится к двум одновременно существующим механизмам. Согласно теории Слончевского-Берже при прохождении спин-поляризованного тока через границу двух неколлинеарных ферромагнетиков происходит передача вращательного момента (*spin torque*) от электронов магнитной решетке, поскольку электроны переходя в среду с другим направлением оси квантования спина, теряют поперечную, перпендикулярную к новому направлению этой оси, компоненту спина, которая и отдается решетке. При достаточно большой плотности тока флуктуации становятся неустойчивыми и их нарастание приводит к перемагничиванию свободного слоя.

Другая модель учитывает инжекцию неравновесных спинов из закрепленного слоя в свободный слой. Поэтому вблизи границы между этими слоями, а также вблизи границы между свободным и немагнитным слоями возникают области неравновесной спиновой поляризации. Ширина этих областей определяется длиной спиновой диффузии и на порядок превышает длину релаксации поперечной составляющей намагниченности. Вследствие существования *sd*-обменного взаимодействия между электронами и магнитной решеткой наличие таких областей может, как уменьшать, так и увеличивать магнитную энергию перехода в зависимости от ориентации намагниченностей и от параметров магнитных слоев. Как следствие, при достаточно большой плотности тока исходная магнитная конфигурация может оказаться энергетически невыгодной, что приведет к неравновесному фазовому переходу с изменением конфигурации структуры. Практическое применение описанного эффекта может быть связано с магнитной записью информации спин-поляризованным током.

Системное изложение физических процессов и моделей в спинтронных элементах изложено в работе [8.39].

Важным этапом в развитии спинтроники стало изучение туннельного магнетосопротивления в магнитных туннельных переходах, образованных ферромагнитными электродами с

различными электрическими сопротивлениями с параллельной и антипараллельной ориентацией магнитных моментов. Магнитные туннельные переходы, способные перемагничиваться за счет переноса спина, могут служить основой для перспективной магнитной памяти с малыми временем доступа и низким энергопотреблением. Развитие полупроводниковой электроники связано [8.39] с разработкой гибридных структур, объединяющих ферромагнитные металлы с немагнитными полупроводниками, получением ферромагнитных полупроводников (в частности, $Ga_{1-x}Mn_xAs$, который, однако, обладает низкой температурой Кюри – 170К), а также структур, использующих спиновый эффект Холла в немагнитных проводниках.

Одной из проблем спинтроники является вопрос о времени спиновой когерентности, так как, если это время слишком мало, то разориентация спинов приводит к потере информации, переносимой каждым спином [8.40]. Измерения времени спиновой когерентности при комнатной температуре показали, что в полупроводниках оно на несколько порядков больше (> 100 нс), чем в металлах (0.1–20 нс), обычно используемых в спиновых мультислоях. Поэтому „полупроводниковая“ спинтроника более привлекательна и реалистична, чем „металлическая“. В настоящее время идет активная работа над созданием *DMS* (*Diluted Magnetic Semiconductors* – разбавленных магнитных полупроводников). Несмотря на выигрыш во времени спиновой когерентности, большинство исследуемых *DMS* имеют весьма низкие температуры Кюри, что ограничивает их практическое применение. Этот недостаток может быть устранен при переходе к наноструктурам *DMS*. Наноструктурирование полупроводника увеличивает растворимость примеси переходных металлов в его кристаллической решетке до желаемых 1–2%, обеспечивающих магнитоупорядоченное состояние.

Как уже отмечалось в работе [8.2] можно использовать в качестве спинтронных компонентов сочетание полупроводниковых и молекулярных структур. В качестве бистабильной системы может использоваться молекула ротаксана (рис. 8.32), структура которой может рассматриваться как аналог нанопроводника. В середине молекулы имеется одномерная компонента с π -связью, связанная с кольцом. Механизм переключения предполагает обмен двумя электронами и переход через фазу окисления из стабильного состояния (с низкой проводимостью) в состояние с высокой проводимостью. При этом вследствие кулоновского отталкивания происходит перемещение кольца. Процессы переключения могут стимулироваться, например, внешним электрическим полем.

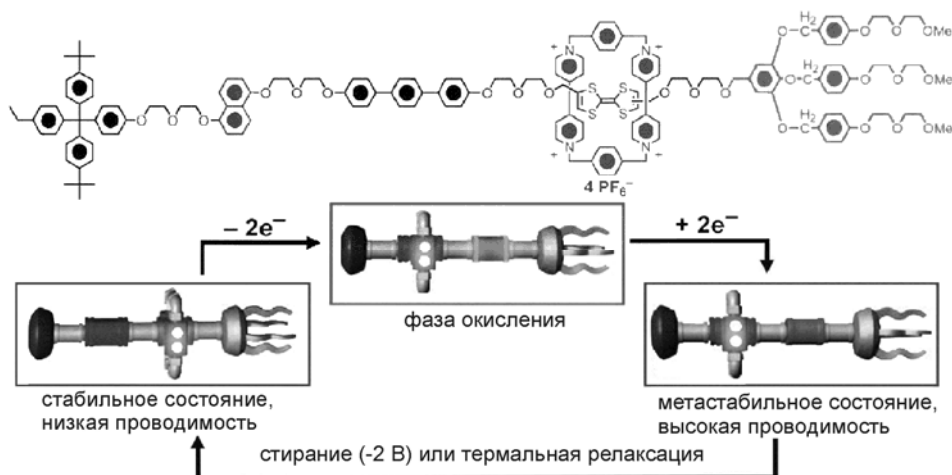


Рис. 8.32. Механизм переключения молекулы ротаксана

Минимальное электрическое смещение, стимулирующее процессы переключения составляет 0,8 В, а время – 0,1 пс. Существует, однако, множество ограничений, препятствующих созданию подобных наноконструкций, среди которых влияние температурных и паразитных эффектов при встраивании молекул ротаксана в функциональные наносистемы.

8.7. ЯЧЕЙКИ ПАМЯТИ НА НАНОКОМПОНЕНТАХ

Как уже отмечалось, одноэлектронный транзистор может находиться в двух состояниях. Одно из них связано с кулоновской блокадой, второе – с протеканием тока. Это позволяет рассматривать их как перспективные ячейки быстродействующей и энергетически экономичной памяти для электронных вычислительных систем.

Ячейка памяти на двойных одноэлектронных транзисторах приводится в работе [8.41]. Структура ячейки памяти на одноэлектронных транзисторах, реализующих три квантовых бита (*qubits*) показана на рис. 8.33. В такой структуре три одноэлектронных транзистора соединены последовательно, а их затворы управляют электронным состоянием квантовых островков и, соответственно, переключением их в состояния «0» или «1», поскольку каждая пара островков имеет два устойчивых состояния (две поляризации). При этом ток переключения составляет 0,1-0,5 нА, при уровнях управляющих воздействий около 1 В. Рассмотренные квантовые системы выполнены по технологии «кремний на изоляторе». Однако, надежные характеристики достигаются при очень низких температурах (4 К).

Вентильные свойства одноэлектронных структур проявляются для различных сочетаний материалов и значений температур [8.42]. Однако, процессы одноэлектронного тунелирования имеют случайную природу и зависят от множества факторов, включая процессы термализации электронов (в том числе при приложении внешнего электрического поля), что при построении квантовых ячеек памяти ограничивает надежность их переключения.

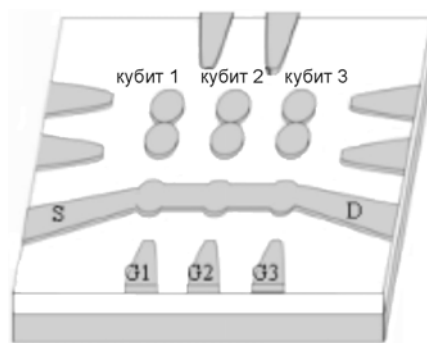


Рис.8. 33. Структура ячейки памяти (3 кубита) на одноэлектронных транзисторах

Наряду с поисками перспективных квантовых структур памяти, продолжается совершенствование субмикронных структур «кремний-на-изоляторе». В частности, разработана полуизолированная однострансторная ячейка памяти [8.43], стабильно работающая в широком диапазоне температур.

Ячейка памяти [8.44] в виде вертикального МДП транзистора на основе нитрида кремния с длиной активной области 0,1 мкм позволяет получать времена переключения в области наносекунд при токе потребления около 10 мкА.

Приведенные примеры ячейки памяти на традиционных транзисторных структурах позволяют получать матрицы быстродействующей памяти большой емкости при сравнительно небольших технологических затратах для электронных устройств широкого применения.

Рассмотренные в разделе структуры не исчерпывают весь возможный спектр компонентов, которые могут быть созданы на основе применения нанотехнологий. Однако, представленные структуры различной размерности позволяют выделить группу физических эффектов, на основе которых уже в ближайшем будущем возможна реализация и массовое изготовление наноконструкций для широкого применения. Приобретают актуальность вопросы разработки теории и моделирования приборных характеристик и параметров новейших наноконструкций, а также измерений и идентификации их параметров. Возможно создание компонентов с совершенно новыми характеристиками, а появление новых функциональных элементов приведет к пересмотру технической базы во многих областях электроники.

Литература к гл. 8

- 8.1. *Curreli M., Zhang R., Ishikawa F.N., Chang H.K., Cote R.J., Zhou C., Thompson M.E.* Real-Time, Label-Free Detection of Biological Entities Using Nanowire-Based FETs // IEEE Transactions on Nanotechnology. – 2008 – Vol. 7, No 6. P. 651–668.
- 8.2. *K.Galatsis, A.Khitun, R. Ostroumov, K.L. Wang, W.R. Dichtel, E. Plummer, J.Stoddart, J.Zink, J.Lee, Y-H.Xie, K.W.Kim.* Alternate State Variables for Emerging Nanoelectronic Devices // IEEE Transactions on Nanotechnology.–2009–Vol. 8, No1. P.66–75.
- 8.3. *Tarakeshwar P., Palacios J.J., Kim D.M.* Electrode-Molecule Interface Effects on Molecular Conductance // IEEE Transactions on Nanotechnology. – 2009 – Vol. 8, No 1. P. 16–21.
- 8.4. *Jung Y.C., Cho K.H., Hong B.Y., Son S.Y., Kim D.S., Whang D., Hwang S.W., Yu Y.S., Ahn D.* Fabrication and Characterization of Sidewall Defined Silicon-on-Insulator Single-Electron Transistor // IEEE Transactions on Nanotechnology. – 2008 – Vol. 7, No 5. P. 544–550.
- 8.5. *Miyaji K., Saitoh M. Hiramoto T.* Compact Analytical Model for Room-Temperature-Operating Silicon Single-Electron Transistors With Discrete Quantum Energy Levels // IEEE Transactions on Nanotechnology. – 2006 – Vol. 5, No 3. P. 167–173.
- 8.6. *Y-C Wu, T-C Chang, P-T Liu, etc.* High-Performance Metal-Induced Lateral-Crystallisation Polysilicon Thin-Film Transistors With Multiple Gates // IEEE Transactions On Nanotechnology.–2006.–V.5, №3–P.157-162.
- 8.7. *David L. Miller, Kevin D. Kubista, Gregory M. Rutter, Ming Ruan, Walt A. de Heer, Phillip N. First, Joseph A. Stroscio.* [Observing the Quantization of Zero Mass Carriers in Graphene](#) // Science. 2009. V. 324. P. 924–927.
- 8.8. *Морозов С.В., Новоселов К.С., Гейм А.К.* Электронный транспорт в графене // Успехи физических наук. 2008–Т.178. №7.– С.776-780.
- 8.9. *Rana F.* Graphene Terahertz Plasmon Oscillators // IEEE Transactions On Nanotechnology.–2008.–V.7, №1–P.91-99.
- 8.10. *Ryzhii M., Ryzhii V.* Injection and population inversion in electrically induced p-n junction in graphene with split gates // J.Appl. Phys. –2007. –V.101, P.P.L151–L153.
- 8.11. *Song S., Kim K., Huh J., etc.* Analytical Modeling of Field-Induced Interband Tunneling-Effect Transistors and Its Application // IEEE Transactions On Nanotechnology.–2006.–V.5, №3–P.192-200.
- 8.12. *Latessa L., Pecchia A., Di Carlo A.* DFT Modeling of Bulk-Modulated Carbon Nanotube Field-Effect Transistors // IEEE Transactions On Nanotechnology.–2007.–V.6, №1–P.13-21.

- 8.13. *Bethoux J.-M., Happy H., Siligaris A., Dambrine G., Borghetti J., Derycke V., Bourgoin J.-P.* Active Properties of Carbon Nanotube Field-Effect Transistors Deduced From S-Parameters Measurements // *IEEE Transactions On Nanotechnology.*—2006.—V.5, №4—P.335-342.
- 8.14. *Neophytou N., Guo J., Lundstrom M.* Three-Dimensional Electrostatic Effects of Carbon Nanotube Transistors // *IEEE Transactions On Nanotechnology.*—2006.—V.5, №4—P.385-392.
- 8.15. *Alam K., Lake R.* Role of Doping in Carbon Nanotube Transistors With Source/Drain Underlaps // *IEEE Transactions On Nanotechnology.*—2007.—V.6, №6—P.652-658.
- 8.16. *Lin A., Patil N., Ryu K., Badmaev A., Acro L.G., Zhou C., Mitra S., Wong H-S.P.* Threshold Voltage and On-Off Ratio Tuning for Multiple-Tube Carbon Nanotube FETs // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №1—P.4-9.
- 8.17. *Patil N., Deng J., Mitra S., Wong H-S.P.* Circuit-Level Performance Benchmarking and Scalability Analysis of Carbon Nanotube Transistors Circuits // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №1—P.37-45.
- 8.18. *Zhang M., Chan P.C.H., Chai Y., Tang Z.K.* Novel Local Silicon-Gate Nanotube Transistors Combining Silicon-on-Insulator Technology for Integration // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №2—P.260-268.
- 8.19. *Lee S.H., Yu Y.S., Hwang S.W., Ahn D.* A SPICE-Compatible New Silicon Nanowire Field-Effect Transistors (SNWFETs) Model // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №5—P.643-649.
- 8.20. *Zhang M., Knoch J., Zhao Q. T., Breuer U., Mantl S.* Impact of doping segregation on fully depleted Schottky-barrier SOI MOSFETs. // *Solid-State Electron.*, vol. 50, pp. 594–600, 2006.
- 8.21. *Sarto M.S., Tamburrano A., D'Amore M.* New Electron-Waveguide-Based Modeling for Carbon Nanotube Interconnects // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №2—P.214-225.
- 8.22. *Liu Z., Ci L., Kar S., Ajayan P.M., Lu J.-Q.* Fabrication and Electrical Characterization of Densified Carbon Nanotube Micropillars for IC Interconnection // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №2—P.196-203.
- 8.23. *Ngo Q., Yamada T., Suzuki M., Ominami Y., Cassell A.M., Li J., Meyyappan M., Yang C.Y.* Structural and Electrical Characterization of Carbon Nanofibers for Interconnect Via Applications // *IEEE Transactions On Nanotechnology.*—2009.—V.6, №6—P.688-695.
- 8.24. *Sarto M.S., Tamburrano A., D'Amore M.* New Electron-Waveguide-Based Modeling for Carbon Nanotube Interconnects // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №2—P.214-225.
- 8.25. *D. Akinwande., Nishi Y., Wong H.-S.P.* Carbon Nanotube Quantum Capacitance for Nonlinear Terahertz Circuits // *IEEE Transactions On Nanotechnology.*—2009.—V.8, №1—P.31–36.

- 8.26. *Huang Y., Yin W-Y., Liu Q.H.* Performance Prediction of Carbon Nanotube Bundle Dipole Antennas // IEEE Transactions On Nanotechnology.-2008.-V.7, №3-P.331-337.
- 8.27. *Слюсар В.* Наноантенны: подходы и перспективы // Электроника: Наука, Технология, Бизнес.-2009.-№2-С.58-65
- 8.28. *P.J.Burke, S.Li, and Z.Yu.* Quantitative Theory of Nanowire and Nanotube Antenna Performance // IEEE Transactions on Nanotechnology.- 2006. -Vol. 5-№ 4- P. 314-334.
- 8.29. *P. J. Burke.* An rf circuit model for carbon nanotubes // IEEE Transactions on Nanotechnology.- 2003-Vol. 2- №. 1-P 55-58
- 8.30. *Serge Luryi.* Quantum capacitance devices // Appl. Phys. Letter -1988 -№ 52-P. 501-503.
- 8.31. *Dresselhaus M. S.* et al. Carbon Nanotubes: Synthesis, Structure, Properties, and Applications. -Springer-Verlag, 2000.
- 8.32. *Огнев А.В., Самардак А.С.* Спинтроника: физические принципы, устройства, перспективы // Вестник ДВО РАН. 2006. № 4 С.70-80.
- 8.33. *Sinova J., Culcer D., Niu Q., Sinitsyn N.A., Jungwirth T., MacDonald A. H.* Universal Intrinsic Spin Hall Effect // Physical Review Letters.-2004.-Vol. 92.-126603.
- 8.34. *Kato Y.K., Myers R.C., Gossard A.C., Awschalom D.D.,* Observation of the spin Hall effect in semiconductors // Science. - 2004.-Vol. 306.-P.1910.
- 8.35. *N. Stern. P., Ghosh S., Xiang G., Zhu M., Samarth N., and Awschalom D. D.,* Current-Induced Polarization and the Spin Hall Effect at Room Temperature // Physical Review Letters.-2006.-Vol.97-126603.
- 8.36. *Huang B., Monsma D.J., Appelbaum I.* Coherent Spin Transport through a 350 Micron Thick Silicon Wafer // Physical Review Letters.-2007-Vol.99. -177209.
- 8.37. *Гуляев Ю.В., Зильберман П.Е., Эпштейн Э.М.* Как ток спины переносит. Спинтроника многослойных ферромагнетиков // Успехи физических наук. 2008.-№4.-Т.178.-С.433-436.
- 8.38. *Гуляев Ю. В., Зильберман П. Е., Панас А. И., Эпштейн Э. М.* Спинтроника: обменное переключение ферромагнитных металлических переходов при малой плотности тока // Успехи физических наук.-2009.-№4.-С.359-368.
- 8.39. *Ферт А.* Происхождение, развитие и перспективы спинтроники // Успехи физических наук.-2008.-№12.-Т.178. С.1336-1348.
- 8.40. *Моргунов Р.Б., Мушенок Ф.Б., Дмитриев А.И., Kazakova O.L., Tanimoto Y.* Электронный спиновый резонанс в ориентированных нанопроволоках Ge_{0.99}Cr_{0.01} // Физика твердого тела, 2009.-Т.51. Вып. 8.-С.1613-1618.
- 8.41. *Y. Kawata., Tsuchiya Y., Oda S., Mizuta H.* Study of Single-Charge Polarization on a Pair of Charge Qubits Integrated Onto a

Silicon Double Single-Electron Transistor Readout // IEEE Transactions On Nanotechnology.-2008.-V.7, №1-P.617-623.

8.42. *Погосов В.В., Васютин Е.В., Курбацкий В.П., Бабич А.В., Коротун А.В.* Одноэлектронные эффекты в точечных структурах // Наносистемы, наноматериалы, нанотехнологии, 2007.-Т.5.-С.39-74.

8.43. *Bae D., Kim S., Choi Y.-K.* Low-Cost and Highly Heat Controllable Capacitorless PiFET (Partially Insulated FET) 1T DRAM for Embedded Memory // IEEE Transactions On Nanotechnology.-2009.-V.8, №1-P.100-105.

8.44. *Jeong H., Song K.-W., Park I., Kim T.-H., Lee Y., Kim S., Seo J., Cho K., Lee K., Shin H., Lee J., Park B.-G.* New Capacitorless 1T DRAM Cell: Surrounding Gate MOSFET With Vertical Channel (SGVC Cell) // IEEE Transactions On Nanotechnology.-2007.-V.6, №3-P.352-357.

ПРИЛОЖЕНИЯ

Приложение к главе 2

П2. УСРЕДНЕНИЕ СКОРОСТЕЙ РАССЕЯНИЯ

Времена релаксации находятся как средние для электронного коллектива скорости рассеяния для каждого вида столкновений. Способы усреднения могут быть различными, но для большинства из них конечные результаты отличаются между собой коэффициентом пропорциональности порядка 1.

Основная проблема состоит в том, что в правой части кинетического уравнения (1.12) стоит искомая функция распределения, которая необходима для усреднения и предварительно должна быть найдена. Поскольку обычно это невозможно без решения уравнения Больцмана, то часто принимают допущение о ее форме и усреднение проводят по этой гипотетической функции. Самым распространенным является усреднение по функции распределения $N(k)$

$$\frac{1}{\tau_{pi}} = \int_{V_k} \lambda_i(k) N(\vec{k}) d\vec{k} / \int_{V_k} N(\vec{k}) d\vec{k}, \quad (\text{П2.1})$$

где значения λ_i выбираются для каждого вида рассеяния. Поскольку в знаменателе под интегралом $N(\vec{k}) d\vec{k}$ представляет собой число электронов в элементе объема dV_k \vec{k} -пространства, то, выбирая его лежащим между двумя изоэнергетическими поверхностями, соответствующими приращению энергии на dE , где количество электронов равно $N(E)dE$, можно перейти от усреднения в \vec{k} -пространстве к усреднению по энергетической функции распределения

$$\frac{1}{\tau_{pi}} = \int_0^{\infty} \lambda_i(E) N(E) dE / \int_0^{\infty} N(E) dE. \quad (\text{П2.2})$$

Здесь интегрирование производится по всей энергетической зоне (или соответствующей подзоне) от ее дна до ∞ .

Наиболее простым является предположение о том, что функция распределения сохраняет вид распределения Максвелла-Больцмана, но определяется не температурой кристаллической решетки, а электронной температурой T_e . При квадратичном законе дисперсии $g(E) \sim E^{1/2}$, $f(E) \sim \exp(-E/k_B T_e)$. Используя обозначение $E/k_B T_e \equiv \varepsilon$, получим после сокращения

$$\frac{1}{\tau_{pi}} = \int_0^{\infty} \lambda_i(\varepsilon) \sqrt{\varepsilon} e^{-\varepsilon} d\varepsilon / \int_0^{\infty} \sqrt{\varepsilon} e^{-\varepsilon} d\varepsilon = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \lambda_i(\varepsilon) \sqrt{\varepsilon} e^{-\varepsilon} d\varepsilon. \quad (\text{П2.3})$$

Здесь использовано значение табличного интеграла, стоящего в знаменателе

$$\int_0^{\infty} \sqrt{\varepsilon} e^{-\varepsilon} d\varepsilon = \sqrt{\pi} / 2.$$

Другой способ вычисления времени релаксации импульса состоит в усреднении λ_i с весовым коэффициентом \vec{k} в соответствии с (1.59)

$$\frac{1}{\tau_{pi}} = \int_{V_k} \lambda_i(\vec{k}) \vec{k} N(\vec{k}) d\vec{k} / \int_{V_k} \vec{k} N(\vec{k}) d\vec{k}, \quad (\text{П2.4})$$

Переходя от усреднения по \vec{k} -пространству к усреднению по энергетическим состояниям зоны подобно предыдущему случаю, при квадратичном законе дисперсии получим

$$\frac{1}{\tau_{pi}} = \int_0^{\infty} \lambda_i(E) \sqrt{E} N(E) dE / \int_0^{\infty} \sqrt{E} N(E) dE. \quad (\text{П2.5})$$

Если считать, что электронный газ остается невырожденным, а его функция распределения максвелловской, то при тех же обозначениях и принимая во внимание, что $\int_0^{\infty} \varepsilon e^{-\varepsilon} d\varepsilon = 1$, получим

$$\frac{1}{\tau_{pi}} = \int_0^{\infty} \lambda_i(\varepsilon) \varepsilon e^{-\varepsilon} d\varepsilon \quad (\text{П2.6})$$

Примесное рассеяние. Для рассеяния на ионизированных примесях в модели Конуэлл-Вайскопфа представим (2.18) в развернутом виде, используя приведенные там обозначения:

$$\lambda_{\text{пр}} = 2N_{\text{пр}} \sqrt{\frac{2}{m^*}} \left(\frac{e^2}{4\pi\epsilon\epsilon_0} \right)^2 E^{-3/2} \ln \left(1 + \frac{4\pi\epsilon\epsilon_0}{e^2} \cdot \frac{E}{N_{\text{пр}}^{1/3}} \right). \quad (\text{П2.7})$$

Аналитическое интегрирование этого выражения затруднительно из-за наличия последнего сомножителя. Поэтому, учитывая, что логарифмическая зависимость является слабой по сравнению со степенной, можно без особого ущерба для точности принять в ней $E = 3k_{\text{B}}T_e/2$. Нормируя энергию ($E/k_{\text{B}}T_e \equiv \epsilon$), представим

$$\lambda_{\text{пр}} = N_{\text{пр}} \sqrt{\frac{8}{m^*}} \left(\frac{e^2}{4\pi\epsilon\epsilon_0} \right)^2 (k_{\text{B}}T_e)^{-3/2} \ln \left(1 + \frac{6\pi\epsilon\epsilon_0}{e^2} \cdot \frac{k_{\text{B}}T_e}{N_{\text{пр}}^{1/3}} \right) \epsilon^{-3/2}. \quad (\text{П2.8})$$

Используя второй способ усреднения (П2.6) и принимая во внимание, что $\int_0^\infty \epsilon^{-1/2} e^{-\epsilon} d\epsilon = \sqrt{\pi}$, окончательно получим для времени релаксации импульса при рассеянии на ионизированной примеси

$$\frac{1}{\tau_{\text{рпр}}} = \sqrt{\frac{8\pi}{m^*}} \left(\frac{e^2}{4\pi\epsilon\epsilon_0} \right)^2 \ln \left(1 + \frac{6\pi\epsilon\epsilon_0}{e^2} \cdot \frac{k_{\text{B}}T_e}{N_{\text{пр}}^{1/3}} \right) N_{\text{пр}} (k_{\text{B}}T_e)^{-3/2}. \quad (\text{П2.9})$$

Здесь надо отметить, что в литературе используются и другие способы усреднения, которые дают результат, отличающийся коэффициентом пропорциональности порядка 1. Кроме того, если считать объем $N_{\text{пр}}^{-1/3}$, занимаемый в среднем одним примесным атомом, сферическим и принять в качестве максимального прицельного расстояния его радиус, то результат (П2.9) надо уменьшить в $6/\pi$ раз. Это говорит о том, что при сравнении с экспериментом в (П2.9) можно ввести некоторый подгоночный параметр.

Акустическое рассеяние. Выражение (2.12) для скорости рассеяния на акустических фононах с использованием переменной ϵ имеет вид

$$\lambda_{\text{а}} = \frac{(2m^*)^{3/2} D_{\text{а}}^2 k_{\text{B}}T}{2\pi\rho v_{\text{ЗВ}}^2 \hbar^4} \sqrt{k_{\text{B}}T_e} \sqrt{\epsilon} \equiv A_{\text{а}} \sqrt{\epsilon}. \quad (\text{П2.10})$$

Вычисление времени релаксации импульса по (П2.6) приводит к интегралу

$$\frac{1}{\tau_{pa}} = A_a \int_0^{\infty} \varepsilon^{3/2} e^{-\varepsilon} d\varepsilon = 3 A_a \sqrt{\pi} / 4. \quad (\text{П2.11})$$

Иногда используют выражение, которое получается из (П2.10) при простой замене $E = 3k_B T_e / 2$:

$$1 / \tau_{pa} = \sqrt{3/2} A_a. \quad (\text{П2.12})$$

Оптическое рассеяние. Для удобства записи последующих результатов представим множитель $(N_{\phi} + 1/2 \pm 1/2)$, учитывающий количество участвующих в рассеянии оптических фононов, с учетом формулы Планка (2.3) в форме:

$$N_{\phi} + 1/2 \pm 1/2 = e^{\pm t} / 2 \text{sh} t. \quad (\text{П2.13})$$

Здесь введено обозначение $t = \hbar \omega_0 / 2k_B T$. Используя также обозначение $t_e = \hbar \omega_0 / 2k_B T_e$ и нормируя энергию ($E/k_B T_e \equiv \varepsilon$), запишем выражение (2.13) для скорости рассеяния на **оптическом деформационном потенциале**

$$\lambda_o^{\pm} = \frac{(m^*)^{3/2} D_o^2}{\sqrt{2\pi} \rho \hbar^3 \omega_0} \cdot \frac{e^{\mp t}}{2 \text{sh} t} \sqrt{k T_e} \text{Re} \sqrt{\varepsilon \pm 2t_e} = A_o^{\pm} \text{Re} \sqrt{\varepsilon \pm 2t_e}. \quad (\text{П2.14})$$

Определение времени релаксации импульса в аналитическом виде здесь возможно при использовании усреднения по (П2.3.)

$$\frac{1}{\tau_{po}^{\pm}} = \frac{2A_o^{\pm}}{\sqrt{\pi}} \int_0^{\infty} \sqrt{\varepsilon} \text{Re} \sqrt{\varepsilon \pm 2t_e} e^{-\varepsilon} d\varepsilon. \quad (\text{П2.15})$$

При поглощении фононов (верхний знак) $\varepsilon + 2t_e > 0$, оператор Re теряет смысл и приходим непосредственно к табличному интегралу¹

$$\int_0^{\infty} \sqrt{\varepsilon(\varepsilon + 2t_e)} e^{-\varepsilon} d\varepsilon = t_e \exp t_e K_1(t_e). \quad (\text{П2.16})$$

где $K_1(t_e)$ – модифицированная функция Бесселя от мнимого аргумента второго рода первого порядка.

¹ Г. Н. Ватсон, Теория бесселевых функций, ИЛ, М., 1949

При возбуждении фононов (нижний знак) оператор Re можно опустить, если нижним пределом интегрирования выбрать значение $\varepsilon = 2t_e$, выше которого подкоренное выражение в (П2.15) становится положительным. Привести этот интеграл к виду (П2.16) можно заменой переменной $\varepsilon = \xi + 2t_e$

$$\int_{2t_e}^{\infty} \sqrt{\varepsilon(\varepsilon - 2t_e)} e^{-\varepsilon} d\varepsilon = e^{-2t_e} \int_0^{\infty} \sqrt{\xi(\xi + 2t_e)} e^{-\xi} d\xi = t_e e^{-t_e} K_1(t_e). \quad (\text{П2.17})$$

С учетом двух последних интегралов (П2.15) может быть записано в общем виде

$$\frac{1}{\tau_{po}^{\pm}} = \frac{2A_o^{\pm}}{\sqrt{\pi}} t_e e^{\pm t_e} K_1(t_e) = \frac{(m^*)^{3/2} D_o^2}{2\sqrt{2}\pi^{3/2} \rho \hbar^2} \cdot \frac{e^{\mp(t-t_e)}}{\sqrt{k_B T_e} \text{sh } t} K_1(t_e). \quad (\text{П2.18})$$

Суммирование обратных времен при поглощении и возбуждении приводит к полному обратному времени релаксации импульса для неполярного оптического рассеяния:

$$\frac{1}{\tau_{po}} = \frac{(m^*)^{3/2} D_o^2}{\sqrt{2}\pi^{3/2} \rho \hbar^2} \cdot \frac{\text{ch}(t - t_e)}{\sqrt{k_B T_e} \text{sh } t} K_1(t_e). \quad (\text{П2.19})$$

При **полярном оптическом рассеянии** представим выражение (2.14) с учетом таких же обозначений, как и при неполярном, и использовании так называемой полярной оптической константы

$$\alpha = \frac{e^2 \sqrt{m^*}}{4\sqrt{2} \pi \varepsilon_0 \hbar (\hbar \omega_0)^{1/2}} \left(\frac{1}{\varepsilon_{\infty}} - \frac{1}{\varepsilon_{\text{CT}}} \right)$$

в следующем виде

$$\lambda_{\text{по}}^{\pm} = \frac{\alpha \omega_0}{\sqrt{2}} \cdot \frac{e^{\mp t}}{\text{sh } t} \sqrt{\frac{t_e}{\varepsilon}} \ln \left| \frac{\sqrt{\varepsilon} + \text{Re} \sqrt{\varepsilon \pm 2t_e}}{\sqrt{\varepsilon} - \text{Re} \sqrt{\varepsilon \pm 2t_e}} \right| \equiv \frac{A_{\text{по}}^{\pm}}{\sqrt{\varepsilon}} \ln F^{\pm}(\varepsilon). \quad (\text{П2.20})$$

Использование усреднения (П2.3) дает выражение для обратных времен релаксации при поглощении и возбуждении фонона:

$$\frac{1}{\tau_{\text{по}}^{\pm}} = \frac{2A_{\text{по}}^{\pm}}{\sqrt{\pi}} \int_0^{\infty} \ln F^{\pm}(\varepsilon) e^{-\varepsilon} d\varepsilon. \quad (\text{П2.21})$$

Вычисление этих интегралов лучше производить отдельно. При поглощении фонона подкоренные выражения положительны и оператор Re теряет смысл, поэтому

$$\frac{1}{\tau_{p\text{по}}^+} = \frac{2A_{\text{по}}^+}{\sqrt{\pi}} \int_0^{\infty} \ln \frac{\sqrt{\varepsilon + 2t_e} + \sqrt{\varepsilon}}{\sqrt{\varepsilon + 2t_e} - \sqrt{\varepsilon}} e^{-\varepsilon} d\varepsilon. \quad (\text{П2.22})$$

Дальнейшее интегрирование удобно проводить по частям. Обозначая $u = \ln F^+(\varepsilon)$, $dv = e^{-\varepsilon} d\varepsilon$ и учитывая, что $e^{-\varepsilon} \ln F^+|_0^{\infty} = 0$, а $d(\ln F^+) = d\varepsilon/2\sqrt{\varepsilon(\varepsilon + 2t_e)}$, приходим к интегралу

$$\frac{1}{\tau_{p\text{по}}^+} = \frac{A_{\text{по}}^+}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-\varepsilon} d\varepsilon}{2\sqrt{\varepsilon(\varepsilon + 2t_e)}} = \frac{A_{\text{по}}^+}{\sqrt{\pi}} e^{+t_e} K_0(t_e). \quad (\text{П2.23})$$

Здесь $K_0(t_e)$ – модифицированная функция Бесселя от мнимого аргумента второго рода нулевого порядка.

При возбуждении фонона (нижний знак) выражение (П2.21) имеет смысл только при $\varepsilon > 2t_e$, поэтому

$$\frac{1}{\tau_{p\text{по}}^-} = \frac{2A_{\text{по}}^-}{\sqrt{\pi}} \int_{2t_e}^{\infty} \ln \frac{\sqrt{\varepsilon} + \sqrt{\varepsilon - 2t_e}}{\sqrt{\varepsilon} - \sqrt{\varepsilon - 2t_e}} e^{-\varepsilon} d\varepsilon. \quad (\text{П2.24})$$

Заменой переменной $\varepsilon = \xi + 2t_e$ можно привести этот интеграл к виду (П2.22)

$$\frac{1}{\tau_{p\text{по}}^-} = \frac{2A_{\text{по}}^-}{\sqrt{\pi}} e^{-2t_e} \int_0^{\infty} \ln \frac{\sqrt{\xi + 2t_e} + \sqrt{\xi}}{\sqrt{\xi + 2t_e} - \sqrt{\xi}} e^{-\xi} d\xi = \frac{A_{\text{по}}^-}{\sqrt{\pi}} e^{-t_e} K_0(t_e). \quad (\text{П2.25})$$

Выражения (П2.22) и (П2.23) можно представить одной формой

$$\frac{1}{\tau_{p\text{по}}^{\pm}} = \frac{\alpha \omega_0}{\sqrt{2\pi}} \frac{e^{\mp(t-t_e)}}{\text{sh } t} \sqrt{t_e} K_0(t_e). \quad (\text{П2.26})$$

Полное время релаксации импульса при полярном оптическом рассеянии находится из суммы последних выражений при поглощении и возбуждении

$$\frac{1}{\tau_{p\text{по}}} = \frac{1}{\tau_{p\text{по}}^+} + \frac{1}{\tau_{p\text{по}}^-} = \frac{\alpha \omega_0 \sqrt{2}}{\sqrt{\pi}} \frac{\text{ch}(t - t_e)}{\text{sh } t} \sqrt{t_e} K_0(t_e). \quad (\text{П2.27})$$

Междолинное рассеяние. Выражение (2.16) для скорости междолинного рассеяния во многом сходно с (2.13) для неполярного оптического рассеяния и для нормированной энергии ($E/k_B T_e \equiv \varepsilon$) с учетом (П2.13) для переходов из i -долины в j -ю может быть представлено в виде

$$\begin{aligned} \lambda_{ij}^{\pm} &= \frac{Z_j (m_j^*)^{3/2} D_{ij}^2}{\sqrt{2\pi} \rho \hbar^3 \omega_{ij}} \cdot \frac{e^{\mp t}}{\text{sh } t} \sqrt{k_B T_e} \text{Re} \sqrt{\varepsilon \pm 2t_e - \delta} \equiv \\ &\equiv A_{ij}^{\pm} \text{Re} \sqrt{\varepsilon \pm 2t_e - \delta}. \end{aligned} \quad (\text{П2.28})$$

Здесь надо иметь в виду, что в параметры t_e и t надо подставлять энергию междолинных фононов $\hbar \omega_{ij}$, а величина $\delta \equiv \Delta_{ij}/k_B T_e$ есть нормированное междолинное расстояние Δ_{ij} , отсчитываемое от дна j -долины до дна i -й с учетом знака (тут для сокращения записи индексы возле δ опущены).

Как и для оптического рассеяния определим время релаксации импульса в аналитическом виде, используя усреднение по (П2.3.)

$$\frac{1}{\tau_{pij}^{\pm}} = \frac{2A_{ij}^{\pm}}{\sqrt{\pi}} \int_0^{\infty} \sqrt{\varepsilon} \text{Re} \sqrt{\varepsilon \pm 2t_e - \delta} e^{-\varepsilon} d\varepsilon. \quad (\text{П2.29})$$

Дальнейшая процедура интегрирования зависит от соотношения между энергией фононов и междолинным расстоянием. Практически более важной является ситуация, когда $\hbar \omega_{ij} < \Delta_{ij}$, или $\delta > 2t_e$, так как именно в этом случае проявляются специфические особенности многодолинных полупроводников.

При переходах электронов из нижней долины в верхнюю величина δ , а значит и $\delta \pm 2t_e > 0$, и рассеяние возможно только если энергия электрона превышает величину $\Delta_{ij} \pm \hbar \omega_{ij}$. При этом оператор Re можно опустить, если нижним пределом интегрирования будет величина $\delta \pm 2t_e$. В этом случае интеграл (П2.3.) вычисляется как и при оптическом рассеянии

$$\int_{\delta \mp 2t_e}^{\infty} \sqrt{\varepsilon} \sqrt{\varepsilon - (\delta \mp 2t_e)} e^{-\varepsilon} d\varepsilon = (\delta / 2 \mp t_e) e^{-(\delta / 2 \mp t_e)} K_1(\delta / 2 \mp t_e). \quad (\text{П2.30})$$

Переходы в нижнюю долину ($\delta < 0$) возможны при любой энергии, а так как $\pm 2t_e - \delta > 0$, то оператор Re теряет смысл и

$$\int_0^{\infty} \sqrt{\varepsilon} \sqrt{\varepsilon - \delta \pm 2t_e} e^{-\varepsilon} d\varepsilon = (\pm t_e - \delta / 2) e^{-(\pm t_e - \delta / 2)} K_1(\pm t_e - \delta / 2). \quad (\text{П2.31})$$

Эти два интеграла можно представить одной формой $|\delta / 2 \mp t_e| \exp(-\delta / 2 \pm t_e) K_1|\delta / 2 \mp t_e|$, а время релаксации импульса для рассеяния между **неэквивалентными** долинами при поглощении и возбуждении фононов вычислять по формуле

$$\frac{1}{\tau_{pij}^{\pm}} = \frac{Z_j(m_j^*)^{3/2} D_{ij}^2}{\sqrt{2\pi}^{3/2} \rho \hbar^3 \omega_{ij}} \cdot \frac{e^{-\delta/2}}{\text{sh } t} \sqrt{\kappa_B T_e} e^{\mp(t-t_e)} \left| \frac{\delta}{2} \mp t_e \right| K_1 \left| \frac{\delta}{2} \mp t_e \right|. \quad (\text{П2.32})$$

Выражение для суммарного обратного времени релаксации не получается компактным в общем виде, но для наиболее важного случая при существенном превышении $\Delta_{ij} \gg \hbar \omega_{ij}$ можно пренебречь t_e по сравнению с $\delta/2$ в медленных функциях под знаком модуля и представить

$$\frac{1}{\tau_{pij}} = \frac{Z_j(m_j^*)^{3/2} D_{ij}^2}{2\pi^{3/2} \rho \hbar^2 (\hbar \omega_{ij})^{1/2}} \cdot \frac{\text{ch}(t - t_e)}{\text{sh } t} \cdot \frac{|\delta|}{t_e^{1/2}} e^{-\delta/2} K_1 \left| \frac{\delta}{2} \right|. \quad (\text{П2.33})$$

Если междолинное расстояние невелико и $\Delta_{ij} < \hbar \omega_{ij}$, то для времени релаксации импульса можно использовать формулу (П2.19) для неполярного оптического рассеяния, где к энергии фонона надо внести поправку $\pm \Delta_{ij}$, либо (П2.32), где под знаком модуля подставлять половину общего изменения энергии, которое происходит при том или ином виде рассеяния.

В частном случае при $\Delta_{ij} = 0$ приходим к рассеянию между *эквивалентными* долинами, а интеграл (П2.29) фактически сводится к (П2.15) и соответствующее время релаксации может быть вычислено, как и для рассеяния на оптическом деформационном потенциале

$$\frac{1}{\tau_{pii}^{\pm}} = \frac{(Z_i - 1)(m_i^*)^{3/2} D_{ii}^2}{\sqrt{2\pi^{3/2}} \rho \hbar^3 \omega_{ii}} \cdot \frac{e^{\mp(t-t_e)}}{\text{sh } t} \sqrt{k_B T_e} K_1(t_e), \quad (\text{П2.34})$$

а для полного времени

$$\frac{1}{\tau_{pii}^{\pm}} = \frac{(Z_i - 1)(m_i^*)^{3/2} D_{ii}^2}{\pi^{3/2} \rho \hbar^2 (\hbar \omega_{ii})^{1/2}} \cdot \frac{\text{ch}(t - t_e)}{\text{sh } t} \sqrt{t_e} K_1(t_e). \quad (\text{П2.35})$$

Здесь $(Z_i - 1)$ представляет собой количество «других» эквивалентных долин, в которые происходит переход электрона, а m^* – эффективная масса плотности состояний в одной долине.

Для упрощения полученных расчетных соотношений можно использовать приближенные формулы для функций Бесселя. При малых значениях аргумента ($x \ll 1$) возможны аппроксимации $K_0(x) \approx \ln(2/x)$, $K_1(x) \approx 1/x$, а при больших аргументах ($x \gg 1$) $K_0(x) \approx K_1(x) \approx \sqrt{\pi/2x} \exp(-x)$. Кроме того полезным может быть дифференциальное соотношение $K_n'(x) = -K_{n-1}(x) - nK_n(x)/x$.

В окрестности значений точки x_0 около 1 возможно линейное представление функции $K_1(x)$ при разложении ее в ряд по степеням, обратным аргументу

$$K_1(x) \approx -x_0 K_0(x_0) + [x_0 K_0(x_0) + K_1(x_0)] x_0/x. \quad (\text{П2.36})$$

Здесь использовано предыдущее соотношение для функции первого порядка: $K_1'(x) = -K_0(x) - K_1(x)/x$.

Такое представление позволяет существенно упростить вид функции (П2.33) для междолинного рассеяния. Если в качестве точки разложения принять электронную температуру $T_e = \Delta_{ij}/k_B$, при которой $\delta = 1$, то

$$K_1\left(\frac{\delta}{2}\right) \approx -\frac{K_0(0,5)}{2} + \left[\frac{K_0(0,5)}{2} + K_1(0,5)\right] \frac{k_B T_e}{\Delta_{ij}} = -0,4622 + 2,1186 \frac{k_B T_e}{\Delta_{ij}}$$

Обычно при этих температурах $t_e \ll t$ и $\text{ch}(t - t_e) \approx \text{cht}$, а остальные множители, кроме $K_1(\delta/2)$, представляют собой весьма слабую функцию электронной температуры, имеющую пологий максимум при $\delta = 1$. Заменяя это произведение его максимальным значением

$$[t_e^{-1/2} \cdot \delta \cdot \exp(\delta/2)]_{\max} = \sqrt{2\Delta_{ij}/e \hbar \omega_{ij}},$$

получим, что обратное время междолинной релаксации с хорошим приближением можно считать линейной функцией температуры

$$\frac{1}{\tau_{pij}} = \frac{Z_j(m_j^*)^{3/2} D_{ij}^2}{2\pi^{3/2} \rho \hbar^3 \omega_{ij}} \sqrt{2\Delta_{ij}/e} \cdot \text{cth} t(-0,4622 + 2,1186 k_B T_e / \Delta_{ij}). \quad (\text{П2.37})$$

Надо иметь в виду, что эта аппроксимация пригодна для переходов в верхние j -долины ($\Delta_{ij} > 0$) для области температур (2000...6000 К), где междолинное рассеяние является преобладающим. Для расчета времени релаксации в нижние i -долины ($\Delta_{ji} < 0$) может оказаться удобным использовать соотношение (1.70), принимая во внимание, что $\tau_{nij} \approx \tau_{pij}$:

$$\frac{1}{\tau_{pji}} = \frac{n_i}{n_j} \cdot \frac{1}{\tau_{pij}} = \left(m_i^*/m_j^*\right)^{3/2} \exp(-\Delta_{ji}/k_B T_e) \frac{1}{\tau_{pij}}. \quad (\text{П2.38})$$

Для рассеяния между эквивалентными долинами аргументом функции Бесселя служит отношение $t_e = \hbar\omega_{ij}/2k_B T_e$, которое в области больших электронных температур существенно меньше 1, и возможна аппроксимация $K_1(x) \approx 1/x$, которая позволяет (П2.35) представить в более простом виде

$$\frac{1}{\tau_{pii}} = \frac{(Z_i - 1)(m_i^*)^{3/2} D_{ii}^2}{\pi^{3/2} \rho \hbar^3 \omega_{ii}} \cdot \text{cth} t \cdot \sqrt{2k_B T_e}. \quad (\text{П2.39})$$

Для полярного оптического рассеяния для больших электронных температур ($\hbar\omega_0/2k_B T_e \ll 1$) можно использовать аппроксимацию: $K_0(x) \approx \ln(2/x) - 0,5772$.

Приложение к главе 3

П3. ВЫСОКОЧАСТОТНАЯ ПРОВОДИМОСТЬ

Ограничимся режимом малого сигнала, при котором все переменные величины $a(t)$ представляются в виде суммы постоянной составляющей a_0 и малой гармонически меняющейся во времени переменной с амплитудой a_{\sim} : $a(t) = a_0 + a_{\sim} e^{i\omega t}$. Пренебрегая затем членами второго порядка малости и учитывая, что $d/dt = i\omega$, используемые уравнения разделяем на не

зависящие от времени для постоянных составляющих и линейные уравнения для переменных.

Использование такой процедуры для плотности тока $j = env$ дает для его высокочастотной составляющей соотношение

$$j_{\sim} = en_0 v_{\sim} + ev_0 n_{\sim} \quad (\text{ПЗ.1})$$

Величина v_{\sim} может быть найдена из уравнения баланса импульса (1.57), которое для переменной составляющей примет вид

$$i\omega v_{\sim} = (e/m^*)\mathcal{E}_{\sim} - (\tau_{p0}^{-1}) v_{\sim} - v_0 (\tau_p^{-1})_{\sim},$$

откуда

$$v_{\sim} = (e/m^*)\tau_{p\omega}\mathcal{E}_{\sim} - v_0\tau_{p\omega} (\tau_p^{-1})_{\sim}, \quad (\text{ПЗ.2})$$

где

$$\tau_{p\omega} = \tau_{p0}/(1 + i\omega\tau_{p0}). \quad (\text{ПЗ.3})$$

Величина $(\tau_p^{-1})_{\sim}$ может быть найдена из функции $\tau_p^{-1}(T_e(\mathcal{E}))$ с учетом неявной зависимости от напряженности поля при малом сигнале следующим образом

$$\tau_{p\sim}^{-1} = \frac{d(\tau_p^{-1})}{dT_e} T_{e\sim}. \quad (\text{ПЗ.4})$$

Тогда вклад первого слагаемого из (ПЗ.1) в динамическую проводимость представляется в виде

$$\sigma_{\omega}^{(1)} = \frac{e^2 n_0 \tau_{p\omega}}{m^*} \left[1 - \frac{m^* v_0}{e} \frac{d(\tau_p^{-1})}{dT_e} T_{e\sim} \right]. \quad (\text{ДЗ.5})$$

Производная здесь, как и ранее, заменяется для сильных полей через отношение τ_{p0}^{-1}/T_e . Высокочастотная составляющая температуры $T_{e\sim}$ определяется инерционностью процессов релаксации энергии и может быть найдена из нестационарного уравнения баланса энергии (1.60). Если принять, что при насыщении скорости хаотическое (тепловое) движение преобладает над дрейфовым ($\gamma = E_T/E_d \gg 1$) и соответственно $\tau_E \approx \text{const}(T_e)$, то малосигнальная часть уравнения баланса энергии примет вид

$$i\omega E_{T\sim} = ev_H \mathcal{E}_{\sim} - E_{T\sim}/\tau_{E0}. \quad (\text{ПЗ.6})$$

Отсюда, вводя понятие комплексного высокочастотного времени релаксации энергии

$$\tau_{E\omega} = \tau_{E0}/(1 + i\omega\tau_{E0}), \quad (\text{П3.7})$$

получим для малосигнальной части электронной температуры

$$T_{e\sim} = (2 ev_H \tau_{E\omega}/3k_B) \mathcal{E}_{\sim}. \quad (\text{П3.8})$$

В итоге (П3.5) можно представить в виде

$$\sigma_{\omega}^{(1)} = \frac{\sigma_0}{1 + i\omega\tau_{p0}} \left[1 - \frac{v_0}{v_H} \cdot \frac{1}{1 + i\omega\tau_{E0}} \right]. \quad (\text{П3.9})$$

Величина переменной составляющей концентрации n_{\sim} в (П3.1) находится из уравнений баланса для концентраций электронов в долинах (1.67), (1.68). Записывая их для относительных величин $b_i = n_i/n$ (в двухдолинном приближении) и принимая во внимание, что $b_{i\sim} = -b_{j\sim} \equiv b_{\sim}$, получим для малосигнальной части

$$i\omega b_{\sim} = -b_{\sim}(1/\tau_{ij0} + 1/\tau_{ji0}) - b_{i0}(\tau_{ij}^{-1})_{\sim} + b_{j0}(\tau_{ji}^{-1})_{\sim}. \quad (\text{П3.10})$$

Если ввести понятие суммарного для всех долин времени междолинной релаксации

$$\tau_M = (1/\tau_{ij0} + 1/\tau_{ji0})^{-1}, \quad (\text{П3.11})$$

то из (П3.10) можно переменную составляющую заселенности представить в виде:

$$b_{\sim} = \frac{b_{i0}b_{j0}}{1 + i\omega\tau_M} \left[\tau_{ij0} (\tau_{ij}^{-1})_{\sim} - \tau_{ji0} (\tau_{ji}^{-1})_{\sim} \right], \quad (\text{П3.12})$$

где использованы соотношения $b_{i0} = \tau_M/\tau_{ji0}$ и $b_{j0} = \tau_M/\tau_{ij0}$, следующие непосредственно из определения заселенностей и правых частей уравнений баланса концентраций для их постоянных составляющих. Кроме того соотношение $b_{j0}/b_{i0} = \tau_{ij0}/\tau_{ji0}$ удобно использовать для представления второго сомножителя в (П3.12) в более простой форме

$$b_{\sim} = \frac{b_{i0}b_{j0}}{1 + i\omega\tau_m} \frac{(\tau_{ji} / \tau_{ij})_{\sim}}{\tau_{ji0} / \tau_{ij0}}. \quad (\text{П3.13})$$

Если воспользоваться для отношения времен междолинной релаксации соотношением (1.70), то получим

$$b_{\sim} = -\frac{b_{i0}b_{j0}}{1 + i\omega\tau_m} \frac{\Delta_{ji}}{k_B T_e^2} T_{e\sim}. \quad (\text{П3.14})$$

Используя (П3.8) для $T_{e\sim}$ получим из (П3.1) в дополнение к (П3.5) второе слагаемое для высокочастотной электропроводности

$$\sigma_{\omega}^{(2)} = -\frac{e^2 n_0 \tau_{p0}}{m^*} \cdot \frac{v_0}{v_H} \cdot \frac{\tau_{E\omega}}{\tau_{E0}} \cdot \frac{b_{i0}b_{j0}}{1 + i\omega\tau_m} \cdot \frac{\Delta_{ji}}{k_B T_e}. \quad (\text{П3.15})$$

Если по-прежнему считать, что величина проводимости определяется в основном электронами Γ -долины, то сумма (П3.5) и (П3.15) при $\omega \rightarrow 0$ совпадает с формулой (3.37), описывающей низкочастотную дифференциальную проводимость. Величина высокочастотной электропроводности, нормированная на ее статическое значение, может быть рассчитана из соотношения

$$\frac{\sigma_{\omega}}{\sigma} = \frac{\tau_{p\omega}}{\tau_{p0}} - \frac{v_{0\Gamma}}{v_{H\Gamma}} \cdot \frac{\tau_{E\omega}}{\tau_{E0}} \left(\frac{\tau_{p\omega}}{\tau_{p0}} + \frac{b_{i0}b_{j0}}{1 + i\omega\tau_m} \cdot \frac{\Delta_{ji}}{k_B T_e} \right). \quad (\text{П3.16})$$

Приложения к главе 4

П4.1. Равновесное состояние контакта канал-подложка

На границе эпитаксиального слоя с подложкой за счет контакта двух полупроводников с различной концентрацией донорной примеси из-за диффузии носителей возникает контактное поле, вызывающее режимы обогащения (в подложке) и обеднения (в канале). Уравнение Пуассона в плоскости, поперечной к каналу запишется в виде

$$\frac{d^2 V}{dy^2} = \frac{e}{\epsilon \epsilon_0} [n(y) - N_{дп}],$$

где в равновесной ситуации $n(y) = n_0 \exp[eV(y) / kT]$.

Выбирая начало отсчета потенциала в глубине эпитаксиального слоя (рис. П4.1), имеем $n_0 = N_d$. В глубине подложки, где $n(y) = N_{дп}$, потенциал будет равен $-V_{к0}$. Контактная разность потенциалов определяется соотношением: $V_{к0} = (kT/e) \ln(N_d / N_{дп})$.

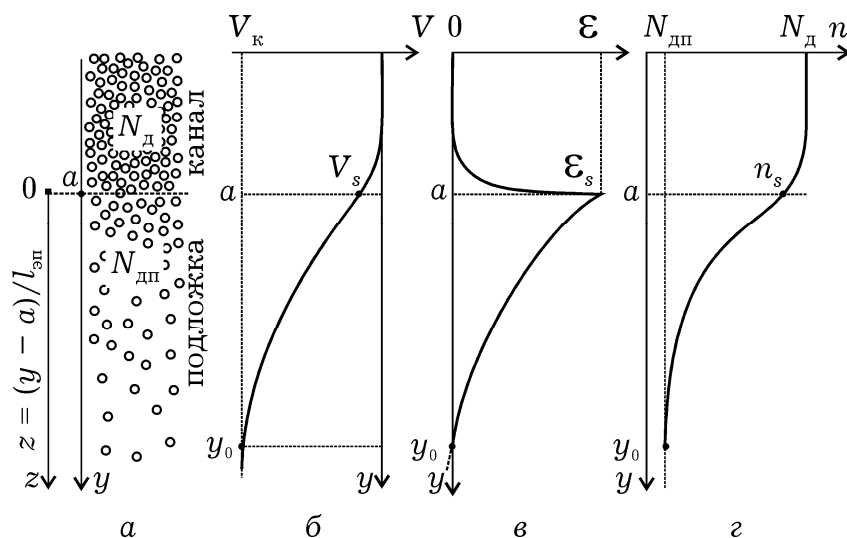


Рис. П4.1. Распределения потенциала, напряженности электрического поля и электронной концентрации вблизи границы донорных полупроводников с разной концентрации примеси

С учетом принятых положений и использованием обозначений: $l_{эп} \equiv \sqrt{2\epsilon\epsilon_0 kT / e^2 N_{дп}}$, $Y \equiv eV/kT$, $z \equiv (y-a)/l_{эп}$, уравнение Пуассона примет вид:

$$d^2 Y / dz^2 = 2[\exp(Y + Y_{к}) - 1]. \quad (\text{П4.1})$$

Умножая обе части уравнения (П4.1) на dY / dz

$$\frac{dY}{dz} \frac{d}{dz} \left(\frac{dY}{dz} \right) = 2[\exp(Y + Y_{к}) - 1] \frac{dY}{dz},$$

получим

$$d \left(\frac{dY}{dz} \right)^2 = 2[\exp(Y + Y_{к}) - 1] dY.$$

Интегрируя это уравнение, получим:

$$dY / dz = \pm \sqrt{2[\exp(Y + Y_{к}) - (Y + Y_{к}) + C]}, \quad (\text{П4.2})$$

где C – константа интегрирования, определяемая из условия $dY/dz \approx 0$ в глубине подложки.

При условии $\exp(Y + Y_k) \gg (Y + Y_k)$, что достаточно хорошо выполняется при $N_d \gg N_{дп}$ и при $dY/dz \approx 0$ в глубине подложки, из (П4.2) можно получить:

$$dY/dz \approx -\sqrt{2} \exp(Y + Y_k)/2. \quad (\text{П4.3})$$

Вводя нормированный потенциал Y_s на границе канал-подложка и интегрируя (П4.3) от Y_s до Y , получим приближенное решение для потенциала:

$$Y \approx Y_s - 2 \ln \left\{ 1 + \frac{z}{\sqrt{2}} \exp[(Y + Y_k)/2] \right\}. \quad (\text{П4.4})$$

Принятое допущение приводит к тому, что потенциал достигает значения, равного контактному, не при $z \rightarrow \infty$, а в точке

$$z_0 = 1 - \exp[-(Y_s + Y_k)/2] \approx 1.$$

Распределение носителей в подложке может быть описано выражением:

$$n(z) = N_d e^Y = N_d e^{Y_s} / \left\{ 1 + \frac{z}{\sqrt{2}} \exp[(Y_s + Y_k)/2] \right\}^2, \quad (\text{П4.5})$$

причем в точке z_0 справедливо равенство $n(z_0) = N_{дп}$.

Кроме того, из (П4.5) следует, что в равновесном состоянии в подложке имеется некоторый избыточный заряд, концентрация которого на границе канал-подложка ($z = 0$) определяется по соотношению:

$$n(0) \equiv n_s = N_d \exp(Y_s).$$

Средний по толщине избыточный заряд электронов определим по формуле:

$$n_{ср} = \sqrt{N_{дп} N_d e^{Y_s}} = \sqrt{N_{дп} n_s}. \quad (\text{П4.6})$$

Поверхностный потенциал Y_s можно рассчитать, решая уравнение Пуассона для обедненного слоя в канале. Принимая, что обеднение не является глубоким, можно распределение потенциала описать экспоненциальной функцией с характерной

константой l_3 , равной длине экранирования для эпитаксиального слоя:

$$V(y) = V_s \exp[(y-a)/l_3]. \quad (\text{П4.7})$$

Соотношение, описывающее распределение напряженности электрического поля в канале, найдем из уравнения (П4.7), принимая $\mathcal{E}_s = |V_s|/l_3$:

$$\mathcal{E}(y) = -(V_s/l_3) \exp[(y-a)/l_3]. \quad (\text{П4.8})$$

С учетом принятых ранее обозначений, переходя от переменной z к переменной y , из уравнения (П4.4) получим соотношение для распределения потенциала в подложке:

$$V(y) = V_s - 2(kT/e) \ln \left\{ 1 + \frac{y-a}{l_{3п}} \exp \left[\frac{e(V_s + V_k)}{2kT} \right] \right\}. \quad (\text{П4.9})$$

Воспользовавшись $dV(y)/dy = \mathcal{E}(y)$ из (П4.9) получим:

$$\mathcal{E}(y) = -\frac{2kT}{el_{3п}} \exp \left[\frac{e(V_s + V_k)}{2kT} \right] / \left\{ 1 + \frac{y-a}{l_{3п}} \exp \left[\frac{e(V_s + V_k)}{2kT} \right] \right\}. \quad (\text{П4.10})$$

Приравнивая между собой напряженности электрического поля на границе ($y = a$), получим:

$$V_s = \frac{\sqrt{2kT}}{e} \sqrt{\frac{N_{дп}}{N_d}} \exp[e(V_s + V_k)/2kT] \quad (\text{П4.11})$$

Переходя к нормированным потенциалам при $V_k = V_{k0}$, запишем:

$$Y_{s0} = \sqrt{2N_{дп}/N_d} \exp[e(Y_{s0} + Y_{k0})/2].$$

Так как $\sqrt{N_{дп}/N_d} = \exp(-Y_{k0}/2)$, то получим трансцендентное уравнение относительно Y_{s0} , откуда $V_{s0} = -0,901kT/e$. Поверхностная концентрация при этом допущении определяется соотношением:

$$n_{s0} = N_d e^{Y_{s0}} = N_d \exp(eV_{s0}/kT) \approx 0,406N_d,$$

а функции, описывающие распределения концентрации в канале и в подложке соответственно запишутся:

$$n(y) = N_d \exp \left[\frac{eV_{s0}}{kT} \exp \left(\frac{y-a}{l_3} \right) \right] \approx N_d \left[1 + \frac{eV_{s0}}{kT} \exp \left(\frac{y-a}{l_3} \right) \right] \quad (\text{П4.12})$$

$$n(y) = N_d \exp \left(\frac{eV_{s0}}{kT} \right) / \left\{ 1 + \frac{y-a}{l_{3п}} \exp \left[\frac{e(V_{s0} + V_k)}{2kT} \right] \right\}^2. \quad (\text{П4.13})$$

Уравнение для напряженности поля в подложке (П4.10) с учетом вычисленного значения V_{s0} запишется:

$$\mathcal{E}(y) = (|V_s| / l_{eп}) \exp \left(\frac{eV_k}{2kT} \right) / \left\{ 1 + \frac{y-a}{l_{eп}} \exp \left[\frac{e(V_{s0} + V_k)}{2kT} \right] \right\}. \quad (\text{П4.14})$$

Распределения потенциала, напряженности электрического поля и концентрации носителей у границы эпитаксиальный слой-подложка в равновесном состоянии показаны на рис. П4.1.

Однако барьер будет определяться таким соотношением для контактной разности потенциалов лишь в отсутствие факторов, нарушающих равновесие. Одним из факторов, приводящих к нарушению равновесия и возникновению явлений, в определенной степени аналогичных термоэлектрическим эффектам, является разогрев электронов электрическим полем. Отличие заключается в том, что греются лишь электроны, а температура решетки остается постоянной и однородной. Эффекты, названные электроградиентными, обусловлены разностью температур электронного газа. Градиенты температур в поперечной к границе эпитаксиального слоя с подложкой плоскости, когда имеет место контакт полупроводников с существенно различной концентрацией примесей, могут возникать вследствие различных условий разогрева носителей электрическим полем. Эти условия связаны в первую очередь с малой ролью примесного рассеяния в подложке.

Термоэдс на nn^+ -переходе может быть определена соотношением², $V_T = V_{k0}(T_e/T - 1)$, где $eV_{k0} = kT_e \ln(n^+/n)$ - высота

² Ашмонтас С. Электроградиентные явления в полупроводниках/ под. ред. Ю. Пожелы.- Вильнюс: Мокслас, 1984 -184с.- (Электроны в полупроводниках; 5).

потенциального барьера pn^+ -перехода, T_e - температура электронов.

Следует отметить, что это соотношение получено в предположении, что концентрация электронов не зависит от напряженности электрического поля. Однако, при примерно одинаковых продольных полях в канале и в подложке транзисторной структуры, вследствие различной степени разогрева носителей, возникает поперечная термоэдс, определяемая разностью температур электронного газа канала и подложки:

$$V_T = V_{k0} \left(\frac{T_k - T}{T} \right) - V_{k0} \left(\frac{T_{\Pi} - T}{T} \right) = V_{k0} \left(\frac{T_k - T_{\Pi}}{T} \right),$$

где T_k , T_{Π} - температура электронного газа в канале и подложке соответственно. Выражение для потенциала барьера канал-подложка с учетом разогрева носителей запишется:

$$V_k = V_{k0} + V_T = V_{k0} \left(1 + \frac{T_k - T_{\Pi}}{T} \right). \quad (\text{П4.15})$$

Анализ этого выражения показывает, что при отсутствии факторов, приводящих к разогреву электронов или при одинаковых условиях разогрева ($T_k = T_{\Pi}$), высота потенциального барьера на границе канал-подложка определяется контактной разностью потенциалов V_{k0} , а при неоднородном разогреве высота потенциального барьера будет изменяться за счет перераспределения электронов между каналом и подложкой.

П4.2. Уравнения непрерывности для канала и подложки

В общем случае между каналом и подложкой хаотическое движение вызывает взаимный обмен носителями за счет встречных потоков, поэтому возникает ток через границу в поперечном направлении $j_{\perp} = j_{кп} - j_{пк}$, где $j_{кп}$, $j_{пк}$ - плотности тока из канала в подложку и из подложки в канал, соответственно.

Проанализируем уравнение непрерывности с учетом влияния подложки. Для стационарного случая ($\partial/\partial t = 0$) уравнение непрерывности тока означает, что полный ток через замкнутую поверхность равен нулю:

$$\oint_S \vec{j} d\vec{s} = 0.$$

Интегрируя по поверхности S_1 , (рис. 5.8, в) ограничивающей элемент канала длиной dx , получим уравнение:

$$j_{\perp} + d(jh)/dx = 0,$$

где j_{\perp} - поперечная составляющая тока, проходящего через границу раздела между каналом и подложкой.

Аналогично уравнение непрерывности для подложки примет вид:

$$d(j_{\pi}h_{\pi})/dx - j_{\perp} = 0.$$

Основным механизмом обмена носителей является тепловое движение, происходящее в условиях разных концентраций и температур в канале и подложке в присутствии электрического поля. Результирующий ток через n^+n -контакт может быть определен из анализа неравновесия между процессами дрейфа и диффузии с учетом электроградиентных эффектов.

Для вывода уравнений, обеспечивающих связь по току между электронами канала и подложки, воспользуемся соотношением для плотности тока дрейфа и диффузии и полученными выражениями для напряженности электрического поля (П4.8) и (П4.10) и концентрации (П4.13) и (П4.14). Плотность тока в подложке с учетом (П4.10) и (П4.14) запишется:

$$j_{др} = en\mu_{\pi} \frac{dV}{dy} = \frac{\sqrt{2}\mu_{\pi}kTn_{s0} \exp[(Y_{s0} + Y_{\kappa})/2]}{\{1 + (z/\sqrt{2})\exp[(Y_{s0} + Y_{\kappa})/2]\}^3 l_{\pi}}. \quad (\text{П4.16})$$

Соответственно для диффузионного тока:

$$j_{D} = eD_{\pi} \frac{dn}{dy} = -\frac{\sqrt{2}eD_{\pi}n_{s0} \exp[(Y_{s0} + Y_{\kappa})/2]}{\{1 + (z/\sqrt{2})\exp[(Y_{s0} + Y_{\kappa})/2]\}^3 l_{\pi}}. \quad (\text{П4.17})$$

В итоге имеем, что токи дрейфа и диффузии уравновешивают друг друга при выполнении соотношения Эйнштейна $\mu kT/e = D$.

Используя выражения (П4.8) и (П4.13) для эпитаксиального слоя, вычислим плотность дрейфовой составляющей тока:

$$j_{др} = en\mu\mathcal{E} = -e\mu N_{\pi} \frac{V_{s0}}{l_{\pi}} \exp\left[\frac{eV_{s0}}{kT} \exp\left(\frac{y-a}{l_{\pi}}\right)\right] \exp\left(\frac{y-a}{l_{\pi}}\right). \quad (\text{П4.18})$$

и диффузионной:

$$j_D = eD_k \frac{dn}{dy} = eD_k N_d \frac{eV_{s0}}{kTl_\vartheta} \exp\left[\frac{eV_{s0}}{kT} \exp\left(\frac{y-a}{l_\vartheta}\right)\right] \exp\left(\frac{y-a}{l_\vartheta}\right). \quad (\text{П4.19})$$

Рассчитаем плотности токов $j_{кп}$, и $j_{пк}$, используя результаты, полученные в равновесном приближении. При этом считаем механизм «инжекции» носителей на границе канала с подложкой диффузионным.

Воспользуемся (П4.18) на границе ($y = a$) и считая, что диффузия из канала обеспечивается за счет превышения текущей концентрации $n(x)$ над концентрацией доноров N_d вследствие эффектов накопления носителей, запишем:

$$j_{Dкп} = (e|V_{s0}|/kT) \exp(eV_{s0}/kT) eD_k [n(x) - N_d] / l_\vartheta. \quad (\text{П4.20})$$

Встречный ток из подложки в канал определится из (П4.17) на границе с каналом ($z = 0$):

$$j_{Dпк} = \sqrt{2}eD_\Pi n_{s0} \exp[e(V_{s0} + V_k)/2kT] / l_{\vartheta\Pi}.$$

Здесь, а также в соотношениях, приведенных выше, D_k и D_Π коэффициенты диффузии для канала и подложки соответственно. Используя вместо поверхностной концентрации n_{s0} усредненное по подложке значение концентрации (П4.6), а также с учетом эффектов накопления в подложке за счет разогрева электрическим полем диффузионную составляющую тока определим по формуле:

$$j_{Dпк} = \sqrt{2}eD_\Pi [n_\Pi(x) - n_{cp}] \exp[e(V_{s0} + V_k)/2kT] / l_{\vartheta\Pi}. \quad (\text{П4.21})$$

Особенность полученного выражения заключается в том, что в соответствии с (П4.15) высота потенциального барьера зависит от разности температур электронного газа в канале и в подложке. Высота потенциального барьера зависит от градиента температур так, что если разогрев в канале преобладает, то потенциальный барьер повышается за счет инжектированных из канала в подложку носителей. Преимущественный разогрев в подложке ($T_\Pi > T_\Pi$) приводит к возврату носителей в канал и понижению потенциального барьера. Возврат электронов из подложки в канал обеспечивается также в соответствии с (П4.20) за затвором, когда

происходит расширение канала и текущее значение концентрации $n(x)$ может стать меньше N_d . Знак тока $j_{\text{кп}}$ при этом изменяется на обратный, что обеспечивает быструю релаксацию тока в канале.

П4.3. Уравнение Пуассона для канала и подложки

Для анализа уравнения Пуассона в канале и подложке, рассмотрим фрагмент области канала и подложки длиной dx (рис. 5.8, з). Применим теорему Гаусса $\oint_s \vec{\mathcal{E}} d\vec{s} = -\rho(x)/\varepsilon\varepsilon_0$ к областям канала (поверхность S_1). Для усредненных по поперечному сечению канала значений напряженности продольного электрического поля и допущения об отсутствии поперечной составляющей поля на границе с подложкой, получим уравнение Пуассона:

$$d(\mathcal{E}h)/dx = eh[n(x) - N_d]/\varepsilon\varepsilon_0. \quad (\text{П4.22})$$

Подобные рассуждения для подложки (поверхность S_2 рис. 5.8, з) приводят к уравнению:

$$d(\mathcal{E}_\pi h_\pi)/dx = eh_\pi[n_\pi(x) - N_{\text{дп}}]/\varepsilon\varepsilon_0 + \mathcal{E}_\perp. \quad (\text{П4.23})$$

Используя полученный из равновесного анализа результат для среднего избыточного заряда электронов по толщине подложки (П4.7), получим уравнение Пуассона для подложки в виде:

$$d(\mathcal{E}_\pi h_\pi)/dx = eh_\pi[n_\pi(x) - \sqrt{N_{\text{дп}}n_s}]/\varepsilon\varepsilon_0. \quad (\text{П4.24})$$

Текущая концентрация в подложке $n_\pi(x)$ представляет собой усредненное по толщине подложки значение концентрации электронов, изменение которой связано как с инжекцией носителей из канала, так и с эффектами накопления из-за разогрева и изменения эффективной массы. Входящая в уравнение (П4.24) толщина подложки, в первом приближении может быть рассчитана как длина экранирования в подложке, если известна концентрация примесей в ней. В качестве некоторой эффективной толщины для практических расчетов выбирается величина порядка 0,15-0,25 мкм, которую можно получить из формулы для длины экранирования в подложке, используя соотношение (П4.6) для среднего избыточного заряда.

Рассмотрим процедуру расчета высоты проводящей части канала и подложки. В приближении плавного канала толщина

обедненной области определяется разностью потенциалов затвора и текущего потенциала канала, поэтому:

$$h(x) = a - \sqrt{2\varepsilon\varepsilon_0[(V_b - V_{gs} + V(x)) / eN_d]}. \quad (\text{П4.25})$$

В режиме перекрытия канала необходимо рассчитывать толщину обедненного слоя, проникающего в подложку. Для этого рассмотрим одномерную задачу о типичном режиме, когда заряд в обедненной области определяется лишь ионизированными донорными атомами (рис. 5.5). Решение уравнения Пуассона в этом случае имеет вид для эпитаксиального слоя:

$$\begin{aligned} dV / dy &= -(eN_d / \varepsilon\varepsilon_0)y + C_1, \\ V(y) &= -(eN_d / 2\varepsilon\varepsilon_0)y^2 + C_1y + V_s, \end{aligned}$$

и для подложки

$$\begin{aligned} dV / dy &= -(eN_{дп} / \varepsilon\varepsilon_0)(y - d_{п}), \\ V(y) &= -(eN_{дп} / 2\varepsilon\varepsilon_0)(y - d_{п})^2. \end{aligned}$$

Здесь V_s – потенциал затвора относительно канала с учетом барьерного потенциала, $d_{п}$ – толщина обеднения. Сшивая полученные решения на границе эпитаксиальный слой–подложка ($y = a$), получим два уравнения относительно C_1 и $d_{п}$, откуда

$$d_{п}^2 = a^2 + \frac{N_d}{N_{дп}}(d_0^2 - a^2) = a^2 + 2\varepsilon\varepsilon_0(|V_s| - V_p) / eN_{дп}, \quad (\text{П4.26})$$

где $d_0 = \sqrt{2\varepsilon\varepsilon_0|V_s| / eN_d}$ – толщина слоя обеднения в однородном материале с концентрацией примеси N_d .

Недостатком полученного из одномерного анализа результата является быстрый рост $d_{п}$ и резкое перекрытие подложки при потенциале $|V_s|$ выше $|V_p|$. Это объясняется тем, что концентрация в подложке в приповерхностном слое принята равной величине $N_{дп}$. Однако, если учесть в (П4.26) обогащенный слой в подложке, распределение концентрации электронов в котором описывается формулой (П4.5), то перекрытие подложки будет плавным при малых значениях z и возрастающим по мере спада функции $n(z)$.

П4.4. Уравнения сохранения энергии и импульса для канала и подложки

Для анализа процессов, связанных с эффектами «всплеска» дрейфовой скорости и разогрева носителей, необходимо в систему исходных уравнений квазидвумерного моделирования включить уравнения баланса энергии и импульса:

$$dE / dt = e\mathcal{E}v - (E - E_0) / \tau_E, \quad (\text{П4.27})$$

$$dp / dt = e\mathcal{E} - p / \tau_p. \quad (\text{П4.28})$$

В квазидвумерном приближении средняя энергия электронов в общем случае зависит от продольной координаты и времени, поэтому

$$dE = \frac{\partial E}{\partial t} dt + \frac{\partial E}{\partial x} dx.$$

В стационарном случае ($\partial/\partial t = 0$) при $dx/dt = v$ уравнение (П4.27) приобретает вид:

$$dE / dx = e\mathcal{E} - (E - E_0) / v\tau_E. \quad (\text{П4.29})$$

Первое слагаемое в правой части (П4.29) характеризует энергию, приобретаемую в продольном поле канала, второе – потери энергии за счет рассеяния. Структура квазидвумерной модели, которая описывает коллективы электронов канала и подложки и их взаимодействие, приводит к включению в уравнения сохранения членов, учитывающих «перемешивание» потоков. Существенным с этой точки зрения является наличие потенциального барьера в подложке. Для анализа этого явления рассмотрим элемент канала длиной dx (рис. 4.8, в).

Через левое сечение элемента поток электронов за единицу времени приносит энергию, пропорциональную $E I_k$, а через правое сечение – $(E + dE)(I_k + dI_k)$. За счет перехода части электронов из канала в подложку через границу уходит энергия $E dI_k$. Для электронов, возвращающихся в канал из подложки, необходимо принять во внимание, что они ускоряются полем контактной разности и их энергия в канале будет равна сумме кинетической энергии в подложке и энергии, приобретенной за счет их «скатывания» с потенциального барьера ΔU . Поэтому эти электроны приносят энергию, пропорциональную $(E_{\text{п}} + \Delta U) dI_{\text{пк}}$. Составляя из этих слагаемых уравнение баланса (на рис. 5.8, в

обозначены «физические» направления токов, соответствующие направлению движения носителей)

$$(E + dE)(I_{\text{к}} + dI_{\text{пк}} - dI_{\text{кп}}) + EI_{\text{кп}} = EI_{\text{к}} + (E_{\text{п}} + \Delta U)dI_{\text{пк}},$$

получаем, что на участке dx канал за счет обмена носителями происходит изменение кинетической энергии на величину

$$dE = (E_{\text{п}} - E_{\text{к}} + \Delta U)dI_{\text{пк}} / I_{\text{к}}. \quad (\text{П4.30})$$

Проводя аналогичные рассуждения для участка подложки длиной dx и учитывая, что электроны, переходя из канала в подложку, уменьшают свою энергию, можно получить:

$$dE_{\text{п}} = -(E_{\text{п}} - E + \Delta U)dI_{\text{кп}} / I_{\text{п}}. \quad (\text{П4.31})$$

В итоге вместо (П4.29) получаем уравнения баланса энергии для канала

$$\frac{dE}{dx} = e\mathcal{E} - \frac{E - E_0}{v\tau_E} + \frac{(E_{\text{п}} - E + \Delta U)}{I_{\text{к}}} \frac{dI_{\text{пк}}}{dx} \quad (\text{П4.32})$$

и для подложки

$$\frac{dE_{\text{п}}}{dx} = e\mathcal{E}_{\text{п}} - \frac{E_{\text{п}} - E_0}{v_{\text{п}}\tau_E} + \frac{(E - E_{\text{п}} - \Delta U)}{I_{\text{п}}} \frac{dI_{\text{кп}}}{dx} \quad (\text{П4.33})$$

В этих выражениях ΔU представляет собой некоторую среднюю потенциальную энергию электронов в подложке, которая определяется как контактным потенциальным барьером эпитаксиальный слой-подложка, так и термоэдс, обусловленной электроградиентными эффектами. Суммарная высота барьера определяется с учетом разогрева в канале и подложке по соотношению (П4.15).

Полученные результаты характеризуют весьма важную особенность транзисторов с полуизолирующей подложкой. Ее удобно проследить, получив из (П4.30) и (П4.31) выражение для изменения средней энергии всех электронов

$$\begin{aligned} dE_{\Sigma} &= (I_{\text{к}}dE + I_{\text{п}}dE_{\text{п}}) / (I_{\text{п}} + I_{\text{к}}) = \\ &= (E_{\text{п}} - E_{\text{к}} + eV_{\text{к}})d(I_{\text{пк}} - I_{\text{кп}}) / (I_{\text{п}} + I_{\text{к}}). \end{aligned} \quad (\text{П4.34})$$

Из этого выражения видно, что при инжекции электронов из канала в подложку ($dI_{\text{кп}} > dI_{\text{пк}}$) происходит уменьшение энергии

носителей, что приводит к охлаждению электронного потока. И, несмотря на то, что при их возврате в канал потенциальная энергия вновь превращается в кинетическую и поток разогревается, все же часть носителей движется в области с малой концентрацией примеси, где их подвижность выше из-за уменьшения примесного рассеяния. Это позволяет получать большие значения дрейфовой скорости, а значит и крутизны, и в конечном итоге улучшает усилительные и шумовые свойства транзистора.

Уравнение сохранения импульса (П4.28) может быть для стационарного случая так же, как и уравнение сохранения энергии, приведено к виду

$$dp / dx = e\mathcal{E} / v - p / (v\tau_p). \quad (\text{П4.35})$$

В этом уравнении p – среднее значение продольной составляющей импульса электронов, которая увеличивается в тянущем электрическом поле и уменьшается за счет рассеяния. При переходе электронов между каналом и подложкой продольная составляющая импульса не изменяется, но в общем случае она имеет разные значения в канале и подложке. Поэтому необходимо в уравнении баланса учесть их изменение за счет «смешивания» потоков. Для этого можно провести для импульсов такие же рассуждения, что и для энергии (не учитывая изменений, происходящих в поле контактной разности потенциалов), и в результате получить

$$dp = (p_{\text{п}} - p)dI_{\text{пк}} / I_{\text{к}} \quad (\text{П4.36})$$

$$dp_{\text{п}} = -(p_{\text{п}} - p)dI_{\text{кп}} / I_{\text{п}}. \quad (\text{П4.37})$$

Соответственно вводя эти изменения в уравнение (П4.35), запишем уравнение баланса импульса для канала и подложки

$$dp / dx = e\mathcal{E} / v - p / v\tau_p + (p_{\text{п}} - p)(dI_{\text{пк}} / dx) / I_{\text{к}} \quad (\text{П4.38})$$

$$dp_{\text{п}} / dx = e\mathcal{E}_{\text{п}} / v_{\text{п}} - p_{\text{п}} / v_{\text{п}}\tau_{\text{п}} - (p_{\text{п}} - p)(dI_{\text{кп}} / dx) / I_{\text{п}}. \quad (\text{П4.39})$$

Обычно в подложке из-за меньшей роли примесного рассеяния скорость дрейфа выше, чем в канале. Поэтому описанное явление приводит к некоторому уменьшению импульса в подложке и увеличению его в канале.

Кроме того, различие в импульсе обусловлено не только различием скоростей электронов, но и эффективных масс

($p = m^* v$, $p_{\Pi} = m_{\Pi}^* v_{\Pi}$). Поэтому указанный эффект ослабляется из-за того, что эффективная масса электронов в канале несколько выше, чем в подложке, так как разогревные явления в канале обычно проявляются сильнее.

П4.5. Расчет малосигнальных параметров

Крутизна и выходная проводимость. Малосигнальные параметры могут быть рассчитаны по результатам квазидвухмерного моделирования по известным приращениям потенциалов, токов и зарядов в различных областях транзистора (рис. П4.2). Для анализа используются приращения, связанные с истоковым и стоковым краями затвора.

Формула для полного дифференциала тока стока может быть записана в виде

$$dI(V_{gs}, V_{ds}) = \frac{\partial I}{\partial V_{gs}} dV_{gs} + \frac{\partial I}{\partial V_{ds}} dV_{ds} = S_0 dV_{gs} + G_d dV_{ds}.$$

Переходя от дифференциалов к конечным приращениям $dV_{gs} \rightarrow \Delta V_{gs}$, $dV_{ds} \rightarrow \Delta V_{ds}$ необходимо, давая приращения ΔI и ΔV_g , следить за изменением напряжения V_{ds} с учетом $V_{gs} = V_g - V_s$ и $V_{gd} = V_g - V_d$.

Перепишем исходное соотношение в конечных разностях:

$$\Delta I = S_0 \Delta V_{gs} + G_d \Delta V_{ds},$$

или $\Delta V_{ds} = \Delta I / G_d - M \Delta V_{gs}$, где $M = S_0 / G_d$.

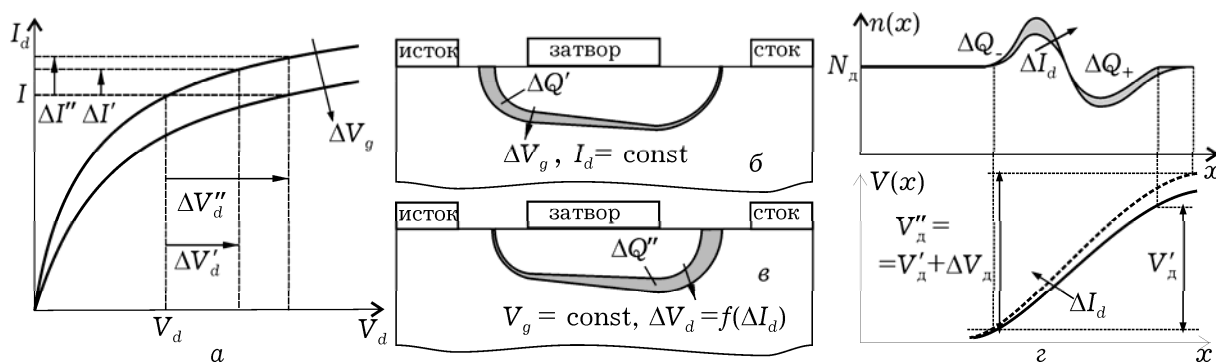


Рис. П4.2. К расчету приращений потенциалов и тока для определения крутизны и выходной проводимости (а), приращений заряда для определения емкостей C_{gs} (б) и C_{gd} (в) и параметров домена (г)

Изменение тока стока на величину ΔI при постоянном потенциале затвора вызовет изменение потенциалов в точках s и d так, что $V'_{gs} = V_g - V'_s$; $V'_{gd} = V_g - V'_d$. Из приведенного выше уравнения получим

$$V'_{ds} - V_{ds} = M(V'_s - V_s) + \Delta I/G_d.$$

Давая приращение ΔV_g при постоянном токе стока ($\Delta I = 0$), получим

$$V''_{ds} - V_{ds} = M(V''_s - V_s - \Delta V_g).$$

Для нахождения параметров S_0 и G_d необходимо совместно решить приведенные выше уравнения.

Динамическая крутизна S , являющаяся параметром схемной модели, должна учитывать запаздывание фазы изменения тока за счет конечного времени пролета электронов в канале:

$$\tau = \sum_{i=1}^n (\Delta x/v_i),$$

а также то обстоятельство, что непосредственно управляющее напряжение источника напряжения V_{\sim} является лишь частью входного напряжения на цепочке $R_i C_{gs}$. Поэтому

$$S = \frac{S_0 e^{-i\omega\tau}}{1 + i\omega R_i C_{gs}},$$

где R_i – сопротивление открытой части канала. Динамическая крутизна может быть также рассчитана по формуле

$$S = S_0 \frac{\sin(\omega\tau/2)}{\omega\tau/2} e^{-i(\omega\tau/2)}.$$

Входная и проходная емкости, параметры домена. Рассмотрим структуру транзистора для расчета потенциалов, как показано на рис. П4.2.

Так как весь заряд ионизированных доноров в обедненной области зависит от потенциалов затвора V_g и стока V_d , то

$$dQ = \frac{\partial Q}{\partial V_d} d(V_g - V_d) + \frac{\partial Q}{\partial V_g} dV_{gs} = C_{gd} d(V_g - V_d) + C_{gs} dV_{gs}.$$

Из уравнения для полного заряда можно получить два уравнения, соответствующих приращениям, соответственно, тока ΔI и входного напряжения:

$$\Delta Q' = C_{gd}\Delta V'_{gd} + C_{gs}\Delta V'_{gs},$$

$$\Delta Q'' = C_{gd}\Delta V''_{gd} + C_{gs}\Delta V''_{gs}.$$

Применяя подход, аналогичный описанному при определении S_0 и G_d для емкостей C_{gs} и C_{gd} , получим

$$\begin{cases} \Delta Q' = C_{gs}\Delta V'_s + C_{gd}\Delta V'_d \\ \Delta Q'' = C_{gs}(\Delta V''_s - \Delta V_g) - C_{gd}(\Delta V_d - \Delta V''_d) \end{cases}$$

Здесь введены следующие обозначения: для приращения тока стока ΔI изменения потенциалов имеют вид

$$\begin{cases} \Delta V'_{gd} = (V_g - V'_d) - (V_g - V_d) = V_d - V'_d \\ \Delta V'_{gs} = (V_g - V'_s) - (V_g - V_s) = V_s - V'_s \end{cases},$$

а для приращения напряжения ΔV_g , соответственно,

$$\begin{cases} \Delta V''_{gd} = (V_g - V''_d) - (V_g + \Delta V_g - V_d) = V_d - V''_d - \Delta V_g \\ \Delta V''_{gs} = (V_g - V''_s) - (V_g + \Delta V_g - V_s) = V_s - V''_s - \Delta V_g \end{cases}.$$

Сопротивление области сильного поля (домена) R_d и емкости C_d определяют через изменение заряда ΔQ_d и изменение падения напряжения на области домена ΔV_d , которые получаются при изменении тока стока на величину ΔI_d при постоянном напряжении затвора:

$$R_d = \Delta V_d / \Delta I_d; \quad C_d = \Delta Q_d / \Delta V_d.$$

Методику расчета соответствующих приращений заряда и потенциалов иллюстрируют рис. П4.2, з.

Приращение потенциала в области домена определяем по формуле

$$V''_d = V'_d + \Delta V_d.$$

Паразитные параметры. Как правило, металлизация затвора описывается комплексным сопротивлением $Z_g = R_g + j\omega L_g$, где $R_g = 3\rho_m W \lambda L / \delta_m$; ρ_m – удельное сопротивление металлизации затвора; δ_m – толщина слоя металлизации.

Для алюминиевого затвора с учетом скин-эффекта сопротивление металлизации

$$R_g = \left[\frac{17W_{\Pi}}{\delta_m L W_{об}} + \frac{1,3W_{\Pi}}{W_{об}} \left(\frac{f}{\delta_m L} \right)^{1/2} \right],$$

где f – частота, ГГц; L – длина затвора, мкм; W_{Π} – ширина одного пальца; $W_{об}$ – общая ширина затвора; δ_s – глубина скин-слоя, рассчитывается по формуле $\delta_s = \sqrt{2\rho_m / \omega \mu_{\Pi}}$; μ_{Π} – магнитная проницаемость на частоте ω . Для алюминиевого затвора глубина скин-слоя составляет 1,4...1,9 мкм на частоте 2 ГГц.

Сопротивление омических контактов стока и истока определим по формулам

$$R_s = \frac{L_{из}}{eN_d \mu_k (a - a_0) W}; \quad R_d = \frac{L_{зс}}{eN_d \mu_k (a - a_0) W},$$

где a_0 – высота слоя обеднения за счет поверхностных состояний на свободных границах, μ_k – подвижность носителей канала, $L_{из}$ – расстояние исток-затвор; $L_{зс}$ – расстояние затвор-сток.

Выходная емкость транзистора может быть определена как

$$C_{ds} = \varepsilon_0 (\varepsilon_r + 1) W K(k') / K(k),$$

где K – полный эллиптический интеграл первого рода, относительная диэлектрическая проницаемость арсенида галлия $\varepsilon_r = 12,9$. Дополнительная функция K' может быть рассчитана по формулам

$$K'(k) = K(k'); \quad k' = \sqrt{1 - k^2}; \quad k = \text{th}(\pi L / 2b),$$

где b – величина зазора конденсатора, образованного контактными площадками истока и стока; $L = L_c = L_{и}$ – длины контактных площадок истока и стока;

$$\frac{K(k)}{K'(k)} = \begin{cases} \left[\frac{1}{\pi} \ln \left(2 \frac{1 + \sqrt{k'}}{1 - \sqrt{k'}} \right) \right]^{-1} & \text{для } 0 \leq k \leq 0,7 \\ \frac{1}{\pi} \ln \left(2 \frac{1 + \sqrt{k}}{1 - \sqrt{k}} \right) & \text{для } 0,7 \leq k \leq 1 \end{cases}.$$

Полное сопротивление выводов, например, из золота, с учетом скин-эффекта может быть оценено по формуле

$$Z_{\text{ВЫВ}} = 3,1 \cdot 10^{-5} (1 + i) \sqrt{f} l / 2r,$$

где f – частота, ГГц; l – длина выводов; r – радиус поперечного сечения.

Приложение к главе 5

П5. АЛГОРИТМ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ ГЕТЕРОСТРУКТУРНЫХ ТРАНЗИСТОРОВ

Уравнения физико-топологических моделей субмикронных структур в двумерном приближении представляют собой уравнения в частных производных различных порядков.

Наибольшую проблему представляет решение дифференциальных уравнений в частных производных второго порядка из-за больших погрешностей их численных аппроксимаций. Уравнения в частных производных второго порядка в нестационарном случае записывается в виде уравнения гиперболического типа:

$$a \frac{\partial^2 u}{\partial t^2} - \nabla (b \nabla u) + c u = f$$

где a , b , c – константы, а f – константа или функция пространственных координат, u – вектор неизвестных.

Уравнения Пуассона и непрерывности в стационарном случае относятся к уравнениям в частных производных эллиптического типа, к которым сводятся гиперболические уравнения при $\partial^2 u / \partial t^2 = 0$.

Каноническая форма эллиптических уравнений представляется как:

$$-\nabla (b \nabla u) + c u = f.$$

Данная форма отвечает структуре уравнений непрерывности тока для электронов и дырок. При $c = 0$ получим

$$-\nabla(b\nabla u) = f,$$

что соответствует уравнению Пуассона для потенциала.

Решение приведенных уравнений в двумерной области со сложной функцией в правой части, задающей распределение плотности заряда, как правило, связано с применением численных методов с заданным шагом дискретизации и неоднородными граничными условиями.

Для анализа субмикронных эффектов в транзисторах с высокой подвижностью электронов используется система двумерных уравнений, полученная с учетом релаксационных процессов, которая включает уравнения сохранения энергии, импульса и частиц. Времена релаксации определяются с учетом преобладающих механизмов рассеяния – междолинного и полярного оптического.

При моделировании распределения потенциала в структуре и решении уравнения Пуассона учитывается разрыв энергетических зон на границе узкозонного и широкозонного полупроводников высотой примерно 0,4 эВ. Высота барьера зависит от свойств полупроводниковых материалов, образующих гетеропереход.

Процедура дискретизации и представление уравнений в конечно-разностном виде позволяет свести дифференциальные уравнения в частных производных к системе алгебраических уравнений:

$$\mathbf{K} \mathbf{V} = \mathbf{F},$$

где \mathbf{K} – матрица коэффициентов, \mathbf{F} – матрица известных величин или функций. При этом в зависимости от сложности топологической структуры моделируемой области субмикронной структуры могут использоваться различные методы генерации конечно-разностной сетки. Для моделируемой области со сложной конфигурацией предпочтительнее метод конечных элементов³, предполагающий триангуляцию. Для достаточно простых областей может использоваться метод конечных разностей с процедурой дискретизации с прямоугольной сеткой. При этом оба подхода, как правило, реализуются с возможностью применения адаптивного

³ *Зенкевич О., Морган К.* Конечные элементы и аппроксимация: Пер. с англ. – М.: Мир, 1986.– 318 с.

Марчук Г.И., Агошков В.И. Введение в проекционно-сеточные методы. – М.: Наука. 1981.–416 с.

изменения шага сетки. Особенно это актуально в областях с большими градиентами моделируемых функций и в приграничных областях. Можно показать⁴, в частности на примере уравнения Пуассона, что разностные уравнения обоих методов отличаются лишь способом учета плотности объемного заряда. В методе конечных разностей плотность заряда интегрируется непосредственно по всей области ячейки, а в методе конечных элементов требуется взвешенное интегрирование в большей области, включая соседние элементы. То есть при одинаковых аппроксимациях оба метода дают практически одинаковые результаты. Выбор метода разбиения обусловлен необходимостью аппроксимации области со сложной конфигурацией, а также тем, что сильные поля локализуются в области с размерами 0,2-0,3 мкм у стокового края затвора и требуется более мелкая сетка.

Метод конечных элементов основан на построении аппроксимации решения u вариационной задачи в конечномерных подпространствах S_D , содержащихся в области D . В основе вариационных методов лежит переход от задачи для уравнения с частными производными к задаче минимизации функционала. Функционалом для уравнения Пуассона в двумерном приближении вида

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$$

является интеграл⁵

$$I(u) = \iint_D \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 - 2uf \right] dx dy = \iint_D F(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}) dx dy.$$

Необходимым условием минимальности интеграла $I(u)$, является то, что функция $u(x, y)$ удовлетворяет уравнению Эйлера-Лагранжа:

$$\frac{\partial}{\partial x} F_{ux} + \frac{\partial}{\partial y} F_{uy} - F_u = 0.$$

Подпространства S_D строятся с помощью деления моделируемой области на конечные элементы – треугольники. С конечными элементами связывается совокупность узлов – вершин треугольника (могут также использоваться середины сторон,

⁴ Бубенников А.Н. Моделирование интегральных микротехнологий, приборов и схем. – М.: Высшая школа, 1989. – 320 с.

⁵ Бабенко К.И. Основы численного анализа. – М.: Наука. 1986. – 744 с.

центры тяжести элементов). На множестве конечных элементов строятся пространства кусочно-полиномиальных однотипных функций – линейных, квадратичных, кубических или сплайнов по каждой переменной в пределах одного элемента. Их коэффициенты выражаются через значения самих функций или их производных в узлах. Эти функции являются финитными, т.е. отличными от нуля только в окрестности, ограниченной несколькими соседними элементами.

Для линейной базисной функции для подобласти в виде треугольника с вершинами P_1, P_2, P_3 с заданными значениями u_1, u_2, u_3 функции $u(x,y)$ аппроксимация представляет собой полином:

$$\tilde{u}(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y, \quad (\text{П5.1})$$

коэффициенты которого могут быть выражены через u_i из решения системы уравнений $\tilde{u}(P_i) = u_i, i = 1, 2, 3$:

$$\alpha_1 = \frac{1}{S_{ijk}} [(x_j x_k - x_k y_i) u_i + (x_k x_i - x_i y_k) u_j + (x_i x_j - x_j y_i) u_k],$$

$$\alpha_2 = \frac{1}{S_{ijk}} [(y_i - y_k) u_i + (y_k - y_i) u_j + (y_i - y_j) u_k],$$

$$\alpha_3 = \frac{1}{S_{ijk}} [(x_k - x_j) u_i + (x_i - x_k) u_j + (x_j - x_i) u_k],$$

где i, j, k – номера вершин, а S_{ijk} – площадь треугольника. Полином (П5.1) может быть выражен через базисные функции как:

$$\tilde{u}(x, y) = \sum_{i=1}^3 u_i p_i(x, y),$$

где p_i – линейный по x, y многочлен, равный единице в i -ом узле и нулю в остальных:

$$p_i(x, y) = (1 / C_{ijk})(\alpha_{ik} + \beta_{ik} x + \gamma_{ik} y) = D_{ik} / C_{ijk}, \quad (\text{П5.2})$$

где $\alpha_{ik} = x_i x_k - y_j x_k$, $\beta_{ik} = x_k - y_j$, $\gamma_{ik} = y_i - y_k$,

$$D_{jk} = \begin{bmatrix} 1 & x & y \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{bmatrix}; \quad C_{ijk} = \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{bmatrix}.$$

Распространяя алгоритм определения $p_i(x, y)$ для каждого l -го узла триангуляции моделируемой области строится кусочно-билинейная функция $\Psi_l(x, y)$, равная единице в l -ом узле и нулю – в остальных. При этом в каждом из примыкающих треугольников эта функция определяется по формуле (П5.2). Совокупность таких функций образует базис пространства S_D кусочно-линейных функций для всех узлов триангуляции, за исключением лежащих на границе области.

Порядок интерполяции повышается, если кроме u_i в вершинах треугольников используется также значения в серединах сторон P_4, P_5, P_6 .

В этом случае

$$\tilde{u}(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 x^2 + \beta_4 xy + \beta_5 y^2 \quad ,$$

где все β_i определяются из уравнений $\tilde{u}_j(x_j, y_j) = u_j$, $j = 1, 2, \dots, 6$. При квадратичной функции $p_i^{(2)}(x, y)$ по каждой переменной, равной единице в p_i и нулю – в остальных, базисная функция может быть записана:

$$\tilde{u}(x, y) = \sum_{i=1}^6 u_i p_i^{(2)}(x, y),$$

где $p_i^{(2)}$ выражается через $p_i(x, y)$ как

$$p_i^{(2)}(x, y) = p_i(2p_i - 1), \quad j = 1, 2, 3,$$

$$p_4^{(2)}(x, y) = 4p_1 p_2 \quad , \quad p_5^{(2)}(x, y) = 4p_2 p_3 \quad , \quad p_6^{(2)}(x, y) = 4p_3 p_1 .$$

Таким образом, совокупность $p_i^{(2)}(x, y)$ в треугольниках с общими узлами образует базисную функцию пространства кусочно непрерывных функций.

Искомое решение представляется в виде:

$$u(x, y) = u_1 \phi_1(x, y) + u_2 \phi_2(x, y) + \dots + u_n \phi_n(x, y),$$

где ϕ_i - кусочно-линейные базисные функции. Триангуляция проводится так, что граница сеточной области совпадает с границей расчётной области. Кроме того, сеточные узлы лежат как на внешней границе, так и на линиях раздела различных сред (металл-полупроводник).

При триангуляции используется адаптивная процедура уменьшения шага сетки до достижения заданной точности. При этом сравниваются потенциалы в соседних точках в области с наибольшим значением градиента потенциала.

Рассмотрим процедуру получения вариационного уравнения для уравнения в частных производных эллиптического типа, адаптированную к уравнению Пуассона. Записанное в матричной форме с учетом описанного ранее алгоритма дискретизации это уравнение определяет алгоритм решения уравнений Пуассона и непрерывности.

В самом общем случае для неоднородных граничных условий 3-го рода

$$\vec{n}(c\nabla u) + qu = g \quad ,$$

где g, q – функции, определенные на области D , \vec{n} – нормаль к поверхности, вариационное уравнение для N базисных функций ($\varphi_i \in V_N$) запишется:

$$\int_D (c\nabla u) \cdot \nabla \varphi_i + au\varphi_i - f\varphi_i dx - \int_{\partial D} (-qu + g)\varphi_i ds = 0, \quad i = 1, \dots, N,$$

где N – размерность пространства V_N .

Тогда

$$u(x) = \sum_{j=1}^N U_j \varphi_j(x).$$

Получим систему уравнений

$$\sum_{j=1}^N \left[\int_D ((c\nabla \varphi_j) \cdot \nabla \varphi_i + a\varphi_j \varphi_i) dx + \int_{\partial D} q\varphi_j \varphi_i ds \right] \cdot U_j = \int_D f\varphi_i dx + \int_{\partial D} g\varphi_i ds,$$

где $I = 1, \dots, N$. Обозначая

$$K_{i,j} = \int_D (c\nabla \varphi_j) \cdot \nabla \varphi_i dx,$$

$$M_{i,j} = \int_D a\varphi_j \varphi_i dx,$$

$$Q_{i,j} = \int_{\partial D} q\varphi_j \varphi_i ds,$$

$$F_i = \int_D f\varphi_i dx,$$

$$G_i = \int_{\partial D} g\varphi_i ds,$$

где матрицы коэффициентов $K_{i,j}$ и $M_{i,j}$ представляют собой так называемые матрицы жесткости и матрицы масс, получим систему уравнений в матричной форме:

$$(\mathbf{K} + \mathbf{M} + \mathbf{Q}) \cdot \mathbf{U} = \mathbf{F} + \mathbf{G}. \quad (\text{П5.3})$$

При этом матрица коэффициентов в левой части системы при \mathbf{U} имеет диагональное преобладание.

Однако, чтобы получить приемлемую точность аппроксимации необходимо брать достаточно много слагаемых, что затрудняет процесс интегрирования.

Важно отметить, что точность приближения в методе конечных элементов зависит не от увеличения числа базисных функций (либо гармоник в приведенном выше базисе), а от размера окрестности, где базисная функция отлична от нуля.

Решение систем нелинейных аппроксимационных уравнений вида (П5.3), как правило, связано с различными модификациями метода Ньютона. В зависимости от системы исходных уравнений для описания субмикронных структур используются различные методы и алгоритмы.

После получения решения с заданной точностью для потенциалов (уравнение Пуассона) и концентрации (уравнение непрерывности) совместно решается система уравнений с расчетом на каждом шаге коэффициента диффузии и подвижности электронов. Алгоритм решения системы уравнений Пуассона и уравнения непрерывности с учетом аппроксимаций и уравнения для напряженности электрического поля $\mathcal{E} = -\text{grad}(\varphi)$ представляет собой несколько вложенных циклов. Каждое из этих уравнений, в свою очередь, является системой конечно-элементных уравнений большой размерности. Разреженные матрицы систем обладают диагональным преобладанием и имеют ленточную структуру. Для решения систем используется метод прогонки в сочетании с итерационным многошаговым методом Ньютона.

Алгоритм решения системы уравнений локальной физико-топологической модели может быть представлен следующими этапами.

Шаг 1. Триангуляция моделируемой области с заданным шагом и расчет профилей легирующей примеси.

Шаг 2. Решение уравнения Пуассона для равновесного распределения концентрации носителей (первая итерация).

Шаг 3. Расчет распределения напряженности электрического поля в структуре $\mathcal{E}(x, y)$ и подвижностей носителей при заданных механизмах рассеяния, а также при заданных значениях концентрации и профилей легирующей примеси.

Шаг 4. Решение уравнений непрерывности тока с заданной точностью и определение распределений концентраций подвижных носителей $n(x,y)$.

Шаг 5. Совместное решение системы уравнений, состоящей из уравнений Пуассона и непрерывности.

Шаг 6. Расчет токов на электродах и проверка выполнения закона Кирхгофа для токов, оценка погрешности вычислений. Если заданной точности не достигнуто, осуществляется дробление шага дискретизации триангулярной сетки и переход к шагу 1.

Таким образом, находятся распределения концентрации носителей и потенциала для заданной структуры. Выход из внешнего цикла осуществляется при условии достижения функцией $\varphi(x,y)$ минимума.

Для нахождения распределения концентрации носителей используется метод Ньютона. В простейшем случае этот метод описывается формулой:

$$n_{i+1} = n_i - F_i'^{(-1)} F_i .$$

Но из-за наличия больших градиентов функций наиболее подходящим является метод Ньютона-Рафсона с дроблением шага, так как в нем заложена адаптация шага для каждого узла сетки:

$$n_{i+1} = n_i - \alpha_i F_i'^{(-1)} F_i ,$$

где целевая функция $F_i = \rho / \varepsilon \varepsilon_0 \varphi$. Величина α_i выбирается так, чтобы обеспечить убывание целевой функции на каждой итерации. Этот выбор связан с проверкой неравенства:

$$f(x_i + \alpha_i p_i) - f(x_i) \leq \delta \cdot \alpha_i (f'(x_i), p_i) ,$$

где $p_i = -(f''(x_i))^{-1} \cdot f'(x_i)$ – направление спуска, а $0 < \delta < S$ – некоторое заданное число, общее для всех итераций. Если это неравенство выполнено при $\alpha_i = 1$, то шаг принимается равным единице и осуществляется следующая итерация. Если нет – дробится шаг до тех пор, пока оно не выполнится.

Если итерационный процесс плохо сходится, то используется ускоренный метод Ньютона третьего порядка⁶ :

⁶ Иванов В.В. Методы вычислений на ЭВМ // Справочное пособие. – Киев: Наук. Думка. – 1986. – 584 с.

$$x_{i+1} = x_i - F_i'^{-1}F_i - \frac{1}{2!}F_i'^{-1}F_i''(F_i'^{-1}F_i)(F_i'^{-1}F_i), \quad i = 0, 1, 2, \dots$$

Метод обладает быстрой сходимостью, однако, добиться с его помощью заданной точности довольно трудно, поэтому он используется для нахождения начального приближения для метода Ньютона-Рафсона с дроблением шага.

Для решения исходной системы уравнений нелокальной двумерной модели ГСПТ применяется метод итераций (последовательных приближений). Так как в данные уравнения входит некая функция от координат $z(x,y)$ и ее первая производная, то можно записать: $z = f(z)$. Тогда z_n будет выражаться через z_{n-1} по следующей рекуррентной формуле⁷: $z_n = f(z_{n-1})$.

Выходной ток транзистора рассчитывается как нормальная составляющая тока к контакту стока по известным двумерным распределениям концентрации электронов, напряженности электрического поля и подвижности по результатам решения исходной системы уравнений при достижении заданной точности решений в узлах сетки.

⁷ П5.6. В.А. Ильин Итерационные методы решения функциональных уравнений, Соросовский Образовательный Журнал – М.: 2001 – No 2 – с. 116–120.

П7.1. Вывод формулы Цу-Эсаки

1. Вычисление тока, создаваемого электронами, вышедшими из левого резервуара $J_{LR}(z)$:

$$\begin{aligned}
\mathbf{J}_{LR}(z) &= \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \mathbf{j}_{LR}(E_z(k_z), z) g(\mathbf{k}) f_{\Phi-D}(\mathbf{k}, E_\Phi) dk_z dk_x dk_y = \\
&= \frac{2}{(2\pi)^3} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{\mathbf{j}_{LR}(E_z(k_z), z) dk_z dk_x dk_y}{\exp\left(\left(\hbar^2(k_x^2 + k_y^2 + k_z^2) / 2m^* + U_{i0} - (E_\Phi + U_1)\right) / k_B T\right) + 1} = \\
&\left| \begin{aligned} dk_z dk_x dk_y &= k_r dk_z dk_r dk_\phi \\ k_r &= \sqrt{k_x^2 + k_y^2} \end{aligned} \right| = \\
&\frac{2}{(2\pi)^3} \int_0^\infty \mathbf{j}_{LR}(E_z(k_z), z) dk_z \int_0^\infty \frac{k_r dk_r}{\exp\left(\left(\hbar^2(k_r^2 + k_z^2) / 2m^* + U_{i0} - (E_\Phi + U_1)\right) / k_B T\right) + 1} \int_0^{2\pi} dk_\phi = \\
&= \frac{1}{(2\pi)^2} \int_0^\infty \mathbf{j}_{LR}(E_z(k_z), z) dk_z \int_0^\infty \frac{dk_r^2}{\exp\left(\left(\hbar^2(k_r^2 + k_z^2) / 2m^* + U_{i0} - (E_\Phi + U_1)\right) / k_B T\right) + 1} = \\
&= \left| \begin{aligned} u &= \left(\hbar^2(k_r^2 + k_z^2) / 2m^* + U_{i0} - (E_\Phi + U_1)\right) / k_B T, \\ du &= \frac{\hbar^2}{2m^* k_B T} dk_r^2, \quad E_z = \frac{\hbar^2 k_z^2}{2m^*} + U_{i0} \end{aligned} \right| = \\
&\frac{2m^* k_B T}{(2\pi\hbar)^2} \int_0^\infty \mathbf{j}_{LR}(E_z(k_z), z) dk_z \int_{\frac{E_z - (E_\Phi + U_1)}{k_B T}}^\infty \frac{du}{e^u + 1} = \left| \int \frac{du}{e^u + 1} = \ln(e^{-u} + 1) \right| = \\
&= \frac{2m^* k_B T}{(2\pi\hbar)^2} \int_0^\infty \mathbf{j}_{LR}(E_z(k_z), z) \ln\left(\exp\left(-\frac{E_z - (E_\Phi + U_1)}{k_B T}\right) + 1\right) dk_z = \\
&\left| dk_z = \frac{\sqrt{m^*}}{\sqrt{2\hbar}} \frac{dE_z}{\sqrt{E_z - U_{i0}}}, \right| = \\
&= \frac{\sqrt{2}(m^*)^{3/2} k_B T}{(2\pi)^2 \hbar^3} \int_{U_{i0}}^\infty \frac{\mathbf{j}_{LR}(E_z(k_z), z)}{\sqrt{E_z - U_{i0}}} \ln\left(1 + \exp\left(-\frac{E_z - (E_\Phi + U_1)}{k_B T}\right)\right) dE_z
\end{aligned}$$

2. Вычисление тока, создаваемого электронами, вышедшими из левого резервуара $J_{RL}(z)$:

$$\begin{aligned}
\mathbf{J}_{RL}(z) &= \int_{-\infty}^0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{j}_{RL}(E_z(k_z), z) g(\mathbf{k}) f(\mathbf{k}, E_{\Phi}) dk_z dk_x dk_y = \\
&= \frac{2}{(2\pi)^3} \int_{-\infty}^0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\mathbf{j}_{RL}(E_z(k_z), z) dk_z dk_x dk_y}{\exp\left(\left(\hbar^2(k_x^2 + k_y^2 + k_z^2) / 2m^* + U_{i5} - (E_{\Phi} + U_N)\right) / k_B T\right) + 1} = \\
&\left| \begin{aligned} dk_z dk_x dk_y &= k_r dk_z dk_r dk_{\varphi} \\ k_r &= \sqrt{k_x^2 + k_y^2} \end{aligned} \right| = \\
&\frac{2}{(2\pi)^3} \int_{-\infty}^0 \mathbf{j}_{RL}(E_z(k_z), z) dk_z \int_0^{\infty} \frac{k_r dk_r}{\exp\left(\left(\hbar^2(k_r^2 + k_z^2) / 2m^* + U_{i5} - (E_{\Phi} + U_N)\right) / k_B T\right) + 1} \int_0^{2\pi} dk_{\varphi} = \\
&= \frac{1}{(2\pi)^2} \int_{-\infty}^0 \mathbf{j}_{RL}(E_z(k_z), z) dk_z \int_0^{\infty} \frac{dk_r^2}{\exp\left(\left(\hbar^2(k_r^2 + k_z^2) / 2m^* + U_{i5} - (E_{\Phi} + U_N)\right) / k_B T\right) + 1} = \\
&\left| \begin{aligned} u &= \left(\hbar^2(k_r^2 + k_z^2) / 2m^* + U_{i5} - (E_{\Phi} + U_N)\right) / k_B T, \\ du &= \frac{\hbar^2}{2m^* k_B T} dk_r^2, \quad E_z = \frac{\hbar^2 k_z^2}{2m^*} + U_{i5} \end{aligned} \right| = \\
&\frac{2m^* k_B T}{(2\pi\hbar)^2} \int_{-\infty}^0 \mathbf{j}_{RL}(E_z(k_z), z) dk_z \int_{\frac{E_z - (E_{\Phi} + U_N)}{k_B T}}^{\infty} \frac{du}{e^u + 1} = \left| \int \frac{du}{e^u + 1} = \ln(e^{-u} + 1) \right| = \\
&= \frac{2m^* k_B T}{(2\pi\hbar)^2} \int_{-\infty}^0 \mathbf{j}_{RL}(E_z(k_z), z) \ln\left(\exp\left(-\frac{E_z - (E_{\Phi} + U_N)}{k_B T}\right) + 1\right) dk_z = \\
&\left| dk_z = -\frac{\sqrt{m^*}}{\sqrt{2\hbar}} \frac{dE_z}{\sqrt{E_z - U_{i5}}} \right| = \\
&= -\frac{\sqrt{2}(m^*)^{3/2} k_B T}{(2\pi)^2 \hbar^3} \int_{\infty}^{U_{i5}} \frac{\mathbf{j}_{RL}(E_z(k_z), z)}{\sqrt{E_z - U_{i5}}} \ln\left(1 + \exp\left(-\frac{E_z - (E_{\Phi} + U_N)}{k_B T}\right)\right) dE_z = \\
&= \frac{\sqrt{2}(m^*)^{3/2} k_B T}{(2\pi)^2 \hbar^3} \int_{U_{i5}}^{\infty} \frac{\mathbf{j}_{RL}(E_z(k_z), z)}{\sqrt{E_z - U_{i5}}} \ln\left(1 + \exp\left(-\frac{E_z - (E_{\Phi} + U_N)}{k_B T}\right)\right) dE_z
\end{aligned}$$

Все обозначения приведены на рис. 7.14

П7.2. Конечно-разностная схема для системы уравнений огибающей волновой функции

Если в переносе заряда участвуют электроны из двух долин (для определенности " Γ " и " X "), для того, чтобы учесть смешивание электронных состояний на гетерограницах необходимо решать следующую систему уравнений относительно огибающих волновых функций ψ_Γ и ψ_X ⁸:

$$\left\{ \begin{array}{l} -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{mm_\Gamma} \frac{d\psi_\Gamma}{dz} + \left(V_\Gamma(z) - E + \underbrace{\frac{\hbar^2 k_\parallel^2}{2mm_\Gamma}}_{-E_z^\Gamma} \right) \psi_\Gamma + \alpha \delta(z_k) \psi_X = 0 \\ -\frac{\hbar^2}{2} \frac{d}{dz} \frac{1}{mm_X} \frac{d\psi_X}{dz} + \left(V_X(z) - E + \underbrace{\frac{\hbar^2 k_\parallel^2}{2mm_X}}_{-E_z^X} \right) \psi_X + \alpha \delta(z_k) \psi_\Gamma = 0 \end{array} \right. \quad (\text{П7.1})$$

Здесь $m_{\Gamma(X)}$ – относительная эффективная масса в $\Gamma(X)$ – долине, m – масса покоя электрона, $k = i_1, i_2, \dots, i_{N_b}$ – номера точек, которые принадлежат гетерограницам (N_b – количество таких границ); α – постоянная междолинного взаимодействия [eV·Å]; $\delta(z)$ – дельта-функция Дирака.

Для большей наглядности выкладок будем считать, что $k_\parallel^2 = 0$, так что $E_z^\Gamma = E_z^X \equiv E_z$ (как обозначено в (П7.1)), в таком предположении $\psi_{\Gamma(X)}$ является функцией лишь координаты z и поперечной энергии E_z .

На гетерограницах эти уравнения должны удовлетворять обобщенным граничным условиям, которые налагаются на огибающие волновых функций ψ_Γ (ψ_X)¹:

$$\psi_\Gamma|_{z_k^-} = \psi_\Gamma|_{z_k^+}, \quad (\text{П7.2})$$

$$\frac{\hbar^2}{2} \left[\left(\frac{1}{mm_\Gamma} \frac{d\psi_\Gamma}{dz} \right)_{z=z_k^+} - \left(\frac{1}{mm_\Gamma} \frac{d\psi_\Gamma}{dz} \right)_{z=z_k^-} \right] + \alpha \psi_X|_{z=z_k} = 0. \quad (\text{П7.3})$$

⁸ Liu H.C. Resonant tunneling through single layer heterostructure // Appl. Phys. Letters – 1987. – Vol. 51, No. 13. – P. 1019–1021.

Перепишем первое из уравнений системы (П7.1) в виде:

$$\frac{d}{dz} \frac{1}{m_\Gamma} \frac{d\psi_\Gamma}{dz} + \frac{2m(E_z - V_\Gamma(z))}{\hbar^2} \psi_\Gamma - \frac{2m}{\hbar^2} \alpha \delta(z_k) \psi_X = 0, \quad (\text{П7.4})$$

и будем искать для него конечно-разностную схему на нерегулярной сетке $\{z_i, i = 1, 2, \dots, N\}$, где N – количество узлов дискретной сетки.

Введем обозначения:

$$\begin{aligned} k(z) &\equiv \frac{1}{m_\Gamma(z)}; \\ q(z) &\equiv \frac{2m}{\hbar^2} (V_\Gamma - E_z); \\ f(x) &\equiv -\frac{2m}{\hbar^2} \alpha \delta(z_k) \psi_X(z). \end{aligned} \quad (\text{П7.5})$$

В них (П7.4) переписывается так:

$$\frac{d}{dz} k(z) \frac{d\psi_\Gamma(z)}{dz} - q(z) \psi_\Gamma(z) = -f(z) \quad (\text{П7.6})$$

Любое трехточечное уравнение, которое соответствует (П7.6) может быть записано в виде⁹:

$$\frac{1}{\bar{h}_i} \left(b_i(z) \frac{\psi_{i+1} - \psi_i}{h_{i+1}} - a_i(z) \frac{\psi_i - \psi_{i-1}}{h_i} \right) - d_i(z) \psi_i(z) + \varphi_i(z) = 0, \quad (\text{П7.7})$$

где $a_i(z)$, $b_i(z)$, $d_i(z)$, $\varphi_i(z)$ – искомые разностные функции, которые могут в общем случае зависеть от трех значений координаты: z_{i-1} , z_i та z_{i+1} ; $h_i = z_i - z_{i-1}$; $\bar{h}_i = (h_i + h_{i-1}) / 2$.

Для того чтобы конечно-разностная схема (П7.7) удовлетворяла условию консервативности и самосопряженности конечно-разностного оператора, необходимо, чтобы $b_i(z) = a_{i+1}(z)$. Тогда она будет описывать уравнение сохранения некоторой физической величины ω в диапазоне $z \in \Omega$, где $\Omega: [z_{i-1/2}; z_{i+1/2}]$:

⁹ Самарский А.А. Введение в численные методы. – М.: «Наука», 1987. – с. 150.

$$\omega_{i+1/2} - \omega_{i-1/2} - \int_{z_{i-1/2}}^{z_{i+1/2}} q(z)\psi(z)dz = - \int_{z_{i-1/2}}^{z_{i+1/2}} f(z)dz, \quad (\text{П7.8})$$

где $\omega = k(z)u'(z)$; уравнение (П7.8) эквивалентно закону сохранения потока вероятности. Легко видеть, что, стягивая область Ω в точку, из (П7.8) получим (П7.3).

Всю совокупность конечно-разностных схем второго порядка точности, соответствующих (П7.6), можно записать в виде¹⁰:

$$\frac{1}{h_i} \left(a_{i+1}(z) \frac{\Psi_{i+1} - \Psi_i}{h_{i+1}} - a_i(z) \frac{\Psi_i - \Psi_{i-1}}{h_i} \right) - d_i(z)\psi_i(z) + \varphi_i(z) = 0, \quad (\text{П7.9})$$

где $a_i = \left(\frac{1}{h_i} \int_{z_{i-1}}^{z_i} \frac{dz}{k(z)} \right)^{-1}$, $\varphi_i = \frac{1}{h_i} \int_{z_{i-1/2}}^{z_{i+1/2}} f(z)dz$, $d_i = \frac{1}{h_i} \int_{x_{i-1/2}}^{x_{i+1/2}} q(z)dz$.

Далее, применяя интегро-интерполяционный метод¹¹, можно заменить соответствующие интегралы их конечно-разностными аппроксимациями. В данном случае важно, чтобы все точки z_k попадали в узлы сетки, поскольку функция $f(z)$ содержит в себе δ -функцию, которая превращается в нуль везде кроме точек z_k . Рассмотрим аппроксимацию по трехточечной схеме в случае, если центральная точка схемы в точности совпадает с гетерограницей: $z_i = z_k$.

Учитывая, что эффективная масса в этой точке обычно изменяется скачкообразно (функция $k(z)$ испытывает разрыв I рода), а в интервалах $z \in [z_{i-1}; z_i]$ и $z \in [z_i; z_{i+1}]$ достаточно гладкая (обычно - просто стала), лучшим приближением для интеграла a_i служит:

$$a_i = k_{i-1/2}. \quad (\text{П7.10})$$

В случае если эффективная масса в левом и правом интервале не изменяется, можно точно положить:

$$a_i = k_{i-1}. \quad (\text{П7.11})$$

¹⁰ Самарский А.А. Введение в численные методы. - М.: «Наука», 1987. - с. 53.

¹¹ Самарский А.А. Введение в численные методы. - М.: «Наука», 1987. - с. 167.

Аналогично, можно положить:

$$a_{i+1} = k_{i+1}. \quad (\text{П7.12})$$

Интеграл φ_i , очевидно, будет равен:

$$\varphi_i = \varphi_k = \frac{1}{\bar{h}_i} \int_{z_{i-1/2}}^{z_{i+1/2}} \varphi(z) dz = -\frac{1}{\bar{h}_i} \frac{2m}{\hbar^2} \alpha \psi_{\text{X}}(z); \quad (\text{П7.13})$$

$$d_i = \int_{z_{i-1/2}}^{z_{i+1/2}} q(z) dz \approx q_i \equiv (q_{i+1} + q_{i-1}) / 2. \quad (\text{П7.14})$$

Все вместе формулы (П7.9)–(П7.14) удовлетворяюь второму порядку точности аппроксимации дифференциального уравнения конечно-разностной схемой. Подставляя аппроксимации интегралов (П7.10)–(П7.14) в (П7.9), получим:

$$\frac{1}{\bar{h}_i} \left(\frac{1}{m_{i+1}} \frac{(\Psi_{\Gamma,i+1} - \Psi_{\Gamma,i})}{h_{i+1}} - \frac{1}{m_{i-1}} \frac{(\Psi_{\Gamma,i} - \Psi_{\Gamma,i-1})}{h_i} \right) - q_i \Psi_{\Gamma,i} + \varphi_i = 0. \quad (\text{П7.15})$$

Проведя эквивалентные преобразования, (П7.15) можно переписать в виде:

$$\Psi_{\Gamma,i+1} + \beta \Psi_{\Gamma,i-1} - (1 + \beta + h_{i+1} \bar{h}_i m_{i+1} q_i) \Psi_{\Gamma,i} + h_{i+1} \bar{h}_i m_{i+1} \varphi_i = 0, \quad (\text{П7.16})$$

где $\beta = \frac{h_{i+1} m_{i+1}}{h_i m_{i-1}}$.

Возвращаясь полностью к изначальным обозначениям в соответствии с (П7.5), получим конечно-разностную схему для всех точек $i = k$:

$$a_k \Psi_{\Gamma,k-1} + b_k \Psi_{\Gamma,k} + c_k \Psi_{\Gamma,k+1} = d_k \Psi_{\text{X},k}, \quad (\text{П7.17})$$

где:

$$a_{\Gamma,k} = \frac{h_{k+1} m_{k+1}}{h_k m_{k-1}};$$

$$b_{\Gamma,k} = -1 - \frac{h_{k+1}}{h_k} \frac{m_{k+1}}{m_{k-1}} + \frac{h_{k+1} \bar{h}_k m_{k+1} 2m(E - V_{\Gamma,k})}{\hbar^2};$$

$$c_{\Gamma,k} = 1;$$

$$d_{\Gamma,k} = -\frac{2n_{k+1}m_{k+1}m\alpha}{\hbar^2}.$$

Еще раз подчеркнем, что схема (П7.17) с коэффициентами, приведенными под ней, должна использоваться лишь для точек с индексами $i \in \{i_1, i_2, \dots, i_{Nb}\}$, то есть для *гетерограниц*. Мы индексируем эти точки через k , чтобы не путать с "обычными" внутренними точками.

Для всех других точек (то есть для $i \notin \{i_1, i_2, \dots, i_{Nb}\}$) в случае, если эффективная масса в пределах слоев остается неизменной, необходимо положить в уравнении (П7.17) $m_{i+1} = m_{i-1} = m_i$, что приведет к выражению:

$$a_i \psi_{\Gamma, i-1} + b_i \psi_{\Gamma, i} + c_i \psi_{\Gamma, i+1} = 0, \quad (\text{П7.18})$$

где:

$$a_i = \frac{h_{i+1}}{h_i};$$

$$b_i = -1 - \frac{h_{i+1}}{h_i} + \frac{h_{i+1} h_i m_i 2m(E - V_{\Gamma, i})}{\hbar^2};$$

$$c_i = 1.$$

Подчеркнем, что эта схема пригодна для всех точек, кроме $i \in \{i_1, i_2, \dots, i_{Nb}\}$.

* * *

Формулы (П7.17) и (П7.18) позволяют сопоставить первому из уравнений (П7.1) его конечно-разностную схему во всех внутренних точках. При ее получении были выполнены все условия, чтобы она имела второй порядок точности, в чем можно непосредственно убедиться.

Для того чтобы получить конечно-разностную схему для второго уравнения системы (П7.1), необходимо в формулах (П7.17) и (П7.18) везде поменять местами индексы "X" и "Г".

Для того чтобы получить конечно-разностную схему с регулярной сеткой, которая имеет шаг h , необходимо в уравнениях (П7.17) и (П7.18) положить $h_{i-1} = h_i = h_{i+1} \equiv h$.

Наконец, если в (П7.17) положить $\alpha = 0$, то это уравнение будет описывать конечно-разностную схему для $\Gamma(X)$ -долины однодолинной модели (формула (П7.17) при этом не изменится).

П7.3. Уравнения, замыкающие конечно-разностную схему

Уравнения, замыкающие конечно-разностную схему, которая сопоставляется уравнению Шредингера для огибающей волновой функции, может быть полученная из условий непрерывности волновой функции и ее первой производной в точках, являющимися левой и правой границей "квантовой" области устройства, z_0 и z_5 соответственно¹²:

$$\begin{cases} \psi^I(z_0) = \psi^{II}(z_0), \\ \frac{1}{m^I(z_{0-})} \frac{d\psi^I(z_0)}{dz} = \frac{1}{m^{II}(z_{0+})} \frac{d\psi^{II}(z_0)}{dz}, \end{cases} \quad (\text{П7.19})$$

$$\begin{cases} \psi^{II}(z_5) = \psi^{III}(z_5), \\ \frac{1}{m^{II}(z_{5-})} \frac{d\psi^{II}(z_5)}{dz} = \frac{1}{m^{III}(z_{5+})} \frac{d\psi^{III}(z_5)}{dz}, \end{cases} \quad (\text{П7.20})$$

где $\psi^{I,II,III}$, $m^{I,II,III}$ – волновые функции и эффективные массы в левом резервуаре, квантовой области и правом резервуаре. В дальнейшем мы не будем указывать, в какой точке берется эффективная масса, поскольку это и так понятно.

В резервуарах вид волновой функции известен: это суперпозиция волн де Бройля; в квантовой области волновая функция обозначается через ψ^{II} , именно она ищется с помощью конечно-разностной схемы, уравнения для крайних точек которой здесь выводятся.

Для электронов из *левого резервуара*:

$$\psi_L(z) = \begin{cases} \psi^I(z) = \exp\{ik_L(z - z_0)\} + r_L \exp\{-ik_L(z - z_0)\}, z < z_0, \\ \psi^{II}(z), z \in [z_0; z_5], \\ \psi^{III}(z) = t_L \exp\{ik_R(z - z_5)\}, z > z_5. \end{cases} \quad (\text{П7.21})$$

где $k_{L(R)} = \sqrt{2m_{L(R)}^*(E_z - U_{i0(i5)})} / \hbar$.

При этом $\psi^{II}(z)$ определяется в дискретных точках $\psi_i = \psi_{II}(z_i) = \psi_{II}\{\Delta(i - 1)\}$ численным методом.

¹² Lent C.S. The quantum transmitting boundary method /C.S. Lent, D.J. Kirkner // Journal of Applied Physics. - 1990. - Vol. 67. - P. 6353-6359

Найдем уравнение для *левой границы* для электронов из *левого резервуара*.

Учитывая (П7.21), (П7.19) переписывается:

$$\begin{cases} 1 + r_L = \psi_1, \\ ik_L(1 - r_L)\alpha = (\psi_2 - \psi_1) / \Delta, \end{cases} \quad (\text{П7.22})$$

где $\alpha = m_{II}/m_I$. Во втором уравнении (П7.22) была использована аппроксимация производной:

$$\frac{d\psi^{II}(z_0)}{dz} \approx \frac{\psi_2 - \psi_1}{\Delta}.$$

Исключая из (П7.22) r_L , находим уравнение, которое замыкает конечно-разностную схему для ψ_L слева:

$$(i\alpha k_L \Delta - 1)\psi_1 + \psi_2 = 2i\alpha k_L \Delta, \quad (\text{П7.23})$$

а также выражение для "комплексного коэффициента отражения" r_L : $r_L = \psi_1 - 1$.

Коэффициент отражения (вероятность отражения) электрона от левой границы активной области равняется¹³:

$$R_{LR} = \frac{|j_{\leftarrow}^I|}{|j_{\rightarrow}^I|},$$

где $j_{\rightarrow(\leftarrow)}^I$ — плотность потока вероятности в падающей (отраженной) волнах:

$$\psi_{\rightarrow}^I(z) = \exp\{ik_L(z - z_0)\} \quad (\psi_{\leftarrow}^I(z) = -r_L \exp\{-ik_L(z - z_0)\})$$

в первой области, и определяется¹⁴:

$$j = \frac{i\hbar}{2m^*} (\psi \nabla \psi^* - \psi^* \nabla \psi). \quad (\text{П7.24})$$

¹³ Ландау Л.Д. Квантовая механика (нерелятивистская теория) / Л.Д. Ландау, Е.М. Лифшиц // - М.: Физматгиз, 1963. - с. 101.

¹⁴ Блохинцев Д.И. Основы квантовой механики: учебное пособие / Д.И. Блохинцев // - М.: Наука, 1976. - с. 122.

Для нахождения $j_{\rightarrow(\leftarrow)}^I$ под ψ в (П7.24) следует понимать $\psi_{\rightarrow(\leftarrow)}^I$. Проведя вычисление, получим:

$$R_{LR} = |r_L|^2 = |\psi_1 - 1|^2. \quad (\text{П7.25})$$

Найдем уравнение для правой границы для электронов из левого резервуара. Учитывая (П7.21), (П7.20) переписывается:

$$\begin{cases} \psi_N = t_L, \\ \frac{m_{III}}{m_{II}} \frac{(\psi_N - \psi_{N-1})}{\Delta} = t_L ik_R, \end{cases}$$

где была использована аппроксимация производной

$$\frac{d\psi^{II}(z_5)}{dz} \approx \frac{\psi_N - \psi_{N-1}}{\Delta}.$$

Из этой системы получим уравнение, которое замыкает конечно-разностную схему для ψ_L справа:

$$(ik_R\Delta - \alpha)\psi_N = \alpha\psi_{N-1} = 0, \quad (\text{П7.26})$$

де $\alpha = m_{III}/m_{II}$.

Коэффициент прохождения (вероятность прохождения) электрона через активную область равняется:

$$T_{LR} = \frac{|j_{\rightarrow}^{III}|}{|j_{\rightarrow}^I|}, \quad (\text{П7.27})$$

где $j_{\rightarrow}^{I(III)}$ — плотность потока вероятности, соответствующая волне, которая падает на активную область $\psi_{\rightarrow}^I(z) = \exp\{ik_L(z - z_0)\}$ (прошла через активную область $\psi_{\rightarrow}^{III} = \psi^{III} = t_L \exp\{ik_R(z - z_5)\}$) и определяется в соответствии с (П7.24). Вычисления дают:

$$T_{LR} = \frac{k_R}{k_L} \frac{m_L}{m_R} |t_L|^2 = \frac{k_R}{k_L} \frac{m_L}{m_R} |\psi_N|^2 \quad (\text{П7.28})$$

Для электронов из *правого* резервуара:

$$\Psi_R(z) = \begin{cases} \Psi^I(z) = t_R \exp\{-ik_L(z - z_0)\}, & z < z_0, \\ \Psi^II(z), & z \in [z_0; z_5], \\ \Psi^III(z) = r_R \exp\{ik_R(z - z_5)\} + \exp\{-ik_R(z - z_5)\}, & z > z_5 \end{cases} \quad (\text{П7.29})$$

Уравнения, замыкающие конечно-разностную схему для Ψ_R , могут быть получены из (П7.19), (П7.20) и (П7.29).

Найдем уравнение, которое замыкает конечно-разностную схему для левой границы для электронов, которые инжектируются из правого резервуара. Учитывая (П7.29), (П7.19) переписывается:

$$\begin{cases} t_R = \Psi_1, \\ \frac{m_I}{m_{II}} \frac{\Psi_2 - \Psi_1}{\Delta} = -ik_L t_R, \end{cases}$$

откуда можно получить уравнение, которое запирает конечно-разностную схему для Ψ_R слева:

$$(ik_L \Delta - \alpha) \Psi_1 + \alpha \Psi_2 = 0, \quad (\text{П7.30})$$

где $\alpha = m_I/m_{II}$.

Аналогично к тому, как было получено (П.28), можно получить выражение для вероятности прохождения электрона через активную область слева направо:

$$T_{RL} = \frac{k_L}{k_R} \frac{m_R}{m_L} |t_R|^2 = \frac{k_L}{k_R} \frac{m_R}{m_L} |\Psi_1|^2. \quad (\text{П7.31})$$

Учитывая (П7.29), (П.20) переписывается:

$$\begin{cases} \Psi_N = 1 + r_R, \\ \frac{m_{III}}{m_{II}} \frac{\Psi_N - \Psi_{N-1}}{\Delta} = ik_R (r_R - 1), \end{cases}$$

где $\alpha = m_{III}/m_{II}$. Исключая r_R , получим уравнение, которое запирает конечно-разностную схему для Ψ_R справа:

$$\alpha \Psi_{N-1} + (ik_R \Delta - \alpha) \Psi_N = 2ik_R \Delta. \quad (\text{П7.32})$$

Вероятности отражения электрона от правой границы квантовой области:

$$R_{RL} = |r_R|^2 = |\psi_N - 1|^2. \quad (\text{П7.33})$$

Для дальнейшего изложения результатов работы необходимо сделать замечание относительно коэффициентов прохождения и отражения от границ квантовой области. Для еще более широкого класса систем, чем рассматриваемая, можно показать, что:

$$T_{RL} = T_{LR} \equiv T, \quad R_{RL} = R_{LR} \equiv R.$$

Это следует из того, что выражения для волновой функции в I и III областях являются формами решения одного и того же дифференциального уравнения (Шредингера).

Заметим, что из (П7.26), (П7.28) и (П7.31), (П7.33) следуют следующие полезные соотношения:

$$\begin{aligned} |t_R|^2 &= \frac{k_R}{k_L} T = \frac{k_R}{k_L} (1 - R), \\ |t_L|^2 &= \frac{k_L}{k_R} T = \frac{k_L}{k_R} (1 - R). \end{aligned}$$