

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	3
1. СРЕДСТВА МОДЕЛИРОВАНИЯ.....	13
1.1. Метод кинетического уравнения	13
1.1.1. Кинетическое уравнение.....	13
1.1.2. Интеграл столкновений.....	18
1.1.3. Приближение времени релаксации.....	21
1.1.4. Примеры решения кинетического уравнения.....	23
1.2. Метод Монте-Карло	29
1.3. Релаксационные уравнения	36
1.3.1. Усреднение кинетического уравнения	36
1.3.2. Релаксационные уравнения	38
1.3.3. Междолинная релаксация	40
1.4 Уравнения дрейфово-диффузационной модели	45
1.5. Квантовые модели переноса заряда	47
1.5.1. Общие положения	47
1.5.2. Метод волновых функций	53
1.6. Модели субмикронных структур	57
Литература к разделу 1.....	67
2. ПОЛУПРОВОДНИКОВЫЕ МАТЕРИАЛЫ	71
2.1. Зонная структура полупроводников	71
2.1.1. Атомарные полупроводники	71
2.1.2. Бинарные полупроводники $A^{III}B^V$	77
2.1.3. Сложные соединения $A^{III}B^V$	86
2.1.4. Спектр фононов	88
2.2. Механизмы и скорости рассеяния	92
2.2.1. Фононное рассеяние	94
2.2.2. Примесное рассеяние	101
2.3. Средние времена релаксации	104
2.3.1. Времена релаксации импульса	104
2.3.2. Времена релаксации энергии, упругость рассеяния	108
2.4. Подвижность электронов	113
Литература к разделу 2	123
3. ГОРЯЧИЕ НОСИТЕЛИ В СИЛЬНОМ ЭЛЕКТРИЧЕСКОМ ПОЛЕ	125
3.1. Разогрев электронного газа	125
3.2. Насыщение дрейфовой скорости	137
3.3. Динамическая проводимость	147
3.3.1. Дифференциальная проводимость	147
3.3.2. Высокочастотная проводимость	150
3.4. «Всплеск» дрейфовой скорости	155
3.5. Диффузия горячих электронов	162
Литература к разделу 3.....	164
4. СУБМИКРОННЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ	165
4.1. Общие свойства	165
4.2. Квазидвумерная модель ПТШ	170
4.3. Схемные модели ПТШ	187

4.3.1. Малосигнальная модель	188
4.3.2. Шумовая модель	192
4.4. Модели ПТШ для режима большого сигнала	198
4.4.1. Субпороговый режим	198
4.4.2. Открытый затвор	204
4.4.3. Инверсный режим	208
4.5. Двумерное моделирование субмикронных ПТШ	209
4.6. Моделирование эффектов третьего измерения	214
4.6.1. Распределенные эффекты в ПТШ	214
4.6.2. Анализ ПТШ с распределенными параметрами во временной области	218
4.6.3. Моделирование размерных эффектов и технологических неоднородностей	222
Литература к разделу 4.....	228

КРАТКОЕ СОДЕРЖАНИЕ ВТОРОЙ ЧАСТИ

5. ГЕТЕРОСТРУКТУРНЫЕ ПОЛЕВЫЕ ТРАНЗИСТОРЫ

5.1. Общие свойства	
5.2. Кинетическое моделирование ГСПТ	
5.3. Двумерное моделирование ГСПТ	
5.4. Гетеротранзисторы с двумя потенциальными ямами	
5.5. Гетероструктурные транзисторы с квантовыми точками	

6. МОЩНЫЕ СВЧ ТРАНЗИСТОРЫ

6.1. Мощные субмикронные ПТШ	
6.2. Мощные гетероструктурные полевые транзисторы	
6.3. Гетеробиполярные транзисторы	
6.4. Мощные МОП-транзисторы	

7. РЕЗОНАНСНО-ТУННЕЛЬНЫЕ ДИОДЫ

7.1. Основные представления	
7.2. Условно-аналитические модели	
7.3. Численные модели	
7.4. Двухдолинное моделирование	
7.5. Экспериментальная верификация моделей	

ПРИЛОЖЕНИЯ

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 ($< 1 \text{ мкм}$) и нанометровые ($< 0,1 \text{ мкм}$) 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.

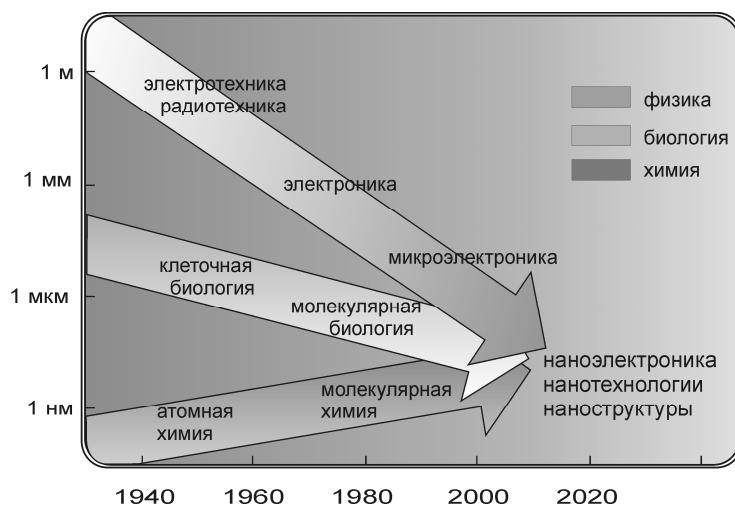


Рис.1. Тенденции развития некоторых направлений науки [1]¹

Сегодня существуют несколько направлений развития субмикронной и наноэлектроники – это поиск материалов и их соединений с принципиально новыми свойствами, совершенствование компонентов с субмикронными размерами

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

активных областей, включая гетероструктуры, а также создание новых наноразмерных приборов, устройств и систем.

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.

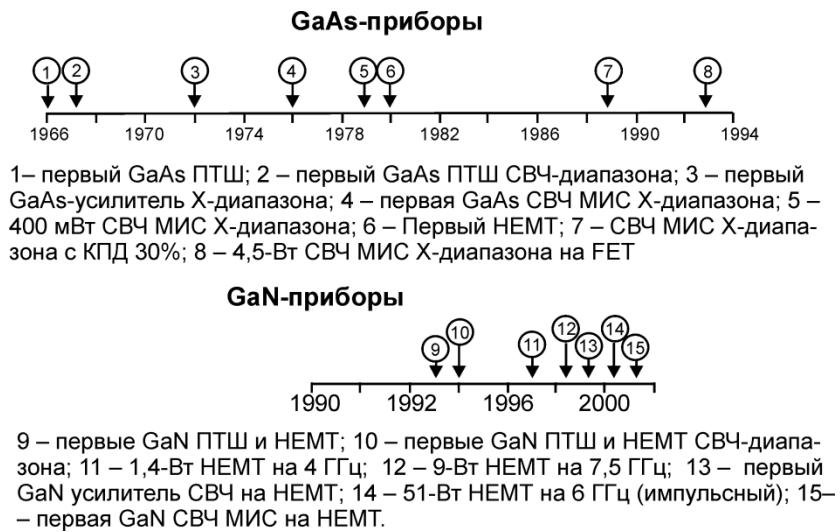


Рис. 2. Хронология создания приборов на арсениде и нитриде галлия [2]²

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»

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.

ультрафиолетовые фотоприемники, мощные транзисторы для мобильных телекоммуникационных систем и других.

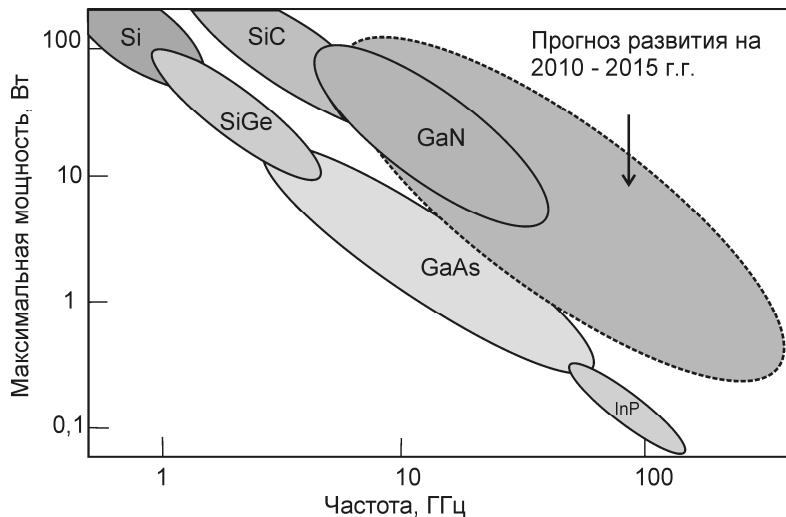


Рис. 3. Зависимость максимальной мощности от частоты для приборов на основе разных полупроводниковых материалов. Прогноз построен на основе работы [3]³ и расчетов разд. 3.

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

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

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

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 *p-n*-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 *tHEMT*. 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.

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 *tHEMT*, 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 AlGaN/GaN, которые уступают по частотным и усиливательным 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 многослойнымиnanoструктурами 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 допплеровскую чувствительность системы в целом.

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

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

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

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 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

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

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 Cellurar and 3G Base Stations. IEEE IEDM. 2003.

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 высокопроизводительных ЦАП, чем другие устройства на основе комплементарных структур металлоксид-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 генера-

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 Cellurar Infrastructure Application. IEEE MTT-S. 2003.

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

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

торы, регистры сдвига, а также статические ЗУПВ с потреблением энергии 50 нВт/бит [13]¹²

Благодаря высокой скорости процесса туннелирования резонансно-тунNELьные 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.

Учебная литература, особенно предназначенная для спецкурсов, не может оставаться в стороне от уровня и тенденций развития электроники. Предлагаемое издание представляет собой попытку отследить эти тенденции и дать будущим специалистам по микро- и наноэлектронике основы знаний, необходимых для понимания и моделирования физических процессов, особенно динамических, происходящих в современных приборах. Авторы, следуя принципу «познание через моделирование», включили в настоящее издание 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.

Данное издание писалось как учебное пособие по однотипному курсу для студентов, обучающихся по направлению «Микро- и наноэлектроника», но авторы надеются, что оно сможет принести пользу для гораздо большего круга читателей, связанных с СВЧ электроникой.

Авторы глубоко благодарны коллегам в лице Минакова В.В., Босого В.И., Иващенко 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.

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 , дает 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 объема $V = 1$, 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 дает 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

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

а значит и распределение плотности тока

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

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

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

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 Г-, 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 коллектипов 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

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

Силы, действующие на электроны, можно условно разделить на две группы. Это, во-первых, силы со стороны внешних полей $\vec{F} = -e(\vec{\mathcal{E}}(\vec{r}, t) + [\vec{v} \times \vec{B}(\vec{r}, t)])$, 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 *согдания*. 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 вакуме, 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 пространственным распределением 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 уравновешивается дрейфовым движением в электрическом поле.

1.1.2. Интеграл столкновений

Для решения кинетического уравнения 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 $w_{kk'}$. 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

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

Аналогично вероятность обратных — $w_{k'k}f(\vec{k}')\left(1 - f(\vec{k})\right)$. 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

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

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

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

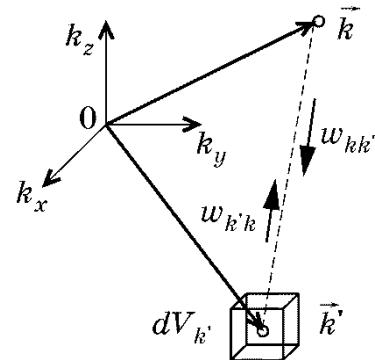


Рис. 1.1. Изображение переходов между k - и k' -состояниями. Выбором элементарного объема в окрестности k' -состояния и последующим интегрированием по всему k' -пространству достигается подсчет изменения числа частиц в выбранном k -состоянии за счет обмена со всеми остальными состояниями.

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

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

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. 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

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

Соответственно в равновесном состоянии

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

Но при подстановке в (1.22) функции распределения Максвелла-Больцмана в виде $f(E(k)) = \exp[(E_\phi - E(k))/k_B T]$ 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 одиночечной 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 состояние \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 штрихованым переменным, то $f(k)$ можно вынести за знак интеграла и получить, что

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

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

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

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, *a*) 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, *b,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 «перескаивают» в случайные места (пример соответствующего перехода обозначен стрелкой).

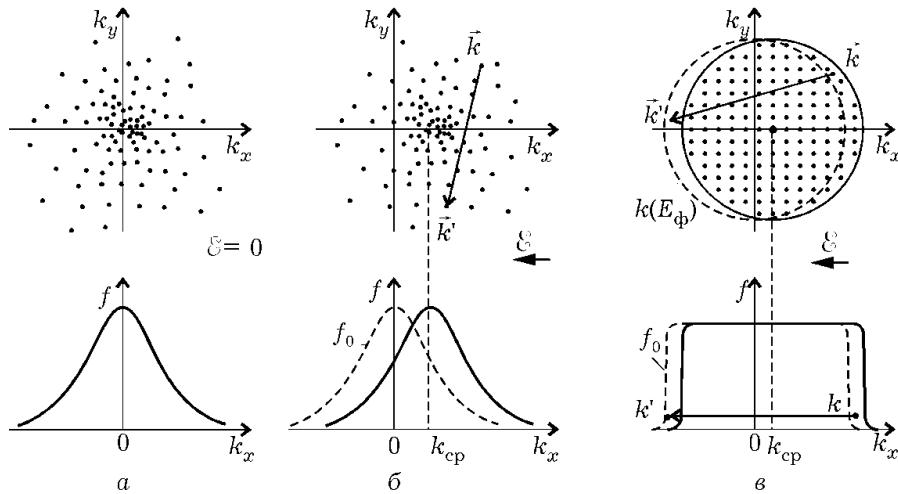


Рис. 1.2. При отсутствии электрического поля (а) равновероятное распределение волновых векторов (вверху) обуславливает симметричный характер функции распределения. В электрическом поле (б, в) смещение волновых векторов в направлении действия силы и рассеивание электронов в случайные состояния уравновешиваются 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.

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 уравновешивают 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 $\Delta N = g(\vec{k}) \cdot \Delta f$. 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 ($\vec{\mathcal{E}} = \text{const}$) и коллектива ($\partial f / \partial \vec{r} = 0$), принимает вид

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

Так как функция распределения f зависит от \vec{k} через закон дисперсии, 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 мало 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

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

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 обычной форме $g(E) = 4\pi(2m_d^*/h^2)^{3/2}(E - E_c)^{1/2}$ [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 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 ($\vec{\mathcal{E}} = 0$) 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.

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 одиночественной модели дрейфа.

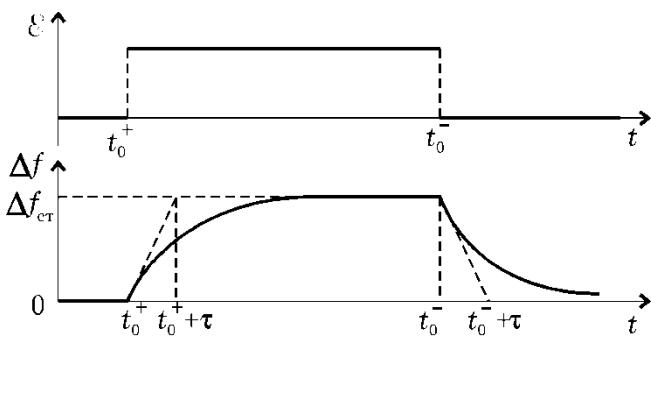


Рис. 1.3. Переходные процессы изменения неравновесной части функции распределения электронов и плотности их тока при включении и выключении электрического поля. Наклонные штриховые линии – касательные к кривым в моменты t_0^+ и t_0^- .

Следует отметить одну особенность релаксации функции распределения, характерную для вырожденных коллективов (рис. 1.2, в). В отличие от невырожденных, где вероятность за-

полнения 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 $\vec{k}(E_\phi)$, 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 $\tau(E_\phi)$ 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 $\arctg \omega t$. 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 ($\omega t \geq 1$). 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 $\text{rot} \vec{H} = (i\omega \epsilon \epsilon_0 + \sigma' - i\sigma'') \vec{\mathcal{E}}$. 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

ге 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 $10^4...10^5$. 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 электронная плазма в полупроводнике представляется некоторым ансамблем модельных частиц, количество которых существенно меньше, чем реальных. Движение частиц моделируется полуклассическими уравнениями движения [1.1]:

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

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

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.

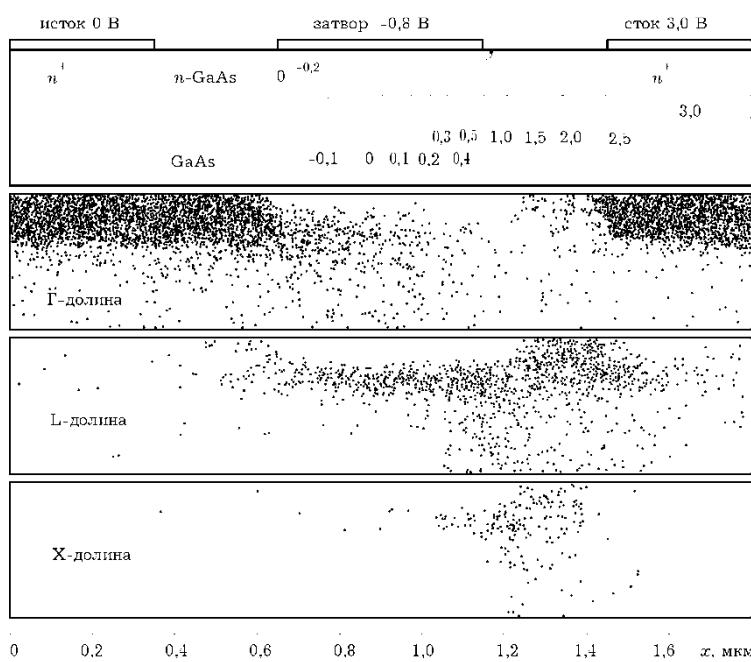


Рис. 1.5. Типичное представление распределения поля с помощью эквипотенциальных линий и распределения частиц в пространстве моделируемой структуры. Пространственное распределение частиц приведено отдельно для каждой долины. Штриховой линией обозначена нижняя граница эпитаксиального слоя с высоким легированием.

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

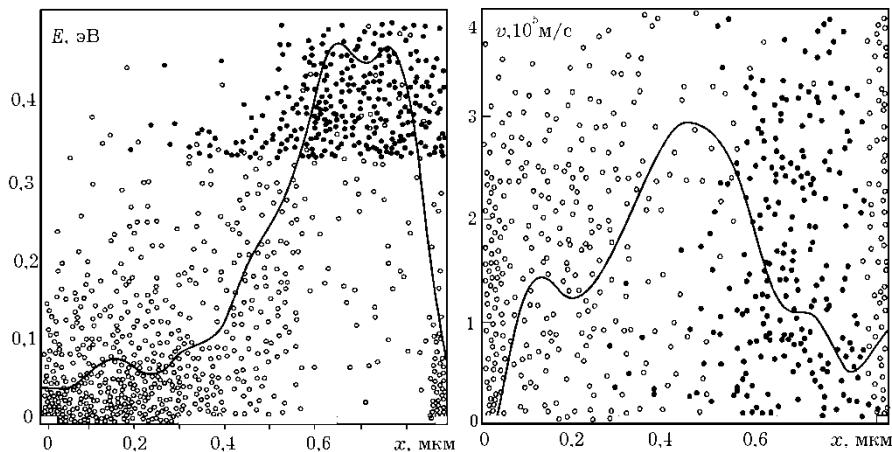


Рис. 1.6. Распределения энергии и скорости электронов вдоль канала полевого транзистора в форме «индивидуального» расположения частиц и усредненных функций. Светлые кружки соответствуют электронам нижней долины, темные — верхних.

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].

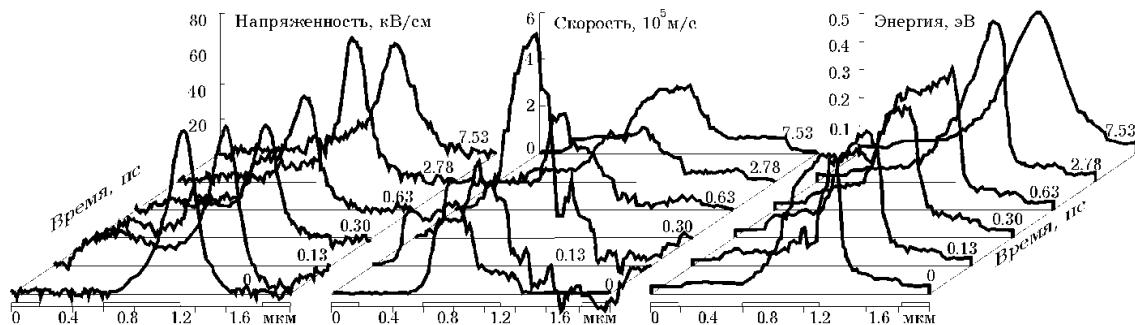


Рис. 1.7. Распределения напряженности поля, средней скорости и энергии частиц вдоль канала транзистора в различные моменты времени переходного процесса.

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 флюктуаций.

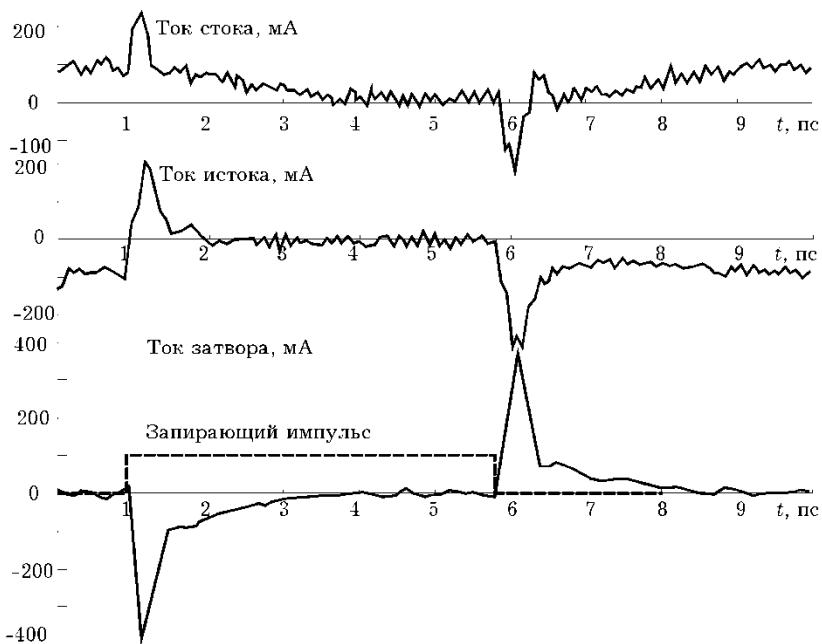


Рис. 1.8. Переходные характеристики токов стока, истока и затвора при подаче на затвор ступенчатого импульса: запирающего и открывающего.

Применительно к твердотельным структурам метод Монте-Карло особенно удобен, когда характерные размеры структур составляют микрометры и менее. Для более крупных требуется моделировать поведение большего 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 назы-

вают 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 $\phi(\vec{k})$ 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 $\phi(\vec{k})$, с учетом (1.6) примет вид:

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

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 входящих в нее 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 $dE/d\vec{k} = \hbar\vec{v}$:

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

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)

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

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,

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

и подвижность определяется временем релаксации импульса.

В соответствии с (1.54) это время находится через усреднение обратного времени релаксации $1/\tau = \lambda_\Sigma = \sum_i \lambda_i$, 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)

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

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 коллектизы 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 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 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

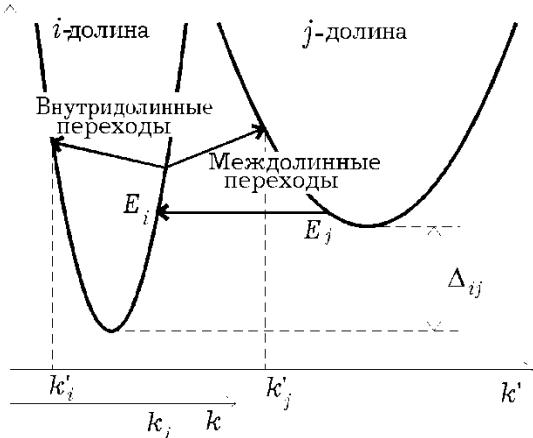


Рис. 1.9. Обозначение внутри- и междолинных переходов электронов из состояния k в k' в двухдолинной модели зонной структуры.

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

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

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

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 ($d n_i / dt = d n_j / dt = 0$) 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

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

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 внутридолинные переходы, так и междолинные

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

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

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

с временем релаксации импульса, учитывающим как внутри- так и междолинные переходы из j -долины 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 долин.

Подобным образом можно провести преобразование интеграла столкновений, умноженного на E , и получить уравнение сохранения энергии для i -долины. По виду оно будет совпадать с (1.51) с правой частью в форме (1.55) со временем релаксации энергии τ_{Eii} , учитывающим только внутридолинные механизмы рассеяния. Междолинное 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 -долине за единицу времени при их переходе в j -долину и связано с междолинным 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 -долины

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

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

При обратных переходах электроны из j -долины 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:

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

Сложив уравнения (1.79), (1.80) и разделив их на суммарную концентрацию $n = n_i + n_j$, вводя средние по долинам значения энергии $E = (n_i E_i + n_j E_j)/n$, скорости $\vec{v} = (n_i \vec{v}_i + n_j \vec{v}_j)/n$ и времени релаксации $1/\tau_E = [n_i(E_i - E_0)/\tau_{Ei} + n_j(E_j - E_0)/\tau_{Ej}]/n$, 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 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 уравновешиваются. Но в дина-

мических ситуациях это явление может быть весьма заметным, так как междолинное расстояние Δ_{ji} 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-ДИФФУЗИОННОЙ 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 $\operatorname{div} \vec{v} = 0$.

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

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

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

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

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

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

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

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

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 диффузационным механизмами

$$\begin{aligned}\vec{j}_n &= e n \mu_n \vec{\mathcal{E}} + e D_n \operatorname{grad} n, \\ \vec{j}_p &= e p \mu_p \vec{\mathcal{E}} - e D_p \operatorname{grad} p.\end{aligned} \quad (1.89)$$

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

$$\nabla^2 V = -\rho(\vec{r}) / \epsilon \epsilon_0 = e(\Delta n - \Delta p) / \epsilon \epsilon_0 \quad (1.90)$$

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-диффузационной 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 текущим 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-диффузационных. 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 полуцелого спина 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. 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 достаточно 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 них - тунNELьный 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 Броиля 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 ровным волновому вектору Ферми: $k = k_\Phi$. Условно считают, что значение длины волны де Броиля в типичных полупроводниковых структурах $\lambda_\Phi = 35$ нм [1.9].

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 как $v_\Phi = \hbar k_\Phi / m$, 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: флюктуационные центры рассеивания вызывают фазовую релаксацию, а стационарные центры - 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

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.

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-тунNELьные 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 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. 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;

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.

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 Лихаревым 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 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 тяготела 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 эффе-

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

тов 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_ϕ , 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 них реализуется *поперечный транспорт*. 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 наноэлектроники в целом. Представления об электронном транспорте в «старых» устройствах, не могут отображать реальные процессы, которые происходят в «новых» устройствах. Поэтому наноэлектроника не может формироваться как «надстройка» микроэлектроники, и с этой точки зрения является революционным, а не эволюционным этапом развития электроники.

Формировать принципы теоретической наноэлектроники приходится на том же фундаменте, на котором была построена микроэлектроника: многочастичная квантовая теория в широком смысле этого понятия, которое включает в себя специальные разделы: 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. Их возникновение объясняется отставанием теоретической наноэлектроники 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, прилагаемым 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 «огибающими волновой функции», а не просто «волновыми функциями». Такое название обусловлено тем, что при $U(\mathbf{r}) = \text{const}$ решением (1.91) являются функции Блоха: $\psi(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\mathbf{r}}$, где функция $u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{a})$ повторяет периодичность решетки и существенно изменяется в течение периода решетки. Множитель

e^{ikr} 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 влияет на $u_{n\mathbf{k}}(\mathbf{r})$; зато изменение отобразится на втором члене [1.21]. В множитель e^{ikr} можно ввести зависимость волнового вектора от координаты $\mathbf{k}(\mathbf{r})$ или же заменить e^{ikr} на медленно изменяющуюся $\chi(\mathbf{r})$: $\Psi(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}(\mathbf{r})\mathbf{r}} = u_{n\mathbf{k}}(\mathbf{r})\chi(\mathbf{r})$. Функция $\chi(\mathbf{r})$ фактически и является искомой в уравнении (1.91) с гамильтонианом (1.93); естественно, что ее тоже называют «огибающей». Дальше, однако, будем обозначать огибающую волновой функции через $\psi(r)$, 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) распадается 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:

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

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.

Границные 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. Кроме того, скачкообразное изменение испытывает также эффективная масса, которая в материалах А и Б в общем случае разная.

На границе этой гетеропары происходит движение электронов проводимости в состояниях Г-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 производ-

ная 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 -проекций 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, внеся 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. Уравнения (1.105) часто называют *уравнением Шредингера с эффективной массой*. Оно является ключевым компонентом так называемого «метода огибающей волновой функции», ко-

торая применяется для моделирования гетероструктур. Название этого метода является синонимом *метода эффективной массы*.

1.6. МОДЕЛИ СУБМИКРОННЫХ СТРУКТУР

Задачи, решаемые в инженерной практике с помощью моделей, зависят, прежде всего, от круга пользователей. К ним относятся разработчики СВЧ устройств, которые используют модель как завершенный 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 тем или иным параметрам. Это позволяет выяв-

лять наиболее сильные связи и проводить поиск более целенаправлено.

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 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 работа разработчиков моделей должна проходить в тесном взаимодействии с разработчиками приборов и технологами, так как процесс моделирования должен соотносится с экспериментальными данными, полученными при испытании приборов и тестовых структур. При этом совершенствование мо-

делей 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 факто-ров, 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 $v(E)$ 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 урав-

нений, 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]. 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.

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-диффузационном 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 изменени-

ем 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 квазимпульсов. 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 прослеживать «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 флюктуации проявляются «естественному» образом. Все это позволяет считать кинетические модели модельями 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 схемотехники, 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 харак-

терными 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, зависящий 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», в той или иной мере хорошо 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.

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.

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 макромодели, 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. Макромодели 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.

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 разделу 1

- 1.1. *Москалюк В. А., Синекоп 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 Харк. 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-тунNELьные 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. «Электроника и связь» – 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 «Электроника и связь», тематический выпуск «Электроника и нанотехнологии». – 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 за-

творм 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 non-stationary 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. *Іващук А.В., Тимофєєв В.І.* Вплив морфології омічних контактів на надвисокочастотні параметри польових транзисторів // Оптоелектроника и полупроводниковая техника.– 2000.–№35.–С.157–161.

1.36. *Іващук А.В., Матуленіс А., Тимофєєв В.І.* Вплив морфології омічних 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 P, мышьяк As, сурьма Sb, сера S, селен (красный) Se, теллур Te, йод J. К полупроводникам относятся также бинарные соединения типа A_xB_{8-x} , где A – элемент группы x , а B – элемент группы $(8-x)$. Наиболее распространены и применимы из них – вещества типа A_3B_5 , к которым относятся соединения азота (нитриды), фосфора (фосфиры), 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

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 отличием в методах вычисления этих кривых, так и разной степенью их детализации. Не останавливаясь на этих от-

личиях, 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,

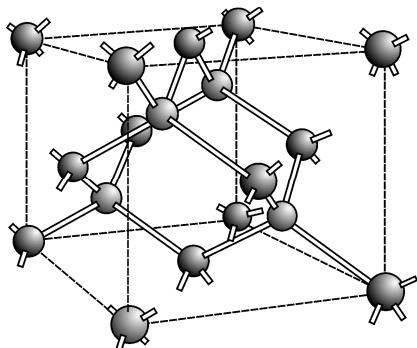


Рис. 2.1. Кристаллическая структура алмаза, состоящая из двух гранецентрированных кубических подрешеток.

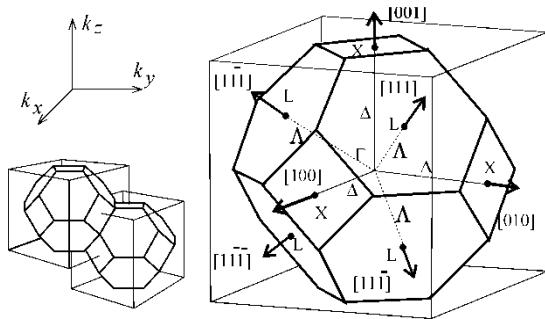


Рис. 2.2. Форма первой зоны Бриллюэна для кристаллической решетки типа алмаза и обозначение характерных точек и направлений симметрии [2.1]

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 Г, 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.

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

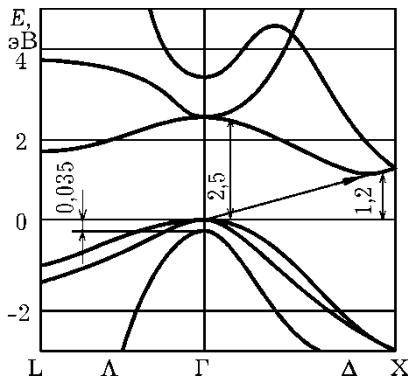


Рис. 2.3. Дисперсионные кривые в зоне проводимости и валентной зоне кремния в Л- и Δ -направлениях. Расстояние между абсолютными экстремумами кривых соответствует запрещенной зоне.

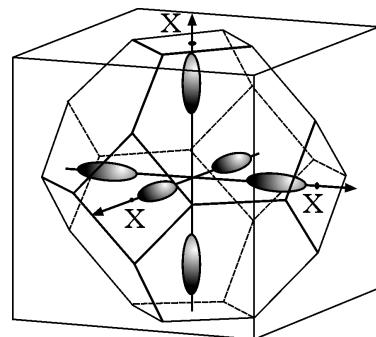


Рис. 2.4. Расположение изоэнергетических поверхностей, соответствующих Δ -долинам у дна зоны проводимости кремния, в первой зоне Бриллюэна.

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 $m_{p\text{Л}}^* \approx 0,16m$, для «тяжелых» $m_{p\text{T}}^* \approx 0,5m$. Для вычисления эффективной плотности состояний у края валентной зоны удобно пользоваться некоторой усредненной массой:

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

Структура зоны проводимости еще более сложная, чем валентной. Абсолютный минимум в зоне проводимости имеет одно из ветвей в Δ -направлении на расстоянии 0,855 отрезка Г-Х. Принимая во внимание 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. 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 Г 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 квазимпульса. Такой переход, обозначенный стрелкой на рис. 2.3, называется *непрямым*. Прямой переход осуществляется без изменения квазимпульса 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 ($m_{p\text{Л}}^* = 0,04m$, $m_{p\text{T}}^* = 0,34m$), что соответствует эффективной массе плотности состояний $m_{pd}^* = (m_{p\text{Л}}^{*3/2} + m_{p\text{T}}^{*3/2})^{2/3} \approx 0,36m$. Третья ветвь отстает от края зоны на величину энергии спин-орбитального взаимодействия, которая равна 0,28 эВ.

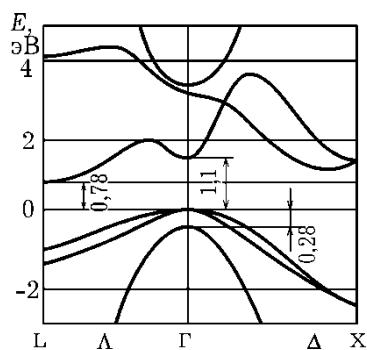


Рис. 2.5. Дисперсионные кривые в зоне проводимости и валентной зоне германия в Λ - и Δ - направлениях. Ширина запрещенной зоны германия – расстояние между абсолютными экстремумами кривых.

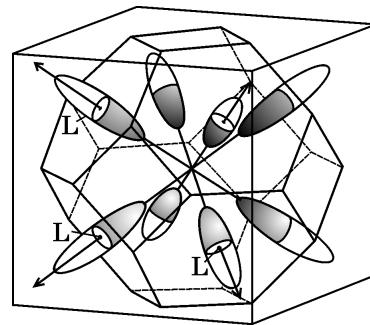


Рис. 2.6. Изоэнергетические поверхности, соответствующие L -долинам у дна зоны проводимости германия. Эллипсоиды вращения лежат лишь наполовину в первой зоне Бриллюэна.

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 Λ -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 $M = 8/2 = 4$.

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, электронное средство 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 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 Г- и L-долин и относительно плавно X-долини. Для $x = 0,15$ положения X- и L-долин выравниваются и ширина запрещенной зоны составляет 0,86 эВ. При увеличении доли кремния она постепенно приближается к значению 1,11 эВ.

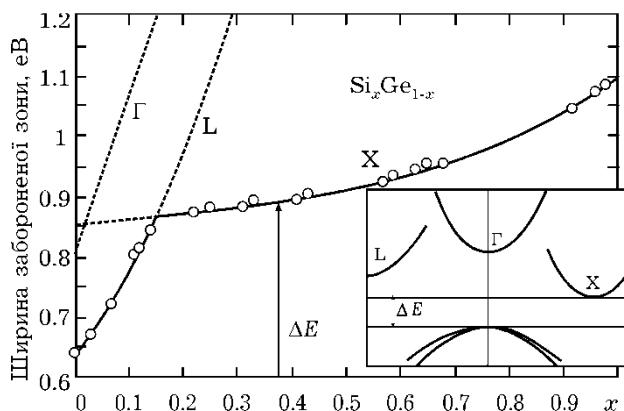


Рис. 2.7. Энергетическое положение долин в зоне проводимости $\text{Si}_x\text{Ge}_{1-x}$, отсчитанное от потолка валентной зоны, в зависимости от доли кремния

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

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

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

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

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

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

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.

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

Соединение	Период решетки, $\times 10$, нм	Плотность, г/см ³	Температура плавле- ния, °C	$10^6 A/(K^2 \cdot m^2)$	Ширина запрещенной зоны, эВ	Подвижность электро- нов, м ² /(В·с)	Подвижность дырок, м ² /(В·с)	Показатель преломления при $h\nu = \Delta E$	Диэлектрическая про- ниаемость
BN _(куб)	3,615	3,49	3000	-	6,0	-	-	2,1	7,1
AlN	3,110	3,28	2400	6,1	5,88	-	-	2,2	9,1
	4,975								
GaN	3,186	6,11	1700	5,65	3,40	0,03	-	2,4	12,2
	5,176								
InN	3,540	6,91	1100	-	1,95	-	-	2,9	-
	5,704								
AlP	5,463	2,37	2000	4,2	2,45	0,008	0,003	3,0	9,8
GaP	5,451	4,07	1467	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	5,2	2,16	0,020,8	-	3,2	10,1
GaAs	5,653	5,32	1237	6,4	1,43	0,95	0,045	3,65	13,1
InAs	6,058	5,67	942	5,3	0,36	3,3	0,046	3,52	14,6
AlSb	6,136	4,28	1060	4,2	1,58	0,02	0,055	3,4	14,4
GaSb	6,096	5,65	710	6,2	0,72	0,4	0,14	3,8	15,7
InSb	6,479	5,78	525	4,9	0,18	7,8	0,075	4,0	17,7

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 [2.5].

Это иллюстрируется рис. 2.8, где для кубической структуры кристаллической решетки представлены поверхности, соответствующие положениям дна зоны проводимости в точ-

ках 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.

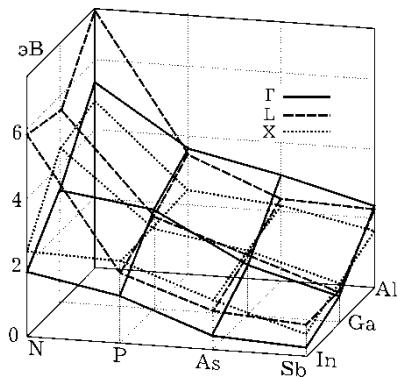


Рис. 2.8. Зависимости ширины запрещенной зоны в разных точках симметрии (Γ , L и X) зоны Бриллюэна от состава соединения.

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.

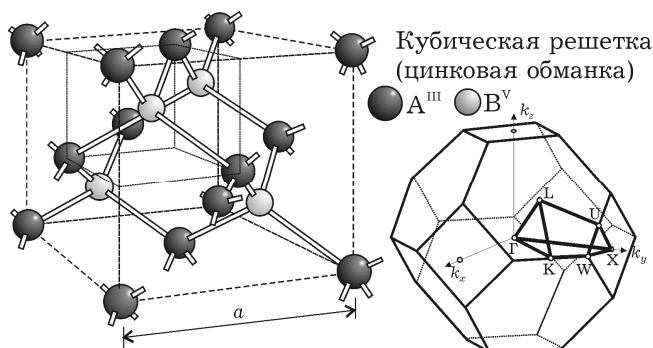


Рис. 2.9. Структура кубической кристаллической решетки типа цинковой обманки и форма первой зоны Бриллюэна. Выделен минимальный элемент симметрии – 1/64 часть зоны

Для нитридов характерна также структура гексагонального 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 [110], полностью состоящих из разноименных атомов. Разное поведение граней сказывается при травлении, окислении, и при выращивании кристаллов.

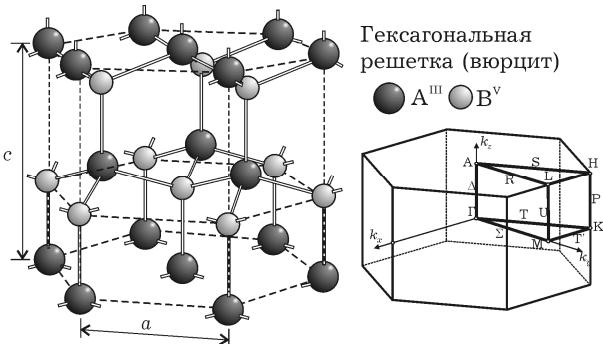


Рис. 2.10. Структура гексагональной кристаллической решетки типа вюрцита и форма первой зоны Бриллюэна. Выделен наименьший элемент симметрии – 1/24 часть зоны

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 $A^{III}B^V$.

В структуре энергетических зон атомарных полупроводников IV группы и соединений $A^{III}B^V$ много общего. В центре валентной зоны (рис. 2.11) находятся максимумы двух ветвей, соответствующих «легким» ($m_{p\text{л}}^* \approx 0,1m$) и «тяжелым» ($m_{p\text{т}}^* \approx 0,5m$) 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 Г). 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 Г-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: $m_l^* = 1,9m$, $m_t^* = 0,075m$.

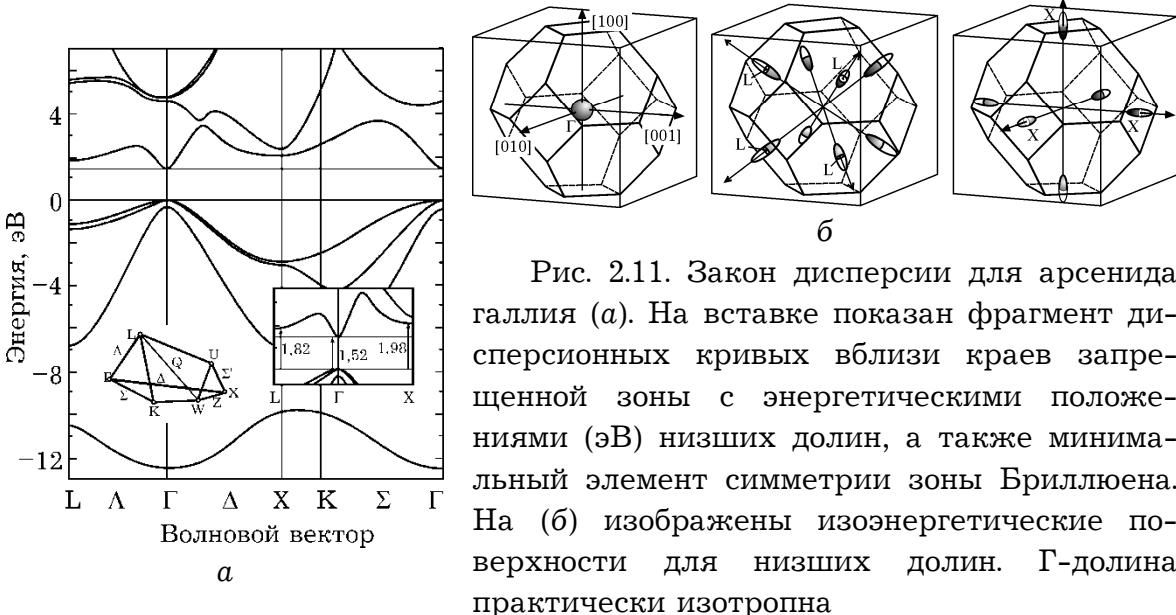


Рис. 2.11. Закон дисперсии для арсенида галлия (а). На вставке показан фрагмент дисперсионных кривых вблизи краев запрещенной зоны с энергетическими положениями (эВ) низших долин, а также минимальный элемент симметрии зоны Бриллюена. На (б) изображены изоэнергетические поверхности для низших долин. Г-долина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

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 Г-, X- и

L-долины (рис. 2.12, а). Дну зоны проводимости соответствует центральная Г-долина с изотропной эффективной массой $m_{\Gamma}^* = 0,13m$. Следующие шесть X-долин находятся на энергетическом расстоянии $\Delta_{GX} = 1,3$ 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 ($M_X = 6/2$) $m_d^* = M^{2/3} \sqrt[3]{m_t^* m_l^*} = = 0,76m$. Более высокие L-долины находятся на 2,3 эВ выше, чем Г-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 Г-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 А-долину – $(1,3...2,1)$ эВ [2.9]. Практически отсутствует информация относительно эффективной массы в верхних долинах, иногда ее принимают ровной обычной массе. Изотропная эффективная масса для верхней Г-долины составляет $(0,9...1)m$ [2.10]. Важным параметром с точки зрения эмиссионных и контактных свойств является энергия электронного средства, которая для нитрида галлия равна 4,1 еВ [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 кристаллографичес-

кому направлению $<100>$ (Х-долина) на расстоянии 4,9 эВ от максимума валентной зоны [2.10].

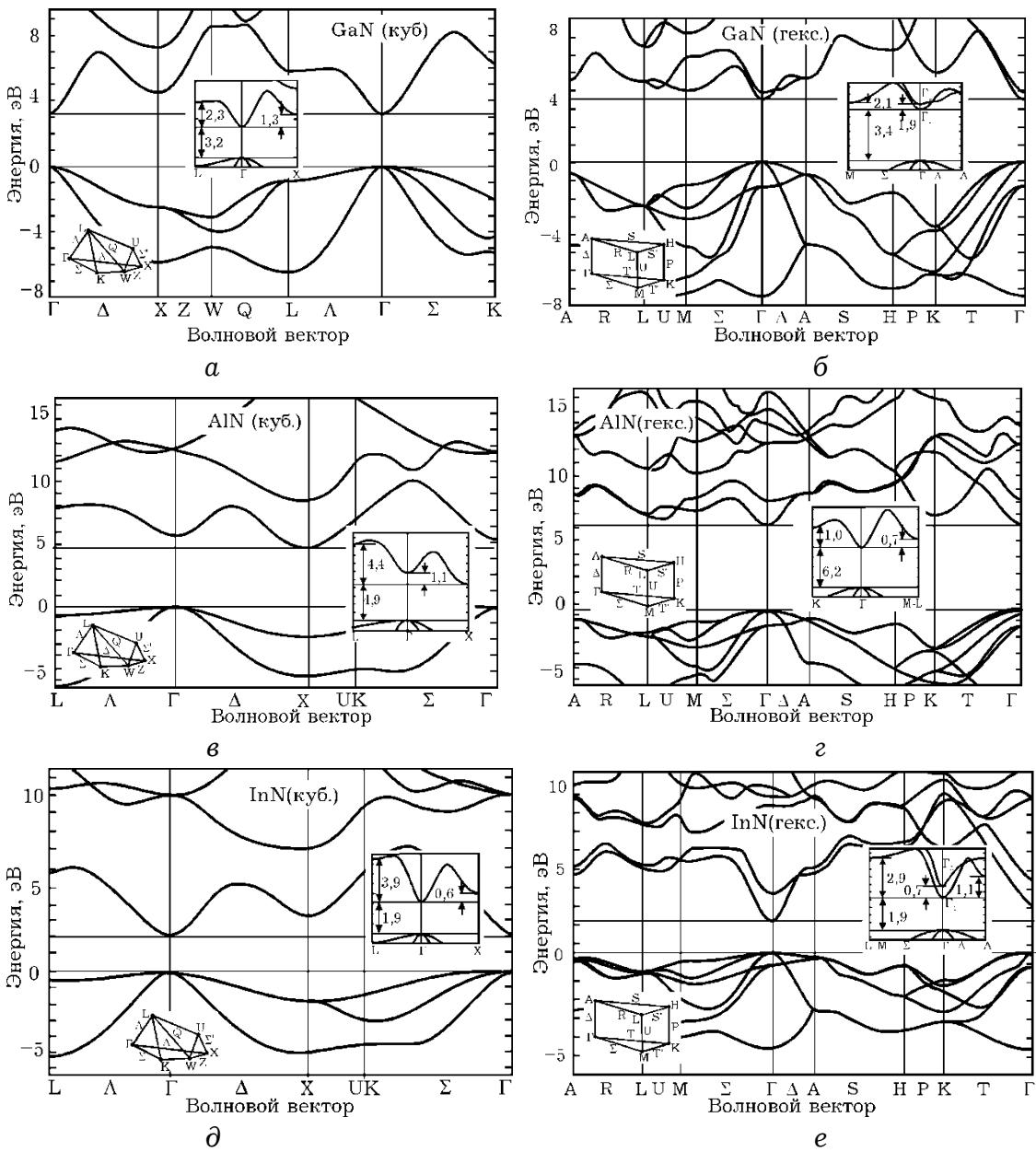


Рис. 2.12. Закон дисперсии для нитридов с кубической (a, c, d) и гексагональной (b, e) структурой кристаллической решетки. На вставках изображены фрагменты дисперсионных кривых у края запрещенной зоны с параметрами, которые использовались при расчетах, а также минимальный элемент симметрии зоны Бриллюэна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

Ширина запрещенной зоны, которая соответствует центральной Г-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 $m_d^* = 4^{2/3} (m_t^{*2} m_l^*)^{1/3} \approx 0,79m$. Энергия электронного сродства составляет лишь 0,6 эВ [2.9].

Нитрид алюминия типа *вюрциита* 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 (Γ -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 0,6 еВ [2.9]. 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 Γ -, X- и L-долин.

Абсолютный минимум зоны проводимости лежит в центре зоны Бриллюэна (точка Γ). 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 Γ -долине является скалярной величиной ($m^* \approx 0,12m$) [2.9].

Следующая X-долина находится на энергетическом расстоянии $\Delta_{\text{GX}} = 0,57$ 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 омическая 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, e) 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 Г₁-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 А-долину, расположенную в Δ-направлении на энергетическом расстоянии (0,7...2,7) эВ. Другие считают таковой вторую Г₂-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 $\Delta_{\text{GML}} = (2,9...3,9)$ эВ [2.9], с эффективной массой плотности состояний с учетом всех долин 1,08 m . 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 Г-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 пло-

тности состояний в одной долине $0,25t$, или с учетом всех долин $0,63t$. 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 $0,32t$ и $0,66t$. Энергия электронного средства составляет 4,38 эВ.

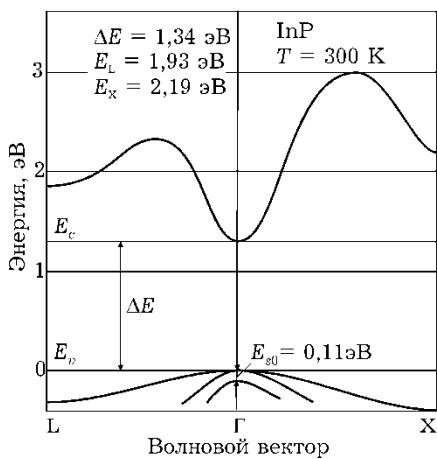


Рис. 2.13. Закон дисперсии для фосфида индия (реконструировано из [2.9]). Характерной является ГЛХ-последовательность долин в зоне проводимости

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

Широкое использование в электронике нашли тройные соединения материалов A^{III} и B^V , например, InGaAs, AlGaN, 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). 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 ГЛХ- в GaAs изменяется на последовательность XLГ- в AlAs. Интересной особенностью соединения является то, что при $x = 0,45$ положения всех долин совпадают. В [2.10] приве-

дены аппроксимации, которые определяют изменение положения соответствующих долин относительно потолка валентной зоны от части алюминия:

$$E_{\Gamma}(x) = 1,424 + 1,247x \quad \text{для } 0 < x < 0,45,$$

$$E_{\Gamma}(x) = 1,424 + 1,247x + 1,147(x - 0,45)^2 \quad \text{для } 0,45 < x < 1,0,$$

$$E_L(x) = 1,708 + 0,642x,$$

$$E_X(x) = 1,900 + 0,125 + 0,143x^2.$$

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]

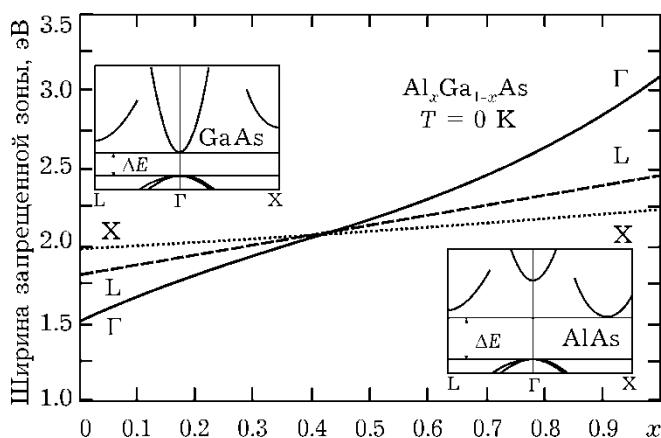


Рис. 2.14. Энергетическое положение долин в зоне проводимости $\text{Al}_x\text{Ga}_{1-x}\text{As}$, отсчитанное от потолка валентной зоны, в зависимости от доли алюминия. На вставках изображены эскизы законов дисперсии для чистых материалов

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

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 С 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 волн .

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 $\omega_m = 2v_{\text{зв}}/a$ – 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 $v_{\text{гр}} = d\omega/dk_{\phi}$ 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: $\omega \approx \omega_m k_{\phi}/2$, поэтому фазовая и групповая скорости одинаковы и равны скорости звука. Такие колебания и соответственно фононы называют *акустическими*.

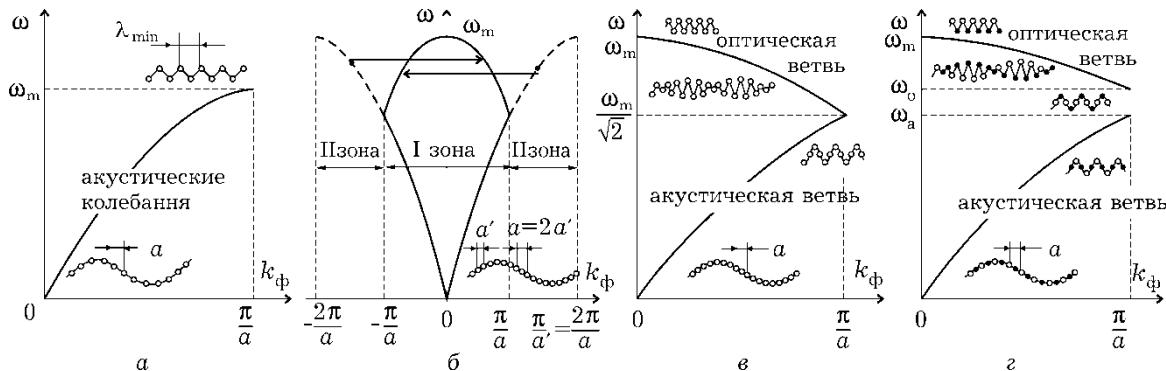


Рис. 2.15. Законы дисперсии волн смещения в одномерной модели кристалла – цепочки атомов: с одноатомным базисом (а), с базисом из двух одинаковых атомов (б, в) и двух разных атомов (г). Показано (б) образование оптической ветви для двухатомного базиса путем приведения верхней части «акустической» ветви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

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 $k_\phi = \pm\pi/a' = \pm2\pi/a$. 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 $2\pi/a$.

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).

В кремнии и германии, элементарные ячейки которых состоят из одинаковых атомов, ветви продольных LA- и LO- мод вырождены на краю зоны Бриллюэна, а энергия междо- линных фононов равна $\sim 0,05$ эВ и $\sim 0,03$ эВ соответственно.

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 ($M_{Ga}/M_N \approx 5$).

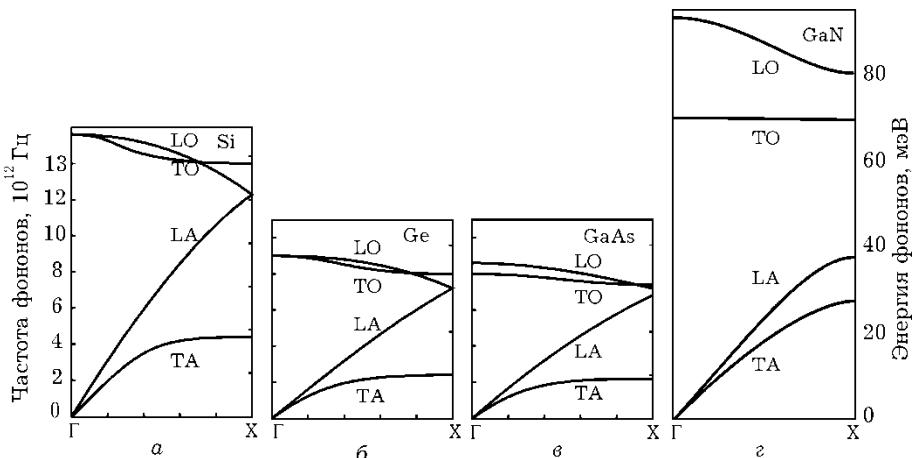


Рис. 2.16. Дисперсионные кривые фононов для кремния (*а*), германия (*б*), арсенида галлия (*в*) и кубического нитрида галлия (*г*) в кристаллографическом направлении [100] (L – продольные, T – поперечные, A – акустические и O – оптические моды). На правых шкалах отложены энергии фононов, на левых –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

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

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

где ϵ_{ct} 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.

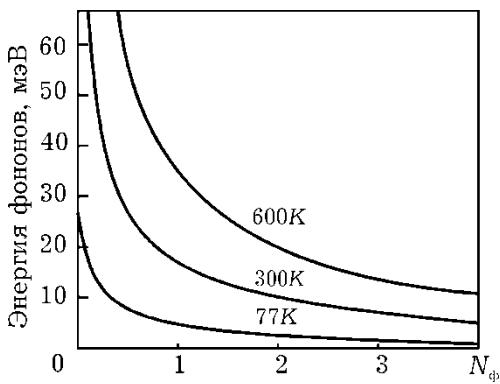


Рис. 2.17. Зависимость между средним числом фононов данной частоты и их энергией при различных температурах.

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 рас-

сияние 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

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 фона. В этом случае говорят о *поглощении* или *возбуждении* фона 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 (квазимпульса), который удобнее выражать через волновые вектора:

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

В приведенных формулах штрихами обозначены энергия и волновой вектор электрона после рассеяния, а знаки “±” относятся соответственно к поглощению фона и к его возбуждению. Квантовомеханический анализ процессов перехода между электронными состояниями дает выражения для вероятностей рассеяния с поглощением фона электроном и с его возбуждением

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

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

Пропорциональность вероятности количеству фононов ($N_\phi + 1/2 \mp 1/2$) говорит о том, что скорость рассеяния с возбуждением фононов в $(N_\phi + 1)/N_\phi$ 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 фона. 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 длинновол-

новых ($k_{\phi} \approx 0$), 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 .

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 фонах.

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 фона на 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 фона в этих условиях практически одинаковы. Это позволяет относительно просто провести вы-

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

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

где m^* – эффективная масса плотности состояний, D_a – деформационный акустический потенциал, v_{3B} – скорость звука, соответствующая наклону акустической ветви продольных колебаний при $k_\phi \approx 0$, 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 $\sigma = e\mu$ 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 пропорционально $T^{-3/2}$.

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 фона на возможно даже при самых низких энергиях (температурах), а для испускания фона на энергия электрона должна быть не ниже энергии фона: $E \geq \hbar\omega_0$. С другой стороны, пропорциональность вероятности рассеяния количеству фононов ($N_\phi + 1/2 \mp 1/2$) приводит к различию скоростей рассеяния: вероятность испускания может быть существенно больше вероятности поглощения, если $N_\phi \ll 1$ (конечно, если энергия электрона достаточна для испускания фона).

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 пригодно выражение [2.16]

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

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]:

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

Наиболее примечательной особенностью полярного оптического рассеяния является зависимость скорости рассеяния от энергии электрона: $\lambda_{po} \sim E^{-1/2}$, 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).

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.

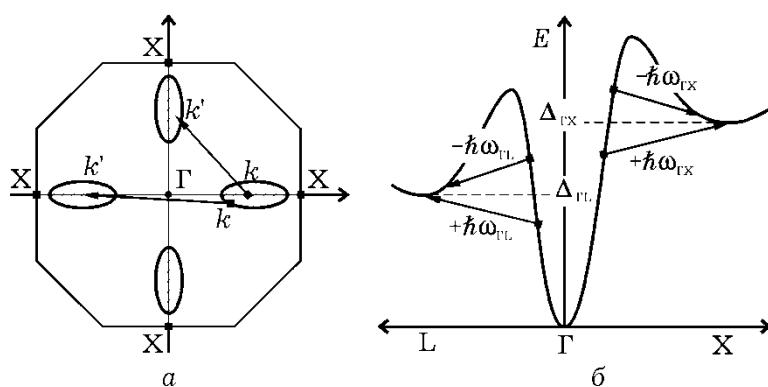


Рис. 2.18. Примеры переходов при междолинном рассеивании между эквивалентными долинами в кремнии (а) и между неэквивалентными долинами в арсениде галлия (б)

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 фона на 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 фона, $\Delta_{ij} = E_{cj} - E_{ci}$ – 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.

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

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

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_{Φ} – 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 меньшее 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 Г- 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.

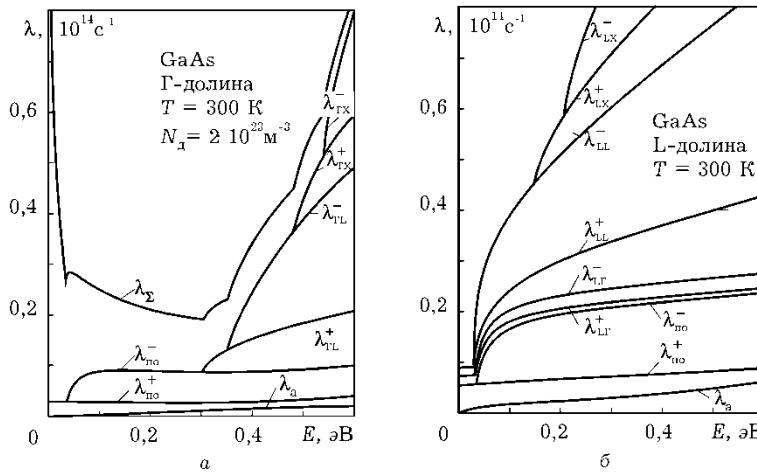


Рис. 2.19. Диаграмма зависимостей скорости рассеивания для разных видов от энергии электрона в Γ -долине (а) и L-долине (б). График для каждого вида суммируется с предыдущими. Верхняя кривая λ_{Σ} суммарной скорости включает также примесное рассеивание.

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.

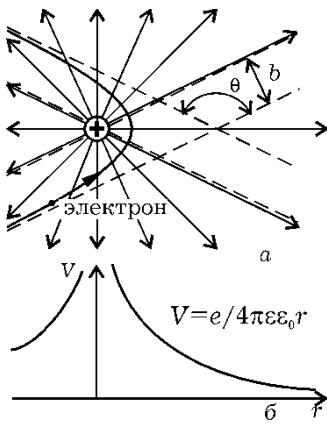


Рис. 2.20. Траектории движения электрона в кулоновском поле иона до-норной примеси.

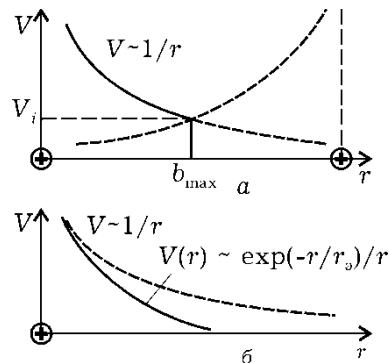


Рис. 2.21. Ограничение потенциальной функции иона примеси при экранировании рассеивающего поля: а – соседними ионами примеси (модель Конуэлл-Вайскопфа) и б – подвижными носителями заряда (модель Брукса-Херринга).

В модели Конуэлл-Вайскопфа [2.16] область действия поля иона примеси ограничиваются 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 $U_i = e^2 / 4 \pi \epsilon_0 b_{\max}$ – потенциальная энергия электрона в поле иона на расстоянии b_{\max} .

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

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

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 сомножитель:

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

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 .

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 $N_{\text{пр}}$ 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, $\text{Al}_x\text{Ga}_{1-x}\text{As}$, $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{P}_{1-y}$. Их можно рассматривать как сплавы бинарных соединений $\text{A}^{\text{III}}\text{B}^{\text{V}}$, 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.

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 . Опуская относительно громоздкие процедуры интегрирования [1.8], 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 диспер-

ции, переходя от интегрирования по k -пространству к интегрированию по энергии, получим для времени релаксации импульса

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

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 $p = \sqrt{3m^* k_B T_e}$, и для модели Брукса-Херринга

$$\eta = (2kr_s)^2 = 12\epsilon\epsilon_0 m^* (k_B T_e)^2 / \hbar^2 e^2 N_\Delta. \quad (2.22)$$

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

$$\eta = 6\pi\epsilon\epsilon_0 k_B T_e / e^2 N_\Delta^{1/3}, \quad (2.23)$$

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 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 $E^{1/2}$. 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 $E = 3k_B T_e/2$:

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

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 [1.8]

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

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 :

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

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

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

и для суммарного времени

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

При этом для «теплых» электронов ($T_e \approx T$) $\operatorname{ch}(t - t_e) \approx 1$ и

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

В последних выражениях стоит так называемая **полярная константа**

$$\alpha = \frac{e^2 \sqrt{m^*}}{4\sqrt{2} \pi \epsilon_0 \hbar (\hbar\omega_0)^{1/2}} \left(\frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{ct}} \right). \quad (2.30)$$

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 [1.8], 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_{pij}^\pm с поглощением (верхний знак) и испусканием (нижний знак) междолинного фона на $\hbar\omega_{ij}$ при переходе электрона из i -долины в j -долину:

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

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 фона. Суммирование обратных времен релаксации с излучением и поглощением фононов дает значения обратных времен релаксации концентрации и импульса:

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

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 фона 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 фо-

нонов около 0,03 эВ, а расстояние $\Delta_{\text{GL}} \approx 0,3$ эВ. Более общие выражения приведены в [1.8].

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 $1/\tau_{E_{\text{пр}}} = 0$.

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

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

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

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

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

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 фона преобладает над рассеянием с их погло-

щением, поэтому $1/\tau_{p\phi}^+ > 1/\tau_{p\phi}^-$ и в числителе (2.34) стоит средняя энергия, передаваемая электроном при фононном рассеянии кристаллической решетки 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

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

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 $w_{kk'}^-$. Отличие в $w_{kk'}^\pm$ определяется множителем $(N_\phi + 1/2 \pm 1/2) = e^{\pm t}/2sht$, 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

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

При междолинном рассеянии, когда энергия электрона изменяется не только на величину энергии фона, но и на величину междолинного расстояния $\pm\Delta_{ij}$, аргументом функции Бесселя в (2.31) является параметр $|\delta_{ij}/2 \pm t_e|$ 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 $\text{th}x \approx x$:

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

и для «горячих» ($T_e \gg T$, $t_e \ll t$)

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

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 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. $\omega_\phi = v_{\text{зв}}|k_\phi|$, поэтому энергия таких фононов $\hbar\omega_\phi \leq v_{\text{зв}}(8m^*E)^{1/2}$. 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

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

Таким образом, в отличие от оптических механизмов упругость акустического рассеяния не зависит от электронной температуры. Подстановка типичных значений ($m^* \approx 0,1m$, $v_{\text{зв}} \approx 10^3$ 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 $\tau_{E0} = 2,5\tau_p$, 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, *a*) происходит постепенный переход из равновесного состояния ($v(t_0^+) = 0$) в стационарное состояние $v(\infty) = \mu\mathcal{E}$ по закону

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

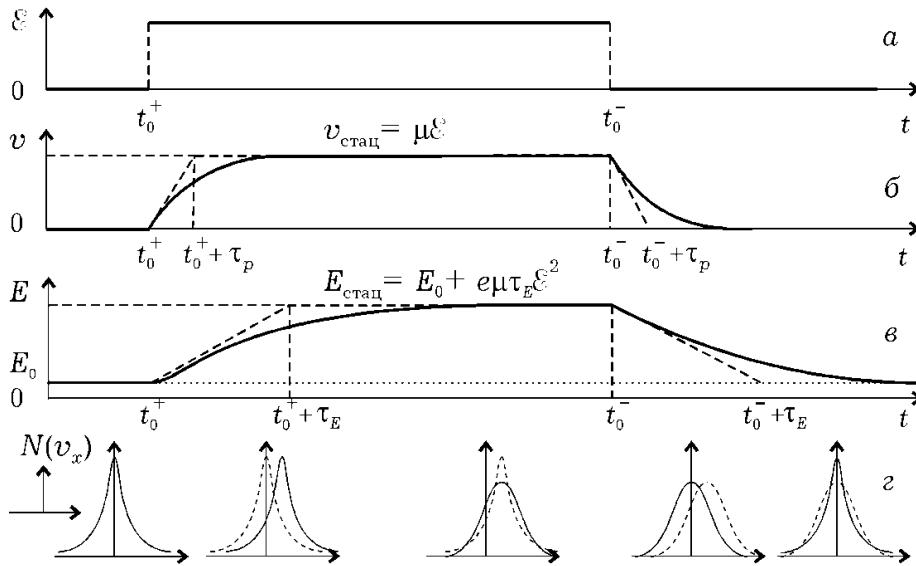


Рис. 2.22. Временные диаграммы переходных процессов, происходящих при включении и выключении электрического поля (*а*). Установление стационарных состояний при включении и релаксация при выключении для скорости (*б*) и энергии (*в*) происходят с существенно разными скоростями. В нижнем ряду (*г*) символически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

После выключения поля ($\vec{\mathcal{E}} = 0$) 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. 2.22, в) состоит в том, что длительность переходных процессов определяется в основном временем τ_E , большим, чем τ_p . При выключении поля релаксация энергии до равновесного значения E_0 происходит по экспоненциальному закону (2.43) с характерной временной константой τ_E . 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 $v = \mu\mathcal{E}$. 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 $v_{\text{зв}} = v_l = 9,0 \cdot 10^3$ 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: $m_d^* = 6^{2/3}(m_t^*)^2 m_l^* = 1,08m$. В (2.39) для расчета времени релак-

сации 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).

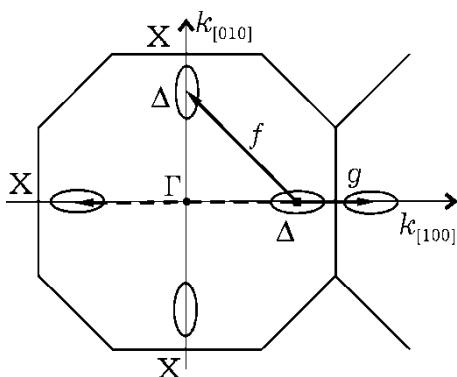


Рис. 2.23. Виды переходов при оптическом рассеянии в кремнии: f —переходы между перпендикулярными долинами, g —переходы между соосными долинами (штриховая линия), которые удобнее рассматривать, как переходы в соседнюю зону Бриллюэна.

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 $m_{d1}^* = (m_t^{*2} m_l^*)^{1/3}$, 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 дальнодействие 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 температуре $T = 300$ К, радиус экранирования $r_s = 0,13$ мкм, а максимальное прицельное расстояние $b_m = 0,05$ 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

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

В области низких температур надо также принимать во внимание «вымораживание» примеси.

Результаты расчетов времен релаксации, поясняющие вклад различных механизмов рассеяния в результатирующее 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.

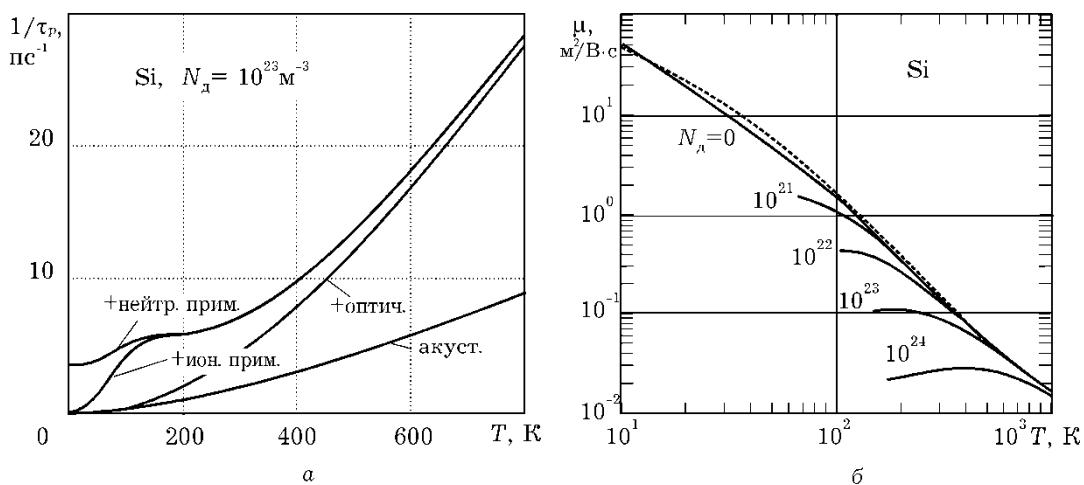


Рис. 2.24. Температурные зависимости обратных времен релаксации импульса при фиксированной концентрации примеси (а) и подвижности при различных концентрациях примеси (б) для кремния. Штриховая линия — экспериментальная. Графики обратных времен релаксации для каждого вида суммируются с предыдущими

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].

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

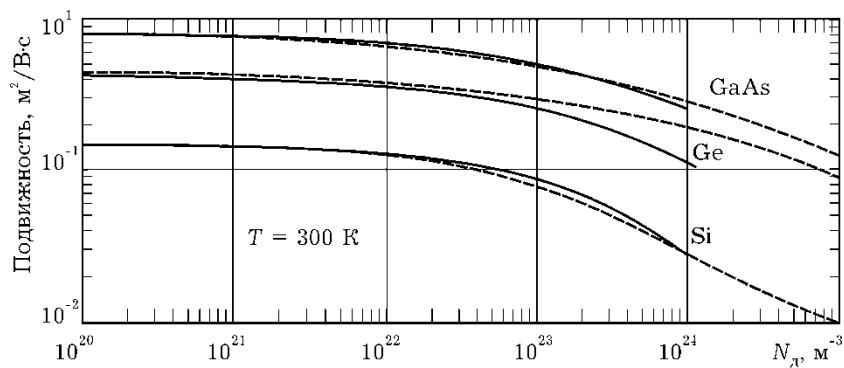


Рис. 2.25. Концентрационные зависимости подвижности для кремния, германия и арсенида галлия при комнатной температуре. Штриховые линии – экспериментальные [2.14]

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 $v_{\text{зв}} = 5,4 \cdot 10^3$ м/с, плотности $\rho = 5,3$ г/см³, диэлектрической проницаемости $\epsilon = 16,0$ и эффективной массы плотности состояний $m_d^* = 0,56m_e$.

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 Л-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 оптиче-

ском деформационном потенциале при $D_0 = 5,5 \cdot 10^8$ эВ/см и участии оптических фононов с энергией $\hbar\omega_0 = 0,034$ эВ. Для анализа примесного рассеяния здесь, как и в кремнии, предпочтительнее модель Конуэлл-Вайскопфа.

Вклад различных механизмов рассеяния представлен 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

. При возрастании температуры возрастает роль фононных 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 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.

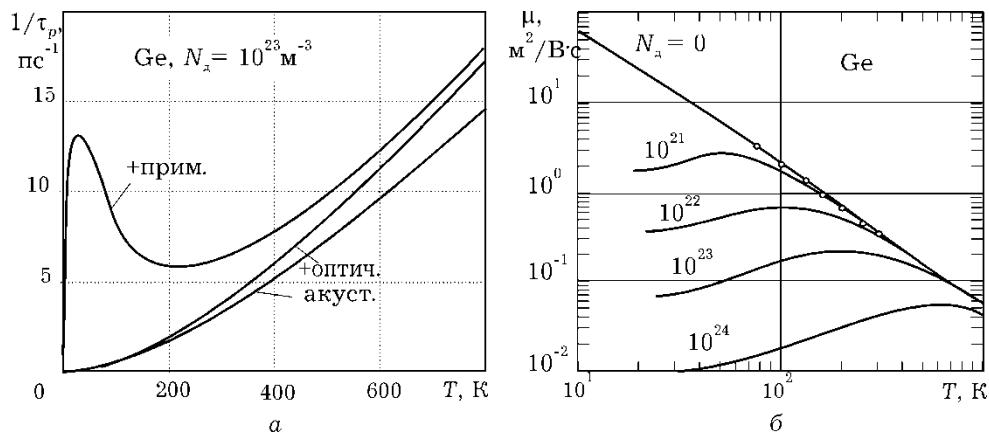


Рис. 2.26. Такие же зависимости, как и на рис. 2.24 для германия. Экспериментальные [2.9] значения для $N_d = 0$ обозначены точками

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 фононных 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 ϵ_{ct} 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. 2.27) приведены результаты для значения $\omega_{LO}/\omega_{TO} = (\epsilon_{ct}/\epsilon_\infty)^{1/2} = 1,15$. Этот вид рассеяния проявляется при температурах выше 100 К и при комнатной температуре является преобладающим (при отсутствии примесного рассеяния).

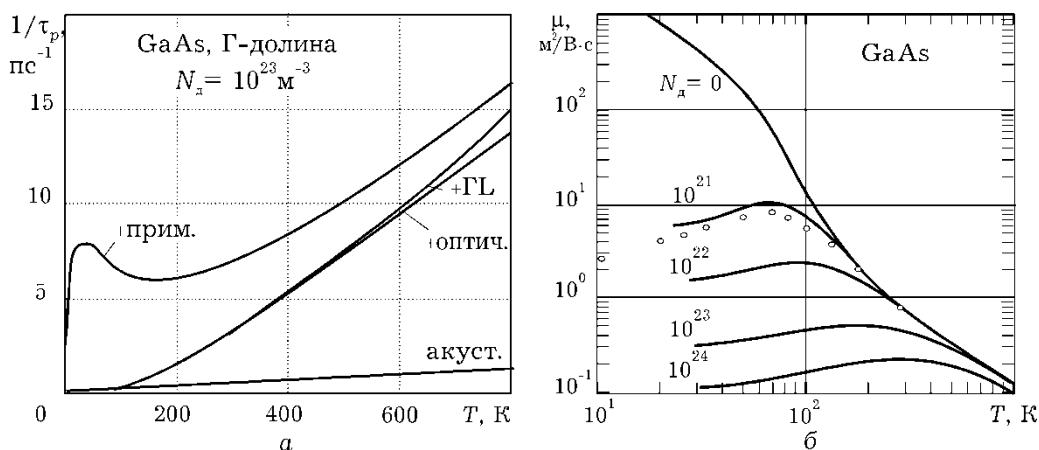


Рис. 2.27. Такие же зависимости, как и на рис. 2.24 для арсенида галлия. Точками обозначены экспериментальные значения для $N_d = 10^{21} \text{ м}^{-3}$

Второе существенное отличие связано с особенностью закона дисперсии GaAs и со структурой его долин [2.9]. Кроме нижней Г-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 из Г-долины в состояния L-долин. 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 мэВ, 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 использована модель Конуелл-Вайскопфа. При самых низких температурах преобладает рассеивание 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 $m(T)$ в логарифмических координатах (рис. 2.28, б). Там же приведены результаты измерений подвижности, взятые из [2.9].

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 рассеиваний. Для акустического рассеивания были использованы значения деформационного потенциала $D_a = 10,1$ эВ, скорости звука $v_{\text{зв}} = 4,57 \cdot 10^3$ 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 прони-

цаемости 9,6 и 5,5 соответственно, энергии полярных оптических фононов 65 мэВ.

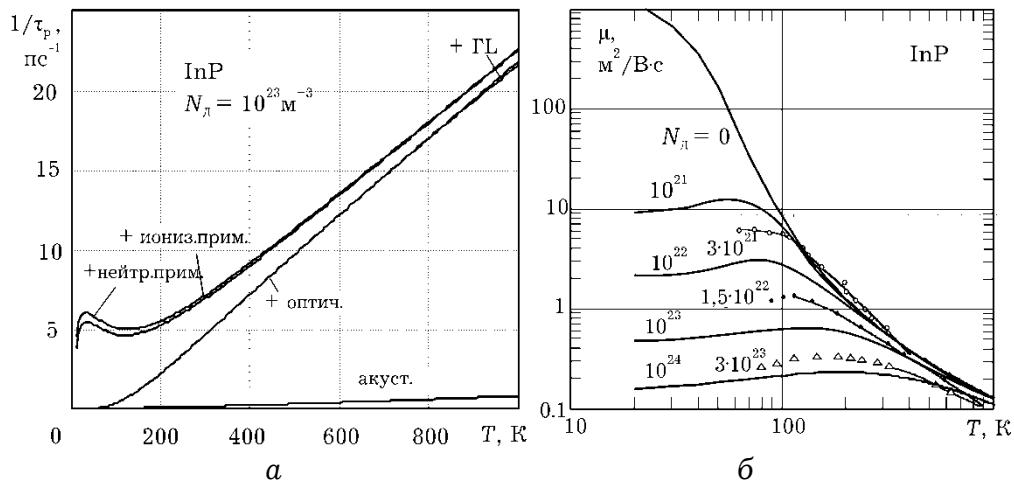


Рис. 2.28. Температурная зависимость обратных времен релаксации импульса при концентрации примеси $N_d = 10^{23} \text{ м}^{-3}$ (а) и подвижности при разных концентрациях примеси (б) для фосфида индия. Графики обратных времен релаксации для каждого вида суммируются с предыдущими

Вклад междолинных механизмов с энергией междолинных фононов около 59 мэВ и константой междолинной связи $1,1 \cdot 10^9 \text{ эВ/см}$ 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.

Значения параметров материалов выбирались в основном из книг [2.15, 2.16].

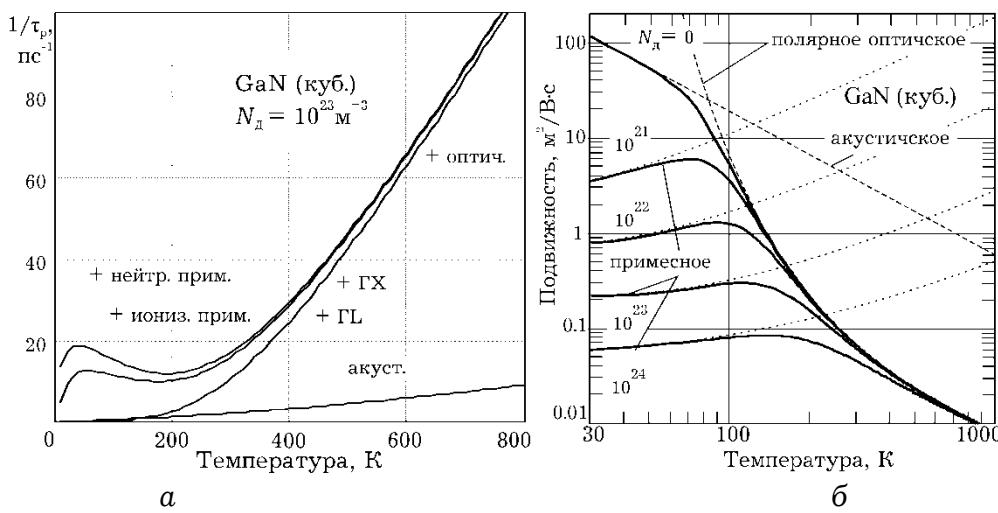


Рис. 2.29. Температурная зависимость обратных времен релаксации импульса при концентрации примеси $N_{\text{д}} = 10^{23} \text{ м}^{-3}$ (а) и подвижности при разных концентрациях примеси (б) для нитрида галлия. Кривые для каждого вида рассеивания суммируются с предыдущими

Литература к разделу 2

- 2.1. *Москалюк В. А., Синекоп 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. Ahuja, B. Johansson // 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** при сохранении в зоне проводимости концентрации электронов ($dn/dt = 0$), используя уравнения (1.43) и (1.47) в форме:

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

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

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

$(E - E_0)/\tau_E$ – теряющую 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

$$dE_{\text{д}}/dt - e\vec{v}\vec{\mathcal{E}} = -2E_{\text{д}}/\tau_p. \quad (3.4)$$

Здесь правая часть представляет собой энергию дрейфа, теряющую 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:

$$\frac{dE_{\text{т}}}{dt} = -\frac{E_{\text{т}} - E_0}{\tau_E} - \frac{E_{\text{д}}}{\tau_E} + \frac{2E_{\text{д}}}{\tau_p} = -\frac{E_{\text{т}} - E_0}{\tau_E} + \frac{E_{\text{д}}}{\tau_E} \left(\frac{2\tau_E}{\tau_p} - 1 \right). \quad (3.5)$$

Слагаемое $2E_{\text{д}}/\tau_p$ равно энергии дрейфа, превращающейся за 1 с непосредственно в тепловую, но ее часть, равная $E_{\text{д}}/\tau_E$, за это же время релаксирует, переходя к кристаллической решетке. В итоге увеличение тепловой энергии за счет дрейфовой за одну секунду составляет $E_{\text{д}}(2\tau_E/\tau_p - 1)/\tau_E$. 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 уравновешивают друг друга. Это наступает, когда электронная температура устанавливается на уровне, соответствующем равенству

$$E_{\text{т}} = E_0 + E_{\text{д}}(\gamma - 1). \quad (3.6)$$

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, расчитанные для кремния при 300 К по формулам, приведенным в разд. 2.3, при тех же значениях параметров, которые использовались при расчете слабополевой подвижности (разд. 2.4). Здесь следует обратить внимание на то, что преобладающим является оптическое рассеяние, акустическое рассеяние из-за большой упругости практически не влияет на время релаксации энергии, а примесное заметно только при очень слабом разогреве ($T \approx T_e$).

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).

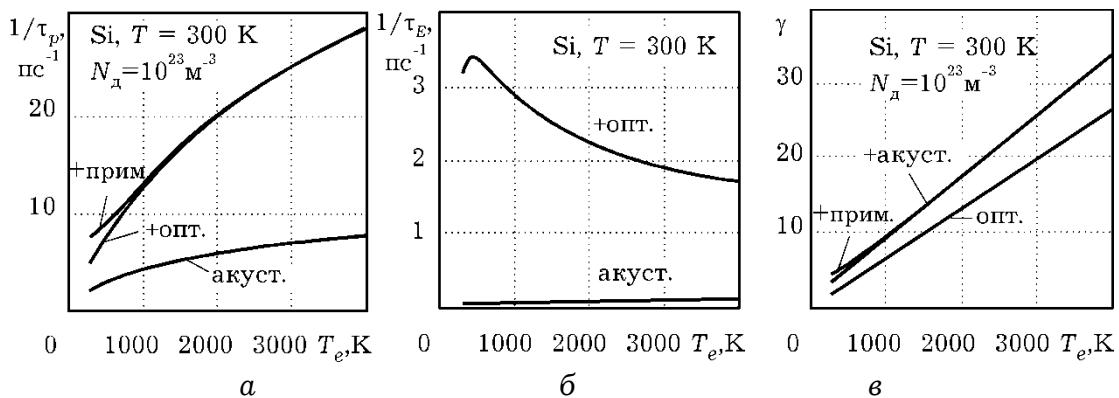


Рис. 3.1. Зависимости обратных времен релаксации импульса (а), энергии (б) и коэффициента неупругости (в) от электронной температуры в кремнии. Кривые для каждого вида рассеяния суммируются с предыдущими.

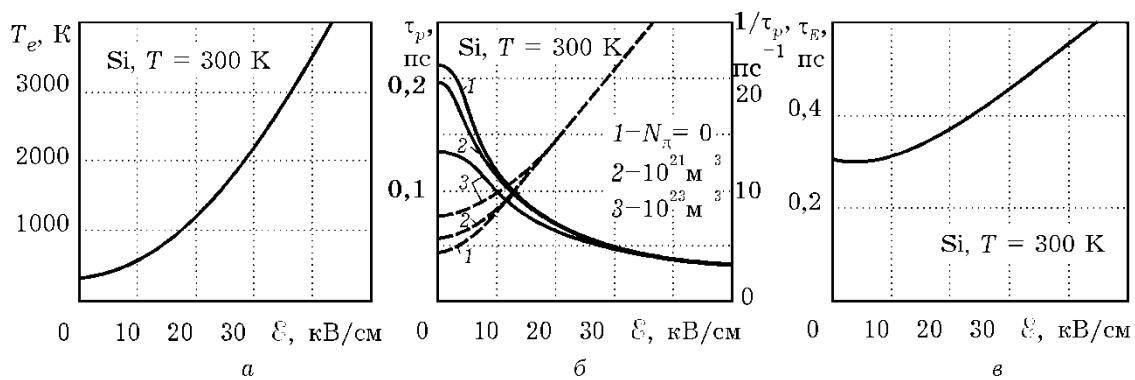


Рис. 3.2. Зависимость электронной температуры (а), времен релаксации импульса (б) и энергии (в) от напряженности электрического поля для кремния. Штриховые линии (б) характеризуют влияние концентрации примеси на обратное время релаксации импульса.

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 $1/\tau_p \sim E$, а τ_E – возрастает, так что неравенство $\tau_E > \tau_p$ усиливается, что означает увеличение степени неупругости рассеяния.

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 осо-

бенно заметно увеличивает время релаксации энергии (рис. 3.3, б, в).

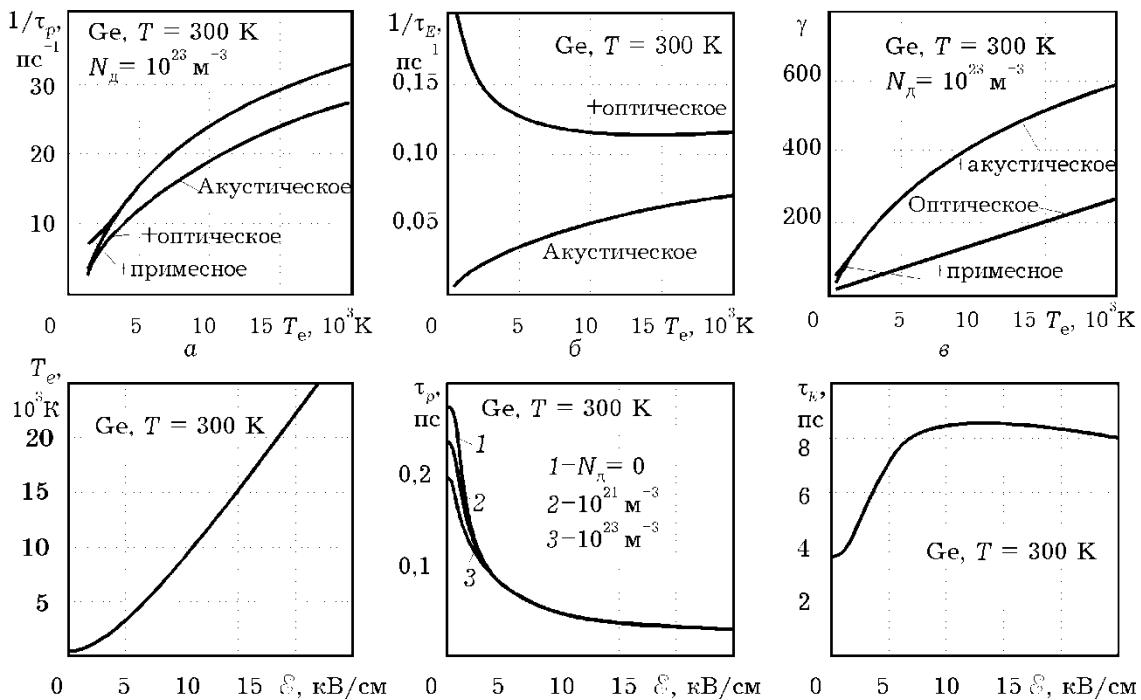


Рис. 3.3. Те же зависимости, что и на рис. 3.1 и 3.2, для германия. Следует обратить внимание на существенно больший вес акустического рассеяния.

Это объясняется прежде всего большой 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.

В многодолинных полупроводниках типа GaAs, кроме Г-долины у дна зоны проводимости, имеются и более высокие L- и X-долины. В каждой из них механизм разогрева в прин-

ципе 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 ($i, j = \Gamma, L, X$):

$b_i(T_e) = n_i/n = n_i/(n_i + n_j) = 1/(1 + n_j/n_i) = 1/(1 + \tau_{nji}/\tau_{nij})$, (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 ГЛ-модель), получим для ее значения в стационарном состоянии

$$E_t = E_0 + b_\Gamma E_{d\Gamma} (\gamma_\Gamma - 1) + b_L E_{dL} (\gamma_L - 1). \quad (3.10)$$

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 Г- и L-долинах в зависимости от электронной температуры при 300 К представлены на рис. 3.4.

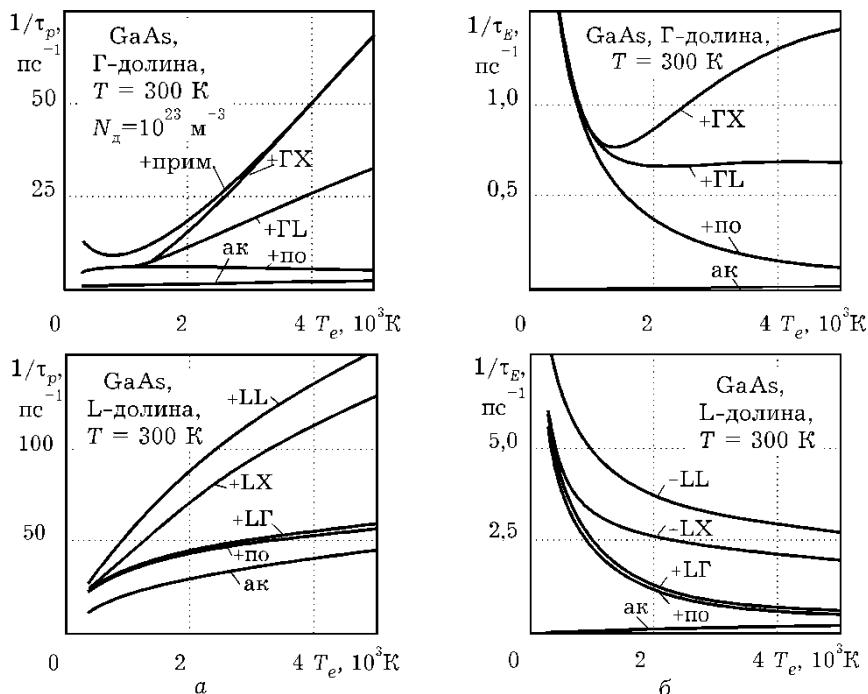


Рис. 3.4. Зависимости обратных времен релаксации импульса (а) и энергии (б) от электронной температуры в Г- (верхний ряд) и L-долинах (нижний ряд). Кривые для каждого вида рассеяния суммируются с предыдущими.

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 Г-долины, причем его роль ослабевает при повышении электронной температуры, а акустическое влияет практически только на время релаксации импульса. Междолинное рассеяние становится заметным в Г-долине лишь при температурах, выше ~1000 К, а в L-долине можно пренебречь рассеянием в Г-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 опреде-

лиять 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 Г-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.

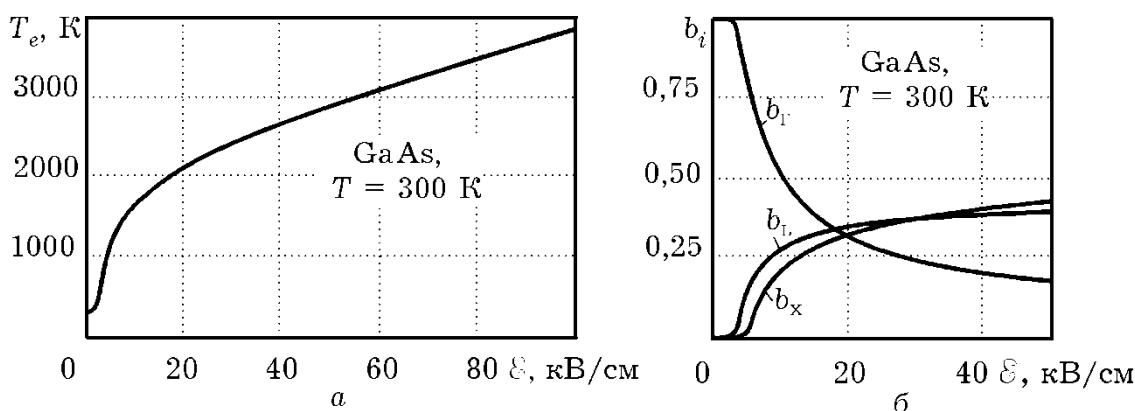


Рис. 3.5. Зависимости электронной температуры (а) и заселенности долин (б) от напряженности электрического поля в арсениде галлия для комнатной температуры.

Фосфид индия. Зависимости времен релаксации для фосфида индия рассчитаны при температуре 300 К для тех же значений параметров, которые были использованы для расчета слабополевой 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 Г-долинах, но существенны для L-долин. Междолинное рассеивание становится заметным лишь при температурах выше ~1200 К. Так же как и для арсенида галлия рассеиванием из L-долины в Г-долину можно пренебречь из-за низкой плотности состояний в ней. При этом достаточно весомым является вклад эквивалентного междолинного LL-рассеивания. Рассеивание на ионах примеси заметно только в Г-долине при низких темпе-

ратурах и с увеличением электронной температуры его вклад уменьшается. Результаты расчетов приведены на рис. 3.6.

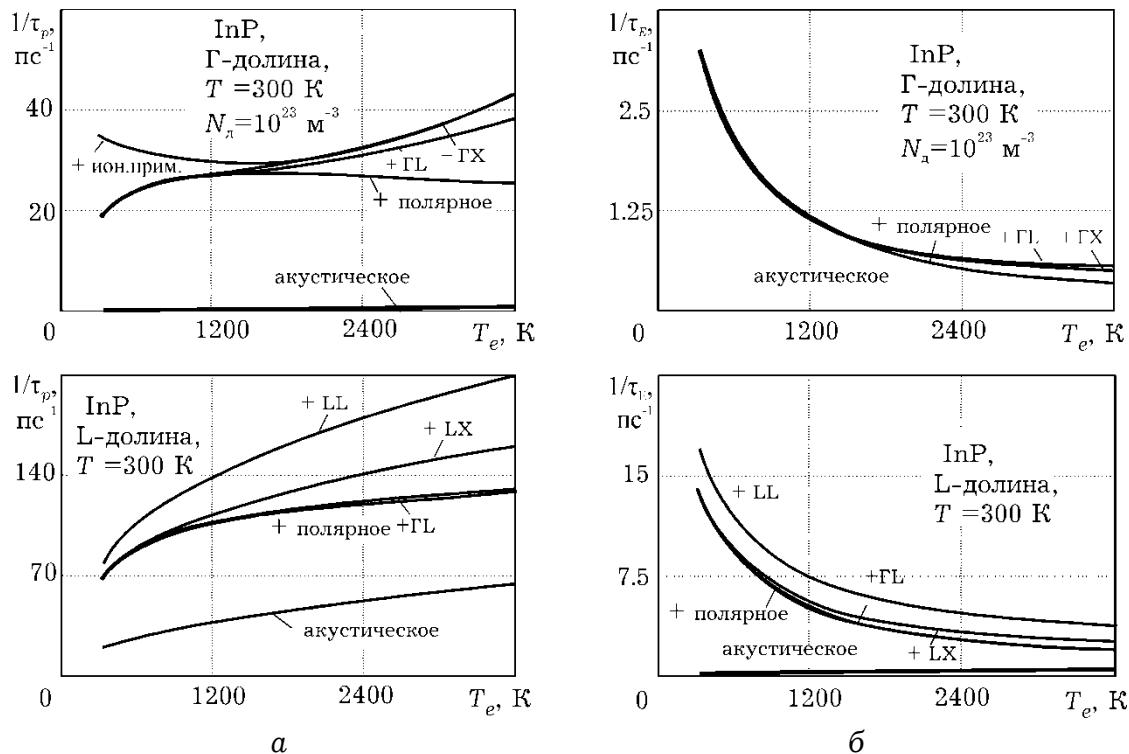


Рис. 3.6 . Зависимость обратных времен релаксации импульса (а) и энергии (б) от электронной температуры в Г- (верхний ряд) и L- (нижний ряд) долинах для фосфида индия. Кривые для каждого вида рассеяния суммируются с предыдущими

Для поле-температурной характеристики фосфида индия характерна аналогичная картина (рис. 3.7). Здесь также можно выделить три участка. Область слабых полей менее 10 кВ/см, в которой, как и в арсениде галлия, разогрев электронного газа, определяется исключительно параметрами Г-долин. Переходная область, где происходит значительное перераспределение электронов по долинам, занимает участок приблизительно до 20 кВ/см. Область очень сильных полей (>20 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).

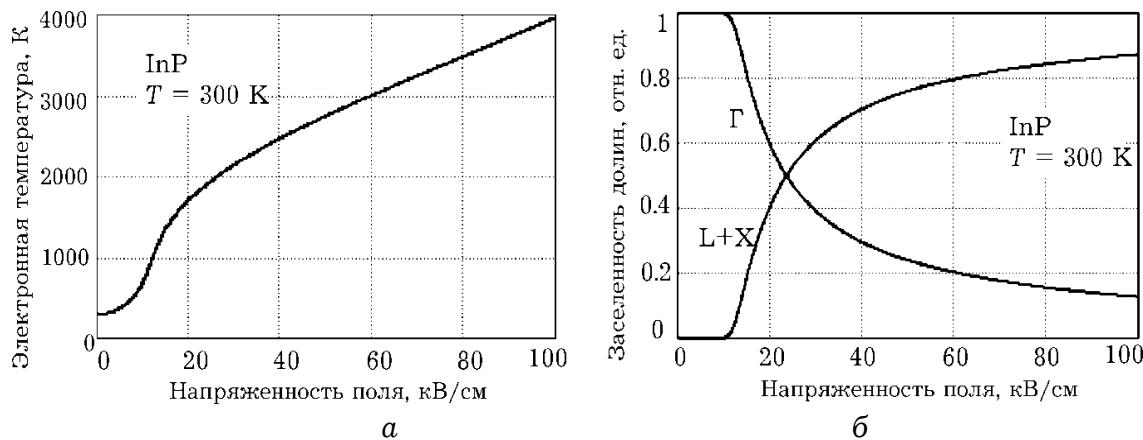


Рис. 3.7. Поле-температурная зависимость (а) и зависимость заселенности долин от напряженности электрического поля (б) в InP для комнатной температуры

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 заметно в Г-долине при повышении температуры до 2500 К, что видно из рис. 3.8:

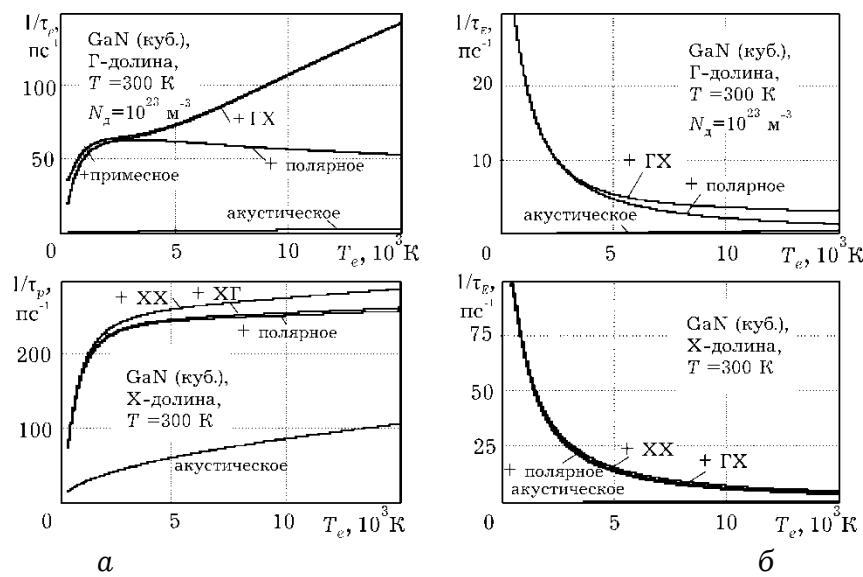


Рис. 3.8. Зависимость обратных времен релаксации импульса (а) и энергии (б) от электронной температуры в Г- (верхний ряд) и X- (нижний ряд) долинах для нитрида галлия.

Для поле-температурной характеристики нитрида галлия 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 10^4 К (рис. 3.9).

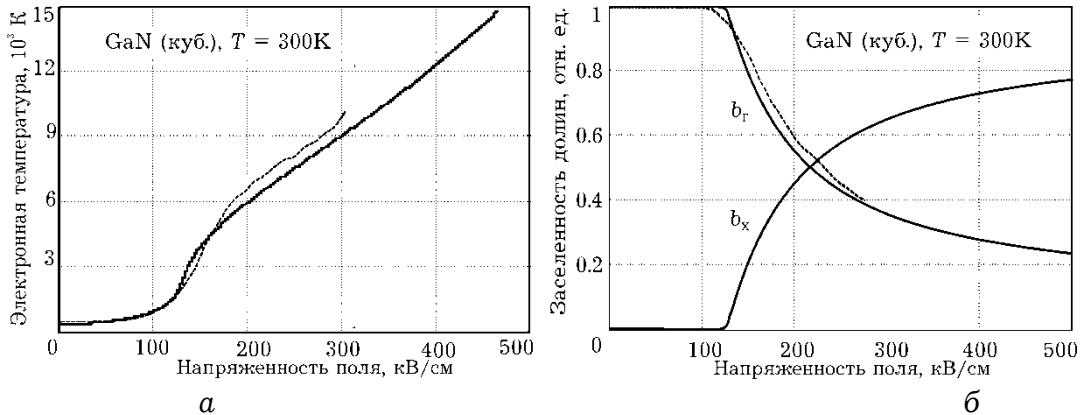


Рис. 3.9. Поле-температурная зависимость (а) и зависимость заселенности долин от напряженности электрического поля (б) в GaN для комнатной температуры. Штриховые линии взяты из [3.3, 3.4]

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)

$$\tau_E = \tau_p / 2 + 3k_B(T_e - T) / 2e\mu\mathcal{E}^2. \quad (3.16)$$

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.

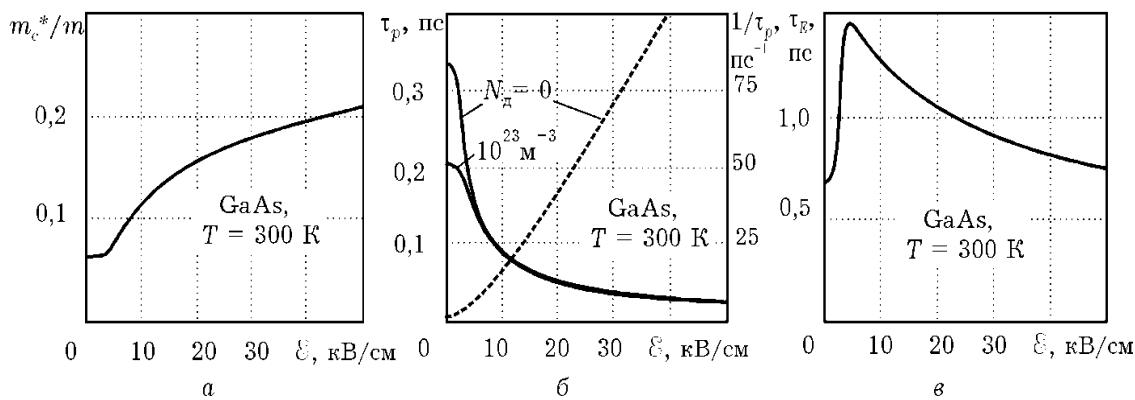


Рис. 3.10. Зависимости усредненных по долинам приведенной эффективной массы (а), времен релаксации импульса (б) и энергии (в) от напряженности электрического поля для арсенида галлия. Штриховая линия (б) – для обратного времени релаксации импульса при $N_d = 0$.

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 $v_{\text{др}} \ll v_{\text{хаот}}$ и средняя скорость определяется только хаотической составляющей. Поэтому подвижность остается постоянной, а вольтамперная характеристика – линейной. Но по мере того, как с ростом 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 $E > 10...20$ 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 ускоряемый 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 $-eE\tau_p$. Соответственно уравнению баланса энергии (3.1) для дрейфового движения приобретаемая от поля за это же время энергия дрейфа $ev_n E \tau_p$ 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

$$v = v_{\tau} \sqrt{\frac{T_e/T - 1}{\gamma - 1}}. \quad (3.18)$$

Это выражение позволяет, рассчитывая времена релаксации, с учетом ранее полученной связи между электронной температурой и напряженностью электрического поля получить так называемую полескоростную 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 остается заметной по сравнению с оптическим. К тому же соотношение между ними остается 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

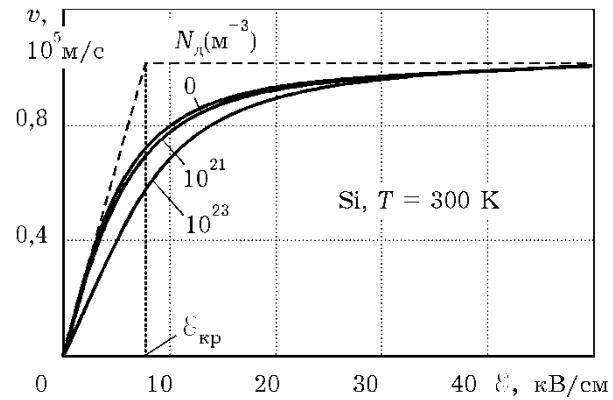


Рис. 3.11. Зависимость дрейфовой скорости от напряженности электрического поля в кремнии при различных концентрациях примеси и ее кусочно-линейная аппроксимация (штриховая линия).

$$\frac{1}{\tau_{po}} \approx \frac{m^{*3/2} D_o^2}{\pi^{3/2} \rho \hbar^2} \cdot \frac{\sqrt{2k_B T_e}}{\hbar \omega_o \operatorname{sh} t} \quad (3.19)$$

Так как время релаксации импульса для акустического рассеяния (2.24) также пропорционально $\sqrt{T_e}$, то их отношение, равное

$$\frac{\tau_{po}}{\tau_{pa}} = \sqrt{\frac{3\pi}{2}} \frac{D_a^2 \hbar \omega_o \operatorname{sh} t}{D_o^2 (\hbar v_{3B})^2} k_B T, \quad (3.20)$$

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 счет 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):

$$\gamma_o = \frac{2\tau_{Eo}}{\tau_{po}} = 6 \frac{T_e}{T} \left(\frac{k_B T}{\hbar \omega_o} \right)^2 = \frac{3}{2t t_e} = \frac{3}{2t^2} \frac{T_e}{T}. \quad (3.23)$$

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 300 К. Это позволяет при больших электронных температурах пренебречь единицей в знаменателе и числитеle (3.18) и получить для скорости насыщения простое выражение

$$v_{\text{н}} = \frac{v_{\text{T}}}{\sqrt{\beta}} = \frac{v_{\text{T}}}{\sqrt{6(1 + \tau_{po}/\tau_{pa})}} \frac{\hbar\omega_0}{k_B T}. \quad (3.25)$$

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 $v_{\text{н}}$ 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:

$$v = v_{\text{н}} \sqrt{\frac{T_e/T - 1}{T_e/T - 1/\beta}}. \quad (3.26)$$

При известных значениях $v_{\text{н}}$ и слабополевой подвижности μ_0 иногда используют кусочно-линейную аппроксимацию (см. рис. 3.11) поле-скоростной характеристики $v(\mathcal{E})$

$$v = \mu_0 \mathcal{E} \text{ при } \mathcal{E} < \mathcal{E}_{\text{кр}}, \quad v = v_{\text{н}} \text{ при } \mathcal{E} > \mathcal{E}_{\text{кр}}, \quad (3.27)$$

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 точно 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 $\mu(N_d)$ для данной концентрации примеси, но \mathcal{E}_{kp} 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

$$\gamma = \gamma_o \frac{1 + \tau_{po} / \tau_{pa}}{1 + \tau_{Eo} / \tau_{Ea}}. \quad (3.29)$$

Коэффициент неупругости оптического рассеяния γ_o , пропорционален T_e и при характерных для германия параметрах равен $\gamma_o \approx 3,7 T_e/T$. 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 $\tau_{Eo}/\tau_{Ea} \approx 0,014 T_e/T$. 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, г).

На рисунках также приведены результаты экспериментов, взятые из книги [3.6], где описаны данные, полученные из различных первоисточников.

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.

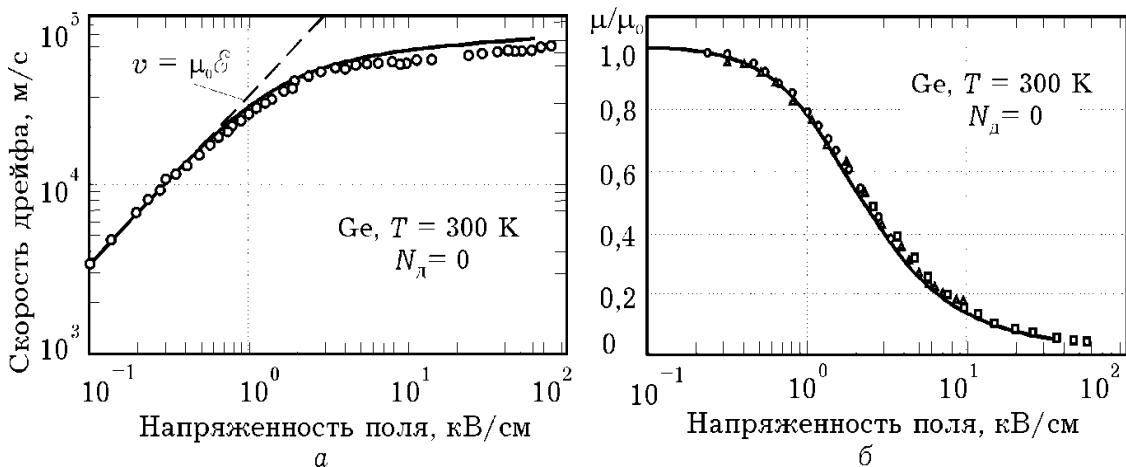


Рис. 3.12. Зависимости дрейфовой скорости (а) и подвижности (б), нормированной на ее слабополевое значение, от напряженности электрического поля в германии при комнатной температуре. Точками обозначены экспериментальные значения различных авторов

В центральной Г-долине при относительно небольшом разогреве ($T_e \approx 1000$ К) из фононных механизмов преобладает полярное оптическое рассеяние (рис. 3.4, а). Акустическое скавивается только на рассеянии импульса, его доля определяется отношением $\tau_{\text{рпо}}/\tau_{\text{pa}}$, 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 $1/\tau_{\text{рпо}}$ падает с ростом T_e и скорость уменьшается, что частично проявляется на верхней кривой рис. 3.13, а. 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

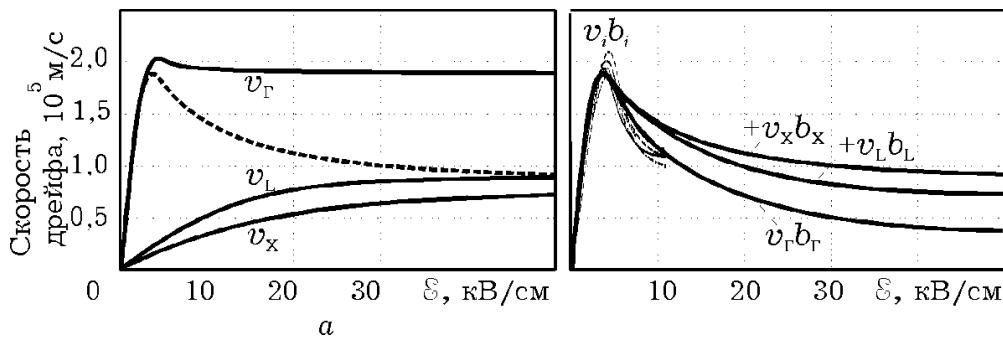


Рис. 3.13. Зависимости от напряженности электрического поля скоростей (а) дрейфа электронов для каждой долины и их усредненная зависимость (штриховая линия). Та же зависимость (б), построенная суммированием скоростей в отдельных долинах с весовыми коэффициентами, равными их заселенности. Штриховые линии (б) от 0 до 10 кВ/см взяты из [3.7, 3.8]

$$\gamma = \gamma_{ij} \frac{1 + \tau_{pij} / \tau_{\text{рпо}}}{1 + \tau_{Eij} / \tau_{\text{Е по}}} \quad (3.30)$$

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 Г-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 Г-долине равна $v_{\text{тГ}} = \sqrt{3k_B T / m_\Gamma^*}$, что при комнатной температуре составляет $4,6 \cdot 10^5$ м/с, то скорость дрейфа достигает значения насыщения $v_{\text{нГ}} \approx 1,8 \cdot 10^5$ м/с.

В итоге поле-скоростная характеристика для Г-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 $v_{\text{нГ}}$. Пиковое значение скорость имеет при значении напряженности электрического поля, не-

сколько 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 $v_{nL} \approx 1,1 \cdot 10^5$ м/с. Соответствующая поле-скоростная характеристика для электронов L-долины представлена на рис. 3.13, 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 $v_{nX} \approx 0,7 \cdot 10^5$ 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 заселеностям (3.9), которые также можно представить функциями напряженности поля, используя предварительно найденную зависимость (3.12) между T_e и 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 [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.

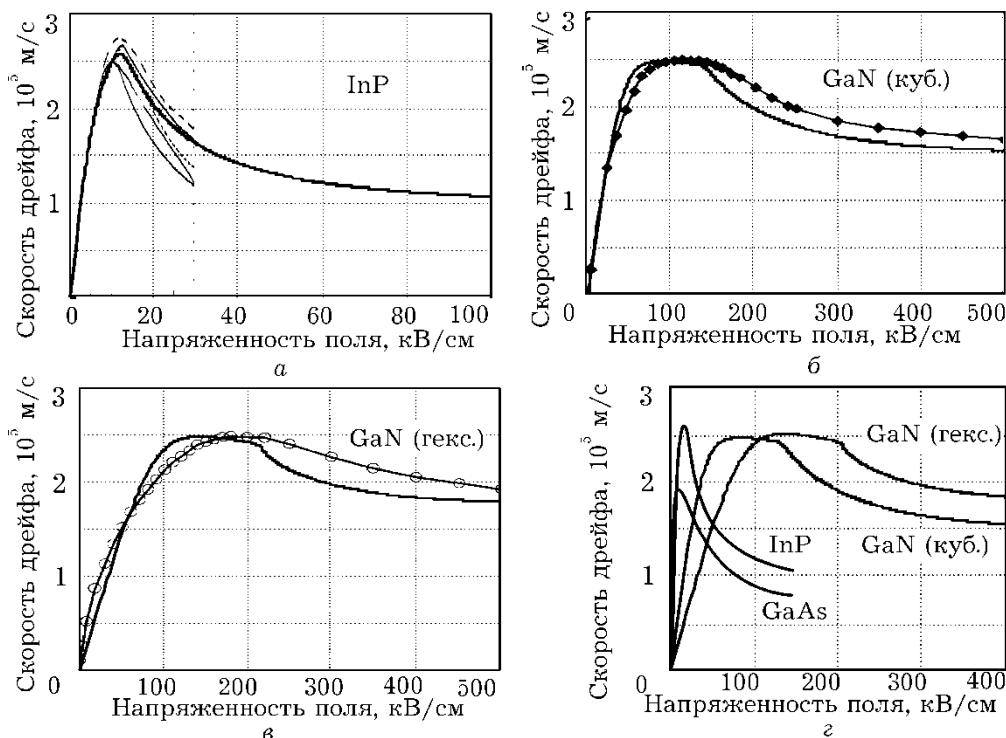


Рис. 3.14. Поле-скоростные характеристики фосфида индия (а), нитрида галлия, с кубической (б) и гексагональной (в) структурой. Для удобства сравнения на (г) эти характеристики приведены в одинаковом масштабе

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 -долини в верхнюю Γ_2 -долину (1,9 эВ), но и в еще более высокие LM-долини (2,1 эВ). Хотя эффективные массы в них почти одинаковы ($m^* \approx m$), за счет большего количества последних будет преобладать Γ_1 -LM 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].

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: $j = \sigma E$. Соответственно определим динамическую (дифференциальную) проводимость как коэффициент пропорциональности между изменениями dj и dE : $\sigma_d = dj/dE$. Так как плотность тока $j = env$, то при неизменной концентрации $dj = = en \cdot dv$. В слабых электрических полях скорость дрейфа пропорциональна напряженности поля $v = \mu E = (e\tau_p/m^*)E$, а так как подвижность постоянна, то $dv = \mu dE$ 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 $(e/m^*)dE = = (\tau_p^{-1}) \cdot dv + v d(\tau_p^{-1})$. В режиме насыщения $dv = 0$ и $d(\tau_p^{-1}) = (e/m^* v_H) dE$. Если эту обратную связь между подвижностью и напряженностью поля экстраполировать на переходную область, то можно найти, что $dv = \mu(1 - v/v_H)dE$ и

$$\sigma_d = \sigma(1 - v/v_H). \quad (3.31)$$

В слабых полях $v \ll v_H$ 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.

Характерной особенностью арсенида галлия и других многодолинных полупроводников является 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 $v(\mathcal{E})$.

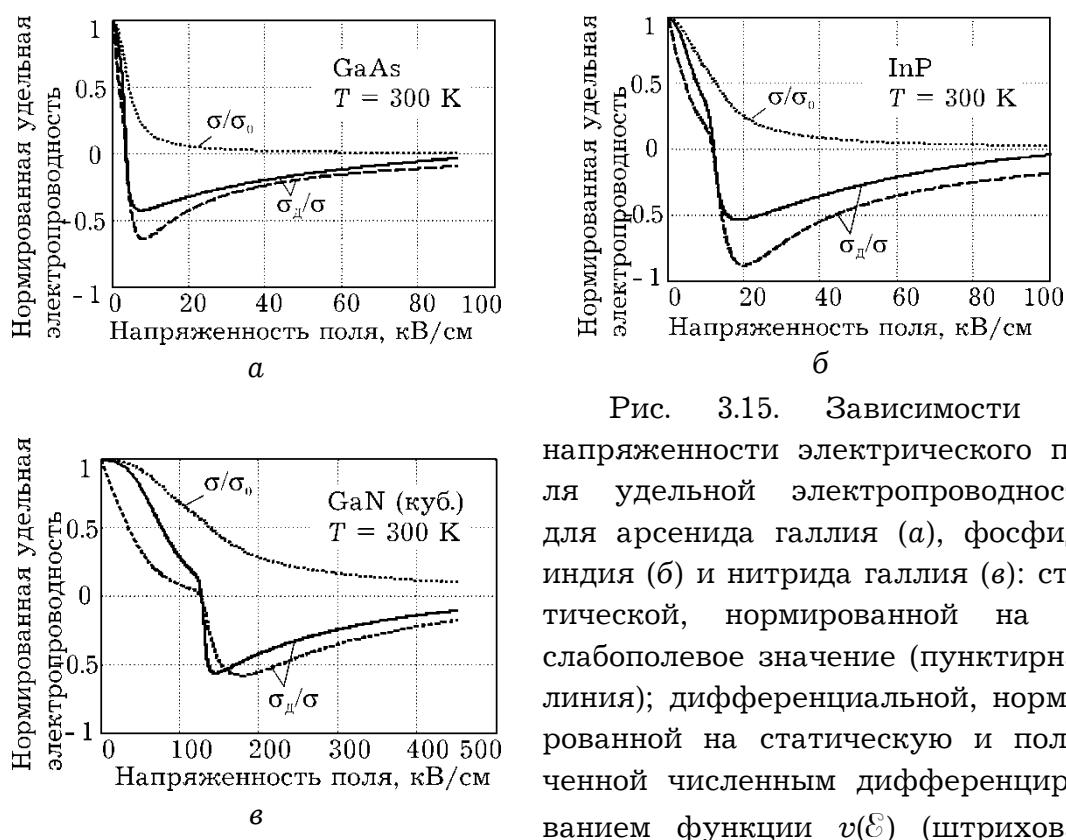


Рис. 3.15. Зависимости от напряженности электрического поля удельной электропроводности для арсенида галлия (а), фосфида индия (б) и нитрида галлия (в): статической, нормированной на ее слабополевое значение (пунктирная линия); дифференциальной, нормированной на статическую и полученной численным дифференцированием функции $v(\mathcal{E})$ (штриховая линия) и рассчитанную по (3.37) (сплошная линия)

Для анализа основных факторов, определяющих ее величину, примем во внимание, что ток создается различными группами электронов, имеющими в разных долинах свои характерные параметры, то есть $j = \sum_i e n_i v_i = e n \sum_i b_i v_i$, и поэтому

$\sigma_d = \sum_i \sigma_{di}$, где суммирование ведется по разным 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).

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 Г-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 $d\mathcal{E} = v_h d(\mu^{-1})$. Отсюда $d\mathcal{E}/dT_e = v_h d(\mu^{-1})/dT_e$, а так как в сильном поле при междолинном рассеянии можно считать, что $1/\tau_p \sim T_e$ (рис. 3.6, б), то

$$d\mathcal{E}/dT_e \approx v_h / \mu T_e. \quad (3.35)$$

В итоге для расчета дифференциальной проводимости, обусловленной электронами только Г-долины, можно получить

$$\sigma_{d\Gamma} = en\mu_\Gamma b_\Gamma \left[1 - \frac{v_\Gamma}{v_{h\Gamma}} - \frac{v_\Gamma}{v_{h\Gamma}} (1 - b_\Gamma) \frac{\Delta_{\Gamma L}}{k_B T_e} \right]. \quad (3.36)$$

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

$$\sigma_d \approx \sigma (1 - v_\Gamma/v_{h\Gamma} - b_\Gamma b_L \Delta_{LG}/k_B T_e), \quad (3.37)$$

в котором вклад электронов верхних долин учтен в суммарной статической электропроводности σ , а в последнее слагае-

мое 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 $v(E)$. Наиболее характерной особенностью данных результатов является наличие области отрицательной дифференциальной проводимости, которая проявляется практически сразу с началом междолинных переходов ($E \approx 4$ кВ/см) и имеет максимальное значение при $E \approx 5\ldots10$ кВ/см, когда концентрации электронов в верхней и нижней долинах приблизительно одинаковы ($b_r b_L \approx 0,25$). 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_- : $a(t) = a_0 + a_- e^{i\omega t}$. 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 $j = ev$ дает для его высокочастотной составляющей соотношение

$$j_- = ev_0 v_- + ev_0 n_- \quad (3.38)$$

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_- и n_- как пропорциональные переменной амплитуде электрического поля \mathcal{E}_- , получим два характерных слагаемых для проводимости. Отсылая читателя к [1.8], здесь приведем лишь окончательный 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 электропроводимости, 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

$$\tau_m = (1/\tau_{i0} + 1/\tau_{j0})^{-1}, \quad (3.41)$$

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). Естественно также, что оно пропорционально энергетическому расстоянию Δ_{ij} между долинами и предопределяет отрицательные значения электропроводности.

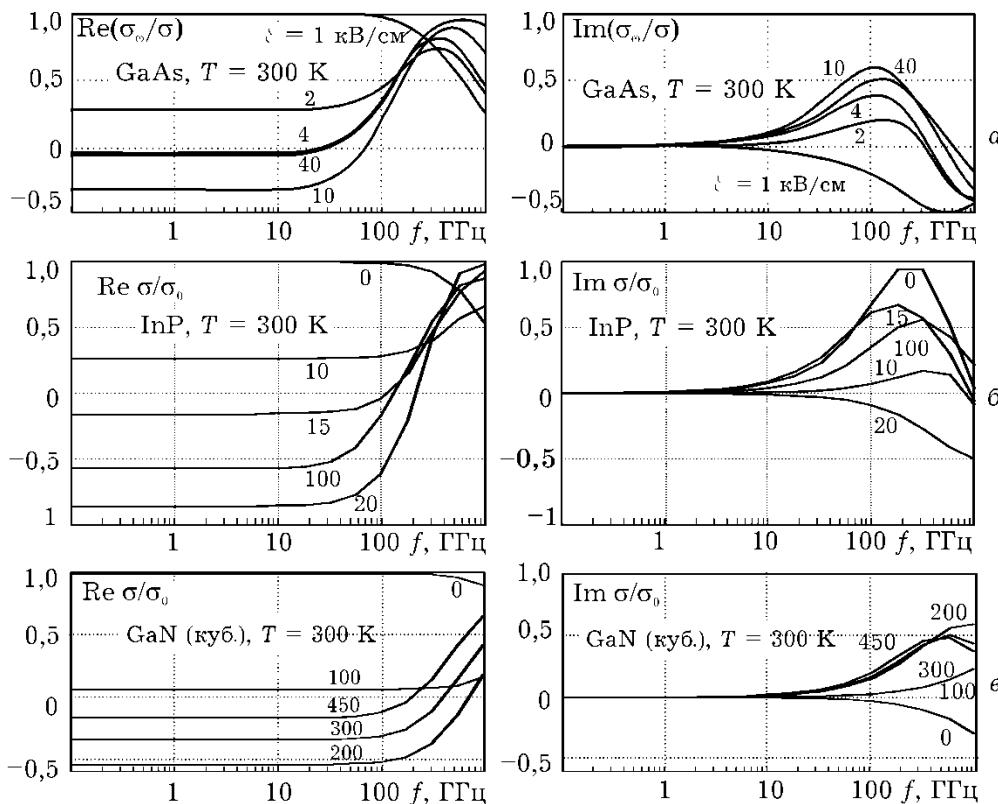


Рис. 3.16. Частотные зависимости действительной и мнимой частей динамической удельной электропроводности, нормированной на статическое значение для арсенида галлия (а), фосфида инидия (б) и нитрида галлия (в)

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 $(1 + i\omega t)^{-1} = (1 - i\omega t)/(1 + \omega^2 t^2)$: реактивная проводимость в отличие от активной пропорциональна первой степени ωt . 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_- 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 $j_{cm} = i\omega\epsilon_0\mathcal{E}$ приводит к появлению в комплексной проводимости σ_ω только мнимой составляющей $i\omega\epsilon_0$, а для нормированной величины σ_ω/σ ее вклад равен $\Delta\sigma_\omega/\sigma = i\omega\epsilon_0/e\mu = i\omega\tau_{mkc}$. Здесь $\tau_{mkc} = \epsilon\epsilon_0/\sigma$ – 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 полях составляет $10^{-1} \dots 10^{-2}$ пс.

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 отри-

цательной динамической проводимости. Его можно провести, пренебрегая инерционностью междолинных переходов ($\omega_m \ll 1$), запаздыванием столкновений ($\omega\tau_p \ll 1$) и считая, что в режиме ОДП скорость электронов достигает насыщения ($v_{0G} = v_{nG}$). 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, предельная частота, определяемая в основном временем релаксации энергии, имеет значение порядка сотен и даже тысяч гигагерц, уменьшаясь с ростом напряженности электрического поля.

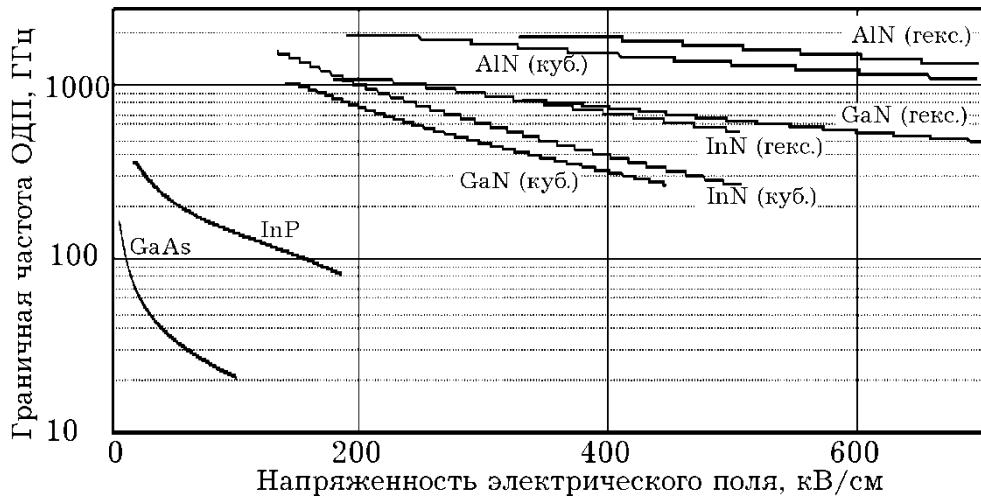


Рис. 3.17. Зависимость максимальной частоты существования ОДП для арсенида галлия, фосфида индия и тринитридов кубической и гексагональной модификаций от напряженности электрического поля

Интересная тенденция наблюдается для нитридов с кубической структурой решетки - зависимости предельной частоты от напряженности электрического поля сдвинуты в сторону меньших значений поля. И хотя для критических полей

предельные частоты для них несколько больше, для режима наибольшей отрицательной дифференциальной проводимости эти частоты практически одинаковы. Стоит также отметить, что тринитриды кубической модификации имеют большие значения отрицательной динамической проводимости.

Полученные результаты расчета предельной частоты для арсенида галлия и фосфида индия в целом соответствуют известным результатам [3.10], но для нитрида галлия существенно превышают результаты прогноза, приведенные в [3.11].

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.

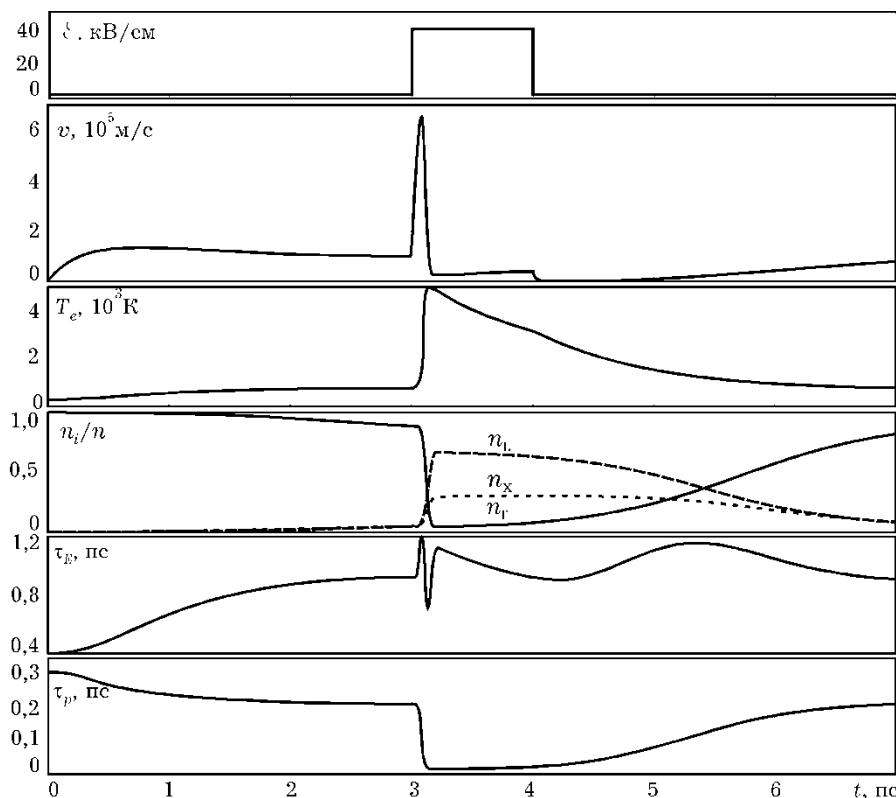


Рис. 3.18. Динамические (временные) зависимости при включении импульса электрического поля от 2 до 40 кВ/см для (сверху вниз): средней скорости дрейфа, электронной температуры, заселенности долин, усредненных по долинам времен релаксации энергии и импульса. Следует обратить внимание на экстремальные 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.

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 достиг-

ла 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 $\tau_{E1} > \tau_{p1}$.

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). Одновре-

менно 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.

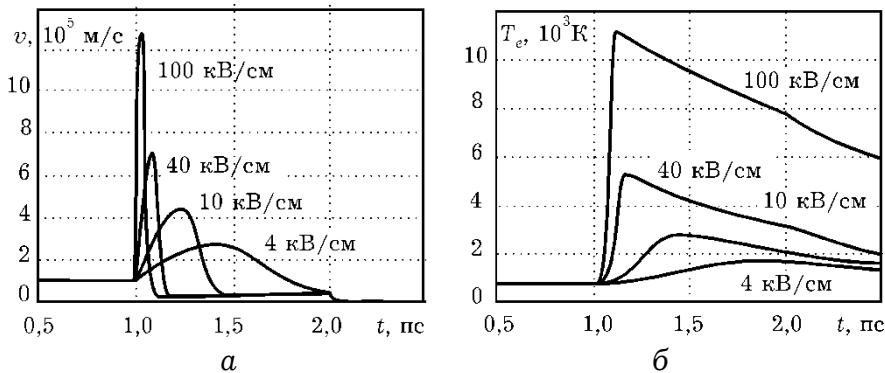


Рис. 3.19. Динамические зависимости дрейфовой скорости и электронной температуры для разных значений амплитуды напряженности поля. Длительность импульса поля 1 пс

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 $\mathcal{E}_1 = 2 \text{ кВ/см}$ со скоростью $v_0 = \mu_1 \mathcal{E}_1$ и влетающих в поле $\mathcal{E}_2 = 40 \text{ кВ/см}$. Там же представлена зависимость $v(x) = \sqrt{v_0^2 + 2(e/m^*)\mathcal{E}_2 x}$, 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

$$\frac{x}{\tau_{p1}} = v_{ct} \ln \left| \frac{v_{ct} - v_0}{v_{ct} + v(x)} \right| + v_0 - v(x), \quad (3.47)$$

являющейся решением уравнения баланса импульса для $\tau_p = \tau_{p1} = \text{const}$. В данном расчете время релаксации импульса соответствует слабому полю, а $v_{\text{ст}} = \mu_1 E_2$ – скорость дрейфа в установившемся режиме в поле с напряженностью 40 кВ/см, но с подвижностью μ_1 , как и перед влетом электрона в сильное поле.

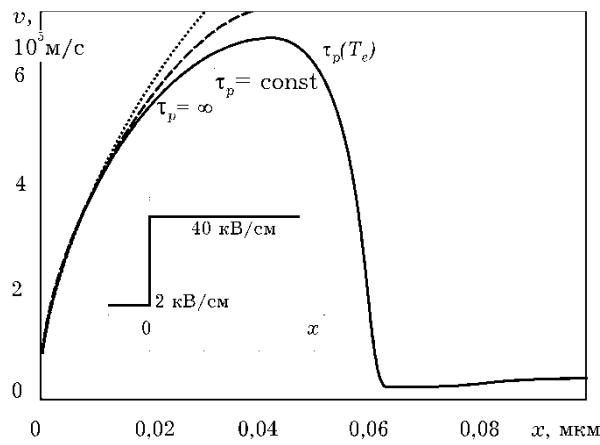


Рис. 3.20. Пространственное изменение скорости электронов, дрейфовавших в поле 2 кВ/см и влетающих в поле 40 кВ/см в точке $x = 0$. Пунктирная линия соответствует баллистическому движению без столкновений, штриховая – движению с теми же параметрами, что и в слабом поле.

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 начидающихся 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.

Результаты моделирования эффекта «всплеска» в нитридах алюминия, галлия и индия приведены на рис. 3.21 для скачкообразного во времени электрического поля, которое изменяется от нулевого до фиксированного значения. Для удобства сравнения эти значения, наибольшие из которых в 3-5 раз превышают критические значения поля, и начальные параметры, кроме энергии фононов, выбирались такими же, как

и в работе [3.12], где подобное моделирование проводилось методом Монте-Карло.

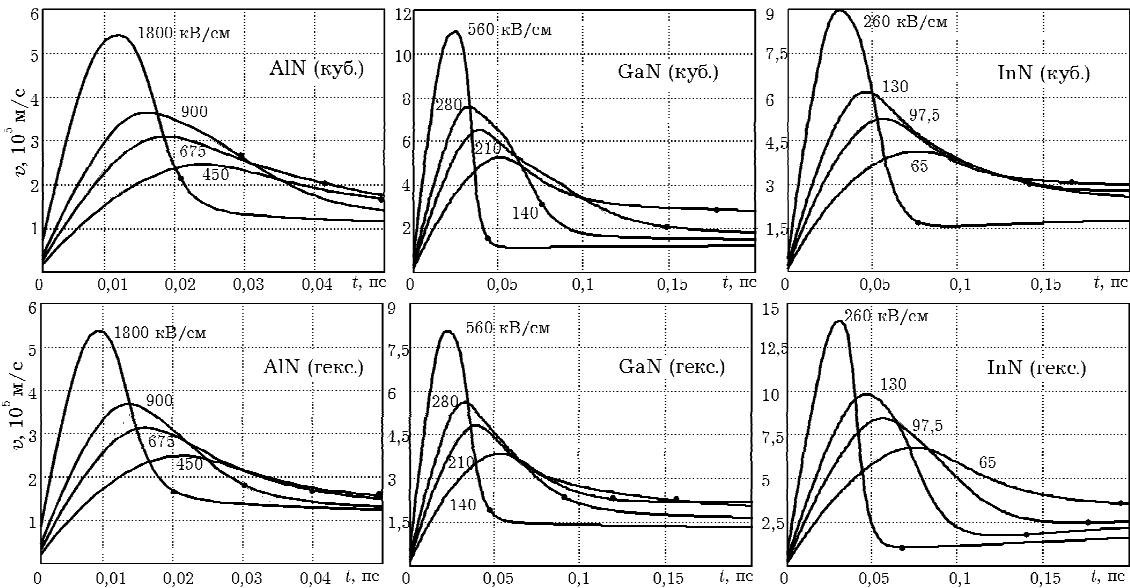


Рис. 3.21. Динамические зависимости скорости дрейфа для три-нитридов при разных значениях амплитуды ступенчатого электрического поля. Точками обозначены моменты времени, которые соответствуют временема релаксации импульса при данной напряженности поля

Динамические зависимости от времени пересчитаны в пространственное распределение скорости (рис. 3.22). Также на вставках к рис. 3.22 приведены некоторые результаты из [3.12] для гексагональных структур.

Естественно, что пиковые значения «всплеска» возрастают с увеличением напряженности. При наибольших значениях поля они в 4-8 раз превышают соответствующие стационарные значения из поле-скоростной характеристики. Наибольшие абсолютные значения получены для нитрида индия - около $15 \cdot 10^5$ м/с при значительно меньшей, чем в других нитридах, напряженности поля. Объяснить это можно тем, что нитрид индия имеет наибольшее время релаксации импульса, а именно за это время длится баллистический рост скорости. Точки на рис. 3.21 показывают, что длительность импульса скорости, а значит и импульса тока, фактически равна времени релаксации импульса при данной напряженности электрического поля.

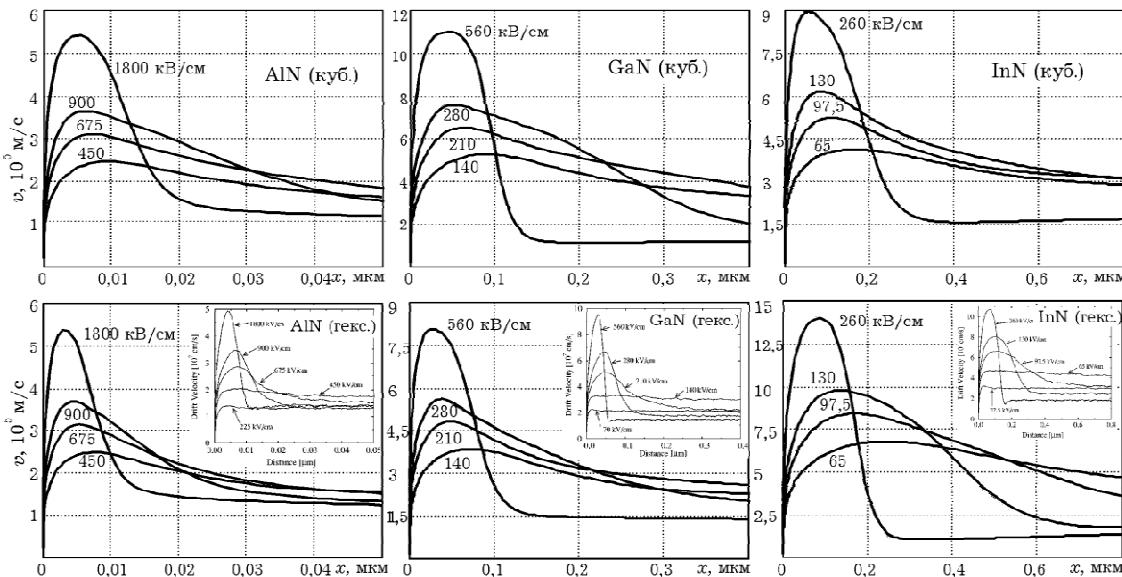


Рис. 3.22. Пространственные зависимости скорости дрейфа для тринитридов при разных значениях амплитуды ступенчатого электрического поля

Фактически по этой же причине длина «баллистического пробега» наибольшая в нитриде индия (рис. 3.22), порядка 0,1 мкм, в то время как для других материалов она почти на порядок меньше.

Полученные результаты находятся в хорошем соответствии с результатами, полученными методом Монте-Карло [3.12], и свидетельствуют о том, что метод релаксационных уравнений пригоден для моделирования как статических, так и динамических свойств полупроводниковых материалов, имея по сравнению с методом Монте-Карло большую временную экономичность и достаточную информативность.

Сравнение полученных зависимостей для гексагональных структур нитридов на рис. 3.22 демонстрирует намного меньшее отличие результатов, полученных методом Монте-Карло и методом релаксационных уравнений, чем тех, которые показаны на рис. 5 в работе [3.12]. Очевидно, что причина такого различия в некорректности использования авторами [3.12] уравнения баланса для полной энергии.

Эффект «всплеска» может проявляться и в том случае, если напряженность поля меняется в пространстве не скачкообразно, а относительно плавно. Он заметен на расстояниях порядка ~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. ДИФФУЗИЯ ГОРЯЧИХ ЭЛЕКТРОНОВ

С особенностями механизма разогрева электронов также связана характерная для многодолинных полупроводников зависимость коэффициента диффузии D от напряженности электрического поля (рис. 3.23). 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 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. 3.23, б) можно теми же механизмами, что и для арсенида галлия (рис. 3.23, 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 около 10 кВ/см

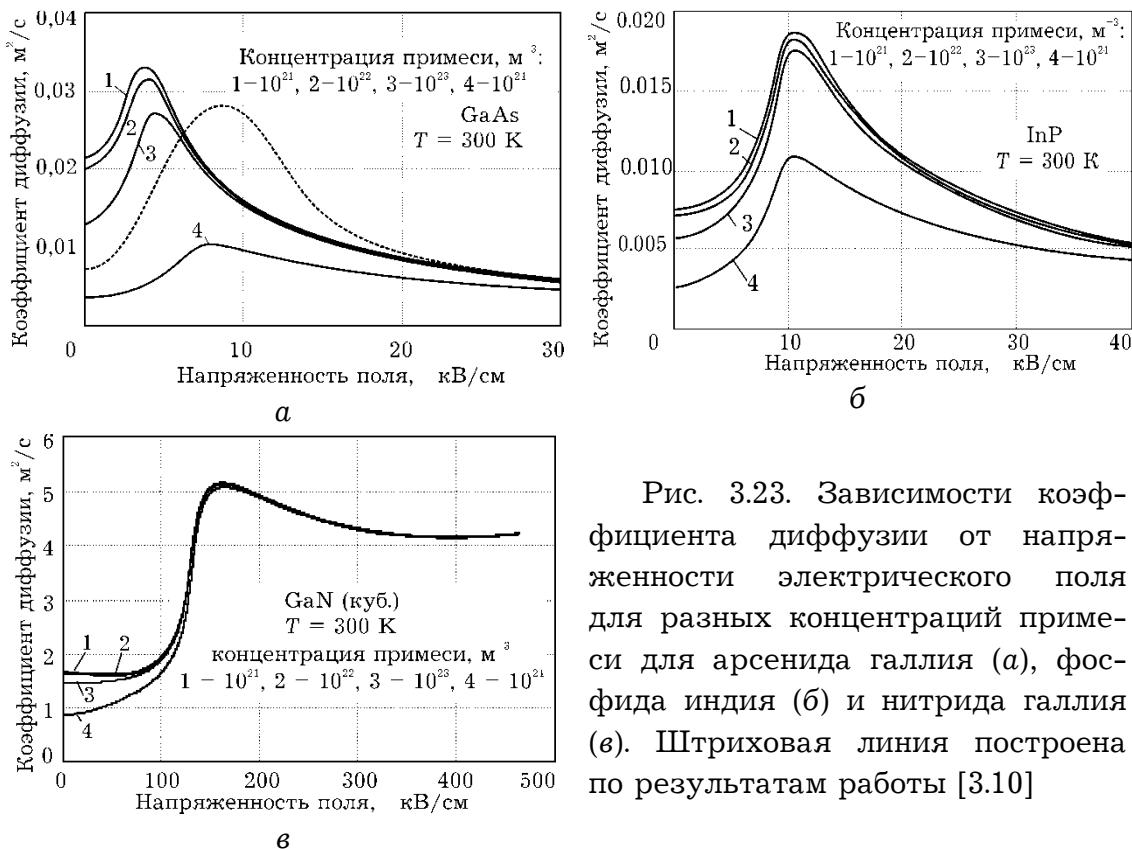


Рис. 3.23. Зависимости коэффициента диффузии от напряженности электрического поля для разных концентраций примеси для арсенида галлия (*а*), фосфида индия (*б*) и нитрида галлия (*в*). Штриховая линия построена по результатам работы [3.10]

Для нитрида галлия (рис. 3.23, *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 $D(E)$, 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.

Литература к разделу 3

- 3.1. *Москалюк В. А., Синекоп 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., Калашников 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.
- 3.12. *Foutz B. E., O'Leary S.K., Shur M.S., Eastman L.F.,* Transient electron transport in wurtzite GaN, InN and AlN, *J. of Appl. Phys.*, v. 85, N. 11, pp.7727-7734.

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, *a*), 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, мало мощные 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

мым 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.

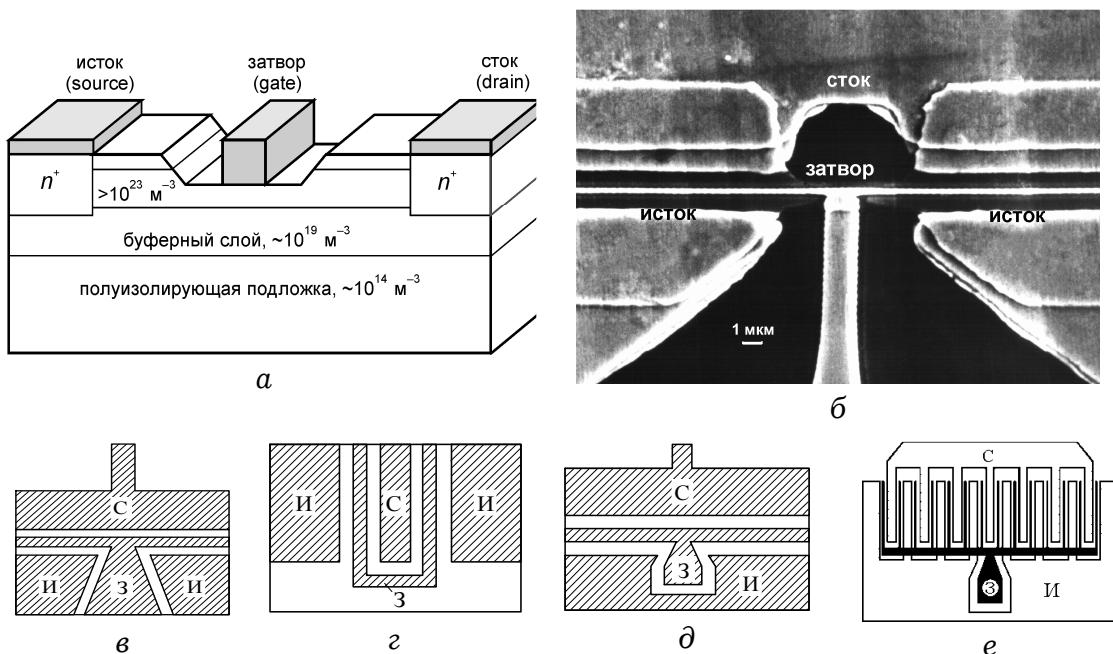


Рис. 4.1. Типичная структура (а) и фотография (б) полевого транзистора с затвором в виде барьера Шоттки и варианты конструкций малошумящих транзисторов: в – Hewlett Packard HP (США); г – Siemens (ФРГ); д – NEC (Япония), а также мощного (е) транзистора

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. Стоконые характеристики полевого транзистора при малых напряжениях стока близки к линейным и хорошо описываются такой «резистивной» моделью управления. При больших стоконых напряжениях при сужении канала напряженность электрического поля возрастает до таких значений, при которых наблюдается насыщение дрейфовой скорости электронов, что вызывает выполнование стоконых 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 $G_0 = e\mu N_d Wa/L$ соответствует проводимости необедненного канала толщиной a , шириной W и длиной L с концентрацией электронов N_d , а V_p – так называемый потенциал перекрытия, равный потенциальному 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:

$$I_{dH} = G_0 \left\{ \frac{1}{3} V_p - \frac{2}{3} \frac{(V_b - V_{gs})^{3/2}}{V_p^{1/2}} - V_b + V_{gs} \right\}. \quad (4.4)$$

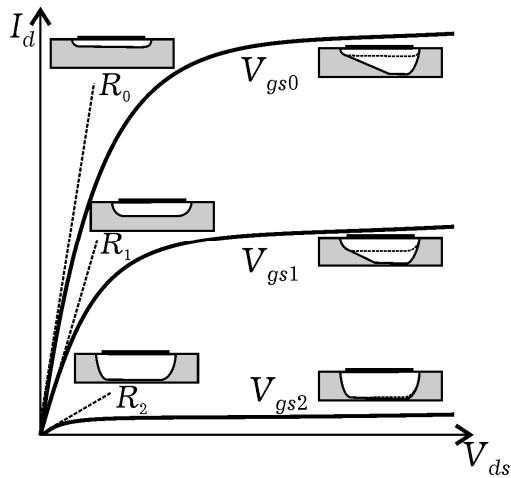


Рис. 4.2. Изменение формы канала при разных напряжениях затвора и стока

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 несостоительными. Это можно проиллюстрировать на примере стоковых характеристик транзистора — укорочение затвора до субмикронных размеров качественно меняет их вид (рис. 4.3)

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 потенциальному 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 проводимо-

сти. Количественные оценки показывают, что такую проводимость могут обеспечить электроны обогащенного слоя, образующегося у границы эпитаксиального слоя и подложки (буферного слоя).

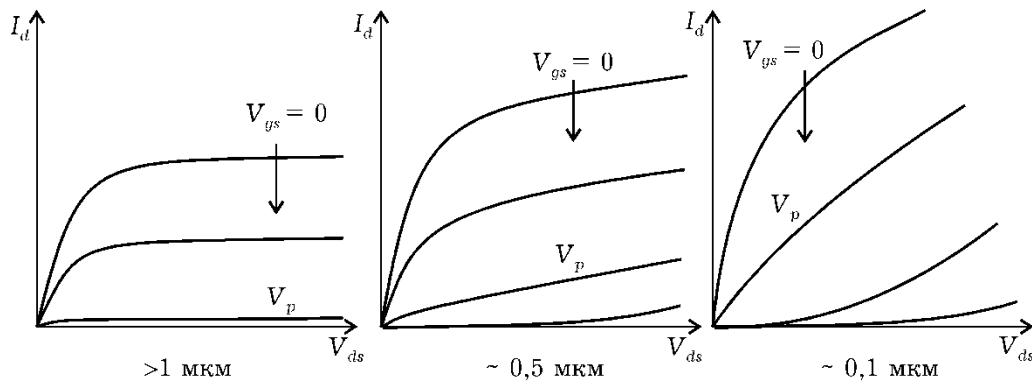


Рис. 4.3. Трансформация формы стоковых характеристик при уменьшении длины затвора

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.

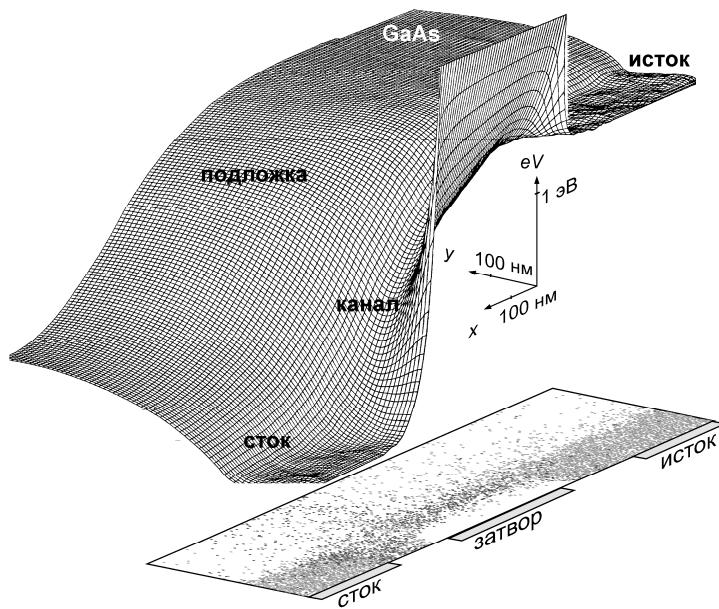


Рис. 4.4. Потенциальный рельеф и распределение электронов в канале открытого ПТШ. Соответствующее распределение электронов по долинам см. на рис. 1.3

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 — $v = \mu E$. В этих условиях необходимо использовать кинетическое уравнение Больцмана, либо релаксационные уравнения сохранения импульса, энергии и концентрации, описанные в первом разделе.

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):

$$d\mathcal{E} / dx = e[n(x) - N_{\text{д}}] / \varepsilon \varepsilon_0, \quad (4.5)$$

$$I = e n(x) v(x) h(x) W, \quad (4.6)$$

а толщину обедненной области $d(x)$ рассчитывать в зависимости от потенциала канала $V(x)$ относительно истока:

$$d(x) = \{2\varepsilon \varepsilon_0 [V_b - V_{gs} + V(x)] / e N_{\text{д}}\}^{1/2}. \quad (4.7)$$

В этих соотношениях e , $n(x)$, $v(x)$ – заряд, концентрация и скорость электронов, $h(x)$ – толщина проводящего канала, W – ширина затвора, $N_{\text{д}}$ – 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

$$V_k = (kT/e) \ln(N_d / N_{dp}). \quad (4.13)$$

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 тянувшего 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.

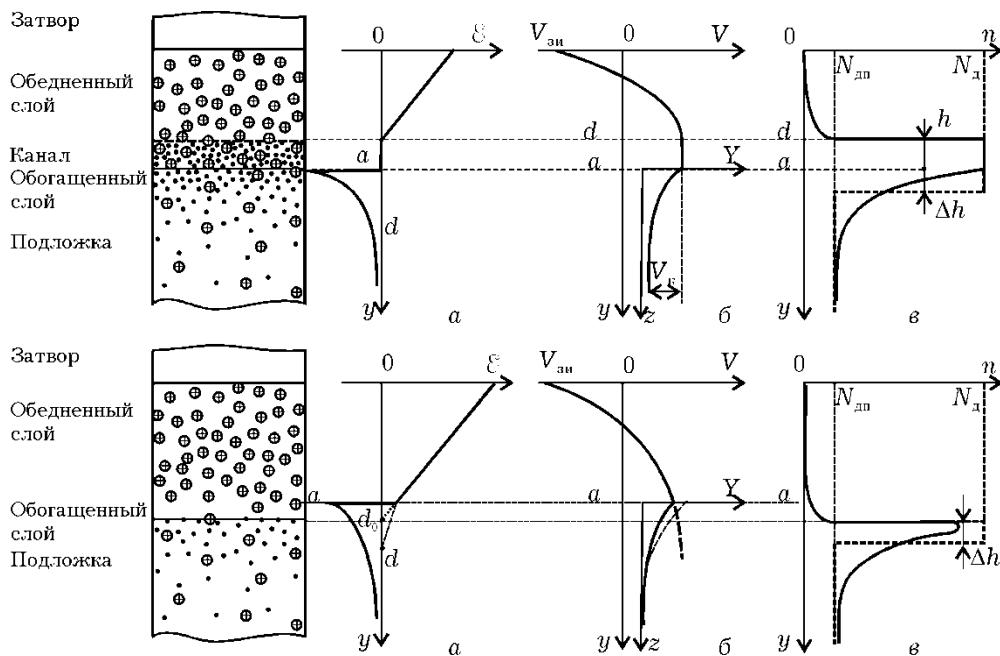


Рис. 4.5. Распределение напряженности поперечного электрического поля (а), потенциала (б) и концентрации электронов (в) при неперекрытом канале (верхний ряд) и перекрытом (нижний)

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

$$\frac{d^2V}{dy^2} = -\rho(y)/\epsilon\epsilon_0 = eN_{дп}/\epsilon\epsilon_0[\exp(eV(y)/kT) - 1]. \quad (4.14)$$

Вводя локальные (рис. 4.5) нормированные координаты $Y = eV/kT$ и $z = y/l_{\text{эп}}$, где $l_{\text{эп}} = \sqrt{\epsilon\epsilon_0 kT/e^2 N_{дп}}$ - дебаева длина экранирования для подложки, получим

$$\frac{d^2Y}{dz^2} = e^Y - 1. \quad (4.15)$$

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:

$$\begin{aligned}\Delta Q_1(\Delta y) &= - \int_{y_1}^{y_2} \rho(y) dy = \epsilon \epsilon_0 \int_{y_1}^{y_2} (d^2 V / d y^2) dy = \\ &= \epsilon \epsilon_0 \int_{y_1}^{y_2} d(dV / dy) = (\epsilon \epsilon_0 kT / e l_{\text{ЭП}})(dY / dz|_{y_2} - dY / dz|_{y_1}).\end{aligned}\quad (4.17)$$

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 при $y = 0$ $dY(0) / dz = \sqrt{2(e^{Y_k} - Y_k - 1)}$ 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:

$$dV / dy = e N_{\text{дп}} (d - y) / \epsilon \epsilon_0. \quad (4.21)$$

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

$$\Delta h_2 = \Delta h_1 - (d - a)N_{dp} / N_d. \quad (4.22)$$

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 и равна

$$h(x) = a - d(x) + \Delta h_1. \quad (4.23)$$

Если под стоковой частью затвора возникает перекрытие ($d > a$), то

$$h(x) = \Delta h_2(x). \quad (4.24)$$

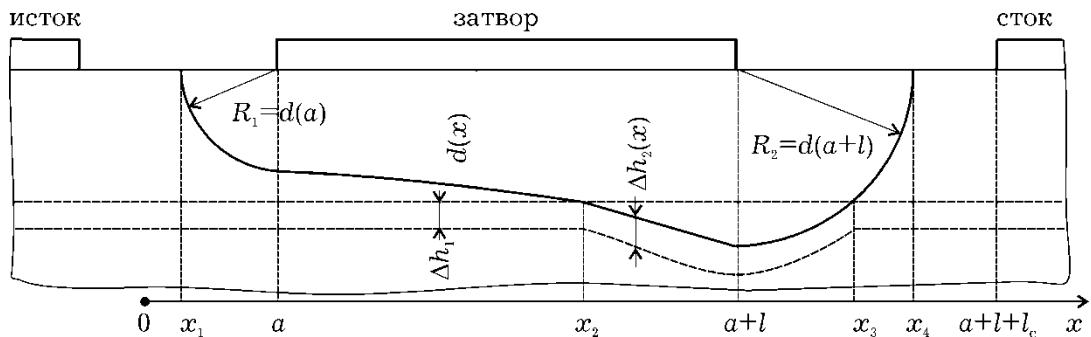


Рис. 4.6. Схема расчета формы обедненной области под затвором и толщины проводящего канала

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 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 тянувшего 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:

$$\begin{aligned} \mathcal{E}_i &= \mathcal{E}_{i-1} - e\Delta x(n_i - N_d) / \epsilon\epsilon_0, \\ I &= e n_i v_i h_i W, \\ V_i &= V_{i-1} + \mathcal{E}_i \Delta x, \\ m_i^* v_i &= m_{i-1}^* v_{i-1} + e\Delta x(\mathcal{E}_i / v_i - \mathcal{E}_{ci} / v_{ci}), \\ E_i &= E_{i-1} + e\Delta x(\mathcal{E}_i - \mathcal{E}_{ci} v_{ci} / v_i), \\ h_i &= a - d_i + \Delta h_i. \end{aligned}$$

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 $N_{dp} = 5 \cdot 10^{14}$ см⁻³, напряжении на затворе $V_{gs} = 0$ В, токе стока 30 мА.

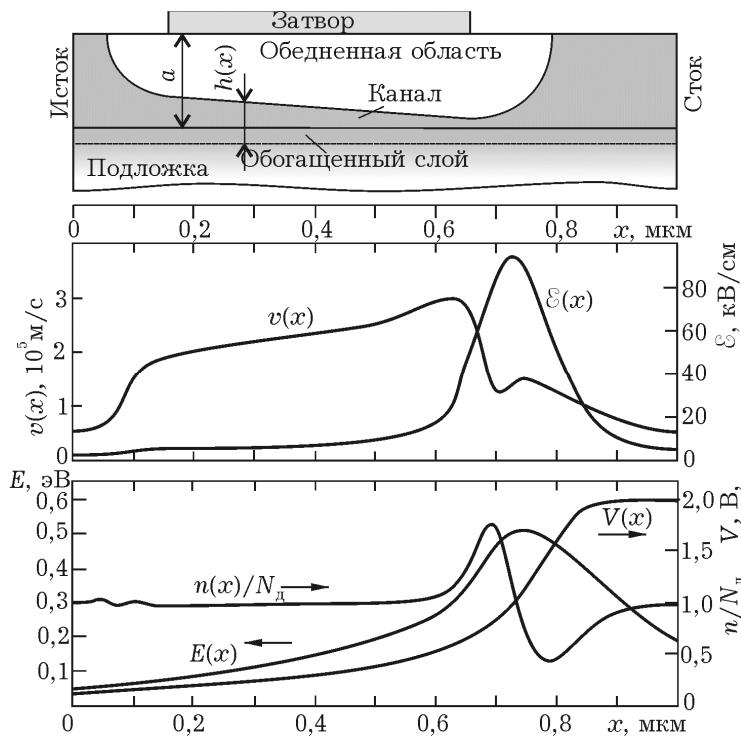


Рис. 4.7. Результаты вычисления формы обедненной области под затвором, распределения скорости электронов, потенциала и энергии, концентрации носителей и напряженности поля вдоль канала транзистора

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.

Данная модель хорошо описывает увеличение выходной проводимости, т. е. наклон стоковой характеристики в транзисторах с длиной затвора около 0,5 мкм, но не адекватно отра-

жает 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 инжеекции 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 инжеекции 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 ин-

жекции 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 НЕМТ. 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:

$dE = (I_k dE_k + I_n dE_n) / (I_k + I_n) = (E_n - E_k + eV_k) d(I_{pk} - I_{kn}) / (I_k + I_n)$,
где dI_{kn} , dI_{pk} – приращения токов на участке dx за счет инжеекции носителей из канала в подложку и их возврата в канал. При инжеекции электронов из канала в подложку ($dI_{pk} > dI_{kn}$) 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:

$$\begin{aligned} d(\mathcal{E}h) &= (e/\epsilon\epsilon_0)h[n(x) - N_d]dx, \\ d(envhW) &= dI_{kn} - dI_{pk}, \\ dE &= e\mathcal{E}dx - [(E - E_0)/v\tau_E]dx + [(E - E_n + eV_k)/I_k]dI_{pk}, \quad (4.26) \\ d(m^*v) &= (e\mathcal{E}/v)dx - (m^*/\tau_p)dx + [(m_n^*v_n - m^*v)/I_k]dI_{pk}, \\ h(x) &= a - \sqrt{2\epsilon\epsilon_0[V_B - V_{gs} + V(x)]/eN_d} \\ dV &= \mathcal{E}dx, \end{aligned}$$

и аналогичные уравнения для подложки:

$$d(\mathcal{E}_n h_n) = (e/\epsilon\epsilon_0)h_n[n_n(x) - n_{cp}]dx,$$

$$\begin{aligned}
d(en_{\text{п}}v_{\text{п}}h_{\text{п}}W) &= -dI_{\text{кп}} + dI_{\text{пк}}, \\
dE_{\text{п}} &= e\mathcal{E}_{\text{п}} dx - [(E_{\text{п}} - E_0)/v_{\text{п}}\tau_{E_{\text{п}}}]dx + [(E - E_{\text{п}} - eV_{\text{k}})/I_{\text{k}}]dI_{\text{кп}}, \quad (4.27) \\
d(m_{\text{п}}^*v_{\text{п}}) &= (e\mathcal{E}_{\text{п}}/v_{\text{п}})dx - (m_{\text{п}}^*/\tau_{p\text{п}})dx - [(m_{\text{п}}^*v_{\text{п}} - m^*v)/I_{\text{k}}]dI_{\text{кп}}, \\
h_{\text{п}}(x) &= a + \sqrt{2l_{\text{эп}}} - \sqrt{a^2 + 2\varepsilon\varepsilon_0[V_B - V_{gs} + V(x) - V_p]/eN_{\text{д}}}, \\
dV_{\text{п}} &= \mathcal{E}_{\text{п}} dx,
\end{aligned}$$

а также уравнения, описывающие токообмен между каналом и подложкой и связывающие две предыдущие системы в единую систему уравнений:

$$dI_{\text{кп}} = 0,901eD_{\text{п}}W[n_{\text{п}} - 0,637\sqrt{N_{\text{д}}N_{\text{дп}}}]dx/l_{\text{эп}}, \quad (4.28, \text{а})$$

$$dI_{\text{пк}} = 0,366eD_{\text{k}}W[n(x) - N_{\text{д}}]dx/l_{\text{э}}. \quad (4.28, \text{б})$$

Здесь \mathcal{E} , $\mathcal{E}_{\text{п}}$ – напряженность продольного электрического поля в канале и в подложке, h , $h_{\text{п}}$ – соответственно толщина канала и эффективная толщина подложки, $N_{\text{д}}$, $N_{\text{дп}}$ – уровни легирования эпитаксиального слоя и подложки, n , $n_{\text{п}}$ – текущие значения концентрации электронов в канале и подложке, v , $v_{\text{п}}$ – средние значения дрейфовой скорости носителей в канале и в подложке, $dI_{\text{кп}}$, $dI_{\text{пк}}$ – составляющие полного тока соответственно из канала в подложку и из подложки в канал, E , $E_{\text{п}}$ – 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^* , $m_{\text{п}}^*$ – средние по долям значения эффективной массы носителей в канале и в подложке, V_b , V_{gs} – потенциал барьера Шоттки и разность потенциалов затвор-исток, $V(x)$, $V_{\text{п}}(x)$ – текущее значение потенциала канала и подложки (потенциал подложки имеет смысл только при полностью перекрытом канале), $n_{\text{ср}}$ – среднее (равновесное) значение концентрации в подложке, V_p – потенциал «отсечки» тока стока, $l_{\text{э}}$, $l_{\text{эп}}$ – длины экранирования в канале и в подложке, D_{k} , $D_{\text{п}}$ – 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.

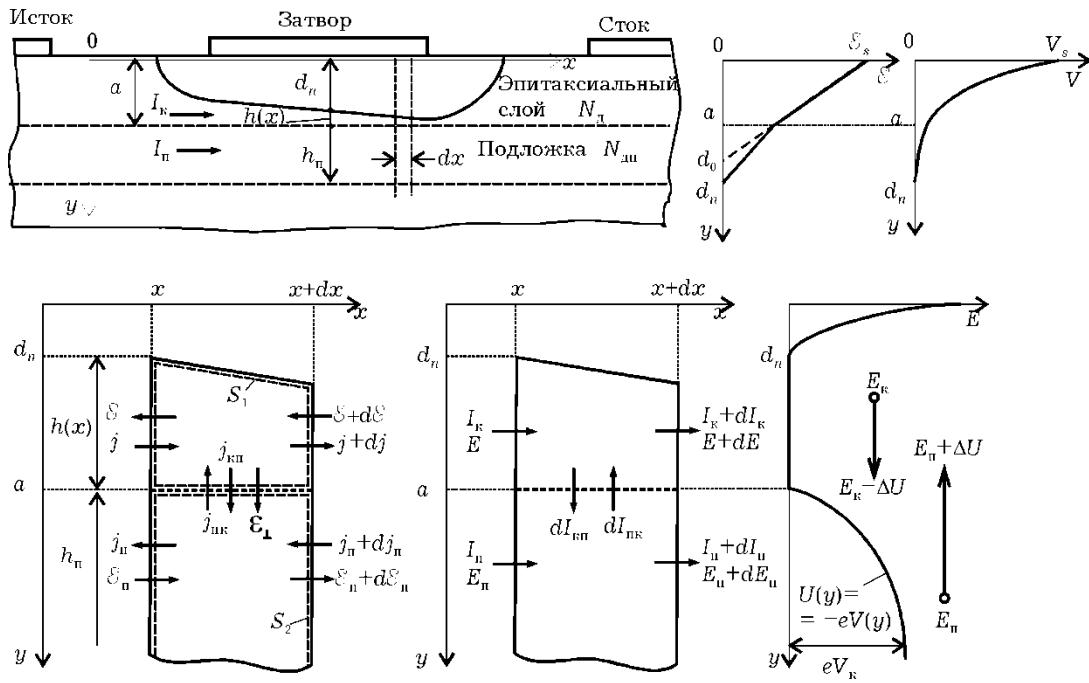


Рис. 4.8. Схемы, поясняющие вывод основных уравнений для канала и подложки

Вид уравнений непрерывности, Пуассона, баланса импульса и энергии, входящих в системы (4.26) – (4.28), приведен в [4.10]. 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 – $n_{cp} = (N_d N_{d\mu})^{1/2}$. 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.

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 $2 \cdot 10^{23} \text{ м}^{-3}$, 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 Г-Л-Х 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, поллярных оптических, междолинных фононах и на ионизированных примесях.

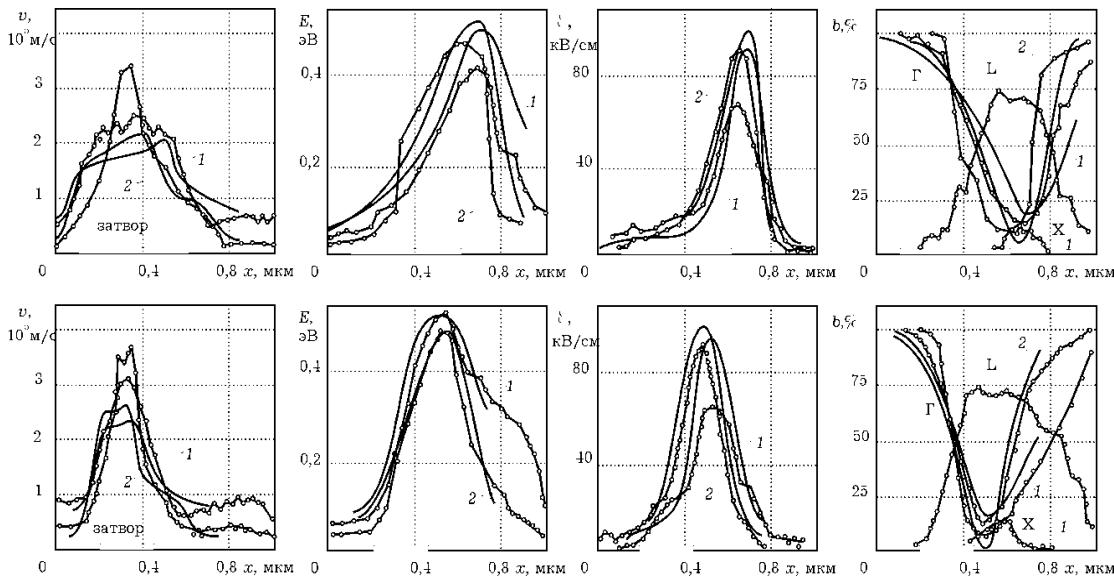


Рис. 4.9. Распределения дрейфовой скорости, энергии, напряженности электрического поля, заселенности долин для потенциалов на затворе 0 В (кривые 1) и $-1,0$ В (кривые 2). Сплошными линиями и линиями с кружками обозначены соответственно результаты расчетов по квазидвумерной и кинетической моделям. Верхний ряд – для длины затвора 0,5 мкм, нижний – 0,25 мкм

Двумерное уравнение Пуассона решалось методом циклической редукции с использованием метода матрицы емкости для учета электродов. Границные условия для уравнения Пуассона – $\phi = \text{const}$ на электродах и $d\phi/dn = 0$ 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 интересующие средние зависимости $n_{\text{ср}}(x)$, $v_{\text{ср}}(x)$, $\mathcal{E}(x)$ 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 $v_{\text{срxi}}$ по формуле

$$v_{\text{срxi}} = (\hbar / N_i) \sum_{j=1}^{N_i} k_{xj} / (m_j^*(1 + 2\alpha_j E_j)),$$

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 флюктуаций проводилось дополнительное усреднение по временному интервалу 2...4 пс. Средний ток подложки $I_{\text{п}}(x)$ определялся по формуле $I_{\text{п}}(x) = = qv_{\text{пх}}(x)N_{\text{пх}}(x)/\Delta x$, где q – заряд частицы; $v_{\text{пх}}(x)$ – средняя дрейфовая скорость в подложке; $N_{\text{пх}}(x)$ – среднее число частиц в подложке в интервале Δx , в котором производится усреднение.

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 затво-

ре, 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.

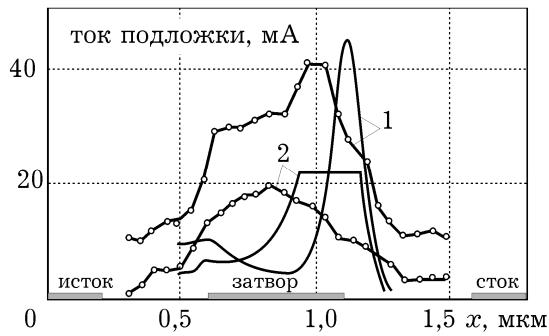


Рис. 4.10. Распределения тока подложки в кинетической (точки) и квазидвумерной (сплошная) модели при разных смещениях на затворе

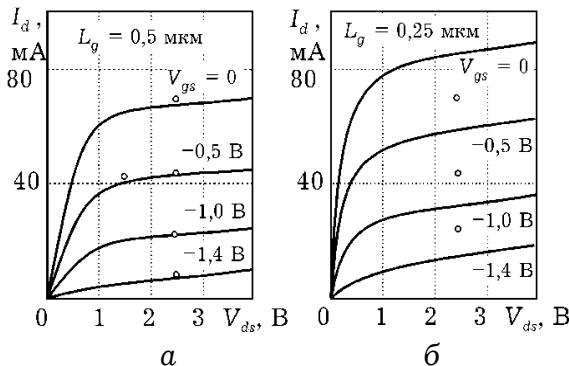


Рис. 4.11. Вольтамперные характеристики ПТШ с различной длиной затвора 0,5 мкм (а) и 0,25 мкм (б)

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 запирания 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 эпитетаксиального слоя 0,12 мкм, концентрация примеси в эпитетаксиальном слое $1,9 \cdot 10^{17}$ см⁻³, в подложке $1 \cdot 10^{14}$ см⁻³, межэлектродные расстояния исток-затвор и затвор-сток по

0,6 мкм, высота металлизации полоски 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 $-1,4$ В, который с учетом барьерной разности потенциалов соответствует потенциальному перекрытия канала.

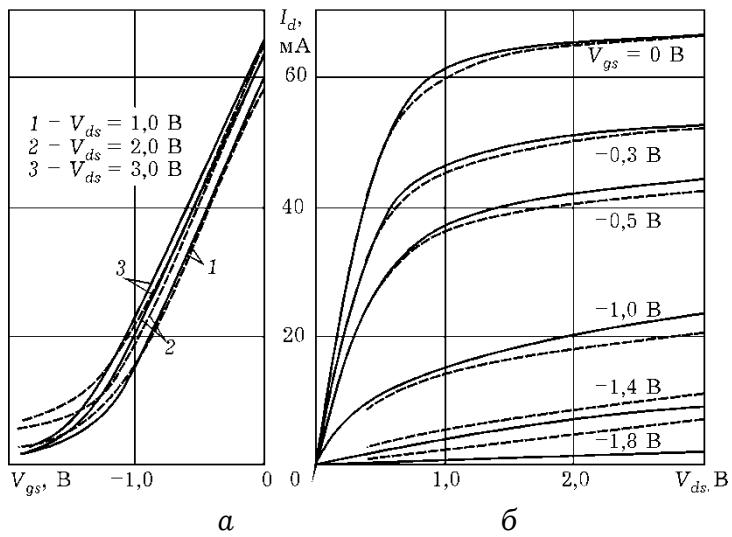


Рис. 4.12. Сток-затворные (а) и стоковые (б) характеристики ПТШ: сплошные линии – экспериментальные, штриховые – расчетные

Заметное отличие характеристик наблюдается в области напряжений затвора, близких к потенциальному перекрытия ($\approx 1,4$ 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 управляет 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 .

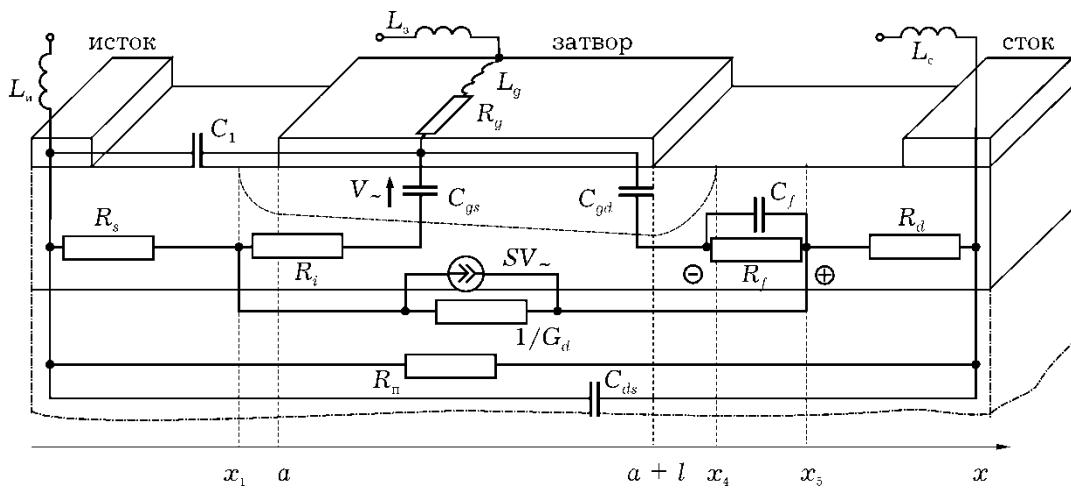


Рис. 4.13. Схемное представление областей субмикронного ПТШ на сверхвысоких частотах

Малосигнальные параметры схемной модели определяются по задаваемым малым приращениям напряжения затвора и тока стока. 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.

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 перехождением 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 потенциальному отсечки, способствует росту скорости, а значит, приводит к увеличению выходной проводимости. Этому потенциальному соответствует максимальный наклон стоковой характеристики и максимальная выходная проводимость.

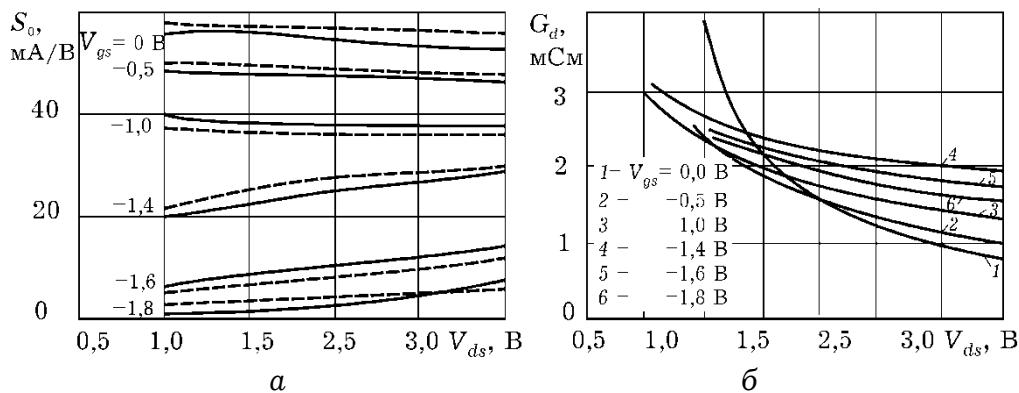


Рис. 4.14. Зависимости крутизны (а) и выходной проводимости (б) от напряжения стока

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 ем-

кости можно отнести ее некоторый рост при запирании транзистора, заметный на экспериментальных характеристиках.

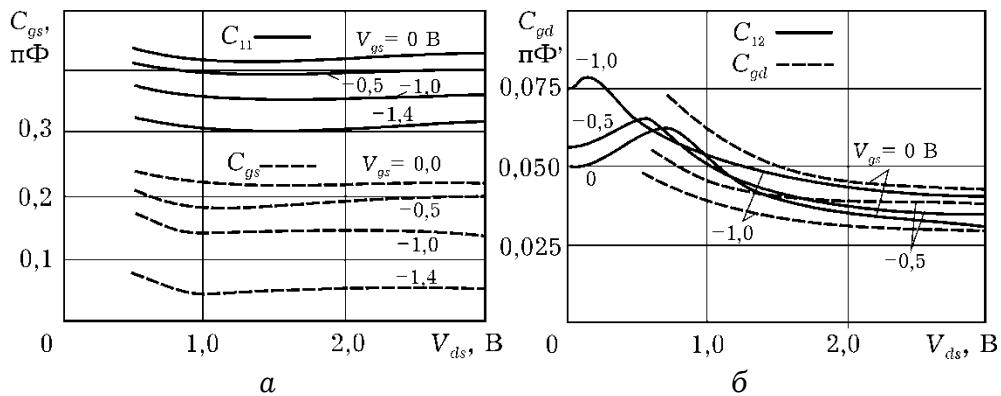


Рис. 4.15. Сравнение экспериментальных (сплошные линии) зависимостей для входной емкости и рассчитанных (штриховые) для емкости C_{gs} (а), а также проходной емкости и емкости C_{gd} (б) от напряжения стока

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: крутизна 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 от электрического режима и определяется краевой емкостью истокового конца затвора.

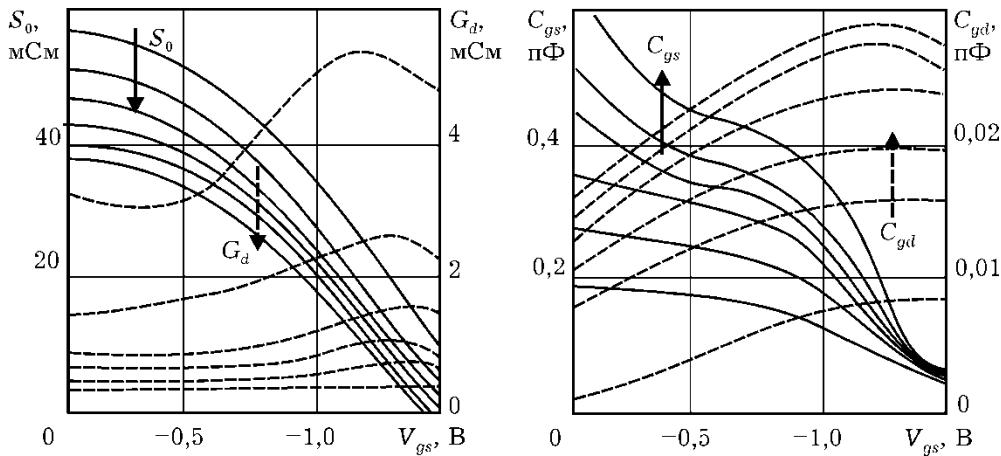


Рис. 4.16. Зависимости малосигнальных крутизны, выходной проводимости, входной и проходной емкостей от напряжения на затворе. Направление стрелок соответствует увеличению длины затвора от 0,25 мкм до 1,5 мкм с шагом 0,25 мкм.

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 \mathbf{Y} -матрицы транзистора, которая затем по известному соотношению пересчитывается в матрицу рассеяния: $[\mathbf{S}] = ([\mathbf{1}] - [\mathbf{Y}]) ([\mathbf{1}] + [\mathbf{Y}])^{-1}$.

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.

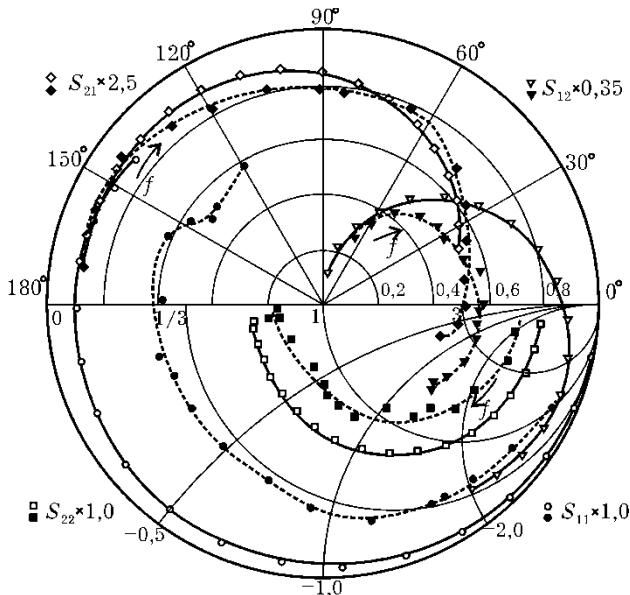


Рис. 4.17. Представление частотных характеристик элементов матрицы рассеяния в полярных координатах (вверху) и на сетке полных проводимостей (внизу). Частотные метки приведены, начиная с частоты 1 ГГц, в направлении стрелок с шагом 1 ГГц. Штриховые линии – экспериментальные. Следует обратить внимание на масштабные множители

4.3.2. Шумовая модель

Известно множество моделей ПТШ (аналитических, квазидвумерных, двумерных, кинетических), отличающихся раз мерностью решаемых уравнений, уровнем строгости и объемом вычислительных затрат. Однако, в этих работах отсутствует подробный анализ составляющих шума и не дается физическая интерпретация малошумящего 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].

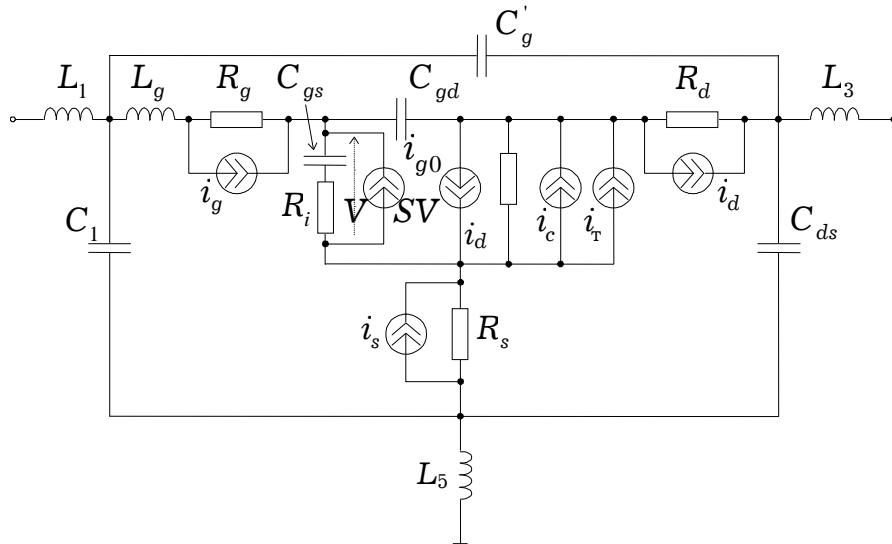


Рис.4.18. Шумовая схемная модель субмикронного ПТШ.

Спектральная плотность теплового шума, связанная с флюктуациями скорости, описывается соотношением:

$$S_i(\omega) = 4k_B T e n(x) \mu \Delta y \Delta z / \Delta x = 4k_B T_B / \Delta R,$$

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 флюктуации тока приводят за счет емкостной связи с затвором к флюктуациям 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 коррели-

рован с источником, описывающим шум канала $\overline{|i_c|^2}$. Степень корреляции определяется по формуле:

$$C = \overline{j i_{g0} \cdot i_c^*} / \sqrt{\overline{i_{g0}^2} \cdot \overline{i_c^2}},$$

$$\text{откуда } \overline{i_c \cdot i_{g0}^*} = -jC\sqrt{\overline{|i_c|^2} \overline{|i_{g0}|^2}} = -j4k_B T_e \Delta f \cdot \omega C_{gs} \cdot C \sqrt{P \cdot R}.$$

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:

$$\overline{|i_t|^2} = 2e(I_k I_n / I_c) \Delta f,$$

где I_d , I_k , I_n – ток стока, составляющие тока канала и подложки. Как видно из последнего соотношения, максимальный шум токораспределения имеет место при $I_k \approx I_n$.

Из анализа составляющих шума субмикронного ПТШ можно сделать следующие выводы.

Малошумящий 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 шума для типичной транзисторной структуры с размерами затвора

0,5×300 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.

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 равны:

$$I_{\text{швх}} = i'_g + i'_{g0} + i'_T + i'_s + i'_d \text{температуры субмикронного ПТШ}$$

$$I_{\text{швых}} = i''_g + i''_{g0} + i''_T + i''_s + i''_d.$$

На втором этапе рассчитываются волновые шумовые τ -параметры.

$$\begin{aligned} \tau_{11} &= \frac{I_{\text{швх}} \cdot I^*_{\text{швх}}}{4k_B T \Delta f R_0 Y_g^2}, & \tau_{22} &= \frac{I_{\text{швых}} \cdot I^*_{\text{швых}}}{4k_B T \Delta f R_0 Y_h^2}, \\ \tau_{12} &= \frac{I_{\text{швх}} \cdot I^*_{\text{швых}}}{4k_B T \Delta f R_0 Y_g Y_h}, & \tau_{12} &= \tau_{21}^*, \end{aligned}$$

где R_0 – стандартное нормировочное сопротивление (волновое сопротивление линий), Y_g и Y_h – 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 выходу с учетом согласования входа четырехполюсника определяется:

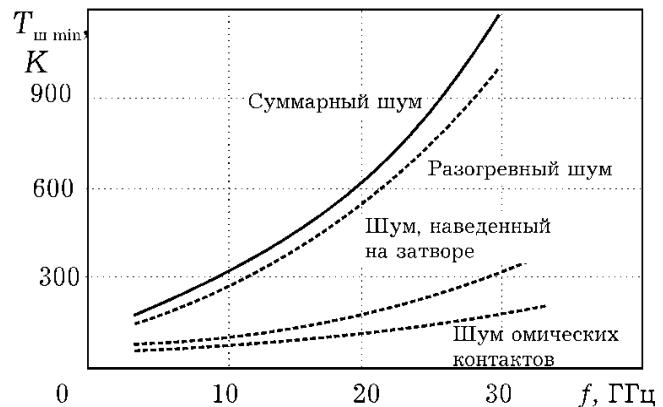


Рис. 4.19. Составляющие шумовой

$$\Gamma_{1\text{opt}} = \frac{2c^*}{\tau_{22} + b + \sqrt{(\tau_{22} + b)^2 - 4|c|^2}}.$$

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:

$$\Gamma_2 = \left[S_{22} + \frac{\Gamma_1 \cdot S_{12} \cdot S_{21}}{1 - \Gamma \cdot S_{11}} \right]^*.$$

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 представлены на рис. 4.20.

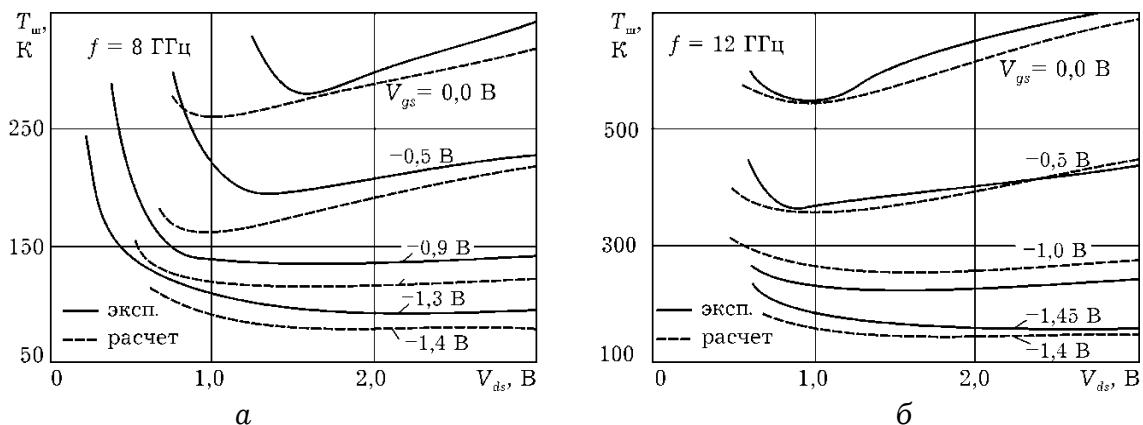


Рис. 4.20. Экспериментальные (сплошные линии) и расчетные (штриховые) зависимости шумовой температуры от напряжения стока на частоте 8 ГГц (а) и 12 ГГц (б)

Из этих результатов видна особенность шумовых свойств субмикронных полевых транзисторов — преимущественное влияние разогревных 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.

Следует отметить, что наличие максимума на зависимости выходной проводимости от напряжения на затворе (рис. 4.21) является характерным для малошумящих транзисторов. На большой выборке малошумящих транзисторов экспериментально установлено, что максимум выходной проводимости достигается вблизи напряжений на затворе, соответствующих малошумящему 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.

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

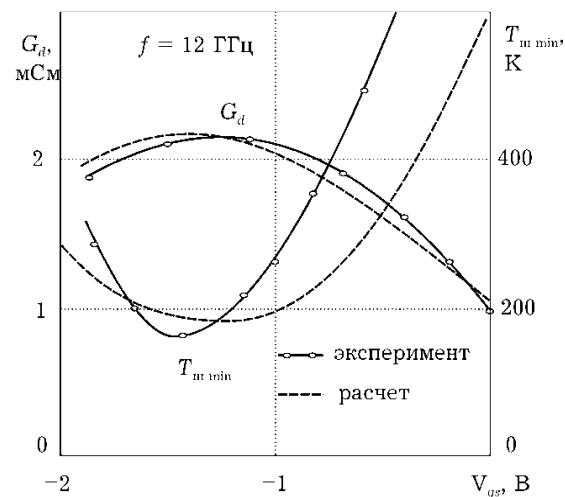


Рис. 4.21. Экспериментальная и расчетная зависимости шумовой температуры и выходной проводимости субмикронного ПТШ от напряжения на затворе

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.

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), расчетанные в широком диапазоне питающих напряжений.

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.

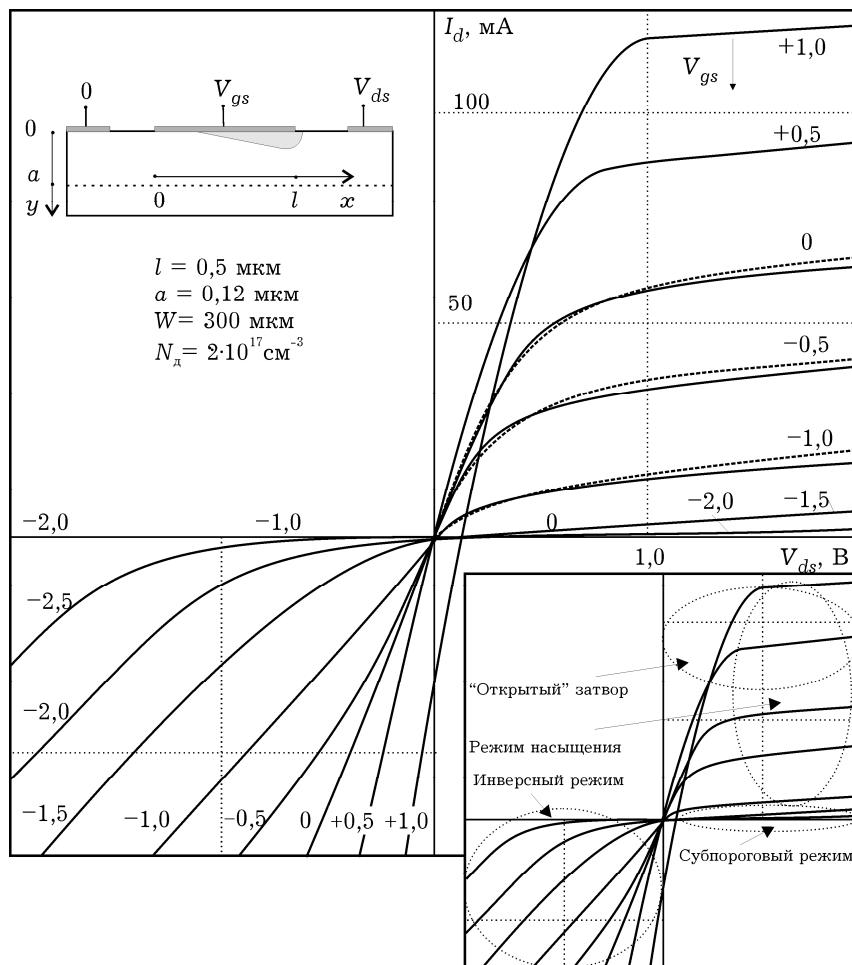


Рис. 4.22. Семейство выходных характеристик, полученное из нелинейной модели ПТШ. Штриховые линии – экспериментальные. На вставке обозначены характеристические нелинейные режимы.

Арсенид-галлиевые полевые транзисторы с барьером Шоттки используются при создании различных СВЧ устройств и сверхбыстро действующих 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 линеаризуется (рис. 4.23,б) в полулогарифмических координатах [4.8], что указывает на ее экспоненциальный характер.

Экспоненциальный характер убывания тока стока при запирании 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 тянувшего поля канала.

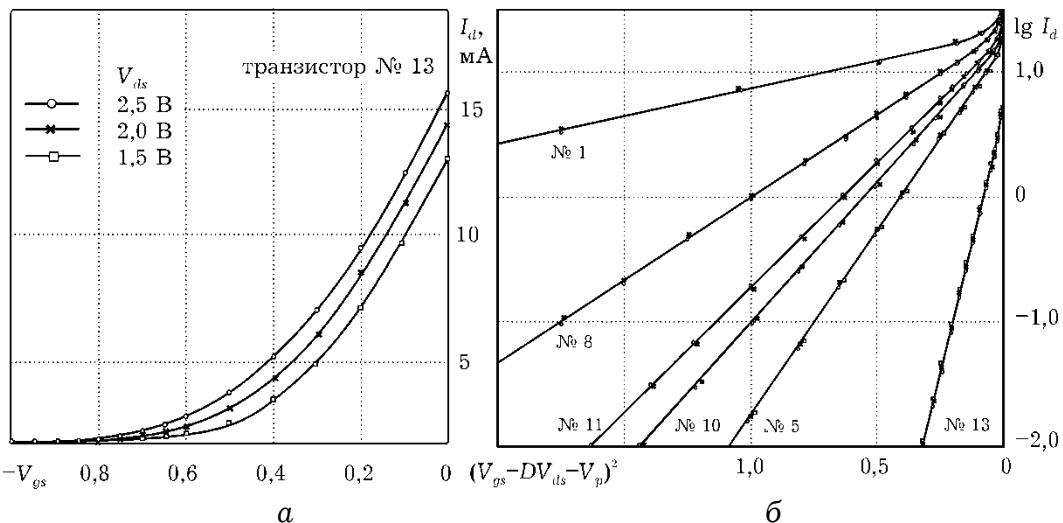


Рис. 4.23. Экспериментальные сток-затворные характеристики ПТШ при разных стоковых напряжениях в естественных и спрямляющих координатах (возле кривых указаны условные номера образцов транзисторов).

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]. Наиболее полно это можно смоделировать с помощью кинетических моделей (Монте-Карло) но их использование не-

рационально 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.

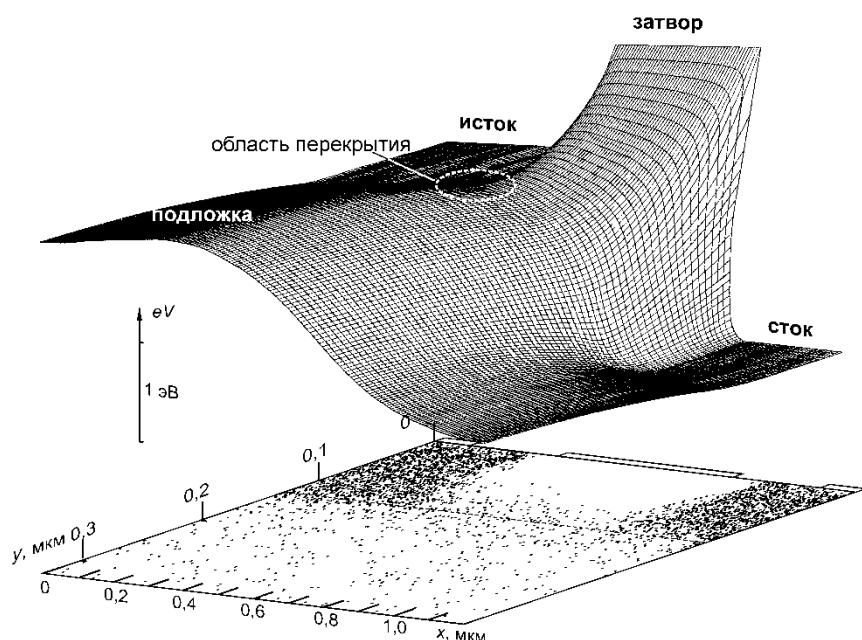


Рис. 4.24. Потенциальный рельеф и распределение электронов в канале ПТШ вблизи режима перекрытия

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_e . Вблизи режима перекрытия распределение потенциала в поперечном сечении канала можно представить (вставка на рис. 4.25) как сумму:

$$V_{\text{подл}}(y) = V_0 \exp [-(y + h/2)/l_e] \quad \text{и}$$

$$V_{\text{обедн}}(y) = V_0 \exp [(y - h/2)/l_e].$$

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

$$\begin{aligned} U(y) = -eV(y) &= 2U_0 e^{-h/2l_s} \operatorname{ch}(y/l_s) = \\ &= U_{\min} \operatorname{ch}(y/l_s) \approx U_{\min}[1 + (y/l_s)^2]. \end{aligned} \quad (4.29)$$

Распределение электронов в поперечном сечении канала можно представить с помощью модели искривленных зон [$E_c(y) = E_{c0} - eV(y)$]:

$$\begin{aligned} n(y) &= N_c \exp[(E_\Phi - E_c)/k_B T] = N_d \exp[eV(y)/k_B T] = \\ &= N_d \exp(-U_{\min}/k_B T) \exp(-U_{\min} y^2/k_B T l_s^2). \end{aligned} \quad (4.30)$$

Использование максвелловского распределения здесь оправдано тем, что в точке максимального перекрытия канала у истока продольное тянувшее поле практически равно нулю и электроны преодолевают потенциальный барьер высотой U_{\min} за счет тепловой скорости.

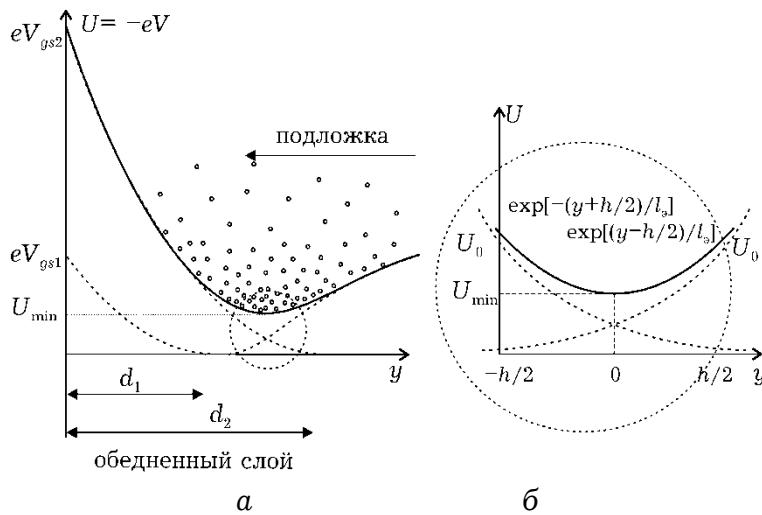


Рис. 4.25. Форма потенциального рельефа в поперечном сечении канала у истока (а), иллюстрирующая надбарьерное прохождение электронов, и его параболическая аппроксимация (б).

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:

$$n_{cp} = (1/h) \int_{-\infty}^{\infty} n(y) dy = (2N_d/h) \int_0^{\infty} \exp[eV(y)/k_B T] dy. \quad (4.31)$$

После интегрирования с использованием (4.29) и табличного интеграла получим функцию

$$n_{\text{cp}} = N_{\text{д}} \sqrt{\pi} (l_{\text{э}} / h) \exp(-U_{\min} / k_{\text{Б}} T) / \sqrt{U_{\min} / k_{\text{Б}} T}. \quad (4.32)$$

С учетом обозначения $U_{\min} = 2U_0 e^{-h/2l_{\text{э}}}$ получим

$$n_{\text{cp}} = N_{\text{д}} \frac{\sqrt{\pi} l_{\text{э}}}{h} \frac{\exp[-(2U_0 / k_{\text{Б}} T) \exp(-h / 2l_{\text{э}})]}{\sqrt{(2U_0 / k_{\text{Б}} T) \exp(-h / 2l_{\text{э}})}}. \quad (4.33)$$

Электроны с такой концентрацией создают надбарьерный ток за счет тепловой скорости $v_{\text{т}} = \sqrt{3k_{\text{Б}} T / m^*}$ с плотностью $j = en_{\text{cp}}v_{\text{т}} / 4$. При ширине транзистора W и толщине канала порядка длины экранирования $l_{\text{э}}$ они обеспечивают при потенциале перекрытия ток порядка 10 мА, что хорошо согла-суется 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 линеаризуются 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 $\Delta V_{\text{си}} / \Delta V_{gs}$ 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 $I_{d0} \approx 25$ 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 $1...10 \text{ В}^{-2}$. 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» 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.

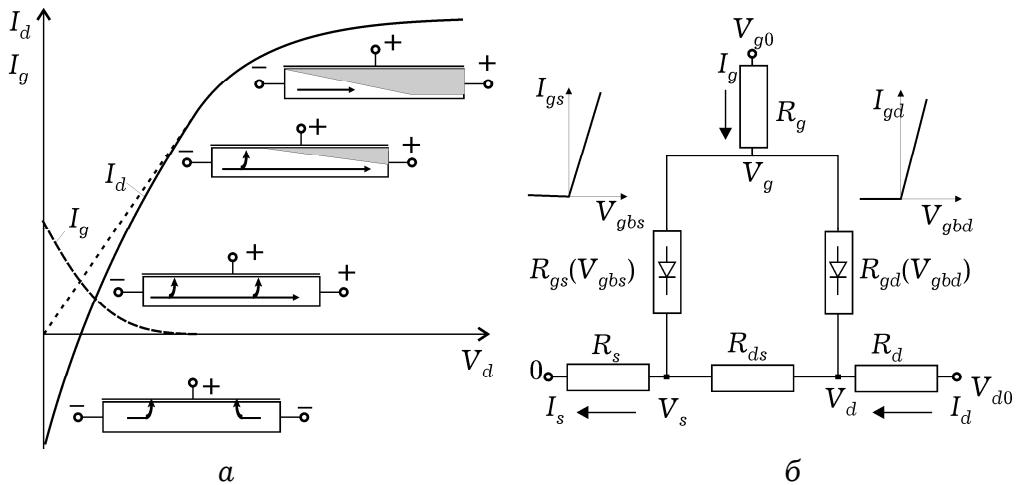


Рис. 4.26.. К расчету начального участка сток-затворной характеристики ПТШ при положительных потенциалах затвора: а – физическая модель, показывающая трансформацию обедненной области и направления токов стока и затвора в разных режимах; б – схемная модель подзатворной области при положительных потенциалах затвора;

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

$$dV = I(x)dx/e\mu Wa. \quad (4.35)$$

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:

$$dI = g[V_{gb} - V(x)]Wdx. \quad (4.36)$$

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 пара-

метр 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:

$$\frac{d^2I}{dx^2} = \alpha^2 I; \quad (4.37)$$

$$\frac{d^2V}{dx^2} = -\alpha^2 [V_{gb} - V(x)]; \quad (4.38)$$

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

$$V_{gb} - V(x) = [V_{gbs}\operatorname{sh}\alpha(L-x) + V_{gbd}\operatorname{sh}\alpha x]/\operatorname{sh}\alpha L, \quad (4.39)$$

$$I(x) = G[V_{gbs}\operatorname{ch}\alpha(L-x) + V_{gbd}\operatorname{ch}\alpha x]/\operatorname{sh}\alpha L, \quad (4.40)$$

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

$$V(x) = V(x_0) + I(x_0)(x - x_0)/en\mu Wa, \quad (4.41)$$

где x_0 — точка, в которой $V(x_0) = V_{gb}$.

Из полученных решений непосредственно следуют выражения для тока каждого из электродов

$$I_s = I(0) = G[V_{gbs}(\operatorname{ch}\alpha L - 1) + V_{ci}]/\operatorname{sh}\alpha L, \quad (4.42)$$

$$I_d = I(L) = G[V_{gs} - V_{gbd}(\operatorname{ch}\alpha L - 1)]/\operatorname{sh}\alpha L, \quad (4.43)$$

$$I_g = I_s - I_d = G(V_{gbs} + V_{gbd})(\operatorname{ch}\alpha L - 1)/\operatorname{sh}\alpha L. \quad (4.44)$$

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 $R_{gs} = R_{gd} = \operatorname{sh}\alpha L/G(\operatorname{ch}\alpha L - 1)$ и сопротивления сток-исток $R_{ds} = \operatorname{sh}\alpha L/G$ (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_{so} , m) при использовании аппроксимации ВАХ затвора в виде $i = i_s [\exp(eV/mk_B T) - 1]$ с током насыщения, равным $i_s = i_{so} \exp(-eV_b/k_B T)$, 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:

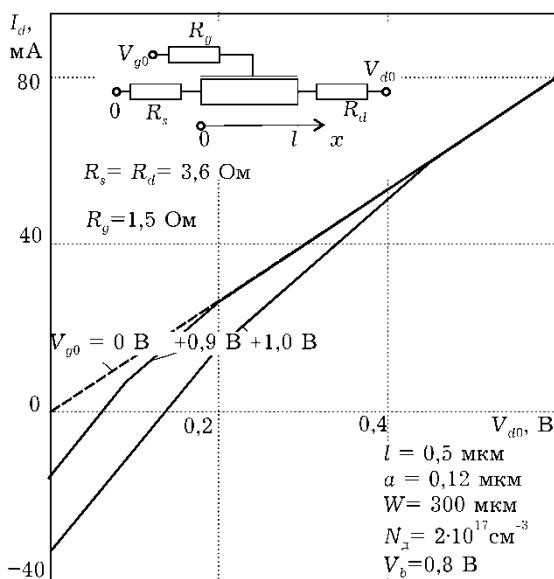


Рис. 4.27. Результаты моделирования стоковой характеристики ПТШ при положительных потенциалах затвора при кусочно-линейной аппроксимации ВАХ барьера Шоттки.

$$V_{d0} = V_d + I_d R_d, V_{g0} = V_g + I_g R_g, V_s = I_s R_s,$$

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.

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 барьера 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 практически 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 системе

макс 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 сигналов, 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 имеет вид:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{e}{\epsilon \epsilon_0} [n(x, y) - N_d(x, y)] \quad (4.45)$$

где $N_d(x, y)$ - функция, задающая профиль легирующей примеси в активной области транзистора.

Уравнение непрерывности тока запишем в виде:

$$\operatorname{div}[D \operatorname{grad} n(x) - \mu \operatorname{grad} V(x)] = 0,$$

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 приведен в гл. 2. При этом времена релаксации рассчитываются по аналитическим соотношениям для отдельных механизмов рассеяния и можно учитывать наиболее существенные из них.

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].

Больший 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 характеристики в широком диапазоне питающих напряжений. Исходная система уравнений нелокальной двумерной модели имеет следующий вид:

$$v_x \frac{\partial n}{\partial x} + v_y \frac{\partial n}{\partial y} = 0, \quad (4.46)$$

$$v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} = \frac{e\mathcal{E}_x}{m^*} - \frac{v_x}{\tau_p}, \quad (4.47)$$

$$v_x \frac{\partial v_y}{\partial x} + v_y \frac{\partial v_y}{\partial y} = \frac{e\mathcal{E}_y}{m^*} - \frac{v_y}{\tau_p}, \quad (4.48)$$

$$v_x \frac{\partial T_e}{\partial x} + v_y \frac{\partial T_e}{\partial y} = -\frac{T_e - T_0}{\tau_E} + \frac{2E_{\text{д}}}{3k_B}(\gamma - 1), \quad (4.49)$$

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{e}{\epsilon \epsilon_0} [n(x, y) - N_{\text{д}}(x, y)] \quad (4.50)$$

где дрейфовая энергия определяется как

$$E_{\text{д}} = m^* v^2 / 2 = \frac{m^*}{2} (v_x^2 + v_y^2).$$

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. 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 $5 \cdot 10^{17}$ см⁻³ и $5 \cdot 10^{15}$ 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.

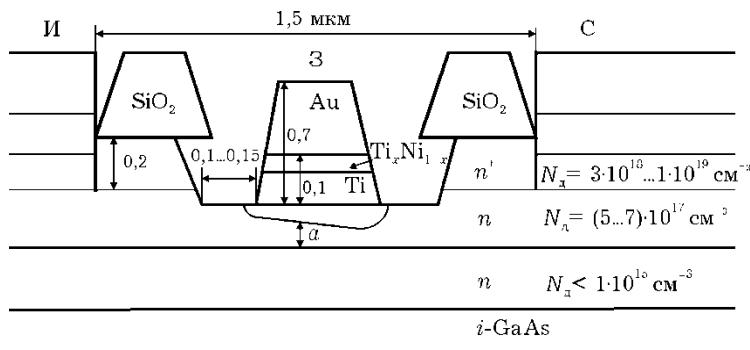


Рис. 4.28. Топология моделируемого субмикронного ПТШ

На всех ниже приведенных рисунках, где представлены двумерные распределения физических величин субмикронного ПТШ с длиной затвора 0,25 мкм, затвор располагается симметрично относительно координаты $x = 0$.

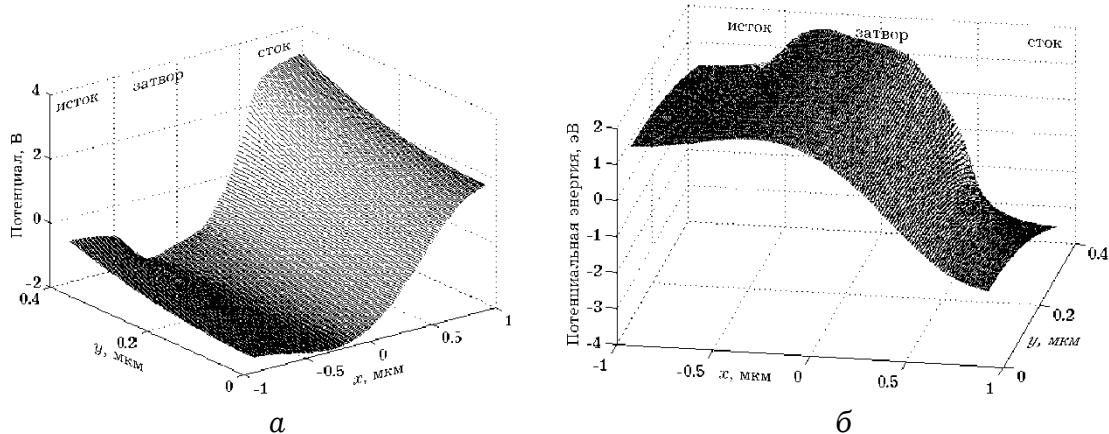


Рис. 4.29. Распределение потенциала (а) и потенциальной энергии (б) в структуре субмикронного ПТШ

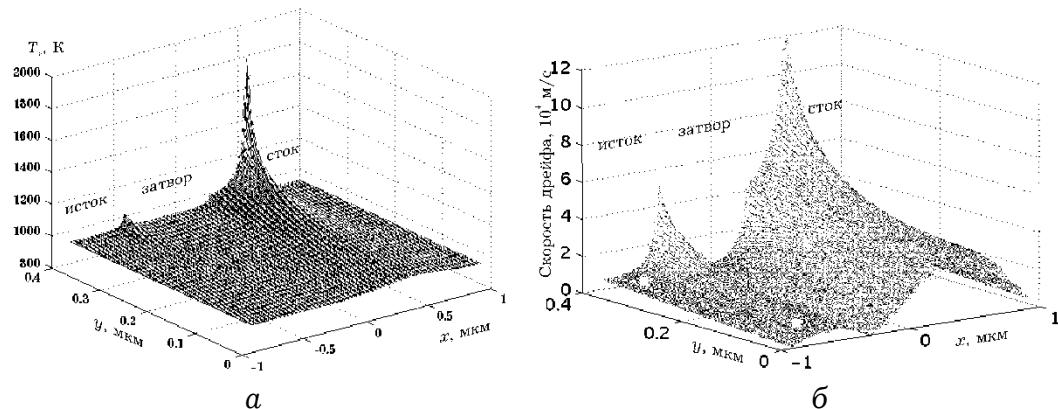


Рис. 4.30. Распределение электронной температуры (а) и продольной скорости (б) для субмикронного ПТШ с длиной затвора 0,25 мкм

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 усиительные и шумовые характеристики активных компонентов [4.30].

Для анализа распределенных эффектов в затворной линии субмикронного ПТШ рассмотрим его структуру (рис. 4.31).

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):

$$\begin{aligned} dV_c(x) &= -I_s(x) z dx; \\ dI_s(x) &= y_{11} dx V_s(x) + y_{12} dx V_c; \end{aligned} \quad (4.52)$$

$$dI_c(x) = y_{21}dxV_s(x) + y_{22}dxV_c.$$

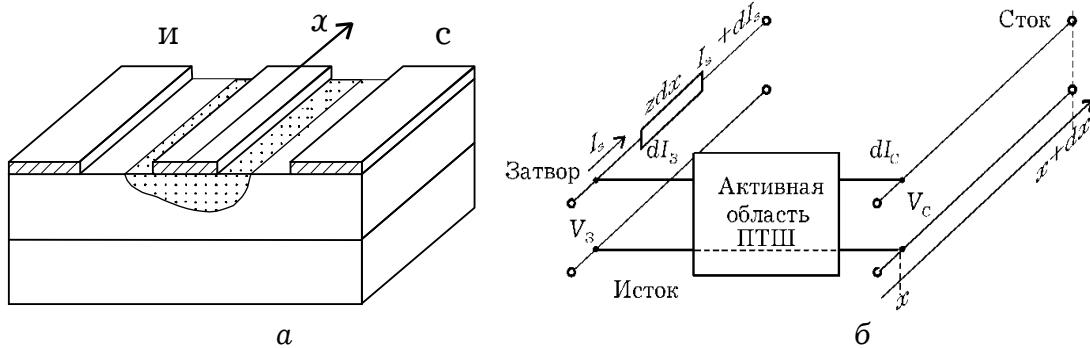


Рис. 4.31. Затворная линия транзистора: а – структура ПТШ; б – схематическая модель дифференциального отрезка линии

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 вдоль затвора и не зависит от x .

Первые два уравнения можно рассматривать как систему уравнений для определения $V_s(x)$ и $I_s(x)$. Разделение переменных достигается повторным дифференцированием и последующей перекрестной подстановкой. Обозначая $\sqrt{zy_{11}} \equiv \dot{\gamma} = \alpha + j\beta$, получим

$$\begin{cases} \frac{d^2I_s(x)}{dx^2} = \gamma^2 I_s(x) \\ \frac{d^2V_s(x)}{dx^2} = \gamma^2 V_s(x) + \gamma^2 (y_{12}/y_{11}) V_c \end{cases}.$$

Решение уравнений можно представить в форме

$$\begin{cases} I_3(x) = I_3^+ e^{-\gamma x} - I_3^- e^{\gamma x} = \frac{V_3^+}{Z_B} e^{-\gamma x} - \frac{V_3^-}{Z_B} e^{\gamma x} \\ V_3(x) = V_3^+ e^{-\gamma x} + V_3^- e^{\gamma x} - (y_{12}/y_{11}) V_c \end{cases}.$$

Подставим эти решения в одно из исходных уравнений первого порядка. Приравнивая 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 $V_3(0) = V_{bx}$ и $I_3(W) = 0$. Подставляя в уравнения предыдущей системы, получим

$$V_3^\pm = \frac{[V_{bx} + (y_{12}/y_{11}) V_c] e^{\pm \gamma W}}{\operatorname{ch} \gamma W}.$$

В итоге выражения, характеризующие распределение амплитуд напряжения и тока, можно представить в форме:

$$I_3(x) = \frac{[V_{bx} + (y_{12}/y_{11}) V_c] \operatorname{sh} \gamma(W-x)}{Z_0 \operatorname{ch} \gamma W};$$

$$V_3(x) = \frac{[V_{bx} + (y_{12}/y_{11}) V_c] \operatorname{ch} \gamma(W-x)}{\operatorname{ch} \gamma W} - \left(\frac{y_{12}}{y_{11}} \right) V_c.$$

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

$$I_{bx} = I_3(0) = [V_{bx} + (y_{12}/y_{11}) V_c] \operatorname{th} \gamma W / Z_0.$$

Подставляя $V_3(x)$ в третье уравнение исходной системы (4.52), интегрированием по всей ширине затвора получим выражение для полного тока стока

$$\begin{aligned} I_{\text{вых}} &= \int_0^W dI_c(x) = y_{21} W (\operatorname{th} \gamma W / \gamma W) V_{bx} + \\ &+ \left[y_{22} W + \frac{y_{21} \cdot y_{12}}{y_{11}} W \left(\frac{\operatorname{th} \gamma W}{\gamma W} - 1 \right) \right] V_c. \end{aligned}$$

Два последних уравнения связи токов $I_{\text{вх}}$ и $I_{\text{вых}}$ с напряжениями $U_{\text{вх}}$ и $U_{\text{вых}}$ 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:

$$\begin{aligned} Y_{11} &= Y_{11}^c F; \quad Y_{12} = Y_{12}^c F; \quad Y_{21} = Y_{21}^c F; \\ Y_{22} &= Y_{22}^c F + \frac{Y_{12}^c \cdot Y_{21}^c - Y_{11}^c \cdot Y_{22}^c}{Y_{11}^c} (F - 1). \end{aligned}$$

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 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.

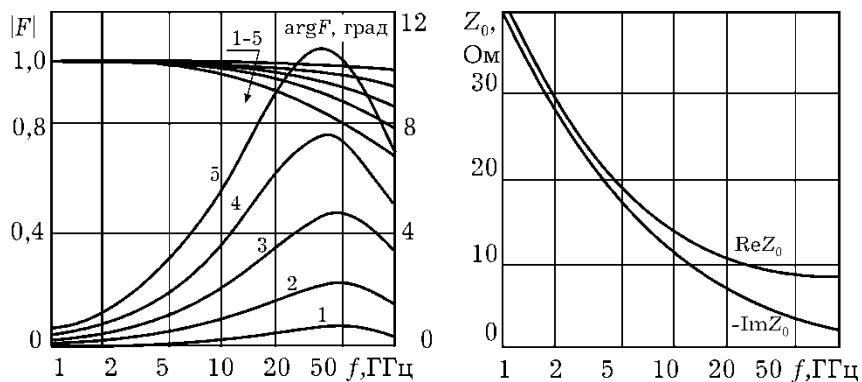


Рис. 4.32. Влияние распределенных эффектов на параметры транзистора

Рассмотренные распределённые эффекты должны влиять не только на усиительные, 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,

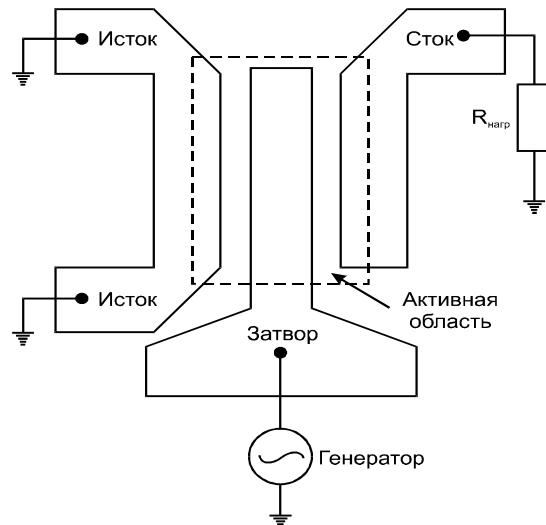


Рис. 4.33. Затворная линия передачи транзистора

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.

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.

Используя закон Кирхгофа для токов, получим систему уравнений, описывающих поведение схемы во временной области:

$$\begin{aligned} \text{узел 1: } & I_{Cgd} - I_{gd} - I_{gs} + I_{Cgs} + I_{Rg} = 0; \\ \text{узел 2: } & I_{Rd} - I_{Cds} - I_{Cgd} + I_{drain} - I_{Rn} + I_{gd} = 0; \\ \text{узел 3: } & I_{Cds} + I_{gs} - I_{Cgs} - I_{drain} + I_{Rs} = 0. \end{aligned} \quad (4.53)$$

Используя зависимость $I = CdU/dt$, запишем токи через емкости:

$$\begin{aligned} C_{gd} \frac{dU_{gd}}{dt} - I_{gd} - I_{gs} + C_{gs} \frac{dU_{gs}}{dt} + I_{Rg} &= 0; \\ I_{Rd} - C_{ds} \frac{dU_{ds}}{dt} - C_{gd} \frac{dU_{gd}}{dt} + I_{drain} - I_{Rn} + I_{gd} &= 0; \\ C_{ds} \frac{dU_{ds}}{dt} + I_{gs} - C_{gs} \frac{dU_{gs}}{dt} - I_{drain} + I_{Rs} &= 0. \end{aligned} \quad (4.54)$$

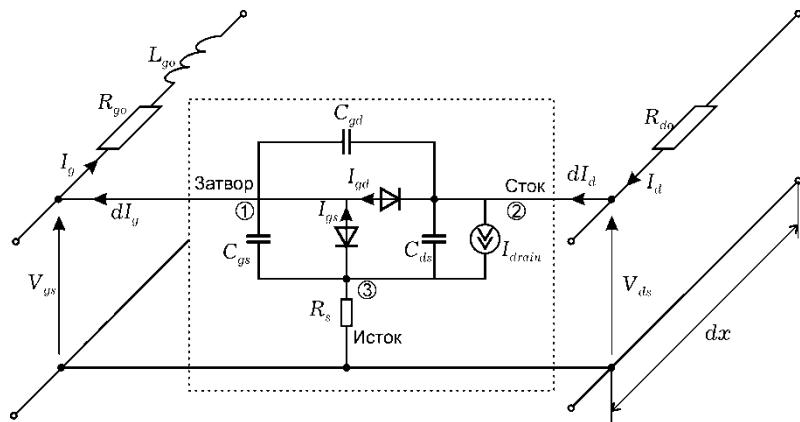


Рис. 4.34. Нелинейная модель дифференциального отрезка активной линии передачи

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:

$$\begin{aligned} \mathbf{M} \frac{d\mathbf{V}}{dt} &= \mathbf{B}, \quad \text{где} \\ \mathbf{M} &= \begin{bmatrix} C_{gs} - C_{gd} & C_{gd} & -C_{gs} \\ C_{gd} & C_{ds} - C_{gd} & -C_{ds} \\ -C_{gs} & -C_{ds} & C_{gs} + C_{ds} \end{bmatrix}, \\ \mathbf{B} &= \begin{bmatrix} I_{gd} \{V_3 - V_2\} + I_{gs} \{V_2 - V_4\} - \frac{V_2 - U_{in}(t)}{R_g} \\ \frac{V_3}{R_n} - I_{drain} \{V_2 - V_4, V_4 - V_2\} - I_{gd} \{V_3 - V_2\} - \frac{V_5 - V_3}{R_d} \\ I_{drain} \{V_2 - V_4, V_4 - V_2\} - \frac{V_4}{R_s} - I_{gs} \{V_2 - V_4\} \end{bmatrix}. \end{aligned}$$

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 и определения распределения токов и напряжений 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:

$$-\frac{\partial V}{\partial x} = Ri + L \frac{\partial i}{\partial t}, \quad -\frac{\partial i}{\partial x} = GV + C \frac{\partial V}{\partial t}.$$

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).

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:

$$Y = \sum_{i=1}^n [Y_i].$$

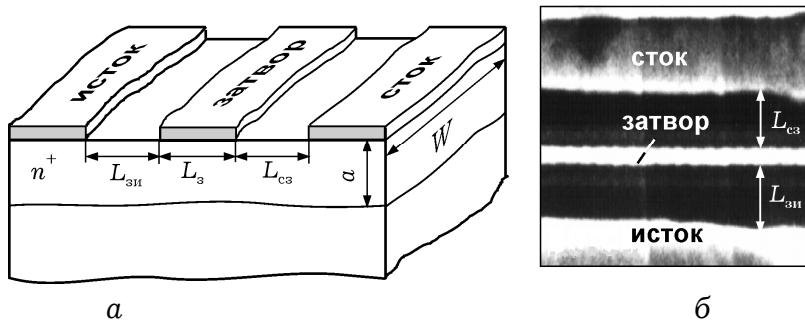


Рис. 4.35. Топология транзистора с технологическими неоднородностями (а) и фрагмент (б) фотографии на рис. 4.16.

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 монолитных интегральных схем, а также расширением рабочего частотного диапазона полевых транзисторов с барьером Шоттки в 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 изготов-

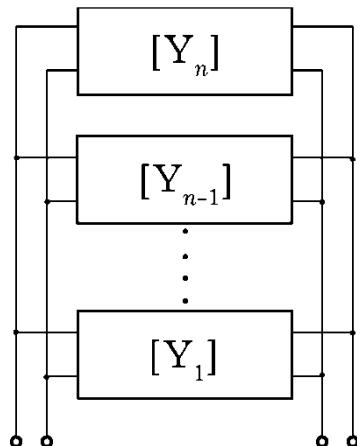


Рис. 4.36. Представление транзистора в виде параллельных четырехполюсников

ления. 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 газофазной эпитаксией металлоорганических соединений на полуизолирующей подложке с буферным слоем. Топология стока и истока в виде омических контактов и затвора в виде барьера 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 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.

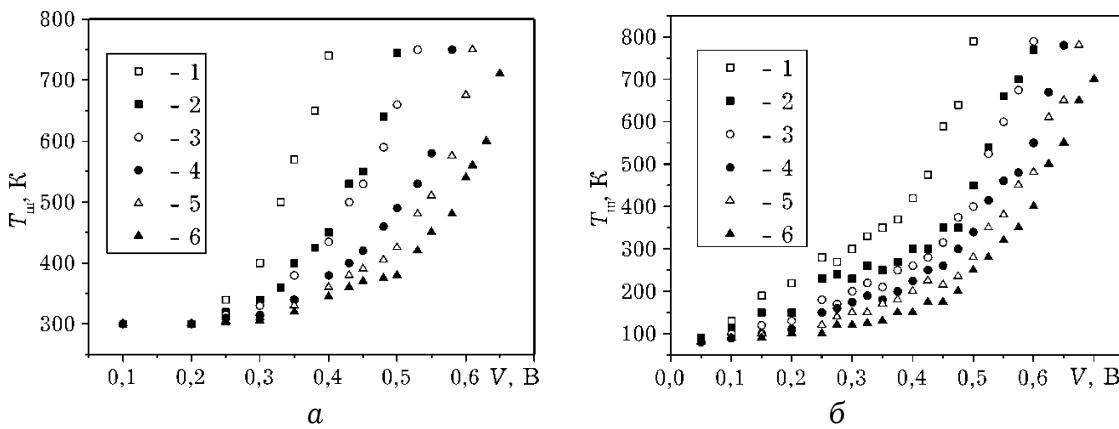


Рис. 4.37. Зависимость шумовой температуры канала от напряжения на активной области образца на частоте 10 ГГц при температуре 300 К (а) и 80 К (б): 1, 2 – $L_g = 0,2$ мкм, 3, 4 – $L_g = 0,6$ мкм, 5, 6 – $L_g = 1,0$ мкм; 1, 3, 5 – стандартная технология, 2, 4, 6 – улучшенная.

Экспериментальные зависимости, приведенные на рис. 4.37 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.

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 малосигнальной схемной 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 исследовалась структуры серийных ПТШ с параметрами: $L_{из} = 0,8 \div 3,8$ мкм – расстояние исток-затвор; $L_3 = 0,3$ 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 возрастать 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. 4.38 приведены результаты расчетов коэффициента усиления по мощности K_{yP} и температуры шума $T_{\text{ш}}$ 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 4.38, K_{yP} уменьшается на 0,8 дБ, а $T_{\text{ш}}$ 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 $T_{\text{ш}}$ 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.

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.

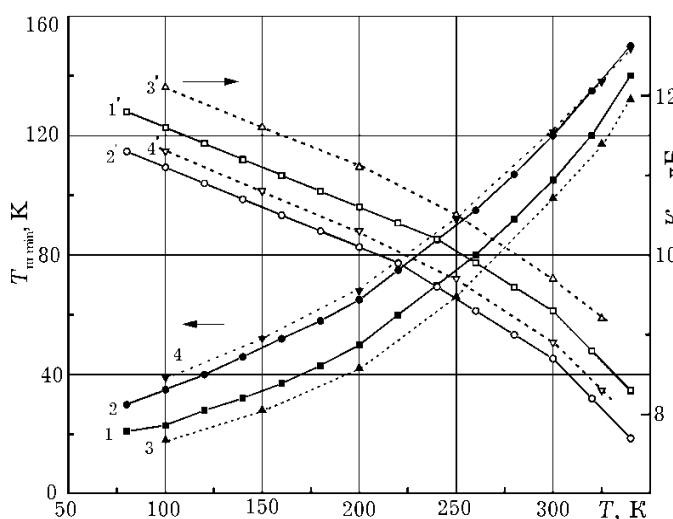


Рис. 4.38. Температурные зависимости K_{yP} и $T_{\text{ш}}$ на частоте 12 ГГц: 1, 1' – ПТШ с улучшенной морфологией омических контактов; 2, 2' – серийный ПТШ, изготовленный по серийной технологии; 3, 3' – рассчитанные параметры модели ПТШ без учета неоднородностей; 4, 4' – с учетом неоднородностей.

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_{yP} при этом увеличивается примерно на 2,5 дБ. Аналогичный характер температурных зависимостей K_{yP} 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 разделу 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.Rauser, 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. -C.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 Compu-

tation 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.–C.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.–C.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 малошумящего 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"– Київ.–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."–Київ.–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.–Київ.–1990.–С.18–19.
- 4.26. *Тимофеев В.И.* Идентификация параметров субмикронных ПТШ для сверхбыстро действующих интегральных схем // Труды Междунар. семинара "Теория идентификации

нелинейных динамических объектов".— г. Торговище (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. *Іващук А.В., Тимофеев В.І.* Вплив морфології омічних контактів на надвисокочастотні параметри польових транзи-

торів // Оптоелектроника и полупроводниковая техника.– 2000.– №35.– С.157–161.

4.37. Іващук А.В., Матуленіс А., Тимофієєв В.І. Вплив морфології омічних контактів на шуми гарячих електронів в транзисторних структурах // Электроника и связь.– 1999.– №7.– С.96–99.

4.38. Москалюк В.А., Тимофієев В.І. Секціонированная модель полевого СВЧ транзистора, учитывающая распределенные эффекты // Автоматизация проектирования в электронике.– 1986.– Вып.33.– С.93–97.

4.39. Тимофієев В.І., Андріец В.В. Шумы полевого СВЧ транзистора с топологическими неоднородностями // Вестник Київського політехнічного інститута. Радіоелектроніка.– 1987.– Вып.23.– С.1671–1708.